TESIS DE LA UNIVERSIDAD
DE ZARAGOZA

2024 106

Miguel Lafuente Blasco

Probabilistic and Statistical properties of delta-record observations

Director/es
Sanz Sáiz, Gerardo
López Lorente, Francisco Javier

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

Universidad de Zaragoza
Servicio de Publicaciones

ISSN 2254-7606

Universidad Zaragoza
1542

Tesis Doctoral

PROBABILISTIC AND STATISTICAL PROPERTIES OF DELTA-RECORD OBSERVATIONS

Autor

Miguel Lafuente Blasco

Director/es

Sanz Sáiz, Gerardo
López Lorente, Francisco Javier

UNIVERSIDAD DE ZARAGOZA
Escuela de Doctorado

Programa de Doctorado en Matemáticas y Estadística

2022

Universidad de Zaragoza

Doctoral Thesis

Probabilistic and Statistical properties of δ-record observations

Thesis submitted to the Faculty of Sciences of the University of Zaragoza for the degree of Doctor in Mathematics and Statistics by:
Miguel Lafuente Blasco

Supervisors:
Dr. F. Javier López Lorente
Dr. Gerardo Sanz Sáiz

A mis padres y a mi gente...

"Quan surts per fer el viatge cap a Ítaca,
has de pregar que el camí sigui llarg,
ple d'aventures, ple de coneixences.
Has de pregar que el camí sigui llarg,
que siguin moltes les matinades
que entraràs en un port que els teus ulls ignoraven,
i vagis a ciutats per aprendre dels que saben.
Tingues sempre al cor la idea d'Ítaca.
Has d'arribar-hi, és el teu destí,
però no forcis gens la travessia.
És preferible que duri molts anys,
que siguis vell quan fondegis l'illa,
ric de tot el que hauràs guanyat fent el camí,
sense esperar que et doni més riqueses.
Ítaca t'ha donat el bell viatge,
sense ella no hauries sortit.
I si la trobes pobra, no és que Ítaca
t'hagi enganyat. Savi, com bé t'has fet,
sabràs el que volen dir les Ítaques."
Lluís Llach, inspired by Konstantinos Kavafis

Acknowledgements / Agradecimientos

El principal agradecimiento que merece esta memoria no puede ser a otras personas que a Gerardo Sanz Sáiz y a Francisco Javier López Lorente.

Gerardo fue en su momento mi principal valedor allá por el 2013, cuando yo aún era un alumno de grado haciendo una estancia en la Universidad de Pau. En todos estos años, he tenido la inmensa suerte de poder compartir con él esta carrera que es mi formación, pero sobretodo una amistad muy profunda. Gerardo es la persona que ha estado en todos los momentos para mi; en los buenos, y muy especialmente en los malos. Todos estos años he encontrado en él a una persona intachable, leal y muy honesta. Mi cariño y aprecio por él me llevarían a escribir varias tesis doctorales acerca de sus virtudes como persona. Entender los inicios de mi carrera, es entender que un día Gerardo depositó en mi su confianza, y jamás me la ha retirado, al contrario, siempre ha redoblado esa confianza. Cada día que comparto con él en la Universidad es un regalo para mi. Y yo, consciente de que dentro de no mucho Gerardo tendrá su merecido descanso en forma de una fantástica jubilación, quiero expresar que Gerardo nunca se irá del todo de la Universidad mientras yo esté en ella. Es mi empeño que las enseñanzas y el buen hacer de un maestro de semejante categoría pervivan. Tratar a las personas y en particular a los estudiantes de la manera más humana y honesta, sacrificarse por sus compañeros y sacar adelante al grupo demostrando que liderar es servir, son el ejemplo que deja esta gran persona. Ese es mi compromiso ahora que empiezo otra etapa, transmitir los valores que yo he vivido contigo, Gerardo, a las nuevas generaciones. Nunca podré agradecer lo suficiente que te hayas cruzado en mi vida.

El agradecimiento que le tengo a Javier es también de dimensiones catedralicias. Para entender la situación, utilizaré estas lineas para hacer una confesión. Tengo que reconocer que en los inicios de mi tesis doctoral me costaba visitar a Javier a su despacho. La razón no era otra que Javier es, estoy convencido de ello, la persona más brillante del edificio de Matemáticas. Y sí, ir a su despacho me costaba porque la cantidad y calidad de las ideas que tenía Javier me abrumaban, me veía desbordado ante semejante derroche de capacidad. Y yo, que aún era más joven que ahora, tenía un inocente e infundado temor a que él pensara que yo no había trabajado lo suficiente. Pero es que lo mejor de todo es que su calidad humana está por lo menos al nivel de su maravillosa capacidad que acabo de describir. En realidad no necesité mucho tiempo para darme cuenta de que era también una persona excepcional. Siempre dispuesto a ayudar, siempre amable y voluntarioso. Estoy muy orgulloso de trabajar con Javier, y disfruto cada momento de ello. Sin embargo, no quiero escribir de momento más lineas acerca de él, ya que estoy convencido de que las lineas que más merezcan la pena aún están por venir. Nada me gustaría más que continuar trabajando con él hasta que él quiera, seguir investigando, aprendiendo, admirándole y compartiendo parte de la vida con él.

También quiero agradecer muy sinceramente al Profesor Raúl Gouet. He tenido la inmensa fortuna de hacer una larga estancia doctoral en el Centro de Modelamiento Matemático de la Universidad de Chile, y de visitarlo en otras ocasiones. Parece que la buena gente se atrae entre sí, y por eso, a través de Gerardo y Javier, yo caí en las manos de Raúl. Raúl es un matemático excepcional, potente en sus ideas, con una exquisita redacción, y sobretodo, un conocimiento muy vasto, casi bibliográfico, de Matemáticas. Allá en Chile, y aquí en España cuando ha venido a visitarnos, no solamente tuve a un profesor excepcional, también me abrió la puertas de su casa, y pude conocer a su bella familia.

Como no puede ser de otra manera, no puedo sino acordarme de mis compañeros del Departamento de Métodos Estadísticos de la Universidad de Zaragoza, y en especial, a todo el Grupo de Investigación de Modelos Estocásticos. Estoy muy orgulloso de poder decir que jamás he tenido problema alguno con nadie en el Departamento, y que desde el primer día que entré lo único que encontré fue respeto hacia mi persona y mi trabajo. Siempre me he sentido arropado por todos ellos. En especial quiero citar a Jesús Asín, con el que compartí despacho durante años y que siempre está velando por mi, y a Carmen Sangüesa, con la que he tenido el inmenso honor de compartir la docencia durante tres años, en los cuales he aprendido mucho de ella. Además, la naturaleza multidisciplinar de la estadística me ha llevado a colaborar con otras personas del Departamento o el grupo de investigación; como Ana Carmen Cebrián, Pedro Mateo, Ana Pérez, Luis Mariano Esteban...

Sin embargo, para hacer una tesis, no todo son los compañeros académicos. A mis amigos; Diego y Cheve por estar siempre cerca y preocupándose por mi, a Antonio por ser incluso parte de mi aún estando últimamente lejos, a Jorge por aceptarme y apreciarme tal como soy. También a Kristen, por, estando tan lejos, hacerme sentir que está tan cerca y darme otra visión del mundo.

Sería incapaz de cerrar una lista de agradecimientos sino cito a mi Inma de Segur, una de las más bellas personas que alguien puede encontrar en este mundo y que me ha tratado siempre como a un hijo. Mi vida siempre irá ligada a los largos veranos con ella y su familia, muy en especial a su brillante hijo Javi, que sigue haciendo que el verano sea la mejor época del año.

También quiero hacer constar el aprecio a los profesores que he tenido a lo largo de mi vida, y que han ido moldeando mis conocimientos hasta el día de hoy. Especial *cariño a profesores y compañeros de clase de mi infancia en Calatayud. Hoy no soy más que un bilbilitano que se forjó entre ellos.*

Y como no puede ser de otra manera tengo que acordarme de mi familia.

Mi abuelo José, que siempre me ha querido. Suelo contar a la gente la anécdota de como de niño recitaba las cotizaciones de las acciones porque me las aprendía al mirar el teletexto contigo. He tenido la oportunidad de disfrutarte como abuelo y lo seguiremos haciendo.

Siguiendo con la familia, durante todos estos años he tenido la suerte de compartir la vida con Irene. A ella esta tesis doctoral le ha robado también mucho tiempo de estar conmigo. Además, algunos de estos años han sido difíciles en cuanto a que hemos estado lejos por estancias muchas veces, he vivido una enfermedad... y ella siempre ha estado a mi lado. Aún aguantando mis rarezas, que supongo que ella dará fe de que son muchas. Solo espero que sea por siempre mi compañera de vida.

En mi familia no puedo olvidarme de mis perros. Para comprender la importancia que tienen los perros en mi vida, no hay más que saber cual fue mi primera palabra de niño: "La Bola", el nombre de mi primera perra. Yo vine al mundo y ya estaba ella aquí. Fue mi hermana mayor, siempre pendiente de mi y protegiéndome. Luego vino Zara, a la que afortunadamente salvamos de una muerte segura y que nos devolvió todo el cariño con creces. Su pérdida fue tan dura que tardamos 8 años en reponernos. Ahora mis galgos, Llamp y Lluvia, son los que nos enseñan como vivir la vida y disfrutarla. Mi corazón está formado por cachitos de alma de perro.

Y reservo el lugar de honor a mis padres. Me dieron la vida y me la siguen dando. He llegado aquí gracias a ellos. La gente dice que soy responsable, pero es gracias a ellos. Mi padre, que es poco hablador y una persona sencilla, comparte conmigo sus pasiones, que es lo más importante que tiene, correr y el Barça. Yo sé que es su forma de disfrutar de la vida, y yo he tenido la fortuna de heredarlas de él. Hemos compartido muchos kilómetros y partidos de fútbol y yo eso lo disfruto mucho. Otras virtudes que pasar a la siguiente generación. Mi madre es mi madre, y con eso podría decir todo. Ella me ha criado, enseñado, y dado un cariño infinito. Sé que vive mis éxitos y fracasos como si fueran suyos, y eso es lo que me lleva a superarme siempre y a alcanzar las metas. Si alguien se pregunta si puede existir una persona carente de toda maldad la respuesta es que sí, y ella es la prueba viviente de ello. La presentación de esta tesis será para ella un momento culminante, y por eso he puesto en esta memoria todo mi empeño. Además, sé que ella no me cree, pero yo siempre voy pregonando a los cuatro vientos que ella es en realidad la persona más brillante de mi familia cuando la gente me señala a mi. Pero no te preocupes mamá, tu potencial se ha desarrollado criándome, y eso es lo que más agradezco de la vida.

En Segur de Calafell, año 2021.

Abstract

Records, defined as observations that exceed all previous observations, are ubiquitous in modern everyday life. They have also attracted much research and attention, due to their intrinsic interest and the mathematical challenges they pose.

In 1952, a seminal paper by Chandler launched what has now grown to become a rich body of literature on the mathematical properties of record observations. The study of the probabilistic properties of records has attained an important degree of maturity and it is therefore natural that significant effort has been devoted to statistical inference with records over recent decades.

The classical probabilistic setting of records in independent and identically distributed (i.i.d.) continuous random variables (r.v.), reflects the scarcity of this kind of observations. Indeed, for sequences of i.i.d. continuous r.v., it is known that the probability that the n-th observation is a record is $1/n$, and the expected number of records is of the order of the logarithm of n, where n is the number of observations. Note however that this universal property is lost when the underlying r.v. are discrete.

The connection of records with many interesting problems led to a considerable interest in the study of record observations, especially from the perspective of physics. Records have proved their worth in many areas such as athletics, risk theory, financial modeling and evolutionary biology. One of the main fields of application is climatology, where the i.i.d. model fails to predict the number of high-temperature records, with these observations being significantly higher than expected.

In this thesis, we are going to consider two distinct generalizations related to the study of usual records with the aim of enabling a greater number of problems to be addressed.

The first concerns the mathematical definition of a record. In this monograph we focus on two of the record-related concepts that have been most studied – near-records introduced in 2005 by Balakrishnan et al. [7] and δ-records proposed by Gouet et al. in 2007 [50]. Given a sequence of observations (X_n), we say that the n-th observation is a δ-record if $X_n > M_{n-1} + \delta$, where M_{n-1} is the maximum among

the first $n-1$ observations, and δ is a real parameter. If $\delta = 0$ records and δ-records are equivalent, while in the case $\delta < 0$ ($\delta > 0$), δ-records are more (less) frequent than records.

An observation is considered to be a near-record if it is not a record but is at a distance of less than a units from the last record. Consequently, the study of δ-records and near-records is closely related, and obtaining properties of one of these notions generally results in obtaining properties of the other.

The other kind of generalization that we consider concerns the model of the underlying variables. Adding a deterministic linear trend to the observations we obtain what is known as a Linear Drift Model (LDM), first introduced by Ballerini and Resnick [8], and studied later by other authors. The LDM has proven particularly useful in the study of global warming to explain the actual number of upper records observed.

In this monograph we address some open problems for near-records and δ-records. Chapter 1 presents the known properties of records and δ-records, as well as establishing the notation that will be used later.

In Chapter 2 we study the point process of near-record values when the observations are discrete, taking values in the integers. This problem was already studied in [55] for continuous distributions. While in the discrete setting the resulting process is also a cluster process, it is no longer a Poisson process, which makes the study of the point process and its characterization more difficult. Laws of large numbers and central limit theorems for the number of near-records with a value in a set are also obtained. Finally, we characterize which discrete distributions fulfill a martingale condition relating the partial maxima and the number of δ-records at time n. We relate this characterization to the open problem of the positivity of the terms in a recurrence relation.

The LDM is studied extensively in Chapter 3 for δ-records. From the basic properties and derivations of the probability of δ-record, we study the asymptotic δ-record probability and its analytic properties. We derive exact expressions for some distributions, some of them also unknown in the case of usual records, and we use these results to assess the effect of the δ parameter when the underlying variables are heavy-tailed. For distributions where an analytic expression is not available, we propose first order approximations to study δ-record probabilities. We also compute the correlation of δ-record observations as a function of the number of observations and the δ parameter. We study the asymptotics of the counting process of δ-records in the LDM. The finiteness of the number of δ-records in the LDM is completely characterized. In particular, this result solves a conjecture posed by Franke et al. [33] for records, proving the result not only for usual records but also in the general setting with $\delta \neq 0$. Finally, we obtain laws of large numbers and a central limit theorem under mild conditions, extending the results in [8] to the case $\delta \neq 0$, and we prove a law of large numbers for a random trend model.

In Chapter 4 we develop statistical inference methods for δ-records in the LDM. We propose two estimators for the variance of the number of δ-records and discuss their properties, proving consistency. We also study Maximum Likelihood Estimation based on δ-records in the LDM. We develop a general framework for Maximum Likelihood Estimation and we find analytic solutions for particular cases. We use Montecarlo simulation to compare the performance of the Maximum Likelihood Estimators using δ-records with those using records only. Finally, the results in Chapter 3 and in this chapter are applied to a real dataset of temperatures, where the LDM is consistent with the findings of other authors and the phenomenon of global warming. In particular, we find good agreement between the theoretical results and the data observed in the example.

Finally, in Chapter 5, we set out some conclusions of the results reached in previous chapters and offer some ideas for future work.

Resumen

Los récords, definidos como observaciones que exceden a todas las anteriores, son un concepto omnipresente en la vida cotidiana. Es por ello que se han dedicado grandes esfuerzos a la investigación de las propiedades de los mismos, tanto debido a su interés intrínseco como a los desafíos matemáticos que este concepto conlleva.

En 1952, Chandler publicó el artículo seminal sobre récords que inició lo que hoy se ha convertido en una rica literatura de resultados y propiedades de estas observaciones. El estudio de las propiedades probabilísticas de los récords goza hoy en día de un importante punto de madurez, que consecuentemente ha derivado asimismo en numerosas aplicaciones estadísticas en las últimas décadas.

El marco clásico de récords en sucesiones de variables aleatorias (v.a.) independientes e idénticamente distribuidas (i.i.d.), refleja la escasez de este tipo de observaciones. En efecto, para v.a. continuas e i.i.d., se conoce que la probabilidad de que la n-ésima observación sea un récord es $1/n$, y el número esperado de los mismos es del orden del logaritmo de n, donde n es el número de observaciones. Sin embargo, esta propiedad universal se pierde en el caso en el que las v.a. son discretas.

La conexión de los récords con una amplia variedad de problemas interesantes ha traído consigo un considerable interés del estudio de estas observaciones, especialmente desde el punto de vista de la física. Los récords se han demostrado útiles en muchas áreas, como el estudio de las marcas en atletismo, ciencias actuariales, aplicaciones financieras o biología evolutiva. Uno de los principales campos de estudio donde los récords juegan un papel crucial es el de la climatología, donde se ha observado que el modelo i.i.d. no logra explicar el número de récords que se observan en la realidad, siendo este número significativamente mayor que el esperado.

En esta tesis, con el objetivo de poder abordar una cantidad de aplicaciones más numerosa, vamos a considerar dos generalizaciones de naturaleza muy distinta respecto al estudio clásico de récords.

La primera de ellas tiene que ver directamente con la definición de récord. En esta monografía nos centramos en dos de los conceptos relacionados con los récords más

estudiados, los near-records (récords cercanos en castellano) introducidos en 2005 por Balakrishnan et al. [7], y los δ-récords, propuestos en 2007 por Gouet et al. [50]. Dada una sucesión de observaciones (X_n), decimos que la n-ésima observacion es un δ-récord si $X_n > M_{n-1} + \delta$, donde M_{n-1} es el máximo de las primeras $n-1$ observaciones, y δ es un parámetro real. Si $\delta = 0$, los récords y los δ-récords son equivalentes, mientras que en el caso $\delta < 0$ ($\delta > 0$), los δ-récords son más (menos) frecuentes que los récords.

Una observación se considera near-record si no es un récord pero está a menos de una distancia de a unidades de serlo. En consecuencia, el estudio de los δ-récords y los near-records está íntimamente relacionado, y obtener propiedades para unos se corresponde habitualmente con obtener propiedades para los otros.

El otro tipo de generalización que vamos a considerar en esta memoria tiene que ver con el modelo subyacente de la variables. Añadiendo una tendencia lineal determinista a la sucesión de observaciones obtenemos lo que se conoce como Modelo con Tendencia Lineal (o LDM por sus siglas en inglés). Este modelo, que fue primero introducido por Ballerini y Resnick [8], y posteriormente desarrollado por otros autores, se ha demostrado particularmente útil en el estudio del calentamiento global para explicar el número de récords de temperaturas que se ha observado en la realidad.

En esta monografía abordamos algunos problemas abiertos para near-records y δ-récords. En el Capítulo 1 se presentan propiedades conocidas para los récords y δ-récords, así como se establece la notación que se utilizará posteriormente.

En el Capítulo 2 se estudia el proceso puntual de valores near-record cuando las observaciones son discretas tomando valores en los números enteros. Este problema ya ha sido estudiado en [55] para distribuciones continuas. Sin embargo, mientras que en el caso discreto el proceso resultante también es un proceso de tipo cluster, en el marco discreto este proceso no es de tipo Poisson, lo que dificulta el estudio y la caracterización de dicho proceso puntual. A partir de dicha caracterización, se obtienen leyes de grandes números y teoremas centrales del límite para el número de near-records con valores en un conjunto dado. Finalmente, también se caracterizan las distribuciones discretas para la cuales su distribución cumple una condición de tipo martingala que relaciona su máximo parcial y el número de δ-récords en el instante n. Para ello, relacionamos este problema con el de garantizar la positividad de las soluciones de las ecuaciones lineales de recurrencia.

En el Capítulo 3 se estudia detalladamente el modelo con tendencia lineal (LDM). Empezando por la obtención de expresiones para las probabilidades de δ-récord y sus propiedades básicas, y continuando con su estudio asintótico y propiedades analíticas. Se obtienen expresiones explícitas para las probabilidades de δ-récord en distintas distribuciones, algunas de ellas desconocidas en la literatura también en el caso de récords usuales, y se utilizan estos resultados para evaluar el efecto del parámetro δ cuando las variables subyacentes tienen colas pesadas. En el caso en el que no se pueda obtener una expresión explícita de estas probabilidades se proponen aproximaciones de primer orden. Además, se calculan las correlaciones entre la ocurrencia de observaciones δ-récord en función del parámetro δ. En el estudio de las propiedades asintóticas del proceso de conteo de δ-récords en el LDM, se obtienen resultados como la caracterización de la finitud del número total de δ-récords. En particular, este resultado resuelve una conjetura planteada por Franke et al. [33] para récords, demostrando el resultado no solamente para los récords usuales sino para el marco general en el que se tiene $\delta \neq 0$. Finalmente, se obtienen leyes de grandes números y un teorema central del límite bajo ciertas condiciones débiles, extendiendo los resultados de [8] al caso $\delta \neq 0$, y se demuestra también una ley de grandes números para un modelo con tendencia aleatoria.

En el Capítulo 4 se desarrollan métodos de inferencia para los δ-récords en el LDM. Se proponen dos estimadores para la varianza del número de δ-récords y se discuten algunas de sus propiedades, demostrando la consistencia de los mismos. Se estudia asimismo la Estimación Máximo Verosímil basada en δ-récords para el LDM, desarrollando un marco general para dicho tipo de estimaciones y encontrando soluciones analíticas en casos particulares. Mediante métodos de tipo Montecarlo se compara el desempeño de los estimadores máximo-verosímiles usando δ-récords con aquellos basados solo en récords. Por último, los resultados del Capítulo 3 y de este mismo, se aplican sobre un conjunto de datos reales de temperaturas, para los cuales el LDM es adecuado en virtud de los hallazgos de otros autores en el marco del calentamiento global. En particular, se concluye que los resultados teóricos y los datos analizados en el ejemplo son consistentes entre sí, y la estimación máximo-verosímil se revela como una herramienta útil en este tipo de problemas.

Para finalizar, en el Capítulo 5 se exponen algunas conclusiones derivadas de los resultados obtenidos en capítulos anteriores, y se plantean ideas para un trabajo futuro.

Contents

List of Tables

List of Figures

Abbreviations

i.i.d.	Independent and identically distributed
r.v.	Random variable
cdf	Cumulative distribution function
pdf	Probability density function
p.g.f.	Probability generating function
p.g.fl.	Probability generating functional
rhs	Right-hand side
lhs	Left-hand side
CRM	Classical Record Model
LDM	Linear Drift Model
MLE	Maximum Likelihood Estimator

"Salid y disfrutad".
Johan Cruyff

Introduction

The Greeks and Romans believed that memory resided in the human heart. As a result, the action of preserving the memory of extraordinary events of all kinds was called *recordor* in Latin, from *re-*, meaning again, and *-cor*, from *cordis*, meaning heart.

Today the term 'record' is ubiquitous. Outstanding achievements and world records in athletics events such as the 100-metre sprint always make the headlines and arise widespread admiration. Similarly, considerable media attention and public concern attaches to record figures (often bad) relating to the economy, the weather or healthcare systems. Crucial social questions arise when we are faced with a steady flow of records, which are presented as ominous signs of dramatic underlying phenomena. It is therefore unsurprising that the term 'record' has become such a constant in our modern everyday life and in a wide range of specialist domains.

Since the dawn of the study of statistics and probability, extreme values and records have attracted much research and attention, due to their intrinsic interest and the mathematical challenges they pose. In 1952, a seminal paper by Chandler [17] launched what has now grown to become a rich body of literature on the mathematical properties of record observations and related concepts. An important motivation for studying records is their connection with other interesting problems and, of course, their countless practical applications in different fields. The classical probabilistic setting of independent and identically distributed (i.i.d.) random observations has been widely studied for decades. The chief results in this framework can be found in dedicated monographs [1, 2, 4, 87]. In recent years, there has been considerable interest in the study of record observations, especially from the perspective of physics.

Early on in research in this field, Foster and Stuart in 1954 [30] and Renyi in 1962 [98] made two key findings reflecting the intrinsic scarcity of records for i.i.d. continuous random variables (r.v.). Using a simple argument that is now known as *stick-shuffling* (i.e. scrambling the observations), Foster and Stuart proved that the probability that the n-th observation is a record is $1/n$. Renyi, in turn, demonstrated that in this context, record occurrences are independent and, furthermore, the number of records is, in the almost sure sense, of the order of the logarithm of the number of observations. Note however that this universal property is lost when the underlying random variables are discrete, where the behaviour in this case depends critically on the tail of the distribution [47, 48].

The study of the probabilistic properties of records has now attained an important degree of maturity and it is therefore natural that significant research effort has been devoted to statistical inference with records over recent decades. Statistical inference based on records is a difficult problem, precisely because of the scarceness of observations of this kind, meaning that samples are small.

Amongst the most influential contributions that laid the foundations of this field were the early works of Foster and Stuart [30] and Foster and Teichroew [31], who proposed and assessed test statistics to detect underlying trends in the data. Some decades later, Samaniego and Whitaker were pioneers in considering maximum likelihood estimation based on records [100], and also in providing non-parametric methods [101]. Feuerverger and Hall [28] made important progress by including information on record times in the estimation process. In 1993, Carlin and Gelfand [15] presented a general framework for Bayesian estimation for record-breaking data, allowing dependence between observations. The work by Gulati and Padgett [59] also marked an important step in opening new lines of research for different models.

The study of a phenomenon based on information from its records corresponds in many cases to the nature of the problem itself, that is, where data is inherently composed of record observations [22, 59, 65, 117]. For example, "stress-testing" consists of assessing the resistance or reliability of a material under stress situations. Glick [37] studied this problem, relating it directly to the problem of inference in records. Indeed, it is natural in this situation to monitor the minimum stimulus needed to break a material. Note that following an initial failure, the next observation will only be collected if the material has broken due to a stimulus of less magnitude than the first entry, and therefore, a (lower) record has been observed. This is performed sequentially, thus obtaining a sequence of record observations from which the inference must be carried out.

There are numerous applications of records in the context of physics. For example, in the field of evolutionary biology, studies have been made of the way in which a mutation spreads in a population from the perspective of records [32, 118, 119] and the timings of those events [66, 73, 74, 107, 109]. A mutation that is advantageous for the survival of the individuals can be considered to have better fitness than the previous one, and it therefore constitutes a record, yielding to a sequence of increasing genotypes arranged by fitness. As argued in [72], the theory of records is a useful tool for studying hard-to-quantify notions such as genotypical fitness, because of the free-distribution properties of records.

Records have also proved useful in risk theory [27], financial data and stock prices [121, 122] and other problems in physics such as their application to the theory of spin-glasses [106, 108], high-temperature superconductors [88], road traffic flow [60], and even in the study of complex systems such as population evolution [91], or the problem of division of labour among social animals [96], where record dynamics have been shown to explain certain features of task-allocation among ants.

One area in which records are especially significant is sport. Historical top scores live on in popular memory, and when they are broken, they become the stuff of headlines and popular conversations. It is therefore unsurprising that the study of records in athletics is another area of particular interest, from the perspective of both modelling and prediction [8, 9, 15, 25, 35, 36, 110, 111], and even in the detection of performing-enhancing drugs among athletes [97].

Records have proved their worth in climatology, one of the most important fields of application given the vast implications of extreme events for human society. Indeed, global warming is one of the greatest challenges faced by humanity this century, and there is a patent need to study climate records.

However, long before the scientific community detected evidence of global warming, the study of records was used precisely as an argument to refute an increase in temperatures. In Glick's 1978 influential article [37], he says that at a meeting of the Royal Statistical Society 25 years before, Foster and Stuart noted that records were more frequent in athletics than in rainfall data. Glick argues that "this is not surprising" given that athletic training has improved over the last century whereas "no one has done much about the weather". He also added that "weather fluctuations over a century are more intuitively random, without a dramatic linear trend".

Currently, many climate studies focus on verifying global warming and climate change. Numerous studies have observed that the i.i.d. model fails to predict the number of high-temperature records, with these observations being found to be significantly higher than expected, revealing a warming effect [11, 18, 84, 93, 94, 124, 125]. As a result, different methods have been developed to detect the non-stationarity of observations based on record occurrences [12, 16, 23]. Nevertheless, the sole aim of climate assessment is not the study of temperatures, and a variety of other subareas of climatology have been studied using extreme and record events. These include the study of precipitation —ranging from simple rainfall models [82] to more complex projections of future extremes [99]; detecting patterns in the occurrence of storms [68], and analyses of hurricanes intensity [58] among many other applications.

Another challenge that we will face as a global society in coming decades has been brought abruptly into the spotlight by the COVID-19 pandemic, which has clearly shown the need for mathematical techniques for early detection of epidemic outbreaks. It is evident that the theory of records can help to address such situations, particularly the collapse of health resources due to high infection rates. For this reason, although epidemiology is not a field that has been extensively studied from the perspective of records, some previous results already existed for the detection of epidemic outbreaks based on records [69], and these results have been applied to the COVID-19 pandemic [70].

In this thesis, we are going to consider two distinct generalizations related to the study of usual records with the aim of enabling a greater number of problems to be addressed.

The first concerns the mathematical definition of a record. Over recent years, different record-related concepts have appeared in the literature as a means of extending the practical applications. Some important notions related to usual records are geometric records [26, 54, 80], records with confirmation [86], δ-exceedance records [6, 92] and weak-records [113, 114, 116].

In this monograph we focus on two of the record-related concepts that have been most studied – near-records and δ-records. Near-records were introduced in 2005 by Balakrhisnan et al. [7] for applications in finance and later studied by other authors (see for instance [55, 90]). In this framework, an observation is considered to be a near-record if it is not a record but is at a distance of less than a units from the last record. That is, a near-record is not a record but it is close to being one.

In 2007, Gouet et al. [50] proposed the so-called δ-records, which merge near-records and usual records in a single mathematical object. Given a sequence of observations, we say that the n-th observation is a δ-record if it is greater than the previous record plus a fixed real quantity δ. It is easy to see that in the case that $\delta < 0$, an observation is a δ-record if it is either a near-record or a record. As a consequence, δ-records are more numerous than records. This reduces the problem of scarcity in records while still maintaining the extreme nature of records. This property has been proven to be advantageous in applications, but it also adds an extra difficulty when studying its properties since the free distribution of usual records in the i.i.d. model is lost. In the case $\delta > 0$, δ-records are less frequent than records, these two concepts being equivalent if $\delta = 0$. In particular, weak-records in discrete distributions coincide with δ-records when the parameter δ is equal to -1.

Great progress has been made in recent years in the study of δ-records in the i.i.d. setting. The seminal paper [50] analyzed the process of counting δ-records, obtaining central limit theorems with a martingale approach for discrete sequences. Later, asymptotic normality was proven for more general distributions [52, 53].

In [53] the authors obtained a classification of the behaviour of the expected number of δ-records. For *heavy-tailed* distributions, the expected number of δ-records is the same as in the case of records, that is, of the order of the logarithm of the number of observations n. For *exponential-like tails*, the expected number of records is proportional to $\log(n)$, while for *light-tailed* distributions the growth rate is faster than $\log(n)$.

The distribution for continuous r.v. was studied by López-Blazquez and Salamanca-Miño [78, 79], who obtained expressions for the density of δ-records and the probability mass function of inter δ-record times. They also noted some interesting applications where δ-records arise naturally, such as queuing theory, blocks in motor traffic and Type-II particle counters, i.e. counters that are unable to detect particles during a dead time following the arrival of new particles.

The process structure of δ-records from continuous parents was studied in [55]. This paper showed that the point process of δ-record values follows a Poisson Cluster Process. One nice feature of this characterization is the role played by each component of δ-records: usual records are the centers for the point process while near-records are points drawn from each cluster conditional on the record values.

Recently, the use of δ-records in statistical inference has been proposed and positively assessed; see [45, 46, 53]. These articles show how information from δ-records can be incorporated successfully into the likelihood of the sample, which is used for computing maximum likelihood and Bayes estimators and predictions of future records. The resulting estimators and predictions outperform those computed using records only; moreover, a slight modification in the sampling scheme for records yields δ-records with a low additional cost. The results are applied to examples of data on rainfall and material strength.

The other kind of generalization we consider in some chapters of this monograph concerns the model of the underlying variables. An interesting departure from the i.i.d. model, which introduces time-dependence between observations, results from adding a deterministic linear trend to the observations, thus obtaining what is known as a Linear Drift Model (LDM), first introduced in [8]. Some important contributions of this paper were the asymptotic results such as laws of large numbers and central limit theorems for the total number of records, and their application to athletics data.

This model was later studied in [20], where the effect of heavy-tailed distributions was assessed for particular distributions. They also considered different kinds of trends to assess its effects on the asymptotic record rate. Borovkov [14] proved the Markovianity of the bivariate process of record times and record values and studied the limiting distributions for the inter-record times and increments between record observations.

The LDM was also studied in a wide range of scenarios in [33], and has proven particularly useful in the study of global warming to explain the actual number of upper records observed [94, 124, 125]. Furthermore, the importance of this model lies not only in its applications but also in its mathematical structure. For instance, the study of records in the LDM model can be helpful in determining whether the underlying distribution is heavy-tailed or not [34, 123].

The usefulness of the LDM has also led to the appearance of statistical inference techniques based on record observations when there is an underlying trend. In 1988, Smith [110] performed simulation-based numerical analyses of record-based maximum likelihood estimates in the LDM. Feuerverger and Hall [28] proposed a non-parametric alternative based on least squares fitting and bootstrapping techniques. Some decades later, Hoayek et al. [64] proposed distribution-free estimators for the increasing variances model (see [126]), for which they showed that it coincides with the LDM in the case where the underlying variables have a Gumbel distribution, and therefore in that case they also propose goodness-of-fit tests for the LDM.

Additionally, some extensions of the LDM were studied in [9], where the independence of the underlying random variables was dropped, and in [56], where the trend of the model was generalized from deterministic to random. The last decade has been especially productive in the study of models with correlated observations, such as moving averages [39], Lévy flights and random walks [40, 41, 67, 83, 85], biased or drifted random walks [77, 81, 122] and random trend models [56].

In this monograph we address some open problems for near-records and δ-records. Chapter 1 presents the known properties of records and δ-records for a better understanding of their properties, as well as establishing the notation that will be used later.

In Chapter 2 we study the point process of near-record values when the observations are discrete, taking values in the integers. This problem was already studied in [55] for continuous distributions. While in the discrete setting the resulting process is also a cluster process, it is no longer a Poisson process, which makes the study of the point process and its characterization more difficult. Laws of large numbers and central limit theorems for the number of near-records with a value in a set are also obtained. Finally, we characterize which discrete distributions fulfil a martingale condition relating the partial maxima and the number of δ-records at time n. We relate this characterization to the open problem of the positivity of the terms in a recurrence relation.

The LDM is studied extensively in Chapter 3 for δ-records. From the basic properties and derivations of the probability of δ-record, we study the asymptotic δ-record probability and its analytic properties. We derive exact expressions for some distributions, some of them also unknown in the case of usual records, and we use these results to assess the effect of the δ parameter when the underlying variables are heavy-tailed. For distributions where an analytic expression is not available, we propose first order approximations to study δ-record probabilities. We also compute the correlation of δ-record observations as a function of the number of observations and the δ parameter. We study the asymptotics of the counting process of δ-records in the LDM. The finiteness of the number of δ-records in the LDM is completely characterized. In particular, this result solves a conjecture posed by Franke et al.

[33] for records, proving the result not only for usual records but also in the general setting with $\delta \neq 0$. Finally, we obtain laws of large numbers and a central limit theorem under mild conditions, extending the results in [8] to the case $\delta \neq 0$, and we prove a law of large numbers for a random trend model.

In Chapter 4 we develop statistical inference methods for δ-records in the LDM. We propose two estimators for the variance of the number of δ-records and discuss their properties, proving consistency. We also study Maximum Likelihood Estimation based on δ-records in the LDM. We develop a general framework for Maximum Likelihood Estimation and we find analytic solutions for particular cases. We use Montecarlo simulation to compare the performance of the Maximum Likelihood Estimators using δ-records with those using records only. Finally, the results in Chapter 3 and in this chapter are applied to a real dataset of temperatures, where the LDM is consistent with the findings of other authors and the phenomenon of global warming. In particular, we find good agreement between the theoretical results and the data observed in the example.

Finally, in Chapter 5, we set out some conclusions of the results reached in previous chapters and offer some ideas for future work.

"Bon viatge per als guerrers
que al seu poble són fidels,
afavoreixi el Déu dels vents
el velam del seu vaixell,
i malgrat llur vell combat
l'amor ompli el seu cos generós,
trobin els camins dels vells anhels,
plens de ventures, plens de coneixences".

Lluís Llach

1

Records and δ-records. Preliminary Results

In this chapter we introduce basic concepts of the theory of records as well as some notation. We also introduce the concepts of near-record and δ-record and review and illustrate some of their properties.

1.1 General notation

Sequences of random variables (r.v.) are indexed by the natural number $\mathbb{N}$ unless otherwise stated, and are written in upper-case, in parentheses, e.g. $(X_n), (Y_k)$. We use $\mathbb{P}()$ and $\mathbb{E}()$ for the probability operator and the mathematical expectation respectively. A random variable can be characterized with its cumulative distribution function (cdf), usually denoted by F. The support of an r.v. with cdf F is denoted by $supp(F)$. Moreover, if the random variable is absolutely continuous it has a probability density function (pdf) f.

For discrete distributions defined on the non-negative integers, we use p_i for the probability that the random variables take the value i, that is $p_i = \mathbb{P}(X = i)$, the survival function $y_k = \sum_{i>k} p_i$ and the discrete hazard rate $r_k = p_k/y_{k-1}$.

We denote $a_n \propto b_n$ if $\lim_{n\to\infty} a_n/b_n = l > 0$, $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$ and $\mathcal{O}(\cdot)$ the standard big-O Landau notation. We will make use of the operator notation $\bigvee_{k=n}^{m} X_k$ for maxima of the r.v. $\max\{X_n, X_{n+1} \ldots, X_m\}$, and $\bigwedge$ is the minima "operator", that is, $\bigwedge_{k=n}^{m} X_k = \min\{X_n, X_{n+1} \ldots, X_m\}$. Also, we denote by $\lfloor\ \rfloor$ and $\lceil\ \rceil$, the floor and ceiling functions respectively.

We occasionally make use of the abbreviations rhs and lhs, which stand for right-hand side and left-hand side respectively.

Convergence of random variables can occur in distribution $\xrightarrow{\mathcal{D}}$, in probability $\xrightarrow{p}$, almost surely, written simply $a.s.$ or $\xrightarrow{a.s.}$, or in L_p.

1.2 Main definitions

As explained in the introduction, record theory is an extensively studied branch of mathematics centring on around the concept of record observation. Mathematically, we consider a sequence of observations $X_1, X_2, \ldots$ which are observed at discrete times. In this framework, the n-th observation of the sequence will be considered an upper record if it is greater than all previous entries, or mathematically, greater than the partial maxima up to time $n-1$.

Definition 1.2.1. To any random sequence (X_n) we associate the sequence of their partial maxima (M_n), defined by

$$M_n = \max\{X_1,\ X_2, \ldots,\ X_n\}, \quad n \geq 1.$$

Definition 1.2.2. Record. Given a sequence of random variables (X_n), we say that the n-th observation with $n \geq 2$ is an upper record, or simply a record, if

$$X_n > M_{n-1}.$$

Also, the n-th observation will be considered a lower record if

$$X_n < \min\{X_1, \ldots X_{n-1}\}.$$

By convention, the first observation is always considered to be an upper and lower record.

Remark 1.2.3. Record theory is usually written for upper records. Nevertheless, results for upper records can be transferred to lower records by considering the opposite sequence. Indeed, if we consider the opposite of the sequence $X_1, X_2, \ldots$, i.e. working with $-X_1, -X_2, \ldots$, it is easy to see that the roles of the upper and lower records are exchanged.

The main objects of interest in this monograph are δ-records and near-records, two concepts which were defined for the purpose of extending the potential applications of records to other problems.

Near-records were defined in 2005 by Balakrhisnan, Pakes and Stepanov [7] as observations that are close to being records. The same paper noted the potential interest of near-records in actuarial mathematics. In particular, this definition blends the well-studied records with a previously introduced concept named near-maxima, which are observations close to the maximum at time n.

Definition 1.2.4. Near-record. Let (X_n) be a sequence of random variables and $a > 0$ a parameter. Then, for $n \geq 2$, X_n is a near-record if

$$M_{n-1} - a < X_n \leq M_{n-1}.$$

In 2007, Gouet, López and Sanz introduced the concept of the δ-record [50], for which an observation is said to be a δ-record if it is greater than the previous record plus a fixed quantity, δ. For applications of δ-records we refer the reader to the introduction. The mathematical definition of δ-record is as follows.

Definition 1.2.5. δ-record. Let (X_n) be a sequence of random variables and $\delta \in \mathbb{R}$ a parameter. Then X_1 is a δ-record and, for $n \geq 2$, X_n is a δ-record if

$$X_n > M_{n-1} + \delta.$$

It is straightforward that for negative δ, a δ-record is either a record or a near-record with parameter $-\delta$. Consequently, the study of δ-records and near-records is closely related, and obtaining the properties of one of these notions generally results in obtaining properties of the other.

In the important case where the underlying variables are discrete, δ-records also extend the notion of weak records, as originally defined by Vervaat in 1973 [116], for which ties with the current maxima are also considered as weak-records.

Definition 1.2.6. Weak-record. Given a sequence of random variables (X_n), the first observation is a weak-record and, for $n \geq 2$, the n-th observation is a weak-record if

$$X_n \geq M_{n-1}.$$

It is easy to see from the definition that, if the sequence of r.v. (X_n) takes values in $\mathbb{Z}$, then weak-records and δ-records with $\delta = -1$ are equivalent.

In order to study δ-records, and thus also records, we introduce the following notation for δ-record indicators, times, cumulative occurrence counting variables, and the sequence of record values.

Definition 1.2.7. Let (X_n) be a sequence of random variables and $\delta \in \mathbb{R}$ a parameter.

(a) Let $1_{n,\delta}$ denote the indicator of the event $\{X_n \text{ is a } \delta\text{-record}\}$. That is, $1_{1,\delta} = 1$, and for $n \geq 2$, $1_{n,\delta} = 1$ if $X_n > M_{n-1} + \delta$ and $1_{n,\delta} = 0$ otherwise.

(b) δ-record times are defined recursively as $L_1(\delta) = 1$ and for $n \geq 2$

$$L_n(\delta) = \inf\{j > L_{n-1}(\delta) : X_j \text{ is a } \delta\text{-record}\}.$$

We occasionally omit δ if $\delta = 0$ to denote the record times, reading L_n.

(c) The number of δ-records up to time $n \geq 1$ is computed as $N_{n,\delta} = \sum_{j=1}^{n} 1_{j,\delta}$.

(d) The sequence of record values (R_n) is given by $R_n = X_{L_n} = M_{L_n}$, for $n \geq 1$.

Example 1.2.8. Given the sequence: $2, 4, 3, 6, 1, 6, 7, 1, 7, 8, 6, 7, 2, 4, 5, 8, 12, \ldots$

We have that X_1, X_2, X_4, X_7, X_{10} and X_{17} are records. So, in particular we have $N_{17,0} = 6$.

The sequence of partial maxima, the sequence of record values, and the sequence of record times are as follows

- $M_1 = 2$, $M_2 = 4$, $M_3 = 4$, $M_4 = 6$, $M_5 = 6 \ldots$
- $R_1 = 2$, $R_2 = 4$, $R_3 = 6$, $R_4 = 7$, $R_5 = 8$, $R_6 = 12$.
- $L_1 = 1$, $L_2 = 2$, $L_3 = 4$, $L_4 = 7$, $L_5 = 10$ and $L_6 = 17$.

Taking $a = 3$ for the near-record parameter, the following values of near-records are observed sequentially: $3, 6, 7, 6, 7, 8$.

Taking near-records with parameter $a = 1$, we observe the values $6, 7, 8$, and then for weak-records, or equivalently for δ-records with $\delta = -1$, we observe the sequence: $2, 4, 6, 6, 7, 7, 8, 8, 12$.

For δ-records with parameter $\delta = -3$, we observe the union of near-records with parameter $a = 3$ and usual records. That is, we observe the following δ-record values sequentially: $2, 4, 3, 6, 6, 7, 7, 7, 8, 6, 7, 8, 12$. The total number of δ-records in the first $n = 17$ observations is $N_{17,\delta} = 12$. The δ-records times are

$$L_1(-3) = 1, L_2(-3) = 2, L_3(-3) = 3, L_4(-3) = 4, L_5(-3) = 6, L_6(-3) = 7, \ldots.$$

1.3 Records from i.i.d. sequences

1.3.1 The Classical Record Model

The most widely studied case in the literature relates to record occurrences arising from sequences of i.i.d. r.v., the study of which had already attained an important degree of maturity some decades ago [1, 2, 4, 87]. As mentioned in the introduction, there are now ever more statistical applications in the literature from multiple fields of science.

In the case of usual records ($\delta = 0$), Foster and Stuart [30] proved in 1954 that the probability that the n-th observation is a record is $1/n$ if the underlying random variables (X_n) are i.i.d. continuous r.v. In fact, they used a simple argument which is now known as *stick-shuffling.* Since the probability of observing ties for continuous i.i.d. r.v. is null, every observation up to time n has an equal probability of being the maxima, and thus the result is straightforward. This result allows us to compute the expected number of records up to time n. Indeed,

$$\mathbb{E}(N_{n,0}) = \sum_{j=1}^{n} \mathbb{E}(1_{j,0}) = \sum_{j=1}^{n} 1/j = \log(n) + \gamma + \mathcal{O}(n^{-1})$$

where $\gamma \approx 0.57721$ is the Euler-Mascheroni constant. Renyi later proved the independence of record indicators and the convergence

$$\frac{N_{n,0}}{\log(n)} \to 1, \quad a.s. \tag{1.1}$$

from which an explicit result for the variance of the observed number of records can be computed as

$$Var(N_{n,0}) = \mathbb{E}\left(N_{n,0}^2\right) - \left(\mathbb{E}\left(N_{n,0}\right)\right)^2 = \sum_{j=1}^{n} \left(\frac{1}{j} - \frac{1}{j^2}\right) = \log(n) + \gamma - \frac{\pi^2}{6} + \mathcal{O}(n^{-1}).$$

Consequently, we only expect to observe a number of records of the order of the logarithm of the total number of observations. Note that these properties are distribution-free, making records a suitable tool in a wide range of scenarios. The study of records arising from sequences of continuous i.i.d. r.v. is therefore often called *Classical Record Model* (CRM).

In addition to the properties of the process of counting the number of records and their universality for continuous distributions, the study of the record values also gives some interesting results.

When the random variables are exponential with parameter 1, $Exp(1)$, it can be proved that the random variable R_n follows a gamma distribution $Gamma(n+1, 1)$.

This elegant result allows us in simple fashion to get the distribution of the n-th record when the underlying random variables have a common continuous cdf F. Thus, in the general case, it turns out that the distribution of the value of the n-th record is

$$F^{-1}\big(1 - e^{-R_n}\big).$$

From these results we obtain one of the main results relating to records. In particular, if the inverse of the hazard function associated with cdf F satisfies a mild condition, the sequence of record values fulfills a Central Limit Theorem [95].

Tata [115] previously obtained conditions on the hazard function to ensure the existence of sequences (a_n) and $(b_n > 0)$ such that

$$\frac{R_n - a_n}{b_n} \xrightarrow{\mathcal{D}} N(0,1).$$

Analogously to what occurs with the limit of the distributions of the maximum of random variables (see Section 1.3.4), Resnick [95] proved a result showing that the only three possible limits for records are the log-normal distribution, the negative-log-normal distribution and the normal itself. Results for the distribution of the record value R_n, and on the joint distribution of different records, can be seen in Arnold et al. (1998) [4].

1.3.2 Records from discrete sequences

Interesting – although more tedious to derive – are the results arising when the underlying distribution, F, of the i.i.d. r.v. X_n, is discrete, taking values in the non-negative integers.

In the first place, while the expected number of records in the CRM is universal, in the sense that it is distribution-free, in the discrete case it depends critically on the tail of the distribution [47, 48].

Moreover, if we assume that there is no right endpoint for the distribution F, that is $F(n) < 1$, $\forall n$, and denoting $\psi_F(x)$ the inverse of the hazard function, $\psi_F^{-1}(x) = -\log\left(1 - F(x)\right)$, Vervaat [116] proved that if we define the discrete version of $\psi_F(x)$ as

$$\psi_d^{-1}(x) = \sum_{j=0}^{\lfloor x \rfloor} r_j \ , \quad \text{with} \quad r_j = \frac{\mathbb{P}(X_n = j)}{\mathbb{P}(X_n \geq j)},$$

and defining

$$I_j = 1_{\{R_n = j \text{ for some } n \geq 1\}},$$

the following convergence towards normal distribution holds

$$\frac{\sum_{j=0}^{n} I_j - \sum_{j=0}^{n} r_j}{\sqrt{\sum_{j=0}^{n} r_j(1-r_j)}} \xrightarrow{\mathcal{D}} N(0,1) \ ,$$

if $\sum_{j=0}^{\infty} r_j \ (1-r_j) = \infty$.

Additionally, if $\psi_d(x)$ satisfies a mild condition then

$$\frac{R_n - \psi_d(n)}{\psi_d(n+\sqrt{n}) - \psi_d(n)} \xrightarrow{\mathcal{D}} N(0, 1-p) \ ,$$

where

$$p = \lim_{n\to\infty} \frac{\sum_{k=0}^{n} r_k^2}{\sum_{k=0}^{n} r_k} \in [0,1).$$

Similar results for weak records as defined in Definition 1.2.6 hold in the discrete case (see the work by Stepanov [113]).

1.3.3 Shorrock's Theorem

An important result which will be used in Chapter 2 to describe the point process of near-record values is Shorrock's Theorem. Originally proved by Shorrock in 1972 [104], a proof of this result using modern notation can be found in [87].

This result describes the process of record values for continuous, discrete, and general distributions. In particular, the decomposition of the counting process of record values as a sum of independent r.v. is particularly useful.

Theorem 1.3.1. *Let F be the cdf of the r.v. (X_n) with supp$(F) = S$ and hazard function $H(x) = -log\,(1 - F(x))$. Let also $D = \{d_j, j \geq 1\}$ denote the set of atoms of F. If $N(t) = card\{j : R_j \leq t\}$, then*

$$N(t) = N_c(t) + N_d(t)$$

where $N_c(t)$ and $N_d(t)$ are two independent point processes on S with $N_d(t)$ a process with independent increments that can be expressed as

$$N_d(t) = \sum_{d_j \leq t} I_j$$

with I_j random variables mutually independent with distribution $Ber(r_j)$, where $r_j = \mathbb{P}(X_n = d_j)/\mathbb{P}(X_n \geq d_j)$ is the discrete hazard rate.

Alternatively, we can write that if X_n are i.i.d. r.v. with values in the non-negative integers, then for $m = 0, 1, \ldots,$ and $n = 1, 2, \ldots$

$$\begin{aligned}\mathbb{P}(R_n > m) &= \mathbb{P}(\eta_0 + \eta_1 + \cdots + \eta_m < n)\mathbb{P}(R_n = m) \\ &= \mathbb{P}(\eta_0 + \eta_1 + \cdots + \eta_{m-1} = n-1)\frac{\mathbb{P}(X = m)}{\mathbb{P}(X \geq m)}\end{aligned}$$

where $\eta_n = 1$ if n is a record value, $\eta_n = 0$ if the sequence of record values, $R_1, R_2, \ldots$ does not contain the value n.

1.3.4 Extreme Value Distributions

From the definition itself, it is obvious the relationship between the record occurrences and the values of the partial maxima in the first n observations. Although the record occurrences in the CRM have already been shown to be independent of one another, and not dependent on the distribution of the cdf, the record values are dependent on the underlying distribution of the random variables, as we have seen in Section 1.3.1. As a consequence, the theory of records and the theory of extreme value statistics are closely related.

The principal result in extreme value theory is possibly the characterization of the three families of extreme-value distributions. Let us now consider that the observations follow the CRM. The problem is to find, when they exist, real sequences (a_n) and $(b_n) > 0$ such that

$$\frac{M_n - a_n}{b_n}$$

converges in distribution to a non-degenerate r.v. as $n \to \infty$. In other words, there exists a cdf G such that

$$\lim_{n\to\infty} F^n(b_n x + a_n) = G(x)$$

for every $x \in \mathbb{R}$. Interestingly, the following result, attributed to Fisher and Tippet in 1928 [29], and to Gnedenko in 1943 [38] gives necessary and sufficient conditions on the parent distribution for the existence of those sequences and when they exist, that G belongs to one of the so-called extreme-value families, these being the Gumbel class (Type-I), Fréchet class (Type-II) and the Weibull class (Type-III). The differences arising from this classification will be used later in Chapter 3.

1. Gumbel class. The limiting distribution is the Gumbel distribution with cdf

   $$G(x) = \exp\left(-\exp\left(-x\right)\right), \quad x \in \mathbb{R}.$$

   The distributions belonging to this class are those with exponential-like tails, i.e., distributions whose pdf decays faster than a power law. In this class we find the Gaussian distribution and the exponential distribution itself. It also includes distributions bounded to the right, provided that they fulfill a certain condition on the decay of the tail towards the right end-point of the support of the cdf.

2. Fréchet class. The limiting distribution is

$$G(x) = \exp\left(-x^{-\alpha}\right), \quad x, \alpha > 0.$$

It comprises distributions unbounded to the right with power-like tails, i.e, heavy-tailed distributions, such as the Cauchy and the Pareto.

3. Weibull Class. The limiting distribution is

$$G(x) = \exp\left(- \mid x \mid^{\alpha}\right), \quad x < 0, \alpha > 0$$

and $G(x) = 1$ if $x > 0$. Distributions in the Weibull class are bounded to the right and violate the condition on the tail mentioned in the Gumbel class. Most right-bounded distributions belong to this class, such as the Beta distribution.

1.4 The Linear Drift Model

The greatest research effort into the properties of records has been made in the study of the CRM. Obviously, the first extension of the CRM is to drop the independence or the identical distribution of the random variables; in other words to introduce some type of dependence in the model. Ballerini and Resnick [8, 9], taking their inspiration from records in athletics, initiated the study of records from observations with linear trend, introducing the so-called Linear Drift Model (LDM). This model is a simple setting where observations enjoy a time-dependence by allowing mean increments over time. As explained in the introduction, this feature allows situations to be modelled with a higher observed number of records than for the CRM, especially in the interesting cases of temperature and athletics data.

In Chapters 3 and 4 of this monograph, we will study the occurrence of δ-records in the LDM. In this model we focus on the random variables Y_n, which are considered to obey the Linear Drift model if Y_n can be represented as

$$Y_n = X_n + cn, \quad n \geq 1, \tag{1.2}$$

where $c \in \mathbb{R}$ is the trend parameter and $(X_n)_{n\geq 1}$ is a sequence of i.i.d. random variables, with (absolutely continuous) cdf F and pdf f. Another important parameter of the model is the right-tail expectation of the X_n, defined as

$$\mu^+ = \int_0^\infty x f(x) dx.$$

For simplicity, we assume the existence of an interval of real numbers $I = (x_-, x_+)$, with $-\infty \leq x_- < x_+ \leq \infty$, such that $f(x) > 0$, for all $x \in I$, and $f(x) = 0$ otherwise. We denote the left and right end-points of the support of F by $x_- = \inf\{x : F(x) > 0\}$ and $x_+ = \sup\{x : F(x) < 1\}$.

In this model, it is clear that the distribution-free property of the CRM is lost. Nevertheless, in their seminal paper, Ballerini and Resnick [8] proved that the asymptotic record rate converges to a constant if $\mu^+ < \infty$. More specifically, they prove that

$$\frac{N_{n,0}}{n} \to p \quad a.s.$$

with $p = \lim_{j\to\infty} \mathbb{P}(Y_j \text{ is a record})$.

Moreover, by extending the sequence (X_n) to a doubly infinite sequence, (X_n^*), they prove a central limit theorem for $N_{n,0}$, the random variable counting the number of records up to time n, under mild conditions:

$$\sqrt{n}\left(n^{-1}N_{n,0} - p\right) \xrightarrow{\mathcal{D}} N(0, \sigma^2)$$

with $\sigma^2 = p - p^2 + 2\sum_{m=1}^{\infty}(r_m - p^2)$, and r_m being the moment of the product of the indicators of record in the doubly infinite sequence for variables X_i^* and X_{i+m}^*.

The model was later studied by different authors, with interesting properties being obtained such as the Markovianity of the bivariate process of the record values and record times [14]. For a brief summary of the advances made and applications of the LDM, see the introduction to this monograph.

In Chapter 3 we also work with a generalization of the LDM introduced by Gouet et al. [55], where they considered the case of a random trend. The authors proved the strong convergence and asymptotic normality for the record rate of observations of the form $Y_n = X_n + T_n,\ n \geq 1$, where $(X_n), n \in \mathbb{Z}$ is a stationary ergodic sequence of random variables and $(T_n), n \geq 1$ is a stochastic trend process with stationary ergodic increments.

The proof of the asymptotic normality relies on the approach of Ballerini and Resnick [9]. However, in order to deal with the random trend, a moment bound for stationary sequences is needed. As an application, the strong convergence and asymptotic normality for the number of ladder epochs in a random walk with stationary ergodic increments is obtained.

1.5 Properties of δ-records

From the Definition 1.2.5 of δ-record it is easy to show that for $\delta < 0$ ($\delta > 0$), δ-records are more (less) frequent than usual records, while for $\delta = 0$ the definitions of

δ-record and record are identical. For this reason, the parameter δ is often set to be negative in applications, in order to mitigate the scarceness of the number of usual records while still keeping the extreme-like behaviour of such observations. Today, δ-records have been extensively studied by different authors obtaining many of their characteristics and proving their principal properties such as the behaviour of the counting process of δ-records, including its asymptotic behaviour, and distributional properties.

Gouet et al. introduced δ-records in [50], where they use a martingale approach to prove a central limit theorem for the number of δ-records $N_{n,\delta}$ with $\delta \neq 0$ for i.i.d. r.v. with common distribution F on the non-negative integers, $\mathbb{Z}_+$. More specifically, they define

$$s_k = \frac{\mathbb{P}(X_1 = k + \delta)}{\mathbb{P}(X_1 \geq k)}$$

and $\theta(k) = \sum_{i=0}^{k} s_i$, the basic martingale for the result is

$$N_{n,\delta} - \theta(M_n).$$

Let us also denote the quantile function $m(t) = min\{j \in \mathbb{Z}_+,\ y_j < 1/t\}$, for $t \geq 0$, and

$$z_k = \sum_{i>k} s_i \big(y_{i+\delta} + y_{i+\delta-1} - y_{i-1}\big)$$

Thus, using the martingale approach and splitting the results according to the value of δ and the behaviour of the hazard rates r_n, we have the following results.

- **Case** $\delta < 0$

  1. If $\limsup_k r_k < 1$, then

  $$\frac{N_{n,\delta} - \theta(m(n))}{\sqrt{\sum_{k=0}^{m(n)} z_k r_k / y_k}} \xrightarrow{\mathcal{D}} N(0,1).$$

  2. If $\lim_k r_k = 1$ and $\lim_k (1 - r_k)/(1 - r_{k-1}) = 1$, then

  $$\frac{N_{n,\delta} - \theta(m(n))}{\sqrt{\sum_{k=0}^{m(n)} (1 - r_k)^{2\delta}}} \xrightarrow{\mathcal{D}} N(0,1).$$

  3. If $\lim_k r_k = r \in [0,1)$, then

  $$\frac{N_{n,\delta} - \theta(m(n))}{\sqrt{\log n}} \xrightarrow{\mathcal{D}} N(0, \sigma_r^2),$$

  where $\sigma_r^2 = -r(1-r)^\delta \big((1-r)^{\delta+1} + (1-r)^\delta - 1\big) / \log(1-r)$ if $r \neq 0$ and $\sigma_0 = 1$. Moreover:

(a) If $r > 0$ and $\sum_{i=0}^{n} |r_i - r|/\sqrt{n} \longrightarrow 0$, then

$$\frac{N_{n,\delta} + r(1-r)^\delta \log n / \log(1-r)}{\sqrt{\log n}} \xrightarrow{\mathcal{D}} N(0, \sigma_r^2).$$

(b) If $r = 0$ and $\sum_{i=0}^{n} r_i^2/\sqrt{n} \longrightarrow 0$, then

$$\frac{N_{n,\delta} - \log n}{\sqrt{\log n}} \xrightarrow{\mathcal{D}} N(0, 1).$$

- **Case** $\delta > 0$ and $\lim_k r_k = r \in [0, 1]$.

1. If $r < 1$, then
$$\frac{N_{n,\delta} - \theta(m(n))}{\sqrt{\log n}} \xrightarrow{\mathcal{D}} N(0, \sigma_r^2),$$
where $\sigma_r^2 = -r(1-r)^\delta\big((1-r)^{\delta+1} - (1+2\delta r)(1-r)^\delta + 1\big)/\log(1-r)$ if $r \neq 0$ and $\sigma_0 = 1$.
2. If $r = 1$, then
$$\frac{N_{n,\delta} - \theta(m(n))}{\sqrt{\sum_{k=0}^{m(n)} e_k}} \xrightarrow{\mathcal{D}} N(0, 1)$$
if $\sum_{k=0}^{\infty} e_k = \infty$ where $e_k = (1 - r_k)(1 - r_{k+1}) \dots (1 - r_{k+\delta-1})$; and $\lim_n N_{n,\delta} < \infty$ (a.s.) if $\sum_{k=0}^{\infty} e_k < \infty$.
3. If $r > 0$ and $\sum_{i=0}^{n} |r_i - r|/\sqrt{n} \to 0$, then
$$(\log n)^{-1/2}\left(N_{n,\delta} + \frac{r(1-r)^\delta \log n}{\log(1-r)}\right) \xrightarrow{\mathcal{D}} N(0, \sigma_r^2).$$
4. If $r = 0$ and $\sum_{i=0}^{n} r_i^2/\sqrt{n} \to 0$, then
$$\frac{N_{n,\delta} - \log n}{\sqrt{\log n}} \xrightarrow{\mathcal{D}} N(0, 1).$$

The authors in [53] consider the more general case where the underlying random variables (X_n) are i.i.d. with common distribution F possibly discontinuous. Strong laws of large numbers for the counting process of δ-records $N_{n,\delta}$ are obtained when $\delta \leq 0$. In that paper, a fundamental result is the relation between the number of δ-records and the sum of partial minima of nonnegative i.i.d. random variables. In particular, it is shown that

$$\frac{N_{n,\delta}}{S_n} \to 1 \;\; a.s.$$

where $S_n = \sum_{k=1}^{n} min\{Y_1, \dots, Y_k\}$, $n \geq 1$ and $Y_n = 1 - F(X_n + \delta) \equiv \overline{F}(X_n + \delta)$, $n \geq 1$.

This result allowed the authors to obtain laws of large numbers for $N_{n,\delta}$ from the corresponding result for the sum of minima S_n of nonnegative i.i.d. r.v. which was studied by Deheuvels in 1974 [21], who established weak and strong convergence results.

Defining $G(y) = \mathbb{P}\big(\overline{F}(X+\delta) \leq y\big)$ and its generalized inverse $G^{\leftarrow}(z) = inf\{y \geq 0 : \ G(y) \geq z\}$, the function $H(x)$ below plays a key role in the results:

$$H(x) = \int_1^{e^x} G^{\leftarrow}(1/t)dt, \ x \geq 0,$$

and, since $G^{\leftarrow}$ is increasing, as $n \to \infty$

$$H(\log n) = \int_1^n G^{\leftarrow}(1/t)dt \sim \sum_{k=2}^{n} G^{\leftarrow}(1/k).$$

We can now describe the main result of the paper, which is namely, weak and strong laws of large numbers for the number of δ-records.

Let (X_n) be an i.i.d. sequence with common (general) distribution function F such that $\overline{F}(x) > 0$, for $x \geq 0$. Let now $\delta \leq 0$.

1. If $\lim_{x\to\infty} H(x+\log x)/H(x) = 1$ and
$$\sum_{n=2}^{\infty} \left(nG^{\leftarrow}(1/n)^2 \Big/ \left(\sum_{k=2}^{n} G^{\leftarrow}(1/k) \right)^2 \right) < \infty$$
holds, then
$$\frac{N_{n,\delta}}{H(\log n)} \to 1 \quad a.s.$$

2. If there exists a strictly increasing sequence of real numbers, $x_n \uparrow \infty$, such that $\lim_{n\to\infty} H(x_n + \log n)/H(\log n) = 1$ and
$$\lim_{n\to\infty} \sum_{k=2}^{n} k \ G^{\leftarrow}(1/k)^2 \Big/ \left(\sum_{k=2}^{n} G^{\leftarrow}(1/k) \right)^2 = 0$$
holds, then
$$\frac{N_{n,\delta}}{H(\log n)} \xrightarrow{p} 1.$$

From this result we can get different laws of large numbers for $N_{n,\delta}$ according to the tail of the distribution. However, for the sake of simplicity and ease of reading, we include here only the results that concern discrete and continuous distributions.

1. **Heavy and exponential light tails**

   (a) Let F be concentrated on $\mathbb{Z}_+$. Then:

      i. If the hazard rates $r_k \to 0$ and $\delta \le 0$, then $N_{n,\delta}/\log n \to 1$ a.s.

      ii. If the hazard rates $r_k \to r \in (0,1)$ and $\delta \le 0$, then $N_{n,\delta}/\log n \to -r(1-r)^\delta/\log(1-r)$ a.s.

   (b) Let F be absolutely continuous. Then:

      i. If the hazard function $\lambda(x) \equiv f(x)/\overline{F}(x) \to 0$, and $\delta \le 0$, then $N_{n,\delta}/\log n \to 1$ a.s.

      ii. If the hazard function $\lambda(x) \to a \in (0,\infty)$, and $\delta \le 0$, then $N_{n,\delta}/\log n \to e^{-a\delta}$ a.s.

2. **Light tails**

   (a) Let F be concentrated on $\mathbb{Z}_+$ with hazard rates $r_k \to 1$ and $\delta < 0$, and let $c_n = \sum_{k=0}^{m(n)}(1-r_k)^\delta$. Then:

      i. If $(1-r_k)/(1-r_{k-1}) \to 1$, then $N_{n,\delta}/c_n \xrightarrow{p} 1$.

      ii. If $k^\alpha(r_k - r_{k-1})/(1-r_{k-1}) \to 0$, for some $\alpha > 1/2$, then $N_{n,\delta}/c_n \to 1$ a.s.

   (b) Let F be absolutely continuous, with differentiable hazard function $\lambda(x) \to \infty$ (F is light tailed) and $\delta < 0$. Let also $c_t = \int_0^{m(t)} \lambda(z)a(z)dz$, for $t \ge 1$ with $a(z) = e^{\int^{z+\delta} z \lambda(u)du}$, then:

      i. If λ' is bounded, then: $N_{n,\delta}/c_n \xrightarrow{p} 1$.

      ii. If $|\lambda'(x)| < 1/x^r$, for some $r > 1/2$ and all x large enough, then $N_{n,\delta}/c_n \to 1$ a.s.

López-Blazquez and Salamanca-Miño [78, 79] provide the basic distribution theory of the δ-record values when the variables are i.i.d. and absolutely continuous. Their approach is based on the derivation of recurring formulas for the density of record values. They also analyzed several aspects of the inter δ-records times.

Finally, the authors in [53] introduce the inference with δ-records when the distribution F is absolutely continuous, computing the likelihood function and showing how δ-records can be used for maximum likelihood estimation. Following these ideas, new inferential procedures for the geometric and the Weibull distributions were later developed in [45, 46]. Maximum likelihood and Bayesian approaches for parameter estimation and prediction of future records based on δ-records were considered. The performance of the estimators was compared with estimations based solely on record-breaking data by means of Montecarlo simulations and showing that the use of δ-records is clearly advantageous.

For the geometric random variable, the distribution of the number of δ-records and of the values of δ-records associated with a record are obtained. The strong consistency of the Maximum Likelihood Estimator (MLE) is also proved and it is shown that this estimator is asymptotically unbiased. On the other hand, for the Weibull distribution, the strong consistency of the scale parameter, λ, when the shape parameter, β, is known, is also established. In both distributions Bayesian inference is also considered. From the results of these papers the conclusion is clear; δ-records improve inference compared with the case based solely on records. However, this conclusion is more noticeable in MLE than in the Bayesian framework, possibly because of the influence of the prior distributions.

One interesting point in these papers is the prediction of future records using δ-records, both using the maximum likelihood approach of Basak and Balakrishnan [10] and the Bayesian approach [112]. In both cases, there is a clear advantage in using δ-records over the exclusive use of records.

2

Near and δ-record values in discrete sequences

In this chapter we study the point process of near-record values when the underlying variables are i.i.d. taking values in the non-negative integers. To that end, we combine theory of point processes with classical results of record theory and new results about near-records. We find that the resultant process is a cluster process. We get the probability generating functional of the point process of near-record values from which we derive explicit expressions of the first moments for the number of near-records in the whole sequence taking values in a set A. We find sufficient conditions to ensure the finiteness of the number of near-records along the whole sequence of observations. If this total number is infinite, we find Laws of Large Numbers and Central Limit Theorems for the number of near-records with value less than or equal to N, as N grows to infinity. Finally, we characterize the distributions satisfying a martingale condition defined through the partial maxima and the number of δ-records, relating this result with the open problem of the positivity of the solution of a recurrence relation.

2.1 The point process of near-record values and notation

We consider the set of record values as a point process on $\mathbb{R}$, which can be described through the random counting measure ξ, defined, for any Borel subset A of $\mathbb{R}$, as

$$\xi(A) = card\{n \in \mathbb{N} \,|\, R_n \in A\}. \tag{2.1}$$

Also, observe that record times L_n, as defined in Definition 1.2.7, are the jump times of the sequence of partial maxima and that record values R_n are the (strictly increasing) subsequence of partial maxima, sampled at those jump times. However, without further probabilistic assumptions on (X_n), it may happen that $L_n = \infty$, from some value of n on, which is equivalent to the existence of a final record. Furthermore, we have to ensure that the counting measure ξ is boundedly finite in the sense of being finite on bounded Borel sets A.

Similarly to the usual record setting, we define the sequences (L_n^a) of near-record times and (R_n^a) of near-record values.

Definition 2.1.1. (a) Near-records times are defined by

$$L_1^a = \min\{k \in \mathbb{N} | k \geq 2,\ M_{k-1} - a < X_k \leq M_{k-1}\}$$

and

$$L_n^a = \min\{k \in \mathbb{N} \,|\, k > L_{n-1}^a, M_{k-1} - a < X_k \leq M_{k-1}\} \quad \text{for } n \geq 2.$$

(b) The sequence of near-record values (R_n^a), is given by $R_n^a = X_{L_n^a}$, for $n \geq 1$.

Now, we define the main process of interest in this chapter, the counting process of near-record values.

Definition 2.1.2. The counting process of near-record values is defined by

$$\eta(A) = card\{n \in \mathbb{N} \,|\, R_n^a \in A\}, \tag{2.2}$$

for any Borel subset A of $\mathbb{R}_+$.

As for records, assumptions are needed in order to ensure that near-record times and values are well defined. Additionally, in order to characterize η as a cluster point process, we consider a classification of near-records in terms of their proximity to records.

Definition 2.1.3. (a) For $m, n \in \mathbb{N}$, the n-th near-record value R_n^a is said to be associated to the m-th record value R_m if $L_m < L_n^a < L_{m+1}$.

(b) The point process $\eta(\cdot \mid R_m)$ of near-record values associated to R_m is defined by the random counting measure

$$\eta(A \mid R_m) = card\{n \in \mathbb{N} \mid R_n^a \in A, L_m < L_n^a < L_{m+1}\}, \tag{2.3}$$

for any Borel subset A of $\mathbb{R}_+$.

(c) Let $S_m = \eta(\mathbb{R}_+ \mid R_m)$ be the number of near-records associated to record R_m, $m \in \mathbb{N}$.

Remark 2.1.4. Note that unlike in the process of record values, where the sequence of records is strictly increasing by definition, in the near-record process there may be different near-record observations with the same value. In the point process η the multiplicity of these observations is included.

Example 2.1.5. Let us consider the sequence of Example 1.2.8. That is, we observe the discrete sequence $2, 4, 3, 6, 1, 6, 7, 1, 7, 8, 6, 7, 2, 4, 5, 8, 12, \ldots$. Taking $a = 3$ as the near-record parameter, we get:

- $L_1^3 = 3$, $L_2^3 = 6$, $L_3^3 = 9$, $L_4^3 = 11$, $L_5^3 = 12$, $L_6^3 = 16$.
- The sequence of near-records is $R_n^a = (3, 6, 7, 6, 7, 8, \ldots)$.
- The observed process of near-record values is $\{3, 6, 6, 7, 7, \ldots\}$.
- $\eta([6, 7]) = 4$.
- The value 3 has multiplicity 1, while, 6 and 7 have multiplicity 2.
- $S_1 = 0$, $S_2 = 1$, $S_3 = 1$, $S_4 = 1$, $S_5 = 3$.
- No near-records are associated to first record ($R_1 = 2$). One near record ($R_1^a = 3$) is associated to R_2. One near record ($R_2^a = 6$) is associated to R_3. One near record ($R_3^a = 7$) is associated to R_4. Three near-records ($R_4^a = 6$, $R_5^a = 7$, $R_6^a = 8$) are associated to R_5.

We finally state here the probabilistic assumptions regarding (X_n), which hold throughout this chapter: (X_n) is a sequence of i.i.d. r.v. defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking non-negative integer values, with probabilities $p_k := \mathbb{P}(X_1 = k)$, $k \in \mathbb{Z}^+ := \{0\} \cup \mathbb{N}$ and, for convenience, we define $p_k = 0$, for $k \in \mathbb{Z}, k < 0$.

In order to ensure that no final record exists and so, all record times L_n are well defined, we suppose that $y_k := \mathbb{P}(X_1 > k) > 0$, $\forall k \in \mathbb{Z}$ (note that $y_k = 1$, $\forall k < 0$). Consequently, ξ is a well-defined point process on $\mathbb{R}_+ := [0, \infty)$, with no multiple points and boundedly finite. Also, to avoid unnecessary complications, we assume $a \in \mathbb{N}$.

2.2 Characterizing the near-record process η

We recall that a point process N on $\mathbb{R}_+$ can be seen as a random measure and has probability generating functional (p.g.fl.) defined by

$$\mathbb{G}_N[h] = \mathbb{E}\left(e^{\int \log h(x) N(dx)}\right),$$

under appropriate conventions regarding the logarithm of 0, acting on measurable functions $h : \mathbb{R}_+ \to [0, 1]$ such that h is equal to 1 outside some bounded subset of $\mathbb{R}_+$. In the rest of this section all the functions h are supposed to be of this form. Alternative formulas for the p.g.fl., in the form of a product-integral or a product are given by

$$\mathbb{G}_N[h] = \mathbb{E}\left(\prod_{x\in\mathbb{R}_+} h(x)^{N(dx)}\right) = \mathbb{E}\left(\prod_{x:N(\{x\})>0} h(x)^{N(\{x\})}\right). \tag{2.4}$$

In this section, we show that the near-record process η is a discrete cluster process. Indeed, since $\eta(A) = \sum_{m=1}^{\infty} \eta(A \mid R_m)$, η can be seen as superposition of a denumerable family of point processes which, by Proposition 2.2.2 (c) below, are conditionally independent.

We characterize η by means of its probability generating functional and compute its first moments and other related quantities of interest. To that end, we first present some useful results about records and near-records.

Lemma 2.2.1. *(a) The point process ξ of record values has its atoms in $\mathbb{Z}_+$ and the random variables $I_n := \xi(\{n\}), n \in \mathbb{Z}_+$, are independent Bernoulli $\{0, 1\}$, with*

$$\mathbb{E}(I_n) = r_n := \frac{p_n}{y_{n-1}} = \mathbb{P}(X_1 = n \mid X_1 \geq n),\ n \in \mathbb{Z}_+.$$

(b) For any h,

$$\mathbb{G}_\xi[h] = \prod_{n=0}^{\infty} (1 - r_n(1 - h(n))). \tag{2.5}$$

Proof. (a) This is a particular case of Shorrock's Theorem presented in Chapter 1 for discrete sequences. For a proof of this result see for instance Theorem 16.1 in

[87]. To prove (b), from (a) and the second formula in (2.4), we obtain

$$\begin{aligned}\mathbb{G}_\xi[h] &= \mathbb{E}\left(\prod_{n=0}^{\infty} h(n)^{\xi(\{n\})}\right)\\ &= \mathbb{E}\left(\prod_{n=0}^{\infty} h(n)^{I_n}\right)\\ &= \prod_{n=0}^{\infty}\mathbb{E}\left(h(n)^{I_n}\right)\\ &= \prod_{n=0}^{\infty}\left(1-r_n+r_n h(n)\right).\end{aligned}$$

□

Proposition 2.2.2. *(a) Let S_m be the number of near-records associated to record R_m, $m\in\mathbb{N}$, according to Definition 2.1.3. Then*

$$\mathbb{P}(S_m = s\,|\,R_m) = (1-q_{R_m})^s q_{R_m},\ s\in\mathbb{Z}_+,$$

where $q_i := \mathbb{P}(X_1 > i\,|\,X_1 > i-a) = y_i/y_{i-a}$. That is, S_m is geometrically distributed (starting at 0), conditionally on R_m.

(b) Let $L^a_{n_1} < \cdots < L^a_{n_{S_m}}$ be the near-record times associated to R_m. Then, conditionally on R_m, S_m, the near-record values $Y_{m,j} := X_{L^a_{n_j}}, j = 1,\ldots,S_m$, are i.i.d. with

$$\mathbb{P}(Y_{m,1} = k_1,\ldots,Y_{m,s} = k_s \mid R_m, S_m) = \prod_{j=1}^{S_m}\pi(k_j, R_m),$$

where $\pi(k,i) := \frac{p_k}{y_{i-a}-y_i}1_{\{k\in(i-a,i]\cap\mathbb{N}\}}$. Moreover, conditionally on R_m, S_m, the random variables $N_m(k) := \eta(\{k\}\,|\,R_m) = \sum_{j=1}^{S_m} 1_{\{Y_{m,j}=k\}}$, for $k\in(R_m-a, R_m]\cap\mathbb{N}$, are multinomially distributed, with parameters $S_n, \pi(k, R_m)$.

(c) Conditionally on $\mathcal{R} := \sigma\{R_m\,|\,m\in\mathbb{N}\}$, $\mathcal{F}_m := \sigma\{R_m, S_m, Y_{m,j}, j = 1,\ldots,m\}$, $m\in\mathbb{N}$, are independent σ-algebras.

Proof. (a) Note that the random variables $X_n, n > L_m$, are independent and identically distributed as X_1. Define de subsequence (X_{k_n}), with $k_1 = \min\{k > L_m\,|\,X_k > R_m - a\}$ and $k_n = \min\{k > k_{n-1}\,|\,X_k > R_m - a\}, n\geq 2$. Then, conditionally on R_m, the sequence (X_{k_n}) is also i.i.d. but their common distribution is $\mathbb{P}(X_1\leq x\,|\,X_1 > R_m - a)$. Last, S_m is the number of terms X_{k_n} up to (but no including) the first $X_{k_n} > R_m$. Hence, conditionally on R_m, S_m is geometrically distributed, as stated.

(b) The near record values $Y_{m,j}$ are precisely the X_{k_n} before the next record. So, conditionally on R_m, S_m, they are i.i.d. with probabilities $\frac{p_k}{y_{R_m-a}-y_{R_m}}$, $R_m - a < k \leq$

R_m. Also, from the arguments above, it is clear that the $N_{m,k}$ are (conditionally) multinomial.

(c) This follows immediately from the fact that the r.v. (X_n) are i.i.d. Once a new record R_m is observed, the number of near-records and their values depend only on R_m and the r.v. before the arrival of a new record. Therefore, they are independent of the random variables before R_m. □

We compute below the p.g.fl. of the point process $\eta(\cdot \mid R_m)$, which is obtained from (2.4), taking conditional expectation. That is,

$$\mathbb{G}_{\eta(\cdot \mid R_m)}[h] := \mathbb{E}\left(\prod_{k=0}^{\infty} h(k)^{\eta(\{k\} \mid R_m)} \middle| R_m \right). \tag{2.6}$$

Proposition 2.2.3. *For any h,*

$$\mathbb{G}_{\eta(\cdot \mid R_m)}[h] = \frac{1}{1+\alpha_{R_m}(h)},$$

where

$$\alpha_i(h) := \frac{1}{y_i} \sum_{k=i-a+1}^{i} p_k(1-h(k)),\ i \in \mathbb{N}. \tag{2.7}$$

Proof. Suppose $R_m = i$, for some $m \in \mathbb{N}$. From the second formula in (2.6) and (b) of Proposition 2.2.2, we get

$$\begin{aligned}
\mathbb{G}_{\eta(\cdot \mid R_m)}[h] &= \mathbb{E}\left(\mathbb{E}\left(\prod_{k=0}^{\infty} h(k)^{\eta(\{k\} \mid R_m)} \middle| R_m, S_m \right) \middle| R_m \right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\prod_{k=0}^{\infty} h(k)^{N_m(k)} \middle| R_m, S_m \right) \middle| R_m \right) \\
&= \mathbb{E}\left(\left(\sum_{k=0}^{\infty} \pi(k, R_m) h(k) \right)^{S_m} \middle| R_m \right) \\
&= \sum_{s=0}^{\infty} \left(\sum_{k=0}^{\infty} \pi(k, R_m) h(k) \right)^{s} (1-q_{R_m})^s q_{R_m} \\
&= \frac{y_{R_m}}{y_{R_m} - \sum_{k=R_m-a+1}^{R_m} h(k) p_k} \\
&= \frac{1}{1+\alpha_{R_m}(h)}.
\end{aligned}$$

□

Definition 2.2.4. (See Section 6.3 in [19]) Let N_c be a point process in $\mathbb{R}_+$ and $\{N(\cdot \mid y) : y \in \mathbb{R}_+\}$ a family of point processes on $\mathbb{R}_+$. A point process N on $\mathbb{R}_+$ is a cluster process with center process N_c and component processes $\{N(\cdot \,|\, y) : y \in \mathbb{R}_+\}$ if, for every bounded A, Borel subset of $\mathbb{R}_+$,

$$N(A) = \int N(A \,|\, y) N_c(dy) = \sum_{y \in \mathbb{R}_+} N(A \,|\, y) 1_{\{N_c(\{y\})>0\}}.$$

Definition 2.2.5. For $i \in \mathbb{Z}_+$, let $\eta(\cdot \,|\, i)$ be the point process with p.g.fl. given by

$$\mathbb{G}_{\eta(\cdot \,|\, i)}[h] = \frac{1}{1+\alpha_i(h)},$$

for any h and with $\alpha_i(h)$ as defined in (2.7).

Theorem 2.2.6. *(a) The point process η of near-records is a cluster process on $\mathbb{Z}^+$, with center process ξ and independent component processes $\{\eta(\cdot \,|\, i), i \in \mathbb{Z}^+\}$, with $\eta(\cdot \,|\, i)$ given in Definition 2.2.5.*

(b) For any h,

$$\mathbb{G}_\eta[h] = \prod_{i=0}^{\infty} \frac{1+(1-r_i)\alpha_i(h)}{1+\alpha_i(h)}. \tag{2.8}$$

In particular, taking $h(k) = t1_A(k), t \in [0,1]$ and A a bounded Borel set, we obtain the probability generating function (p.g.f.) of $\eta(A)$ as

$$\varphi_A(t) := \mathbb{E}\left(t^{\eta(A)}\right) = \prod_{i=0}^{\infty} \frac{1+(1-r_i)(1-t)\alpha_i(A)}{1+(1-t)\alpha_i(A)}. \tag{2.9}$$

(c) Let $A, B \subset \mathbb{Z}_+$, then

1. $\mu(A) := \mathbb{E}(\eta(A)) = \sum_{i=0}^{\infty} \alpha_i(A) r_i$,
2. $Var(\eta(A)) = \sum_{i=0}^{\infty} (\alpha_i(A))^2 r_i(2-r_i) + \mathbb{E}(\eta(A))$,
3. $Cov(\eta(A), \eta(B)) = \sum_{i=0}^{\infty} \alpha_i(A)\alpha_i(B) r_i(2-r_i), \quad$ *for* $A \cap B = \emptyset$,

where $\alpha_i(A) := \alpha_i(1-1_A)$.

Proof. (a) From Proposition 2.2.3, we have

$$\eta(A) = \sum_{m=1}^{\infty} \eta(A \,|\, R_m) = \int \eta(A \,|\, x)\xi(dx) = \sum_{i=0}^{\infty} \eta(A \,|\, i) 1_{\{\xi(\{i\})>0\}}.$$

Moreover, for a bounded set A, $\eta(A)$ is finite a.s. since the sum in the rhs is a finite sum and $\eta(A \mid i)$ are geometric r.v. So, according to Definition 2.2.4, η is cluster point process as asserted. Independence of component processes follows from (c) in Proposition 2.2.2, because $\eta(A \,|\, R_m)$ is $\mathcal{F}_m$-measurable, for any $m \in \mathbb{N}$.

(b) From Lemma 2.2.1 (a), formula (2.4) and (a) in this theorem, we have

$$\begin{aligned}
\mathbb{G}_\eta[h] &= \mathbb{E}\left(\mathbb{E}\left(\prod_{k=0}^{\infty} h(k)^{\sum_{m=1}^{\infty}\eta(\{k\}\,|\,R_m)}\middle|\mathcal{R}\right)\right)\\
&= \mathbb{E}\left(\prod_{m=1}^{\infty}\mathbb{E}\left(\prod_{k=0}^{\infty} h(k)^{\eta(\{k\}\,|\,R_m)}\middle|\mathcal{R}\right)\right)\\
&= \mathbb{E}\left(\prod_{m=1}^{\infty}\mathbb{E}\left(\prod_{k=0}^{\infty} h(k)^{\eta(\{k\}\,|\,R_m)}\middle|R_m\right)\right)\\
&= \mathbb{E}\left(\prod_{m=1}^{\infty}\frac{1}{1+\alpha_{R_m}(h)}\right)\\
&= \mathbb{E}\left(\prod_{i=0}^{\infty}\frac{1}{1+\alpha_i(h)I_i}\right)\\
&= \prod_{i=0}^{\infty}\left(\frac{r_i}{1+\alpha_i(h)}+1-r_i\right).
\end{aligned}$$

For $\varphi_A(t)$ we replace h by $t1_A$ in (2.8) and obtain (2.9), after simple manipulations.

(c1) Observe that $\eta(A) = \sum_{k\in A}\sum_{m=1}^{\infty}\eta(\{k\}\,|\,R_m) = \sum_{k\in A}\sum_{m=1}^{\infty}N_m(k)$ and recall that $N_m(k)$ is binomial, conditional on R_m, S_m, with parameters $S_m, \pi(k, R_m)$ and that S_m is geometric, conditional on R_m, with expectation $\frac{1-q_{R_m}}{q_{R_m}}$.

Moreover, $N_m(A) := \sum_{k\in A} N_m(k)$ is binomial, conditional on R_m, S_m, with parameters $S_m, \pi(A, R_m) := \sum_{k\in A}\pi(k, R_m)$, hence

$$\begin{aligned}
\mathbb{E}(N_m(A)) &= \mathbb{E}(\mathbb{E}(N_m(A)\,|\,R_m, S_m))\\
&= \mathbb{E}(S_m\pi(A, R_m))\\
&= \mathbb{E}(\mathbb{E}(S_m\pi(A, R_m)\,|\,R_m))\\
&= \mathbb{E}(\pi(A, R_m)\mathbb{E}(S_m\,|\,R_m))\\
&= \mathbb{E}\left(\pi(A, R_m)\frac{1-q_{R_m}}{q_{R_m}}\right).
\end{aligned}$$

So, noticing that $\pi(A,i)\frac{1-q_i}{q_i} = \alpha_i(A)$,

$$\begin{aligned}\mathbb{E}(\eta(A)) &= \mathbb{E}\left(\sum_{m=1}^{\infty}\alpha_{R_m}(A)\right)\\ &= \mathbb{E}\left(\sum_{i=0}^{\infty}\alpha_i(A)I_i\right)\\ &= \sum_{i=0}^{\infty}\alpha_i(A)r_i.\end{aligned}$$

(c2) From the computations above, it is clear that

$$\mathbb{E}(\eta(A)\,|\,\mathcal{R}) = \sum_{m=1}^{\infty}\alpha_{R_m}(A) = \sum_{i=0}^{\infty}\alpha_i(A)I_i.$$

Hence, the variance of the conditional expectation is

$$Var(\mathbb{E}(\eta(A)\,|\,\mathcal{R})) = \sum_{i=0}^{\infty}\alpha_i^2(A)r_i(1-r_i).$$

We compute next the expectation of the conditional variance, namely $\mathbb{E}(Var(\eta(A)\,|\,\mathcal{R}))$. Observe that, because of the conditional independence of the $\eta(A\,|\,R_n)$, we have

$$Var(\eta(A)\,|\,\mathcal{R}) = \sum_{m=1}^{\infty}Var(\eta(A\,|\,R_m)\,|\,R_m) = \sum_{m=1}^{\infty}Var(N_m(A)\,|\,R_m).$$

Moreover,

$$\begin{aligned}Var(N_m(A)\,|\,R_m) &= \mathbb{E}(Var(N_m(A)\,|\,R_m,S_m)\,|\,R_m) + Var(\mathbb{E}(N_m(A)\,|\,R_m,S_m)\,|\,R_m)\\ &= \mathbb{E}(S_m\pi(A,R_m)(1-\pi(A,R_m))\,|\,R_m) + Var(S_m\pi(A,R_m)\,|\,R_m)\\ &= \frac{1-q_{R_m}}{q_{R_m}}\pi(A,R_m)(1-\pi(A,R_m)) + \frac{1-q_{R_m}}{q_{R_m}^2}\pi^2(A,R_m)\\ &= \alpha_{R_m}(A)(1-\pi(A,R_m)) + \alpha_{R_m}(A)\frac{\pi(A,R_m)}{q_{R_m}}\\ &= \alpha_{R_m}(A)\left(1+\pi(A,R_m)\left(\frac{1-q_{R_m}}{q_{R_m}}\right)\right)\\ &= \alpha_{R_m}(A)(1+\alpha_{R_m}(A)).\end{aligned}$$

So, $Var(\eta(A)\,|\,\mathcal{R}) = \sum_{m=1}^{\infty}Var(N_{m,A}\,|\,R_m) = \sum_{i=0}^{\infty}\alpha_i(A)(1-\alpha_i(A))I_i$. Collecting terms from the expressions above, we obtain

$$\begin{aligned}Var(\eta(A)) &= \sum_{i=0}^{\infty}\alpha_i^2(A)r_i(1-r_i) + \sum_{i=0}^{\infty}\alpha_i(A)(1+\alpha_i(A))r_i\\ &= \sum_{i=0}^{\infty}(\alpha_i(A))^2\,r_i(2-r_i) + E(\eta(A)).\end{aligned}$$

(c3) The covariance $Cov(\eta(A),\eta(B))$, when $A\cap B=\emptyset$, is obtained from the formula for the variance, noting that $\eta(A\cap B)=\eta(A)+\eta(B)$ a.s. Indeed, $Var(\eta(A)+\eta(B))=Var(\eta(A))+Var(\eta(B))+2Cov(\eta(A),\eta(B)))=Var(\eta(A\cup B))$, and so,

$$Cov(\eta(A),\eta(B)))=\frac{1}{2}(Var(\eta(A\cup B))-Var(\eta(A))-Var(\eta(B))).$$

□

Remark 2.2.7. The p.g.fl. of the process η of near-record values shown in Theorem 2.2.6 for discrete r.v., reflects its similarity with the process of near-record values for continuous variables studied in [55]. Indeed, from (2.8) we may write

$$\mathbb{G}(h)=\exp\left(\sum_{i=0}^{\infty}\log\left(1-r_i\frac{\alpha_i(h)}{1+\alpha_i(h)}\right)\right).$$

Consider now that each realization of a r.v. X_n with value i could arise as gathering a realization over the interval $[i,i+1)$ of a continuous r.v. X'_n. If we repeatedly split the values taken by the r.v. X_n when $X_n=i$ into different discrete points in the interval $[i,i+1)$ in such a way that the X_n approximates X'_n, then we can consider $r_i\ll 1$. Using $\log(1+x)\sim x$ for $x\ll 1$ we have

$$\begin{aligned}\mathbb{G}(h)&\sim\exp\left(-\sum_{i=0}^{\infty}r_i\frac{\alpha_i(h)}{1+\alpha_i(h)}\right)\\&=\exp\left(-\sum_{i=0}^{\infty}\left(1-\frac{\sum_{j=i+1}^{\infty}p_j}{\sum_{j=i-a+1}^{\infty}p_j-\sum_{j=i-a+1}^{i}p_jh(j)}\right)r_i\right),\end{aligned}$$

which is an analogous expression to that of result Theorem 1 (b) in [55] by substituting the sum for an integral.

Example 2.2.8. (Geometric distribution) Let (X_n) be a sequence of i.i.d. r.v. with geometric distribution with success parameter p. This distribution has $p_k=p(1-p)^k,\ k\in\mathbb{Z}^+$, $y_i=(1-p)^{i+1}$ and $r_i=p_i/y_{i-1}=p$. Applying Theorem 2.2.6 we obtain the expected number of near-records with value smaller than or equal to $N\in\mathbb{Z}^+$ in the whole sequence (X_n). If $a<N$ we have

$$\begin{aligned}\mathbb{E}\left(\eta([0,N])\right)=&p\sum_{i=0}^{a-1}\frac{1-(1-p)^{i+1}}{(1-p)^{i+1}}+p\sum_{i=a}^{N}\frac{(1-p)^{i-a+1}-(1-p)^{i+1}}{(1-p)^{i+1}}\\&+p\sum_{i=N+1}^{N+a-1}\frac{(1-p)^{i-a+1}-(1-p)^{N+1}}{(1-p)^{i+1}}\\=&p(-a+N+1)\left((1-p)^{-a}-1\right)+(ap-1)(1-p)^{-a}+(1-p)^{-a}-ap\\=&(N+1)p\left((1-p)^{-a}-1\right).\end{aligned}$$

Note that for the geometric distribution the number of near-records grows linearly with N. Also, after cumbersome computations,

$$\begin{aligned} Var\left(\eta([0,N])\right) = &- (p-2)p(1-p)^{-2a}\left(-a+N+p^2+2\right) \\ &+ p(1-p)^{-a}(2p(-a+N+2)+4a-3N-7) - 2(p-1)^{-2a} \\ &+ p(p-1)^{-2a}\left(-2(p-2)(1-p)^a - a(p-2)+(p-1)^2p\right) \\ &- p(N(p-1)+p+1)+2, \end{aligned}$$

which also has a linear growth with N.

Example 2.2.9. Let (X_n) be a sequence of i.i.d. r.v. with common hazard function $r_i = 1 - \frac{1}{i+1}$. We have a converging hazard function $r_i \to 1$ as $i \to \infty$ and explicit expressions for the probability mass function and the survival function

$$\begin{aligned} p_i &= \frac{1}{i!}\left(1-\frac{1}{i+1}\right), \\ y_i &= \prod_{k=0}^{i}(1-r_k) = \prod_{k=0}^{i}\frac{1}{k+1} = \frac{1}{(i+1)!}. \end{aligned}$$

For $i \in \{0,\ldots,a-1\}$ we compute the growth rate of p_N/y_{N+i} as $N \gg 1$ increases

$$\frac{p_N}{y_{N+i}} = \frac{\frac{1}{N!}\left(1-\frac{1}{N+1}\right)}{\frac{1}{(N+i+1)!}} \sim N^{i+1}.$$

We can compute the growth rate for the expectation and variance for the number of near-records taking value equal to N, $\eta(\{N\})$, in the whole sequence (X_n). First note that $\alpha_i(\{N\}) = p_N/y_i$ if $i \in \{N-a+1,\ldots,N\}$ and 0 otherwise. Now, from Theorem 2.2.6, we have

$$\begin{aligned} \mathbb{E}\left(\eta(\{N\})\right) &= p_N \sum_{i=N}^{N+a-1}\frac{r_i}{y_i}, \\ Var\left(\eta(\{N\})\right) &= p_N \sum_{i=N}^{N+a-1}\frac{r_i(2-r_i)}{y_i^2} + E\left(\eta(\{N\})\right). \end{aligned}$$

Recalling that $r_i \to 1$ as $i \to \infty$

$$\mathbb{E}\left(\eta(\{N\})\right) \sim \sum_{i=0}^{a-1} N^{i+1} \sim N^a.$$

For the variance we also have $r_i(2-r_i) \to 1$ which results in

$$Var\left(\eta(\{N\})\right) \sim \sum_{i=0}^{a-1} N^{2(i+1)} + N^a \sim N^{2a}.$$

Additionally, from Theorem 2.2.6 we can compute the p.g.f. of $\eta(\{N\})$

$$\begin{aligned}\varphi_{\{N\}}(t) &= \prod_{i=N-a+1}^{N} \left(\frac{1+(1-r_i)(1-t)\frac{p_N}{y_i}}{1+(1-t)\frac{p_N}{y_i}} \right) \\ &= \prod_{i=N-a+1}^{N} \left(\frac{1+(1-t)\frac{p_N}{y_{i-1}}}{1+(1-t)\frac{p_N}{y_i}} \right) \\ &= \frac{1+(1-t)\frac{p_N}{y_{N-a}}}{1+(1-t)\frac{p_N}{y_N}} \\ &= \frac{1+(1-t)r_N(1-r_{N-1})\ldots(1-r_{N-a+1})}{1+(1-t)\frac{r_N}{1-r_N}}.\end{aligned}$$

2.3 Finiteness of the number of near-records

In this section we analyze the situation where the total number of near-records along the whole sequence of observations is finite.

Theorem 2.3.1. *If $\sum_{i=0}^{\infty} r_i^2 < \infty$ then $\eta(\mathbb{R}_+) < \infty$ a.s. That is, the number of near-records in the whole sequence (X_n) is finite a.s. Moreover, $\eta(\mathbb{R}_+)$ has finite expectation*

$$\mathbb{E}(\eta(\mathbb{R}_+)) = \sum_{i=0}^{\infty} \alpha_i(\mathbb{R}_+) r_i \tag{2.10}$$

and probability generating function given by

$$\varphi_{\mathbb{R}_+}(t) := \prod_{i=0}^{\infty} \frac{1+(1-r_i)(1-t)\alpha_i(\mathbb{R}_+)}{1+(1-t)\alpha_i(\mathbb{R}_+)},$$

where $\alpha_i(\mathbb{R}_+) = \sum_{j=i-a+1}^{i} p_j/y_i$.

Proof. From Proposition 2.2.2 we have

$$\sum_{m=1}^{\infty} \mathbb{P}(S_m > 0 \,|\, \mathcal{R}) = \sum_{m=1}^{\infty} (1-q_{R_m}) = \sum_{i=0}^{\infty} \left(1 - \frac{y_i}{y_{i-a}}\right) I_i.$$

Taking expectations above we obtain

$$\begin{aligned}\sum_{m=1}^{\infty}\mathbb{P}(S_m>0) &= \sum_{i=0}^{\infty}\left(1-\frac{y_i}{y_{i-a}}\right)r_i\\ &= \sum_{i=0}^{\infty}\left(1-\prod_{j=i-a+1}^{i}(1-r_j)\right)r_i\\ &\le \sum_{i=0}^{\infty}\sum_{j=i-a+1}^{i} r_j r_i \qquad (2.11)\\ &= \sum_{j=0}^{a-1}\sum_{i=j}^{\infty} r_{i-j}r_i\\ &\le a\sum_{i=0}^{\infty} r_i^2.\end{aligned}$$

Therefore, by the Borel-Cantelli lemma, $\mathbb{P}(S_m > 0 \text{ i.o.}) = 0$, which yields the result.

In order to compute the p.g.f. of $\eta(\mathbb{R}_+)$, we observe that $\eta(n) := \eta([0,n]) \to \eta(\mathbb{R}_+)$ a.s. and so, by the monotone convergence theorem, $\varphi_{[0,n]}(t) = \mathbb{E}(t^{\eta(n)}) \to \varphi_{\mathbb{R}_+}(t)$, for $t \in [0,1]$, as $n \to \infty$. Furthermore, from (2.9) we have

$$\varphi_{\mathbb{R}_+}(t) = \lim_{n\to\infty}\prod_{i=0}^{\infty}\frac{1+(1-r_i)(1-t)\alpha_i(n)}{1+(1-t)\alpha_i(n)} = \prod_{i=0}^{\infty}\frac{1+(1-r_i)(1-t)\alpha_i(\mathbb{R}_+)}{1+(1-t)\alpha_i(\mathbb{R}_+)}. \quad (2.12)$$

The interchange of limit and product above is justified by the monotone convergence theorem, after taking logarithms, since the sequence inside the product decreases with n.

Finally, (2.10) is obtained, for example, from the derivative of $\varphi_{\mathbb{R}_+}$ at $t = 1^-$ or as the limit of $\mathbb{E}(\eta(n))$. Finiteness follows from the bound $1 - q_i \le \sum_{j=i-a+1}^{i} r_j$, used in (2.11), which implies $q_i \to 1$, as $i \to \infty$. Indeed, for sufficiently large i, we have $q_i \ge 1/2$ and

$$\alpha_i(\mathbb{R}_+) = \sum_{j=i-a+1}^{i} p_j/y_i = \frac{y_{i-a}}{y_i} - 1 = \frac{1-q_i}{q_i} \le 2(1-q_i).$$

The conclusion $\alpha_i(\mathbb{R}_+) < \infty$ is obtained after arguing as in (2.11). □

Example 2.3.2. Let (X_n) with $p_k = 1/(k(k+1)), y_k = (k+1)^{-1}$ and $r_k = (k+1)^{-1}$, $\forall k \in \mathbb{N}$. We have $\sum_{i=0}^{\infty} r_i^2 < \infty$ and, from Theorem 2.3.1, $\eta(\mathbb{R}_+) < \infty$ a.s.

To compute the expectation note that $\alpha_i(\mathbb{R}_+) = i$, for $i < a$, and $\alpha_i(\mathbb{R}_+) = \frac{i}{i-a+1}$, for $i \geq a$. So, from (2.10), we obtain

$$\begin{aligned}\mathbb{E}(\eta(\mathbb{R}_+)) &= \sum_{i=0}^{a-1}\left(1-\frac{1}{i+1}\right)+\sum_{i=a}^{\infty}\left(\frac{1}{i-a+1}-\frac{1}{i+1}\right)\\ &= a-\sum_{i=0}^{a-1}\frac{1}{i+1}+\sum_{i=0}^{a-1}\frac{1}{i+1}\\ &= a.\end{aligned}$$

It is interesting to see that the expected total number of near-records turns out to be equal to the near-record parameter a. For the variance we use Theorem 2.2.6 (c2), and taking limits as in the derivation of (2.10) we obtain

$$\begin{aligned}Var\,(\eta(\mathbb{R}_+)) =& a+2\sum_{i=0}^{a-1}\frac{i^2}{i+1}-\sum_{i=0}^{a-1}\left(\frac{i}{i+1}\right)^2+2\sum_{i=a}^{\infty}\frac{1}{i+1}\left(\frac{a}{i-a+1}\right)^2\\ &-\sum_{i=a}^{\infty}\left(\frac{a}{(i+1)(i-a+1)}\right)^2\end{aligned}$$

which we write as $a+A-B+C-D$.

We now compute each term. For A

$$\begin{aligned}A &= 2\sum_{i=0}^{a-1}\frac{i^2}{i+1}\\ &= 2\left(\sum_{i=1}^{a} i-2\sum_{i=1}^{a}1+\sum_{i=1}^{a}\frac{1}{i}\right)\\ &= 2\left(\frac{a(a+1)}{2}-2a+H(a)\right)\\ &= a^2-3a+2H(a),\end{aligned}$$

where $H(n)$ represents the n-th harmonic number.

In a similar way for B we have

$$B=\sum_{i=0}^{a-1}\left(\frac{i}{i+1}\right)^2=\sum_{i=1}^{a}1-2\sum_{i=1}^{a}\frac{1}{i}+\sum_{i=1}^{a}\frac{1}{i^2}=a-2H(a)+\sum_{i=1}^{a}\frac{1}{i^2}.$$

The sum of C can be computed via simple fractions as follows

$$\begin{aligned}
C &= 2a^2 \sum_{i=a}^{\infty} \frac{1}{(i-a+1)^2(i+1)} \\
&= 2a^2 \sum_{i=a}^{\infty} \left(\frac{1}{a^2(i+1)} + \frac{1}{a(i-a+1)^2} - \frac{1}{a^2(i-a+1)} \right) \\
&= -2 \sum_{i=1}^{a} \frac{1}{i} + 2a \sum_{i=1}^{\infty} \frac{1}{i^2} \\
&= -2H(a) + a\frac{\pi^2}{3}.
\end{aligned}$$

And again for D we can split the sum and compute it

$$\begin{aligned}
D &= \sum_{i=a}^{\infty} \left(\frac{a}{(i+1)(i-a+1)} \right)^2 \\
&= a^2 \sum_{i=a}^{\infty} \left(\frac{1}{a^2(i+1)^2} + \frac{2}{a^3(i+1)} + \frac{1}{a^2(i-a+1)^2} - \frac{2}{a^3(i-a+1)} \right) \\
&= \sum_{i=a+1}^{\infty} \frac{1}{i^2} - \frac{2}{a}H(a) + \frac{\pi^2}{6}.
\end{aligned}$$

Therefore

$$Var(\eta(\mathbb{R}_+)) = a^2 - 3a + 2\left(1 + \frac{1}{a}\right) H(a) + (a-1)\frac{\pi^2}{3},$$

that is, the variance is a quadratic function of a.

2.4 Asymptotic behaviour

We now focus on the asymptotic behaviour of $\eta([0,n])$ as $n \to \infty$. We know that $\lim_{n\to\infty} \eta([0,n])$ is finite a.s. if $\sum_{i=1}^{\infty} r_i^2 < \infty$. In this section we obtain laws of large numbers and a central limit theorem for $\eta([0,n])$ under the assumption $\sum_{i=1}^{\infty} r_i^2 = \infty$.

By the definition of near-record we have that $Z_i := \sum_{m=1}^{\infty} S_m 1_{\{R_m = i\}}$ is the total number of near-records associated to a record with value i. The r.v. Z_i's are mutually independent by Proposition 2.2.2 and the total number of near-records with value smaller than or equal to n can be bounded as

$$\sum_{i=0}^{n} Z_i \le \eta([0,n]) \le \sum_{i=0}^{n+a-1} Z_i. \tag{2.13}$$

The strategy now is to show the desired asymptotic results to the sum of Z_i's and then transfer its behaviour to η. For this purpose, we require to introduce some minimal conditions on the sequence of hazard rates (r_n).

Note first that Z_i has a geometric (starting at 0) distribution by Proposition 2.2.2 with parameter $q_i = y_i/y_{i-a}$ if i is a record in the sequence (X_n), and 0 otherwise. Its expectation is

$$\mathbb{E}\left(Z_i\right) = \mathbb{E}\left(\mathbb{E}\left(Z_i \mid I_i\right)\right) = \mathbb{E}\left(\left(\frac{1-q_i}{q_i}\right) I_i\right) = \frac{1-q_i}{q_i} r_i, \tag{2.14}$$

and the variance of Z_i is also easily computed as

$$\begin{aligned} Var\left(Z_i\right) &= Var\left(\mathbb{E}\left(Z_i \mid I_i\right)\right) + \mathbb{E}\left(Var\left(Z_i \mid I_i\right)\right) \\ &= Var\left(\left(\frac{1-q_i}{q_i}\right) I_i\right) + \mathbb{E}\left(\left(\frac{1-q_i}{q_i^2}\right) I_i\right) \\ &= \frac{1-q_i}{q_i^2} r_i\left((1-q_i)(1-r_i)+1\right). \end{aligned} \tag{2.15}$$

The following result establishes some properties of the random variables Z_n.

Proposition 2.4.1. *If one of the following conditions holds*

1.- $\limsup_{n\to\infty} r_n < 1$,

2.- $\lim_{n\to\infty} r_n = 1$ *and* $\lim_{n\to\infty}(1-r_n)/(1-r_{n-1}) = 1$.

then

a) $\sum_{i=0}^{\infty} \mathbb{E}\left(Z_i\right) = \infty$ *and* $\sum_{i=0}^{\infty} Var\left(Z_i\right) = \infty$.

b) For fixed $k \in \mathbb{N}$

$$\frac{\mathbb{E}\left(Z_{n+k}\right)}{\sum_{i=0}^{n} \mathbb{E}\left(Z_i\right)} \to 0 \text{ as } n \to \infty.$$

c) For fixed $k \in \mathbb{N}$

$$\frac{Z_{n+k}}{\sqrt{\sum_{i=0}^{n} Var\left(Z_i\right)}} \xrightarrow{p} 0 \text{ as } n \to \infty.$$

d) For fixed $k \in \mathbb{N}$

$$\frac{Var\left(Z_{n+k}\right)}{\sum_{i=0}^{n} Var\left(Z_i\right)} \to 0 \text{ as } n \to \infty.$$

Proof. **Case 1.-** $\limsup_{n\to\infty} r_n < 1$.

a) If $\limsup_{n\to\infty} r_n < 1$, then there exists $\gamma \in (0,1)$ with $\gamma < 1 - r_i$ for all i and we have

$$1 > q_i = \prod_{k=i-a+1}^{i} (1 - r_k) > \gamma^a. \tag{2.16}$$

This bound assures the equivalence

$$\sum_{i=0}^{\infty} \mathbb{E}(Z_i) = \sum_{i=0}^{\infty} \left(\frac{1-q_i}{q_i}\right) r_i < \infty \Longleftrightarrow \sum_{i=0}^{\infty} (1-q_i)\, r_i < \infty, \tag{2.17}$$

and since $1 - q_i > r_i$ and $\sum_{i=0}^{\infty} r_i^2 = \infty$ then $\sum_{i=0}^{\infty} \mathbb{E}(Z_i)$ diverges.

For the sum of variances we note that

$$1 < (1-q_i)(1-r_i) + 1 < 2, \tag{2.18}$$

and thus

$$\sum_{i=0}^{\infty} Var(Z_i) < \infty \Longleftrightarrow \sum_{i=0}^{\infty} (1-q_i) r_i < \infty,$$

which diverges under $\sum_{i=0}^{\infty} r_i^2 = \infty$.

b) Applying (2.16) and the divergence of the infinite sum of expectations we have

$$\frac{\mathbb{E}(Z_{n+k})}{\sum_{i=0}^{n} \mathbb{E}(Z_i)} = \frac{\left(\frac{1-q_{n+k}}{q_{n+k}}\right) r_{n+k}}{\sum_{i=0}^{n} \mathbb{E}(Z_i)} \le \frac{\gamma^{-a}}{\sum_{i=0}^{n} \mathbb{E}(Z_i)} \to 0, \text{ as } n \to \infty$$

and thus the result holds.

c) We prove the convergence in L^1 of the desired random variables using the same bound as in part (b) and the divergence of the sum of variances as follows

$$\frac{\mathbb{E}(Z_{n+k})}{\sqrt{\sum_{i=0}^{n} Var(Z_i)}} = \frac{\left(\frac{1-q_{n+k}}{q_{n+k}}\right) r_{n+k}}{\sqrt{\sum_{i=0}^{n} Var(Z_i)}} \le \frac{\gamma^{-a}}{\sqrt{\sum_{i=0}^{n} Var(Z_i)}} \to 0, \text{ as } n \to \infty.$$

Since convergence in L^1 implies convergence in probability the result is proved.

d) From (2.15), (2.16) and (2.18) we have $Var(Z_{n+k}) \le 2\gamma^{-a}$, which is a constant, and thus the result holds since $\sum_{i=0}^{n} Var(Z_i)$ diverges as seen in part (a).

Case 2.- $\lim_{n\to\infty} r_n = 1$ and $\lim_{n\to\infty}(1-r_n)/(1-r_{n-1}) = 1$.

a) In this setting, we have

$$q_i = \prod_{k=i-a+1}^{i} (1 - r_k) \to 0 \text{ as } i \to \infty.$$

and from (2.14) and (2.15) it is readily seen that the growth of the first moments of the Z_i random variables is

$$\mathbb{E}(Z_i) \sim \frac{1}{q_i} := \alpha_i, \tag{2.19}$$

$$Var(Z_i) \sim \frac{1}{q_i^2} = \alpha_i^2, \tag{2.20}$$

as $i \to \infty$. Since $\alpha_i \to \infty$ as $i \to \infty$, we obtain the divergence of both $\sum_{i=0}^{\infty} \mathbb{E}(Z_i)$ and $\sum_{i=0}^{\infty} Var(Z_i)$.

b) Note first that

$$\frac{\alpha_n}{\alpha_{n+k}} = \frac{q_{n+k}}{q_n} = \frac{\prod_{i=n+k-a+1}^{n+k}(1-r_i)}{\prod_{i=n-a+1}^{n}(1-r_i)}$$

and then it is straightforward that

$$\lim_{n\to\infty} \frac{1-r_n}{1-r_{n-1}} = 1 \Longrightarrow \lim_{n\to\infty} \alpha_n/\alpha_{n+k} = 1.$$

Now,

$$\frac{\mathbb{E}(Z_{n+k})}{\sum_{i=0}^{n}\mathbb{E}(Z_i)} \sim \frac{\alpha_{n+k}}{\sum_{i=0}^{n}\alpha_i} \sim \frac{\alpha_n}{\sum_{i=0}^{n}\alpha_i} \to 0, \text{ as } n \to \infty,$$

where the convergence to 0 holds as a consequence of Lemma A1 in [51].

c) We prove convergence in L^1 using the above equivalences in order to obtain convergence in probability. Thus,

$$\frac{\mathbb{E}(Z_{n+k})}{\sqrt{\sum_{i=0}^{n} Var(Z_i)}} \sim \frac{\alpha_{n+k}}{\sqrt{\sum_{i=0}^{n}\alpha_i^2}} \sim \sqrt{\frac{\alpha_n^2}{\sum_{i=0}^{n}\alpha_i^2}} \to 0 \text{ as } n \to \infty,$$

guaranteed again by Lemma A1 in [51].

d) From (2.20) and Lemma A1 in [51] we get

$$\frac{Var(Z_{n+k})}{\sum_{i=0}^{n} Var(Z_i)} \sim \frac{\alpha_{n+k}^2}{\sum_{i=0}^{n}\alpha_i^2} \to 0 \text{ as } n \to \infty.$$

□

Theorem 2.4.2. *If one of the following conditions holds*

1.- $\limsup_{n\to\infty} r_n < 1$,

2.- $\lim_{n\to\infty} r_n = 1$ *and* $k^\theta\left(\frac{r_k - r_{k-1}}{1-r_k}\right) \to 0$ *for some* $\theta > 1/2$ *as* $k \to \infty$, *then $\eta([0,n])$ obeys the following Strong Law of Large Numbers*

$$\frac{\eta([0,n])}{\mathbb{E}\left(\eta([0,n])\right)} \to 1 \text{ a.s. when } n \to \infty.$$

Proof. As stated in the introduction of this section, we first establish the Strong Law of Large Numbers for the sequence of partial sums of Z_n such that

$$\frac{\sum_{i=0}^{n} Z_i}{\sum_{i=0}^{n} \mathbb{E}\left(Z_i\right)} \to 1 \text{ a.s. when } n \to \infty, \tag{2.21}$$

for which it is enough to prove

$$\sum_{n=0}^{\infty} \frac{Var\left(Z_n\right)}{\left(\sum_{i=1}^{n} \mathbb{E}\left(Z_i\right)\right)^2} < \infty,$$

since we already proved that $\sum_{i=0}^{\infty} \mathbb{E}\left(Z_i\right) = \infty$ in Proposition 2.4.1. In order to do that, we proceed differently in the two scenarios.

Case 1.- $\limsup_{n\to\infty} r_n < 1$. Expressions (2.16) and (2.18) allow us to bound the expectation and variance

$$\mathbb{E}\left(Z_i\right) > (1-q_i)r_i,$$
$$Var\left(Z_i\right) < 2\gamma^{-2a}(1-q_i)r_i.$$

Then

$$\sum_{n=0}^{\infty} \frac{Var\left(Z_n\right)}{\left(\sum_{i=0}^{n} \mathbb{E}\left(Z_i\right)\right)^2} < 2\gamma^{-2a} \sum_{n=0}^{\infty} \frac{(1-q_n)r_n}{\left(\sum_{i=0}^{n}(1-q_i)r_i\right)^2}. \tag{2.22}$$

The series in the rhs of (2.22) converges by the Abel-Dini Theorem of convergence of series. Indeed, taking $d_n = (1-q_n)r_n$ and $\alpha = 1$, we apply Abel-Dini Theorem as stated in [63] pg. 441, since $d_n > 0$ and $\sum_{n=0}^{\infty} d_n$ is divergent following the reasoning in (2.17) and Proposition 2.4.1.

Case 2.- $\lim_{n\to\infty} r_n = 1$ and $k^{\theta}\left(\frac{r_k - r_{k-1}}{1-r_k}\right) \to 0$ for some $\theta > 1/2$ as $k \to \infty$. Note that this condition is stronger than $\lim_{n\to\infty}(1-r_n)/(1-r_{n-1}) = 1$. Recall that this hypothesis implies $1 - r_k \sim 1 - r_{k-1}$.

From (2.19) and (2.20), it suffices to check

$$\sum_{n=0}^{\infty} \frac{\alpha_n^2}{\left(\sum_{i=1}^{n} \alpha_i\right)^2} < \infty.$$

This series was studied in [42], we gather here the elements of that paper that allow us to prove its convergence. Let us suppose that $k^{\theta}\left(1 - \frac{\alpha_{k-1}}{\alpha_k}\right) \to 0$ as $k \to \infty$.

Taking $\varepsilon = 1$, there exists a $k_0 \in \mathbb{Z}^+$ such that , $\forall k > k_0$, it holds

$$\left| k^\theta \left(1 - \frac{\alpha_{k-1}}{\alpha_k}\right)\right| < \varepsilon \;\Rightarrow\; k^\theta \left(1 - \frac{\alpha_{k-1}}{\alpha_k}\right) < \varepsilon = 1.$$

which yields

$$\alpha_{k-1} > \left(1 - \frac{1}{k^\theta}\right)\alpha_k, \quad \forall k > k_0.$$

Iterating the last inequality we have

$$\alpha_l > \left(1 - \frac{1}{(l+1)^\theta}\right)\alpha_{l+1} > \left(1 - \frac{1}{(l+1)^\theta}\right)\left(1 - \frac{1}{(l+2)^\theta}\right)\alpha_{l+2}$$

and so we get

$$\alpha_l > \prod_{i=l+1}^{k}\left(1 - \frac{1}{i^\theta}\right)\alpha_k, \quad \text{for } l = k_0, \ldots, k-1,$$

and then

$$\frac{\alpha_n}{\sum_{k=k_0}^n \alpha_k} < \frac{\alpha_n}{\sum_{k=k_0}^{n-1}\prod_{i=k+1}^n \left(1 - \frac{1}{i^\theta}\right)\alpha_n}.$$

Moreover, note that $\log(1-x) \geq -2x$, $\forall\ 0 < x \leq 1/\sqrt{2}$, so $\log\left(1 - 1/i^\theta\right) \geq -2\frac{1}{i^\theta}$ $\forall i \geq 2$ and since $\theta > 1/2$, we get

$$\begin{aligned}
\prod_{i=k+1}^n \left(1 - \frac{1}{i^\theta}\right) &= \exp\left(\sum_{i=k+1}^n \log\left(1 - \frac{1}{i^\theta}\right)\right) \\
&\geq \exp\left(-2\sum_{i=k+1}^n \frac{1}{i^\theta}\right) \\
&\geq \left(-2\int_k^n \frac{1}{x^\theta}dx\right) \\
&= \exp\left(-\frac{2\left(n^{(1-\theta)} - k^{(1-\theta)}\right)}{1-\theta}\right), \quad \forall k > 1.
\end{aligned}$$

Let now $\beta = \frac{2}{1-\theta}$, $\theta \neq 1$, we have

$$\begin{aligned}
\sum_{k=k_0}^{n-1}\prod_{i=k+1}^n \left(1 - \frac{1}{i^\theta}\right) &\geq e^{-\beta n^{(1-\theta)}}\sum_{k=k_0}^{n-1} e^{\beta k^{(1-\theta)}} \\
&\geq e^{-\beta n^{(1-\theta)}}\int_{k_0-1}^{n-1} e^{\beta x^{(1-\theta)}}dx \\
&\geq e^{-\beta n^{(1-\theta)}}\frac{(n-1)^\theta e^{\beta(n-1)^{(1-\theta)}}}{3} \\
&= \frac{(n-1)^\theta e^{\beta\left[(n-1)^{(1-\theta)} - n^{(1-\theta)}\right]}}{3},
\end{aligned} \tag{2.23}$$

where for the last inequality we have used the relation

$$\int_0^y e^{\beta x^{(1-\theta)}} dx \sim \frac{e^{\beta y^{(1-\theta)}} y^\theta}{2} \quad as\ y \to \infty.$$

Moreover,

$$n^{(1-\theta)} - (n-1)^{(1-\theta)} \leq \frac{1}{(n-1)^\theta} \to 0 \quad as \quad n \to \infty,$$

and so from (2.23) and $e^{-\beta/(n-1)^\theta} \to 1,\ as\ n \to \infty$, we have

$$\begin{aligned}\sum_{k=k_0}^{n-1} \prod_{i=k+1}^{n} \left(1 - \frac{1}{i^\theta}\right) &= \frac{(n-1)^\theta e^{\beta\left[(n-1)^{(1-\theta)} - n^{(1-\theta)}\right]}}{3} \\ &\geq \frac{(n-1)^\theta e^{-\beta \frac{1}{(n-1)^\theta}}}{3} \\ &\geq \frac{(n-1)^\theta (1-\varepsilon)}{3}.\end{aligned}$$

Finally,

$$\frac{(n-1)^\theta(1-\varepsilon)}{3} \geq \frac{n^\theta}{4} \Leftrightarrow \frac{(n-1)^\theta}{n^\theta} \geq \frac{3}{4}\frac{1}{(1-\varepsilon)} \Leftrightarrow \frac{n-1}{n} \geq \left(\frac{3}{4}\frac{1}{(1-\varepsilon)}\right)^{\frac{1}{\theta}},$$

where the last inequality is true for n big enough if $\frac{3}{4}\frac{1}{(1-\varepsilon)} < 1$ and that is true if we chose ε such that $1 - \varepsilon > 3/4$.

Summarizing,

$$\frac{\alpha_n}{\sum_{k=k_0}^{n} \alpha_k} < \frac{4}{n^\theta},$$

which finally yields the result since $\theta > \frac{1}{2}$ and

$$\sum_{n=0}^{\infty} \frac{\alpha_n^2}{\left(\sum_{k=k_0}^{n} \alpha_k\right)^2} < \infty \Longleftrightarrow \sum_{n=0}^{\infty} \frac{1}{n^{2\theta}} < \infty.$$

Now, given that $\alpha_{k-1}/\alpha_k = (1 - r_k)/(1 - r_{k-a})$, the result holds under the condition

$$\lim_{k\to\infty} k^\theta \left(\frac{r_k - r_{k-a}}{1 - r_{k-a}}\right) = 0,$$

which is a weaker condition than

$$\lim_{k\to\infty} k^\theta \left(\frac{r_k - r_{k-1}}{1 - r_k}\right) = 0,$$

since $1 - r_k \sim 1 - r_{k-1}$.

Cases 1. and **2.-** It only remains to transfer the strong convergence of the sum of Z_n to $\eta([0,n])$. In both scenarios, Proposition 2.4.1 (b) implies that

$$\lim_{n\to\infty} \frac{\sum_{i=0}^{n+a-1} \mathbb{E}\left(Z_i\right)}{\sum_{i=0}^{n} \mathbb{E}\left(Z_i\right)} = 1.$$

Now, from (2.21), we have

$$\frac{\sum_{i=0}^{n+a-1} Z_i}{\sum_{i=0}^{n} \mathbb{E}\left(Z_i\right)} \to 1 \text{ a.s. as } n \to \infty. \tag{2.24}$$

Finally, from (2.13), (2.21) and (2.24),

$$\frac{\eta([0,n])}{\sum_{i=0}^{n} \mathbb{E}\left(Z_i\right)} \to 1 \text{ a.s. as } n \to \infty$$

and

$$\lim_{n\to\infty} \frac{\mathbb{E}(\eta([0,n]))}{\sum_{i=0}^{n} \mathbb{E}\left(Z_i\right)} = 1,$$

proving that the Strong Law of Large Numbers holds. □

In order to prove that a central limit theorem holds, we will make use of the next elementary lemma on the sum of a random number of independent Bernoulli r.v.

Lemma 2.4.3. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $A \in \mathcal{F}$ and let $p_A = \mathbb{P}(A)$. Let (B_k) be a sequence of independent events in $\mathcal{F}$, independent of A, with $\mathbb{P}(B_k) = p_B$, for all $k \geq 1$. Let N be a r.v. which takes values in $\{0, 1, \ldots\}$ such that N is independent of $1_A, 1_{B_k}$, $k \geq 1$ and $\mathbb{E}\left(N^2\right) < \infty$. Define $X = 1_A N$ and $Y = 1_A \sum_{k=1}^{N} 1_{B_k}$. Then, $Var(Y) \leq Var(X) + \mathbb{E}(X)$.*

Proof. We have $Var(X) = Var\left(\mathbb{E}\left(X|1_A\right)\right) + \mathbb{E}\left(Var\left(X|1_A\right)\right)$. Since $\mathbb{E}\left(X|1_A\right) = \mathbb{E}(N)1_A$, the first term is equal to $\left(\mathbb{E}(N)\right)^2 p_A(1-p_A)$. Also, the second term is $\mathbb{E}\left(Var(N)1_A\right) = Var(N)p_A$. Thus.

$$Var(X) = \left(\mathbb{E}(N)\right)^2 p_A(1-p_A) + Var(N)p_A.$$

We compute $Var(Y)$ in a similar way. First, we note that $\mathbb{E}\left(Y|1_A\right) = \mathbb{E}(N)p_B 1_A$, so $Var\left(\mathbb{E}\left(Y|1_A\right)\right) = \left(\mathbb{E}(N)\right)^2 p_B^2 p_A(1-p_A)$. For $Var\left(Y|1_A\right)$ we use the formula for the variance of the sum of a random number of elements, yielding

$$Var\left(Y|1_A\right) = \left(\mathbb{E}(N)p_B(1-p_B) + Var(N)p_B^2\right) 1_A.$$

Therefore,

$$Var(Y) = \left(\mathbb{E}(N)\right)^2 p_B^2 p_A(1-p_A) + \mathbb{E}(N)p_A p_B(1-p_B) + Var(N)p_A p_B^2.$$

Since $p_B \leq 1$ and $\mathbb{E}(X) = \mathbb{E}(N)p_A$, the result is proved. □

Theorem 2.4.4. *If one of the following conditions holds*

1.- $\limsup_{n\to\infty} r_n < 1$,

2.- $\lim_{n\to\infty} r_n = 1$ *and* $\lim_{n\to\infty}(1-r_n)/(1-r_{n-1}) = 1$.

then $\eta([0,n])$ *obeys the following Central Limit Theorem*

$$\frac{\eta([0,n]) - \mathbb{E}\left(\eta([0,n])\right)}{\sqrt{Var\left(\eta([0,n])\right)}} \xrightarrow{\mathcal{D}} N(0,1) \text{ as } n \to \infty.$$

Proof. As in the proof of Theorem 2.4.2, we first prove the asymptotic normality for the sequence of the partial sums of Z_i.

To that end, let us see that $\sum_{i=0}^{n} Z_i$ satisfies the following Lyapunov condition

$$\frac{1}{s_n^3}\sum_{i=0}^{n}\mathbb{E}\left(|Z_i - \mathbb{E}(Z_i)|^3\right) \to 0, \text{ as } n \to \infty, \tag{2.25}$$

where $s_n^2 = \sum_{i=0}^{n} Var(Z_i)$, which implies the asymptotic normality of the sum of Z_i

$$\sum_{i=0}^{n}\left(Z_i - \mathbb{E}(Z_i)\right)/s_n \xrightarrow{\mathcal{D}} N(0,1) \text{ as } n \to \infty.$$

From the elementary inequality $\mid a - b \mid^3 < a^3 + b^3$ with $a, b \geq 0$, we get

$$\begin{aligned}
\mathbb{E}\left(|Z_i - \mathbb{E}(Z_i)|^3\right) &= \mathbb{E}\left(\mid Z_i - \mathbb{E}(Z_i)\mid^3 \mid I_i = 1\right) r_i + \mathbb{E}\left(\mid Z_i - \mathbb{E}(Z_i)\mid^3 \mid I_i = 0\right)(1-r_i)\\
&\leq (\mathbb{E}\left(Z_i^3 \mid I_i = 1\right) + \mathbb{E}(Z_i)^3) r_i + \mathbb{E}(Z_i)^3(1-r_i)\\
&= \mathbb{E}\left(Z_i^3 \mid I_i = 1\right) r_i + \mathbb{E}(Z_i)^3\\
&= r_i\frac{(1-q_i)}{q_i^3}(6 + q_i^2 - 6q_i) + r_i^3\left(\frac{1-q_i}{q_i}\right)^3.
\end{aligned} \tag{2.26}$$

Case 1.- $\limsup_{n\to\infty} r_n < 1$. From (2.16) and (2.26)

$$\mathbb{E}\left(|Z_i - \mathbb{E}(Z_i)|^3\right) \leq K_1(1-q_i)r_i \tag{2.27}$$

for a positive constant K_1.

We bound the sequence (s_n) using (2.15), (2.16) and (2.18)

$$s_n^3 = \left(\sum_{i=0}^{n} Var(Z_i)\right)^{3/2} \geq K_2\left(\sum_{i=0}^{n}(1-q_i)r_i\right)^{3/2}, \tag{2.28}$$

where K_2 is a positive constant. Finally from (2.27) and (2.28) the Lyapunov condition (2.25) is satisfied since $\sum_{i=0}^{n}(1-q_i)r_i$ diverges as $n \to \infty$.

Case 2.- $\lim_{n\to\infty} r_n = 1$ and $\lim_{n\to\infty}(1-r_n)/(1-r_{n-1}) = 1$. For the numerator in the Lyapunov condition we proceed from (2.26) taking equivalences under $\lim r_i = 1$, yielding

$$\mathbb{E}\left(|Z_i - \mathbb{E}(Z_i)|^3\right) \leq K_3 \alpha_i^3,$$

with K_3 a positive constant.

From (2.20) and (2.21), the Lyapunov condition in (2.25) will be implied by

$$\lim_{n\to\infty} \frac{\sum_{i=0}^{n} \alpha_i^3}{\left(\sum_{i=0}^{n} \alpha_i^2\right)^{3/2}} = 0.$$

Applying Cauchy-Schwartz's Inequality we get

$$\begin{aligned}
\frac{\sum_{i=0}^{n} \alpha_i^3}{\left(\sum_{i=0}^{n} \alpha_i^2\right)^{3/2}} &= \frac{\sum_{i=0}^{n} \alpha_i^2 \alpha_i}{\left(\sum_{i=0}^{n} \alpha_i^2\right)^{3/2}} \\
&\leq \frac{\left(\sum_{i=0}^{n} \alpha_i^4\right)^{1/2}\left(\sum_{i=0}^{n} \alpha_i^2\right)^{1/2}}{\left(\sum_{i=0}^{n} \alpha_i^2\right)^{3/2}} \\
&= \left(\frac{\sum_{i=0}^{n} \alpha_i^4}{\left(\sum_{i=0}^{n} \alpha_i^2\right)^{2}}\right)^{1/2} \\
&\to 0, \text{ as } n \to \infty,
\end{aligned}$$

where the last limit holds if $(1-r_k)/(1-r_{k-1}) \to 1$ by means of Lemma A1 in [51]. Indeed, taking α_n^2 as a_n in the cited lemma, note that $a_n \to \infty$ and

$$\frac{a_n}{a_{n-1}} = \left(\frac{q_{n-1}}{q_n}\right)^2 = \left(\frac{1-r_{n-a}}{1-r_n}\right)^2 \to 1$$

if $(1-r_n)/(1-r_{n-1}) \to 1$.

Cases 1. and **2.-** We have proved

$$\frac{\sum_{i=0}^{n} Z_i - \sum_{i=0}^{n} \mathbb{E}(Z_i)}{\sqrt{\sum_{i=0}^{n} Var(Z_i)}} \xrightarrow{\mathcal{D}} N(0,1) \text{ as } n \to \infty.$$

This implies, by (2.13) and Proposition 2.4.1 (c), that

$$\frac{\eta([0,n]) - \sum_{i=0}^{n} \mathbb{E}(Z_i)}{\sqrt{\sum_{i=0}^{n} Var(Z_i)}} \xrightarrow{\mathcal{D}} N(0,1). \tag{2.29}$$

Now, taking expectations in (2.13) we have

$$\sum_{i=0}^{n} \mathbb{E}(Z_i) \leq \mathbb{E}(\eta([0,n])) \leq \sum_{i=0}^{n+a-1} \mathbb{E}(Z_i),$$

which, together with Proposition 2.4.1 (b) and (2.29) yields

$$\frac{\eta([0,n]) - \mathbb{E}(\eta([0,n]))}{\sqrt{\sum_{i=0}^{n} Var(Z_i)}} \xrightarrow{\mathcal{D}} N(0,1) \text{ as } n \to \infty. \tag{2.30}$$

The last step is to prove

$$\frac{Var(\eta([0,n]))}{\sum_{i=0}^{n} Var(Z_i)} \to 1 \text{ as } n \to \infty.$$

To that end, we note that $\eta(n) = \sum_{i=0}^{n} Z_i + \sum_{i=n+1}^{n+a-1} Z_i'$, where

$$Z_i' = \sum_{m=1}^{\infty} \left(\sum_{j=1}^{S_m} 1_{\{Y_{m,j} \in [0,n]\}} \right) 1_{\{R_m = i\}},$$

is the number of near-records associated to a record with value i which are less than or equal to n. Since the random variables $Z'_{n+1}, \ldots, Z'_{n+a-1}$ are independent, and also independent of $Z_1, \ldots, Z_n$, we have

$$Var\left(\sum_{i=0}^{n} Z_i\right) \leq Var\left(\sum_{i=0}^{n} Z_i\right) + \sum_{i=n+1}^{n+a-1} Var(Z_i') = Var(\eta([0,n])). \tag{2.31}$$

We now find an upper bound for $Var(Z_i')$ by using Lemma 2.4.3. The distribution of Z_i can be written as $N1_A$ where $A = \{\text{a record takes the value } i\}$ and N has a Geometric distribution with parameter q_i, so we can identify Z_i with X in Lemma 2.4.3, that is, $1_A = \xi(\{n\})$. Also, given that the m-th record has taken the value i, the value of each of the S_m near-records associated to it is in $[0,n]$ with probability $(y_{i-a} - y_n)/(y_{i-a} - y_i)$, independently of other near-records. That is, writing B_k for the event $\{\text{the } k\text{-th near-record associated to the record with value } i \text{ is in } [0,n]\}$, we can identify Z_i' with Y. Therefore, we obtain $Var(Z_i') \leq Var(Z_i) + E(Z_i)$ for all $i \geq 0$. This inequality, together with (2.31) and the mutual independence of the variables $Z_1, \ldots, Z_{n+a-1}$ yields

$$\sum_{i=0}^{n} Var(Z_i) \leq Var(\eta([0,n])) \leq \sum_{i=0}^{n+a-1} Var(Z_i) + \sum_{i=n+1}^{n+a-1} \mathbb{E}(Z_i). \tag{2.32}$$

The last term in the rhs of (2.32) is negligible, since, under condition (1) it is uniformly bounded in n and, under condition (2) $\mathbb{E}(Z_i) \leq Var(Z_i)$ for all large enough i by (2.19) and (2.20).

Finally, from Proposition 2.4.1 (d), we have that

$$\frac{\sum_{i=0}^{n} Var\,(Z_i)}{\sum_{i=0}^{n+a-1} Var\,(Z_i)} \to 1 \text{ as } n \to \infty,$$

which, together with (2.30) and (2.32) proves the result. □

2.5 A martingale related to the counting of δ-records

Martingales have been used successfully in the study of the counting process of records and δ-records [48, 49, 50, 52, 53]. For usual records, the sequence $N_n - cM_n$ is a martingale when the underlying distribution is exponential, where $c > 0$ is a positive constant. This fact was taken as a starting point in the paper [49], which found all the probability distributions such that $N_n - cM_n$ is a martingale. In the particular case of discrete random variables taking values in the integers, the geometric distribution is essentially the only distribution with this property when $0 < c < 1$, and there is no solution for $c \geq 1$. In this section we pose the same problem for δ-records. Here, the solution will depend both on the sign of δ and the value of $c > 0$.

Let us recall some definitions. Given a probability space $(\Omega, \mathcal{F}, P)$, a sequence $(\mathcal{F}_n)$ of sub-σ-fields of $\mathcal{F}$ is called a filtration if $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$, for all $n \geq 1$.

Definition 2.5.1. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_n)$ a filtration. A sequence (Z_n) of random variables is a martingale with respect to the filtration $(\mathcal{F}_n)$ if it satisfies:

(a) $\mathbb{E}(|Z_n|) < \infty$, for all $n \geq 1$,

(b) (Z_n) is adapted to $(\mathcal{F}_n)$, that is, Z_n is $\mathcal{F}_n$ measurable, for all $n \geq 1$,

(c) $\mathbb{E}(Z_n|\mathcal{F}_{n-1}) = Z_{n-1}$ almost surely, for all $n \geq 2$.

In this section we study the following problem. Let $\delta \neq 0$, $c > 0$ be fixed: find all the distributions F with support in $\mathbb{Z}$ having $\mathbb{E}(|X_1|) < \infty$ and such that the sequence (Z_n) is a martingale with respect to the filtration $(\sigma(X_1, \ldots, X_n))$, where

$$Z_n = N_{n,\delta} - cM_n, \qquad n \geq 1. \tag{2.33}$$

We impose that $E(|X_1|) < \infty$ because this is equivalent to $E(|M_n|) < \infty$; note that $N_{n,\delta} \leq n$, so the integrability of Z_n for $n \geq 1$ is equivalent to the integrability of X_1.

We begin with a proposition giving an equivalent condition for our problem.

Proposition 2.5.2. *Let (X_n) be a sequence of i.i.d. r.v. taking values in the integers with common cdf F such that $\mathbb{E}(|X_1|) < \infty$. Then (Z_n) defined in (2.33) is a martingale if and only if, for all $n \geq 2$,*

$$G(k+\delta) = c\sum_{j=k}^{\infty} G(j), \tag{2.34}$$

for every $k \in supp(F)$ where $G(x) := 1 - F(x)$ denotes the survival function.

Proof. Note first that, for $n \geq 2$,

$$\mathbb{E}(N_{n,\delta} - cM_n \mid \mathcal{F}_{n-1}) = N_{n-1,\delta} - cM_{n-1} + \mathbb{E}(1_{n,\delta} \mid \mathcal{F}_{n-1}) - c\mathbb{E}(M_n - M_{n-1} \mid \mathcal{F}_{n-1}).$$

Therefore, (Z_n) is a martingale if and only if

$$\mathbb{E}(1_{n,\delta} \mid \mathcal{F}_{n-1}) = c\mathbb{E}(M_n - M_{n-1} \mid \mathcal{F}_{n-1}), \tag{2.35}$$

for all $n \geq 2$. Note that the lhs of (2.35) is $\mathbb{P}(X_n > M_{n-1} + \delta)$. For the rhs we have

$$c\mathbb{E}\left(M_n - M_{n-1} \mid \mathcal{F}_{n-1}\right) = c\mathbb{E}\left((X_n - M_{n-1})^+ \mid \mathcal{F}_{n-1}\right),$$

where $x^+ = x \vee 0$. It is now obvious that

$$\begin{aligned}\mathbb{E}\left((X_n - M_{n-1})^+ \mid \mathcal{F}_{n-1}\right) &= \sum_{i=1}^{\infty} i\mathbb{P}\left(X_n = M_{n-1} + i\right) \\ &= \sum_{k=1}^{\infty}\sum_{i=k}^{\infty} \mathbb{P}\left(X_n = M_{n-1} + i\right) \\ &= \sum_{k=0}^{\infty} \left(1 - F(M_{n-1} + k)\right),\end{aligned}$$

so (2.35) can be written as

$$G(M_{n-1} + \delta) = c\sum_{k=0}^{\infty} G(M_{n-1} + k),$$

and thus the result holds. □

Let $x_- = \inf\{k : G(k) < 1\} \geq -\infty$ and $x_+ = \sup\{k : G(k) < 1\} \leq +\infty$ be leftmost and rightmost points of the distribution. The next proposition gives conditions on x_-, x_+ for (Z_n) to be a martingale and find a family of solutions in the case $\delta > 0$

Proposition 2.5.3. *In the setting of Proposition 2.5.2, we have:*

(a) If (Z_n) is a martingale, then $x_- > -\infty$;

(b) If $\delta < 0$ and (Z_n) is a martingale, then $x_+ = +\infty$.

Proof. (a) Suppose that (Z_n) is a martingale and $x_- = -\infty$. There exists a decreasing sequence (k_j) with $k_j \in supp(F)$, with $k_j \to -\infty$ as $j \to \infty$. By Proposition 2.5.2, we have

$$G(k_{j+1} + \delta) - G(k_j + \delta) = c\left(G(k_{j+1}) + \cdots + G(k_j - 1)\right). \tag{2.36}$$

Since G is a survival function, $G(k_j) \to 1$ as $j \to +\infty$. So the lhs of (2.36) converges to 0 while the rhs does not, which is a contradiction.

(b) Note that $x_+ \in supp(F)$, and that $G(x_+ - 1) > 0$ and $G(x_+) = 0$. Then, by Proposition (2.5.2), since $\delta < 0$:

$$0 > G(x_+ - 1) \geq G(x_+ + \delta) = \sum_{j=x_+}^{\infty} G(j) = 0,$$

which proves the result.

□

The last proposition shows that F must be bounded below for (Z_n) to be a martingale. Besides, for $\delta < 0$, the distribution cannot be bounded above. However, it does not rule out the posibility that the distribution can be bounded above in the case $\delta > 0$. In fact, under some conditions on the support, and on δ and c, there are bounded distributions for which (Z_n) is a martingale, when $\delta > 0$. We give a complete answer to this situation in the following result.

Proposition 2.5.4. *In the setting of Proposition 2.5.2, let $\delta > 0$ and $-\infty < x_- < x_+ < +\infty$. Let $x_- = k_1 < \cdots < k_m = x_+$ be the support of F. Then, (Z_n) is a martingale if and only if the four following conditions hold:*

1. *$k_i - k_{i-1} > \delta$, for all $i = 2, \ldots, m$,*
2. *$k_m - k_{m-1} = 1/c$,*
3. *$k_i - k_{i-1} < 1/c$, for all $i = 2, \ldots, m-1$,*
4. *$G(k_i) = G(k_1)\prod_{j=2}^{i}(1 - c(k_j - k_{j-1}))$ for all $i = 2, \ldots, m-1$.*

Proof. Applying (2.34) to k_{m-1} we get

$$G(k_{m-1} + \delta) = c(k_m - k_{m-1})G(k_{m-1}).$$

If $k_{m-1} + \delta \geq k_m$ then the lhs is 0 while the rhs is not, which is a contradiction. Therefore, k_m must be at least $k_{m-1} + \delta + 1$, so we assume this in the rest of the proof. In this case, $G(k_{m-1} + \delta) = G(k_{m-1})$ and we conclude $k_m - k_{m-1} = 1/c$.

Applying (2.34) to k_{m-2}, we get

$$G(k_{m-2} + \delta) = c(k_{m-1} - k_{m-2})G(k_{m-2}) + G(k_{m-1}). \tag{2.37}$$

Now, if $k_{m-2} + \delta \geq k_{m-1}$ then $G(k_{m-2} + \delta) = G(k_{m-1})$, leading to $G(k_{m-2}) = 0$, which is a contradiction. Then $k_{m-1} - k_{m-2}$ must be greater than δ and, in that case, (2.37) is equivalent to

$$G_{m-2}\left(1 - c(k_{m-1} - k_{m-2})\right) = G(k_{m-1}),$$

which implies $k_{m-1} - k_{m-2} < 1/c$.

Iterating the above procedure, we get

$$G(k_{i-1}) = c(k_i - k_{i-1})G(k_{i-1}) + G(k_i),$$

for $i = 2, \ldots, m-1$ and the result follows. □

Remark 2.5.5. (a) In Proposition 2.5.4 we leave out the trivial case $x_- = x_+$, that is, the random variables are all equal to the constant x_-. In this case, condition (2.34) is clearly satisfied so (Z_n) is a martingale. In fact $Z_n = 1 - x_-$ for all $n \geq 1$.

(b) The distribution given in Proposition 2.5.4 coincides with Example 3.4 in [49], where the problem was studied for records. Now, we need the extra assumption $k_i - k_{i-1} > \delta$, for $i = 2, \ldots, m$. Note that this condition asserts that there is no bounded distribution taking values in consecutive integers such that (Z_n) is a martingale.

Summarizing the above, we have proved that $x_- > -\infty$ is a necessary condition for (Z_n) to be a martingale; moreover, for $\delta < 0$ it is also necessary that $x_+ = +\infty$, while for $\delta > 0$, it is possible that $x_+ < +\infty$ and the necessary and sufficient conditions, together with the explicit expression of F are given in Proposition 2.5.4. Therefore, we are left with the case $-\infty < x_- < x_+ = +\infty$. In the rest of the section we assume that the support of F is $\{x_- + k : k \geq 0\}$; without loss of generality we can take $x_- = 0$, so the support of F is $\{0, 1, \ldots\}$. Note that if (Z_n) is a martingale for a distribution F, then it is also a martingale for the shifted distribution $\tilde{F}$, with $\tilde{F}(x) = F(x + b)$, $x \in \mathbb{Z}$.

We first consider the case $\delta < 0$.

Theorem 2.5.6. *In the setting of Proposition 2.5.2, if $\delta < 0$ and the support of F is $\{0, 1, \ldots\}$, we have:*

a) If $\delta < -1$ then (Z_n) is not a martingale;

b) If $\delta = -1$, then (Z_n) is a martingale if and only if F is the geometric distribution (starting at 0) with success parameter $1/(1+c)$.

Proof. a) When $\delta < -1$, applying condition (2.34) for $k = 1$ we get $G(1+\delta) = c\sum_{k=1}^{\infty} G(k)$. Since $\delta < -1$, the lhs is 0 while the rhs is not, so no G can satisfy (2.34).

(b) For $\delta = 0$, condition (2.34) gives, for $k \geq 0$,

$$G(k-1) = c\sum_{j=k}^{\infty} G(j),$$

$$G(k) = c\sum_{j=k+1}^{\infty} G(j),$$

and then, taking differences we get

$$G(k-1) = (c+1)G(k). \tag{2.38}$$

Since we have the initial condition $G(-1) = -1$, the solution for (2.38) is

$$G(k) = (c+1)^{-(k+1)},$$

and the result is proved.

□

Remark 2.5.7. Recall that δ-records for $\delta = -1$ are weak-records. Theorem 2.5.6 states that, while (Z_n) can be a martingale for weak-records, this cannot happen for any other negative δ.

We now turn to the case $\delta > 0$. Condition (2.34) generates a recurrent relation for $G(k)$, $k \geq 0$. This recurrence implies that the value of $G(k)$ can be written in terms of $G(0), \ldots, G(\delta)$. However, it might be the case that, depending on the values of $G(0), \ldots, G(\delta)$, the resulting G is not a survival function (that is, a nonnegative decreasing to 0 function). Note that the problem of deciding if the sequence of real numbers $(u_n)_n$ arising as a solution of a recurrence relation is non-negative is still an open problem in the literature [89]. Therefore, we will focus on the case $\delta = 1$ here.

Theorem 2.5.8. *In the setting of Proposition 2.5.2, if $\delta = 1$ and the support of F is $\{0, 1, \ldots\}$, we have:*

a) If $c > 1/4$, then (Z_n) is not a martingale;

b) If $c = 1/4$, then (Z_n) is a martingale if and only if $G(0) \in (0,1)$, $G(1) \in [G(0)/2, G(0))$ and

$$G(k) = 2^{-k}\big(G(0) + (2G(1) - G(0))k\big),$$

for $k \geq 2$;

c) If $0 < c < 1/4$, then (Z_n) is a martingale if and only if $G(0) \in (0,1)$, $G(1) \in [\lambda_2 G(0), G(0))$ and

$$G(k) = \frac{G(1) - G(0)\lambda_2}{\lambda_1 - \lambda_2}\lambda_1^k + \frac{G(0)\lambda_1 - G(1)}{\lambda_1 - \lambda_2}\lambda_2^k, \tag{2.39}$$

for $k \geq 2$, where

$$\lambda_1 = \frac{1}{2}\left(1 + \sqrt{1-4c}\right), \qquad \lambda_2 = \frac{1}{2}\left(1 - \sqrt{1-4c}\right). \tag{2.40}$$

Proof. Applying Proposition 2.5.2 with $\delta = 1$, for $k \geq 0$, we get that (Z_n) is a martingale if and only if

$$G(k+2) - G(k+1) + cG(k) = 0, \tag{2.41}$$

for all $k \geq 0$. That is, (Z_n) is a martingale if and only if the function G defined by (2.41), with $0 < G(1) < G(0) < 1$ is a survival function on $\{0, 1, \ldots\}$. Recurrence equations have been profusely studied in mathematics, and the general term can be computed as a function of the roots of their characteristic polynomial. In (2.41), the characteristic polynomial is $z^2 - z + c$, with discriminant $1 - 4c$. We study the cases (a), (b) and (c) separately.

a) If $c > 1/4$, then the roots of the characteristic polynomial are complex numbers. In this case, the proof of Lemma 5 in [61] remains valid in our setting, guaranteeing that the sequence $G(k)$ will be negative for some $n \in \mathbb{N}$ so no survival function can satisfy (2.41).

b) If $c = 1/4$, then $1/2$ is a double root of the characteristic polynomial. The solution of (2.41) is $G(k) = 2^{-k}(A + Bk)$, $k \geq 0$, with $A, B \in \mathbb{R}$. Taking $k = 0, 1$ we get $G(0) = A$, $G(1) = (A + B)/2$. We now look for conditions on $G(0), G(1)$ such that G is a survival function.

   First, $G(0) \in (0,1)$ if and only if $A \in (0,1)$; also $G(1) > 0 \Leftrightarrow A + B > 0$ and $G(1) < G(0) \Leftrightarrow A > B$. Moreover, $G(k+1) < G(k)$ is equivalent to $A > B - Bk$; since this must hold for every $k \geq 0$, it is equivalent to $B \geq 0$ (since $A > 0$). Also, it is clear that $G(k) \to 0$ as $k \to \infty$. Therefore, G is a survival function if and only if $0 \leq B < A < 1$, which is equivalent to $G(0) \in (0,1)$, $G(1) \in [G(0)/2, G(0))$, as stated.

c) The roots of the characteristic polynomial are λ_1 and λ_2 in (2.40) and the solution of (2.41) is $G(k) = A\lambda_1^k + B\lambda_2^k$, for $k \geq 1$, where

$$A = \frac{G(1) - G(0)\lambda_2}{\lambda_1 - \lambda_2}, \qquad B = \frac{G(0)\lambda_1 - G(1)}{\lambda_1 - \lambda_2}. \tag{2.42}$$

We now check under which conditions on A, B, the solution (2.39) is a survival function. We analyze the different cases depending on the sign of A and B. Note that $0 < \lambda_2 < \lambda_1 < 1$ so, by (2.42), A and B cannot be both equal to 0, so this case is excluded in the analysis below.

- ($A, B \geq 0$.) Since $G(0)$ must be smaller than 1, we have $A + B < 1$. It is clear that $A\lambda_1^k + B\lambda_2^k$ is positive, decreasing in k and $\lim_{k\to\infty} G(k) = 0$ because $\lambda_1, \lambda_2 \in (0, 1)$ so G is a survival function. Therefore, in this case, the solution of (2.41) is a survival function if and only if $A + B < 1$. Using the expressions (2.42), the conditions on A and B are equivalent to
$$0 < G(0)\lambda_2 \leq G(1) \leq G(0)\lambda_1 < G(0) < 1.$$
- ($A \geq 0, B \leq 0$.) Note first that $G(0) < 1$ if and only if $A + B < 1$. Also, the condition $G(k) > 0$ is equivalent to $A > |B|(\lambda_2/\lambda_1)^k$, for all $k \geq 1$; since $\lambda_2 < \lambda_1$, this is equivalent to $A + B > 0$. Also, the condition for G to be a decreasing function is
$$A\lambda_1^{k+1} + B\lambda_2^{k+1} < A\lambda_1^k + B\lambda_2^k \Longleftrightarrow A > |B| \left(\frac{\lambda_2}{\lambda_1}\right)^k \frac{1 - \lambda_2}{1 - \lambda_1}, \tag{2.43}$$
for every $k \geq 0$. Since $\lambda_2 < \lambda_1$, condition (2.43) needs only be checked for $k = 0$, which is
$$A > |B| \frac{1 + \sqrt{1 - 4c}}{1 - \sqrt{1 - 4c}}. \tag{2.44}$$
As $A + B > 0$ is implied by (2.44), the conditions on A, B in this case are $A + B < 1$ and (2.44). Using the expressions (2.42), we note that $A \geq 0$, $B \leq 0$ are equivalent to $G(1) \geq \lambda_2 G(0)$ and $G(1) \geq \lambda_1 G(0)$, where the former condition is implied by the latter. Also condition (2.44) is easily seen to be equivalent to $G(0) > G(1)$. Therefore, the conditions on $G(0), G(1)$ in this case are
$$0 < G(0)\lambda_1 \leq G(1) < G(0) < 1.$$
- ($A \leq 0, B \geq 0$.) $G(k) > 0$ is equivalent to $B > |A|(\lambda_1/\lambda_2)^k$. Since $\lambda_1 > \lambda_2$, this condition cannot hold for every $k \geq 1$ so G cannot be a survival function in this case.
- ($A, B \leq 0$.) In this case $G(0) = A + B \leq 0$, so G cannot be a survival function.

Summarizing the four cases above, we get that G is a survival function if and only if condition (2.39) holds.

□

Remark 2.5.9. In [49], where the problem was solved for usual records ($\delta = 0$) it was proved that the distributions F with support $\{0, 1, \ldots\}$ such that (Z_n) is a martingale are convex combinations of a Dirac delta on 0 and a Geometric distribution with parameter c. In this section we have proved that, for $\delta < -1$, no F with support $\{0, 1, \ldots\}$ has (Z_n) as a martingale. For $\delta = -1$, the only F is the Geometric distribution with parameter $1/(1 + c)$. For $\delta = 1$ a variety of situations arises: for $c < 1/4$ the general solution is a convex combination of the Dirac delta on 0 and two Geometric distributions with parameters λ_1 and λ_2; for $c = 1/4$, convex combinations of a Dirac delta and a Geometric distribution are obtained for $G(1) = G(0)/2$ but different distributions are found when $G(1) > G(0)/2$.

3

Probabilistic properties of δ-records in the Linear Drift Model

The study of records in the Linear Drift Model (LDM) has attracted much attention recently due to applications in several fields. In this chapter we study δ-records in the LDM. We give analytical properties of the probability of δ-records and study the correlation between δ-record events. We propose a first order approximation as a function of both the values of δ and the trend to study the δ-record probability. We assess our results via Montecarlo simulations finding that the approximations are accurate for a small-moderate number of observations. We also analyze the asymptotic behaviour of the number of δ-records among the first n observations and give conditions for convergence to the Gaussian distribution. As a consequence of our results, we solve a conjecture posed in the Physics literature regarding the total number of records in a LDM with negative drift. Examples of application to particular distributions, such as Uniform, Gumbel or Pareto are also provided. A generalization of the LDM where the underlying trend is random with a linear-growing expectation is also considered. Most part of these results have been published in [43], [44] and [76]

3.1 First steps in the study of δ-records in the Linear Drift Model

Throughout this chapter we assume that (Y_n) are random variables obeying the LDM as defined in equation (1.2) of Section 1.4. We briefly recall that we consider observations Y_n that follow

$$Y_n = X_n + cn, \quad n \geq 1,$$

where $c \in \mathbb{R}$ is the trend parameter and (X_n) is a sequence of i.i.d. random variables, with absolutely continuous cdf F and probability density function pdf f for which it exists an interval $I = (x_-, x_+)$, with $-\infty \leq x_- < x_+ \leq \infty$, such that $f(x) > 0$, for all $x \in I$, and $f(x) = 0$ otherwise. Also, we recall that the right-tail expectation of the X_j, that is $\mu^+ = \int_0^\infty x f(x)dx$, plays a critical role in the record occurrences in the LDM.

Our first aim is to compute the probability of an observation to be a δ-record, denoted by $p_{j,\delta}$, for the sequence of r.v. (Y_n) following the LDM. Note that the probability of the event $\{Y_j \text{ is a } \delta\text{-record}\}$ is the expectation of the indicator r.v. $1_{j,\delta}$.

Theorem 3.1.1. *Under the LDM, the probability of δ-record $p_{j,\delta}$ is*

$$p_{j,\delta} = \int_{-\infty}^{\infty} \prod_{i=1}^{j-1} F(x + ci - \delta) f(x)dx. \tag{3.1}$$

Moreover, the asymptotic δ-record probability, denoted by p_δ, is

$$p_\delta = \int_{-\infty}^{\infty} \prod_{i=1}^{\infty} F(x + ci - \delta) f(x)dx. \tag{3.2}$$

Proof. From the definition of δ-record in the LDM and by conditioning on the value of X_j we have

$$\begin{aligned} p_{j,\delta} &= \mathbb{P}\left(X_j + cj > \bigvee_{i=1}^{j-1} (X_i + ci + \delta) \right) \\ &= \int_{-\infty}^{\infty} \mathbb{P}\left(x > \bigvee_{i=1}^{j-1} (X_i - c(j-i) + \delta) \right) f(x)dx \\ &= \int_{-\infty}^{\infty} \mathbb{P}\left(\bigcap_{i=1}^{j-1} \{X_i < x + c(j-i) - \delta\} \right) f(x)dx \\ &= \int_{-\infty}^{\infty} \prod_{i=1}^{j-1} F(x + ci - \delta) f(x)dx. \end{aligned}$$

Moreover, the asymptotic δ-record probability is given by the formula

$$p_\delta := \lim_{n\to\infty} p_{n,\delta} = \int_{-\infty}^{\infty} \prod_{i=1}^{\infty} F(x + ci - \delta) f(x) dx, \tag{3.3}$$

which is mathematically justified by the monotone convergence theorem for integrals; see [13, Theorem 2.8.2]. □

In what follows we occasionally write $1_{j,\delta}(c), N_{n,\delta}(c), p_{j,\delta}(c), p_\delta(c)$, etc. to emphasize the dependence on the trend parameter c.

3.2 Properties of the δ-record probabilities

We begin with a simple property about the asymptotic δ-record probability of an affine transformation of the LDM, which can be easily checked from the δ-record definition.

Proposition 3.2.1. *Let $\tilde{X}_n = bX_n + a$, with $b > 0$, $a \in \mathbb{R}$, and $\tilde{Y}_n = \tilde{X}_n + cn, n \geq 1$. If $\tilde{p}_{j,\delta}(c)$ and $\tilde{p}_\delta(c)$ are the analogous δ-record probability and asymptotic δ-record probability in this model, then it holds*

$$\tilde{p}_{j,\delta}(c) = p_{j,\frac{\delta}{b}}(\tfrac{c}{b}).$$

and

$$\tilde{p}_\delta(c) = p_{\frac{\delta}{b}}(\tfrac{c}{b}).$$

Proof. For fixed j, $\tilde{Y}_j$ is a δ-record if

$$\tilde{X}_j + cj > \bigvee_{i=1}^{j-1} \left(\tilde{X}_i + ci + \delta\right) \Leftrightarrow X_j + \frac{c}{b} j > \bigvee_{i=1}^{j-1} \left(X_i + \frac{c}{b} i + \frac{\delta}{b}\right),$$

which is the δ-record condition for δ/b in a LDM with trend parameter c/b, and so both results hold. □

We consider next some analytical properties of $p_{j,\delta}(c)$ and $p_\delta(c)$, as functions of c and δ. We note first that both are increasing in c and decreasing in δ. Moreover, it is easy to see that $p_{j,\delta}(c)$ is decreasing in j and continuous in c, converging to 1 as $c \to \infty$. The continuity of $p_\delta(c)$ is less clear because of the infinite product within the integral in (3.2).

3.2.1 Positivity of $p_\delta(c)$

We show that the positivity of $p_\delta(c)$ depends on c and δ and on the right-tail behaviour of F. The characterization of the positivity of these quantities reveals the importance of the parameter μ^+ in the δ-record behaviour in the LDM.

Theorem 3.2.2. *$p_\delta(c) > 0$ if and only if $\mu^+ < \infty$ and one of the following conditions holds*

1. *$c > 0$ and $\delta < x_+ - x_- + c$,*
2. *$c = 0$, $\delta < 0$ and $x_+ < \infty$.*

Proof. **1**. $\mu^+ = \infty$. In this case $p_\delta(c) = 0$, for all $\delta, c \in \mathbb{R}$.

To justify this claim, we show that $\prod_{j=1}^{\infty} F(x + cj - \delta) = 0$, for all $x \in (x_-, x_+)$.

a) If $c < 0$ the conclusion is immediate because $F(x + cj - \delta) \to 0$, as $j \to \infty$.

b) If $c = 0$, we note that $\mu^+ = \infty$ implies $x_+ = \infty$ and so, $F(x - \delta) < 1$. Thus $\prod_{j=1}^{\infty} F(x + cj - \delta) = 0$.

c) Finally, if $c > 0$, we note that $\mu^+ = \infty$ implies $\sum_{j=1}^{\infty}(1 - F(x + cj - \delta)) = \infty$, which in turn implies $\prod_{j=1}^{\infty} F(x + cj - \delta) = 0$. This follows from the definition of μ^+ and from Taylor's expansion of $\log(1 + x)$.

2. $\mu^+ < \infty$. As in the previous case, we have three situations depending on the sign of c.

a) For $c < 0$, $p_\delta(c) = 0$, for all $\delta \in \mathbb{R}$, since $\prod_{j=1}^{\infty} F(x + cj - \delta) = 0$, for all $x \in (x_-, x_+)$.

b) If $c = 0$,

$$p_\delta(0) = \int_{-\infty}^{\infty} \prod_{j=1}^{\infty} F(x - \delta) f(x) dx = \int_{x_+ + \delta}^{\infty} f(x) dx, \tag{3.4}$$

which is positive if and only if $x_+ < \infty$ and $\delta < 0$.

c) Finally, if $c > 0$, then $p_\delta(c) = 0$ if and only if $x_+ - x_- \leq \delta - c$. Indeed, note that, if $x_+ - x_- \leq \delta - c$, then $\mathbb{P}(Y_n > Y_{n-1} + \delta) = 0$, for all n, and so, only the first observation (by convention) is a δ-record. Conversely, if $x_+ - x_- > \delta - c$, then the interval $J := (x_-, x_+) \cap (x_- - c + \delta, \infty)$ is nonempty and, for every $x \in J$, we have $F(x + cj - \delta) \geq F(x + c - \delta) > 0$, for all j. Now, since $F(x + cj - \delta) \to 1$ as $j \to \infty$, and $\mu^+ < \infty$, we have $\sum_{j=1}^{\infty}(1 - F(x + cj - \delta)) < \infty$, which implies $\prod_{j=1}^{\infty} F(x + cj - \delta) > 0$ and, so $p_\delta(c) > 0$.

□

Remark 3.2.3. Distributions with $\mu^+ = \infty$ can be considered as "right-heavy-tailed" and we observe that, for such distributions, the linear trend has no impact on the asymptotic probability of a δ-record. This class of distributions includes the Pareto and Fréchet, with shape parameter $\alpha \in (0, 1]$.

3.2.2 Continuity of $p_\delta(c)$

As commented at the beginning of this section, the continuity of $p_\delta(c)$ as a function of δ and c is not obvious. However, thanks to Theorem 3.2.2 we can restrict our attention to distributions F with finite right-tail expectation since, otherwise, $p_\delta(c)$ is identically zero and continuity is trivial. Thus, we assume throughout this subsection that $\mu^+ < \infty$.

It is obvious that the continuity of $p_\delta(c)$ will depend critically on the behaviour of the infinite product appearing in its expression (3.2). Proposition 3.2.4 shows the continuity of the infinite product under certain hypotheses that will be used later in the proof of Theorem 3.2.5, where a full characterization of the continuity of $p_\delta(c)$ is stated.

Proposition 3.2.4. $\prod_{j=1}^{\infty} F(x + cj - \delta)$, *as a function of* c *is continuous at* $c \in \mathbb{R} \setminus \{0\}$, *for every* $x \in (x_-, x_+)$, $x \neq x_- + \delta - c$.

Proof. Let $(c_n)_{n\geq 1}$ be a real sequence converging to $c > 0$. We show that

$$\prod_{j=1}^{\infty} F(x + c_n j - \delta) \to \prod_{j=1}^{\infty} F(x + cj - \delta), \tag{3.5}$$

as $n \to \infty$, for fixed $x \in (x_-, x_+)$, $x \neq x_- + \delta - c$.

Let $x \in (x_-, x_+)$ be such that $x < x_- + \delta - c$ (this can only happen if $x_- > -\infty$ and $\delta - c > 0$). In this case $F(x + c - \delta) = 0$, so the rhs of (3.5) is 0. Also, since $c_n \to c$, $F(x + c_n - \delta) = 0$, for n large enough, the lhs of (3.5) is also 0 and (3.5) is proved.

Let now $x > x_- + \delta - c$, then $F(x + cj - \delta) > 0$ for all $j \geq 1$. Let $\epsilon > 0$ such that $x + c - \epsilon - \delta > x_-$ and let $n_0 \geq 1$, such that $|c_n - c| < \epsilon$, for all $n \geq n_0$. We have, for $n \geq n_0$,

$$-\log F(x + c_n j - \delta) \leq -\log F(x + (c - \epsilon)j - \delta).$$

Since $x > x_- + \delta - (c - \epsilon)$ and $\mu^+ < \infty$, we have $-\sum_{j=1}^{\infty} \log F(x + (c - \epsilon)j - \delta) < \infty$, so the dominated convergence theorem yields

$$\sum_{j=1}^{\infty} \log F(x + c_n j - \delta) \to \sum_{j=1}^{\infty} \log F(x + cj - \delta),$$

as $n \to \infty$, so (3.5) also holds for $x > x_- + \delta - c$. Finally, for $c < 0$, we have $\prod_{j=1}^{\infty} F(x + cj - \delta) = 0, \forall x \in \mathbb{R}$, since $F(x + cj - \delta) \to 0$, as $j \to \infty$. □

In the following theorem we summarize the conditions for the continuity of $p_\delta(c)$.

Theorem 3.2.5. *The asymptotic δ-record probability $p_\delta(c)$, as a function of c, δ, is*

(a) continuous at every $c \neq 0$ and right-continuous at $c = 0$, for all δ;

(b) discontinuous at $c = 0$ if and only if $x_+ < \infty, \delta < 0$, and

(c) continuous in δ, for all c.

Proof.

a) For $c \neq 0$, Proposition 3.2.4 states that $\prod_{j=1}^{\infty} F(x + cj - \delta)$ is continuous at every $c \neq 0$, for every $x \in (x_-, x_+)$, such that $x \neq x_- + \delta - c$. Then, thanks to the dominated convergence theorem, we conclude that $p_\delta(c)$ is continuous, at every $c \neq 0$.

The continuity at $c = 0$ is subtler to establish and it depends on the sign of δ and the finiteness of x_+, the right-end point of F. Note that, for every $c > 0$ and $N \geq 1$, we have

$$\prod_{j=1}^{\infty} F(x - \delta) \leq \prod_{j=1}^{\infty} F(x + cj - \delta) \leq \prod_{j=1}^{N} F(x + cj - \delta).$$

Then, taking the limit as $c \to 0^+$ in the above inequalities,

$$\prod_{j=1}^{\infty} F(x - \delta) \leq \lim_{c \to 0^+} \prod_{j=1}^{\infty} F(x + cj - \delta) \leq F(x - \delta)^N.$$

Therefore, $\lim_{c \to 0^+} \prod_{j=1}^{\infty} F(x + cj - \delta)$ is 0 if $x < x_+ + \delta$, and 1 otherwise. Then, by the dominated convergence theorem,

$$\lim_{c \to 0^+} p_\delta(c) = \int_{-\infty}^{\infty} \lim_{c \to 0^+} \prod_{j=1}^{\infty} F(x + cj - \delta) f(x) dx = \int_{x_+ + \delta}^{\infty} f(x) dx.$$

Thus, $p_\delta(c)$ is right-continuous at $c = 0$ by (3.4).

b) Regarding left-continuity at 0, recall that $p_\delta(c) = 0$ for $c < 0$. So, $p_\delta(c)$ is discontinuous at 0 if and only if $x_+ < \infty$ and $\delta < 0$.

c) We now show the continuity of $p_\delta(c)$ as a function of δ. The result is trivial if $c < 0$, since $p_\delta(c) = 0$, for all $\delta \in \mathbb{R}$. For $c = 0$ note that, by (3.4), $p_\delta(0) = 1 - F(x_+ + \delta)$, which is continuous since F is a continuous function.

If $c > 0$ and $(\delta_n)_{n\geq 1}$ is a sequence converging to δ, we prove that

$$\lim_{n\to\infty} \prod_{j=1}^{\infty} F(x + cj - \delta_n) = \prod_{j=1}^{\infty} F(x + cj - \delta), \tag{3.6}$$

for all $x \in (x_-, x_+)$, $x \neq x_- + \delta - c$. Indeed, let $x < x_- + \delta - c$, then $F(x+c-\delta) = 0$ yielding $\prod_{j=1}^{\infty} F(x + cj - \delta) = 0$. Also $F(x + c - \delta_n) = 0$ for n large enough and (3.6) follows. Let now $x > x_- + \delta - c$ and $\varepsilon > 0$ such that $x + c - \delta - \varepsilon > x_-$. Then, for n large enough, we have $| \delta_n - \delta | < \varepsilon$ and

$$-\sum_{j=1}^{\infty} \log F(x + cj - \delta_n) \leq -\sum_{j=1}^{\infty} \log F(x + cj - (\delta + \varepsilon)) < \infty,$$

since $\mu^+ < \infty$. So (3.6) holds, and continuity follows. □

3.3 Exactly solvable models

In general it is not possible to compute the values $p_{j,\delta}$ or p_δ exactly. We show below explicit results for the Gumbel distribution, the uniform and an extreme value distribution in the Weibull class (see Section 1.3.4), and for particular instances of the Dagum family of distributions [71] in the Fréchet class.

3.3.1 The Gumbel distribution

The Gumbel distribution is a key probability distribution in the context of extreme-value theory, as it arises naturally to describe the behaviour of the partial maxima of i.i.d. observations drawn from distributions in the Gumbel-domain, as it was pointed in Section 1.3.4.

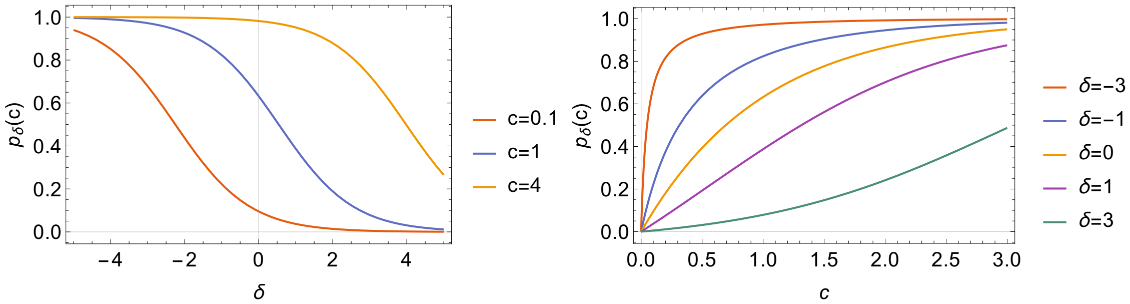
Example 3.3.1. Let $F(x) = \exp(-\exp(-x))$, for $x \in \mathbb{R}$, be the cdf of the (standard) Gumbel distribution. Note first that we have

$$F(x + cj - \delta) = F(x)^{e^{-cj+\delta}}.$$

Then, if $c \neq 0$,

$$\prod_{j=1}^{n-1} F(x + cj - \delta) = F(x)^{\sum_{j=1}^{n-1} e^{-cj+\delta}} = F(x)^{e^{\delta} \frac{e^{-c} - e^{-nc}}{1 - e^{-c}}}.$$

Figure 3.1
Asymptotic δ-record probability $p_\delta(c)$ for the Gumbel distribution as a function of δ and c.



So, from (3.1) we get

$$p_{n,\delta}(c) = \int_{-\infty}^{\infty} F(x)^{e^\delta \frac{e^{-c}-e^{-nc}}{1-e^{-c}}} f(x)dx$$
$$= \frac{1-e^{-c}}{1-e^{-c}+e^\delta(e^{-c}-e^{-nc})}.$$

If $c=0$,

$$\prod_{j=1}^{n-1} F(x+cj-\delta) = F(x)^{(n-1)e^\delta},$$

which yields

$$p_{n,\delta}(0) = \frac{1}{(n-1)e^\delta+1}.$$

Taking limits as $n \to \infty$ in the above formulas, we obtain

$$p_\delta(c) = \frac{1-e^{-c}}{e^\delta e^{-c}+1-e^{-c}} = \frac{1}{1+\frac{e^{-c}}{1-e^{-c}}e^\delta},$$

if $c>0$ and $p_\delta(c)=0$, if $c \leq 0$, as expected from Theorem 3.2.2.

Also, for every $c>0$, $p_\delta(c)$ decreases with δ as a logistic function of $-\delta$. Figure 3.1 shows the behaviour of $p_\delta(c)$ as a function of δ and c.

3.3.2 Distributions in the Weibull class

In this subsection we analyze two distributions in the Weibull class of extreme value distributions. These distributions have $x_+ < \infty$.

Example 3.3.2. The Type III max-stable distribution has cdf given by $F(x)=e^x$, for $x<0$ and $F(x)=1$ if $x \geq 0$, so the pdf is $f(x)=e^x$ if $x<0$ and 0 otherwise. Note that this is the limiting distribution of the maxima of i.i.d. r.v. in the Weibull class with $\alpha=1$, as we saw in Section 1.3.4.

The probability $p_{n,\delta}(c)$ can be computed from (3.1) by splitting the integral depending on the support of the cdfs in the product. Indeed, taking a positive trend $c > 0$, and $\delta \leq c$, we have

$$\begin{aligned}
p_{n,\delta}(c) &= \int_{-\infty}^{\infty} \prod_{j=1}^{n-1} \left(e^{x+cj-\delta} 1_{\{x+cj-\delta<0\}} + 1_{\{x+cj-\delta\geq 0\}} \right) e^x 1_{\{x<0\}} dx \\
&= \int_{\delta-c}^{0} e^x dx + \sum_{k=1}^{n-2} \left(\int_{\delta-c(k+1)}^{\delta-ck} \prod_{j=1}^{k} \left(e^{x+cj-\delta} \right) e^x dx \right) \\
&\quad + \int_{-\infty}^{\delta-c(n-1)} \prod_{j=1}^{n-1} \left(e^{x+cj-\delta} \right) e^x dx \\
&= \int_{\delta-c}^{0} e^x dx + \sum_{k=1}^{n-2} \left(e^{-k\delta} e^{c\frac{k(k+1)}{2}} \int_{\delta-c(k+1)}^{\delta-ck} e^{(k+1)x} dx \right) \\
&\quad + e^{c\frac{n(n-1)}{2}} e^{-\delta(n-1)} \int_{-\infty}^{\delta-c(n-1)} e^{nx} dx \\
&= 1 - e^{\delta-c} + \sum_{k=1}^{n-2} \left(\frac{1}{k+1} e^{-k\delta} e^{c\frac{k(k+1)}{2}} \left(e^{(k+1)(\delta-ck)} - e^{(k+1)(\delta-c(k+1))} \right) \right) \\
&\quad + \frac{1}{n} e^{c\frac{n(n-1)}{2}} e^{-\delta(n-1)} e^{n(\delta-c(n-1))} \\
&= 1 - e^{\delta} e^{-c} + \sum_{k=1}^{n-2} \left(\frac{e^{\delta} e^{-c\frac{k(k+1)}{2}} \left(1 - e^{-c(k+1)}\right)}{k+1} \right) + \frac{1}{n} e^{\delta} e^{-c\frac{n(n-1)}{2}} \\
&= 1 - e^{\delta} \left(e^{-c} - \sum_{k=1}^{n-2} \left(\frac{e^{-c\frac{k(k+1)}{2}} \left(1 - e^{-c(k+1)}\right)}{k+1} \right) - \frac{1}{n} e^{-c\frac{n(n-1)}{2}} \right).
\end{aligned}$$

As it can be seen in the last expression, not only can $p_{n,\delta}(c)$ be explicitly computed, but also the effect of the δ parameter can be isolated to assess its influence on the δ-record probability. Now, the computation of the asymptotic δ-record probability is straightforward by taking the limit as n goes to infinity, yielding

$$p_\delta(c) = 1 - e^{\delta} \left(e^{-c} - \sum_{k=1}^{\infty} \left(\frac{e^{-c\frac{k(k+1)}{2}} \left(1 - e^{-c(k+1)}\right)}{k+1} \right) \right).$$

Example 3.3.3. Let us consider the case where X_n is a sequence of uniform r.v. For simplicity we will take $X_n \sim U(0,1)$, $-1 < \delta \leq 0$ and $c \in (0,1)$, although the computation is easily adapted for other values of the parameters. In this case we have, $F(x) = x1_{\{x\in(0,1)\}}$, if $x < 1$, and $F(x) = 1$, if $x \geq 1$, while the pdf is $f(x) = 1_{\{x\in(0,1)\}}$.

We first note that if $\delta \leq c - 1$, then every observation will be a δ-record, and thus $p_{n,\delta}(c) = 1$ for all $n \in \mathbb{N}$, so let us choose $\delta > c - 1$. We use (3.1) by splitting the integration domain conveniently, which for $n \geq 2$ yields

$$p_{n,\delta}(c) = \int_{1+\delta-c}^{1} 1dx + \sum_{k=1}^{\min\{n-2,\lfloor\frac{1+\delta}{c}\rfloor-1\}} \left(\int_{1+\delta-c(k+1)}^{1+\delta-ck} \prod_{j=1}^{k}(x+c-\delta)dx \right)$$
$$+ \int_{0}^{1+\delta-c(1+\min\{n-2,\lfloor\frac{1+\delta}{c}\rfloor-1\})} \prod_{j=1}^{1+\min\{n-2,\lfloor\frac{1+\delta}{c}\rfloor-1\}} (x+c-\delta)dx. \quad (3.7)$$

We notice that the products appearing in the integrals can be written as rising factorials (or Pochhammer Function), where

$$\prod_{j=1}^{n}(x+cj-\delta) = \prod_{j=1}^{n}((x-\delta)+cj) = \sum_{j=1}^{n+1} \begin{bmatrix} n+1 \\ j \end{bmatrix} (x-\delta)^{j-1}c^{n+1-j},$$

and $\begin{bmatrix} n \\ m \end{bmatrix}$ represents the (unsigned) Stirling numbers of first kind. Now, from (3.7), and writing $m(c,\delta,n) := \min\{n-2, \lfloor\frac{1+\delta}{c}\rfloor - 1\}$, we have

$$p_{n,\delta}(c) = \int_{1+\delta-c}^{1} 1dx + \sum_{k=1}^{m(c,\delta,n)} \left(\int_{1+\delta-c(k+1)}^{1+\delta-ck} \sum_{j=1}^{k+1} \left(\begin{bmatrix} k+1 \\ j \end{bmatrix} (x-\delta)^{j-1}c^{k+1-j} \right) dx \right)$$
$$+ \int_{0}^{1+\delta-c(1+m(c,\delta,n))} \sum_{j=1}^{2+m(c,\delta,n)} \left(\begin{bmatrix} 2+m(c,\delta,n) \\ j \end{bmatrix} (x-\delta)^{j-1}c^{2+m(c,\delta,n)-j} \right) dx$$
$$= \int_{1+\delta-c}^{1} 1dx + \sum_{k=1}^{m(c,\delta,n)} \left(\sum_{j=1}^{k+1} \left(\begin{bmatrix} k+1 \\ j \end{bmatrix} c^{k+1-j} \int_{1+\delta-c(k+1)}^{1+\delta-ck} (x-\delta)^{j-1}dx \right) \right)$$
$$+ \sum_{j=1}^{2+m(c,\delta,n)} \left(\begin{bmatrix} 2+m(c,\delta,n) \\ j \end{bmatrix} c^{2+m(c,\delta,n)-j} \int_{0}^{1+\delta-c(1+m(c,\delta,n))} (x-\delta)^{j-1}dx \right).$$

The last step is to compute the integrals to get a closed expression for $p_{n,\delta}(c)$:

$$p_{n,\delta}(c) = c - \delta + \sum_{k=1}^{m(c,\delta,n)} \left(\sum_{j=1}^{k+1} \left(\begin{bmatrix} k+1 \\ j \end{bmatrix} \frac{(1-ck)^j - (1-c(k+1))^j}{j} c^{k+1-j} \right) \right)$$
$$+ \sum_{j=1}^{2+m(c,\delta,n)} \left(\begin{bmatrix} 2+m(c,\delta,n) \\ j \end{bmatrix} \frac{(1-c\,m(c,\delta,n))^j - (-d)^j}{j} c^{2+m(c,\delta,n)-j} \right).$$

Finally, note that a general feature for r.v. bounded both to the left and to the right, which can be derived directly from (3.7), is that $p_{n,\delta}(c) = p_\delta(c)$ for all n large enough. In particular, for the uniform distribution we have $p_{n,\delta}(c) = p_\delta(c)$ for all $n \geq \lceil(1+\delta)/c\rceil$.

3.3.3 The Dagum family of distributions

In this subsection, we obtain the growth rate and explicit probabilities for distributions in the Dagum Family. The interest is to prove novel explicit results, not only for the case of δ-records but also for usual records. Furthermore, this example allows us to conclude that in the case $\mu^+ = \infty$ we still find a different qualitative behaviour of δ-records in the long term as a function of δ. Note that $\mu^+ = \infty$ implies that the variables X_n take large values, and so a negligible impact of the δ parameter could be expected.

Example 3.3.4. The Dagum distribution has cdf given by

$$F(x) = \left(1 + \left(\frac{x}{b}\right)^{-a}\right)^{-q} 1_{\{x\geq 0\}}, \tag{3.8}$$

where a, b, q are positive parameters. Note that if $q = 1$, the distribution is referred to as log-logistic [105]. Also, the Pareto distribution [3] with cdf

$$F(x) = (1 - 1/x)1_{\{x\geq 1\}}, \tag{3.9}$$

can be seen as a shifted version of the Dagum family, with $a = b = q = 1$. For simplicity, in this section we limit our attention to the case $a = 1$, which has $\mu^+ = \infty$.

By Theorem 3.2.2 we know that $p_\delta(c) = 0$, for every $c, \delta \in \mathbb{R}$, so we chose to analyze the speed of convergence of $p_{n,\delta}(c)$ to 0, for some values of c, δ. To that end, observe that the formula for $p_{n,\delta}$ takes the manageable form

$$p_{n,\delta}(c) = \int_{(\delta-c)^+}^{\infty} \prod_{j=1}^{n-1} \left(\frac{x + cj - \delta}{x + b + cj - \delta}\right)^q f(x)dx, \tag{3.10}$$

which becomes simpler if we further assume that $c = b$ (that is, the trend parameter of the LDM is equal to the scale parameter of the distribution). From (3.10) we get

$$p_{n,\delta}(c) = \int_{(\delta-c)^+}^{\infty} \left(\frac{x + c - \delta}{x + cn - \delta}\right)^q f(x)dx. \tag{3.11}$$

We introduce the notation $p^{(q)}_{n,\delta}(c)$ to make explicit the dependence of $p_{n,\delta}(c)$ on q. First, for records ($\delta = 0$) we have,

$$\begin{aligned} p^{(q)}_{n,0}(c) =& cq \int_0^{\infty} x^{q-1} (x + cn)^{-q} (x + c)^{-1} dx \\ =& qn^{-q} \int_0^1 t^{q-1}(1 - t(n-1)/n)^{-q} dt \qquad (3.12) \\ =& \frac{q}{(n-1)^q} \int_1^n \frac{(y-1)^{q-1}}{y} dy, \qquad (3.13) \end{aligned}$$

where the second equality follows from the change of variable $x = ct/(1-t)$ and the third from $1 - t(n-1)/n = 1/y$.

Observe that (3.12) and (3.13) do not depend on c and so, for the sake of simplicity, we write $p_{n,0}^{(q)}$. Moreover, from (3.12) we see that

$$p_{n,0}^{(q)} = n^{-q} \ {}_2F_1\left(q, q; q+1; (n-1)/n\right),$$

where ${}_2F_1$ is the Gauss hypergeometric function.

Also, from (3.13) and using the binomial expansion, for $q = 1, 2, \ldots$, we readily obtain

$$p_{n,0}^{(q)} = \frac{q}{(n-1)^q}\left((-1)^{q-1}\log n + \sum_{k=1}^{q-1}\binom{q-1}{k}\frac{(-1)^{q-1-k}}{k}(n^k - 1)\right). \tag{3.14}$$

The asymptotic behaviour of $p_{n,0}^{(q)}$, for any $q \in (0, \infty)$, can be obtained from (3.13). For $q = 1$, (3.14) yields $p_{n,0}^{(1)} = \frac{1}{n-1}\log n$. For $q > 1$, the leading term in the integral in (3.13) is y^{q-2}, so $p_{n,0}^{(q)} \sim \frac{q}{q-1}\frac{1}{n}$. For $q \in (0,1)$, the integral in (3.13) converges and, using formula 3.191.2 in [57], we get

$$p_{n,0}^{(q)} \sim n^{-q} q \int_1^\infty \frac{(y-1)^{q-1}}{y} dy = n^{-q} q\Gamma(1-q)\Gamma(q).$$

Thus,

$$p_{n,0}^{(q)} \sim \begin{cases} n^{-q} q\Gamma(1-q)\Gamma(q), & \text{if } 0 < q < 1, \\ \log(n)/n, & \text{if } q = 1, \\ n^{-1}\frac{q}{q-1}, & \text{if } q > 1. \end{cases} \tag{3.15}$$

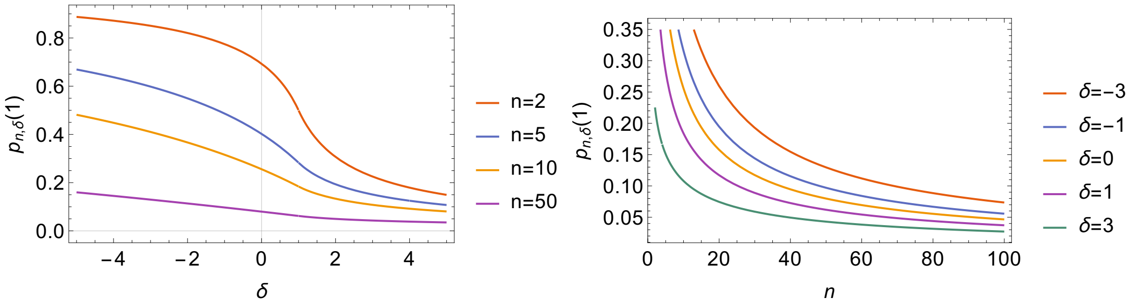
It is interesting to observe that the limiting behaviour of $p_{n,0}^{(q)}$, as a function of the power of the tail q, seems to match the asymptotic behaviour of $p_{n,0}(c)$ when F is the Fréchet distribution ($F(x) = \exp(-x^{-1})$, $x > 0$) and the tuning parameter is the trend c, studied in [20].

We now consider $\delta \neq 0$ and investigate whether $p_{n,\delta}^{(q)}/p_{n,0}^{(q)} \to 1$, as $n \to \infty$. This result can be expected since, as $\mu^+ = \infty$, the variables X_n take very large values, so δ may have little influence on the probability of δ-record, in the long term.

From (3.11) we may evaluate $p_{n,\delta}^{(q)}$, for any $q \in \mathbb{N}$, although the computation becomes lengthy as q grows. We have carried out the computation with values of q from 1 to 7, and obtained

$$p_{n,\delta}^{(1)} \sim \frac{\log(n)}{n}, \qquad p_{n,\delta}^{(q)} \sim \frac{q}{q-1}\frac{1}{n}, \quad q = 2, \ldots, 7.$$

Figure 3.2
δ-record probability $p_{n,\delta}(c)$ for the Pareto distribution as a function of δ and n with $c = 1$.



So, from (3.15) we have $p^{(q)}_{n,\delta}/p^{(q)}_{n,0} \to 1$, at least for $q = 1, \dots, 7$.

For noninteger values of $q \in (0, \infty)$, the limit behaviour of (3.11) is harder to analyze. To get a tractable expression, we impose $\delta = c$. Proceeding as above, we have, for $n > 2$,

$$p^{(q)}_{n,\delta} = \frac{q(n-1)^q}{(n-2)^{2q}} \int_1^{n-1} \frac{(y-1)^{2q-1}}{y^{q+1}} dy.$$

Therefore, we have

$$p^{(q)}_{n,\delta} \sim \begin{cases} n^{-q} \frac{\Gamma(2q)\Gamma(1-q)}{\Gamma(q)}, & \text{if } 0 < q < 1, \\ \log(n)/n, & \text{if } q = 1, \\ n^{-1} \frac{q}{q-1}, & \text{if } q > 1. \end{cases}$$

So, under the above stated conditions, $p^{(q)}_{n,\delta} \sim p^{(q)}_{n,0}$, for $q \geq 1$, but this is not the case if $q \in (0, 1)$.

To conclude this example we study the Pareto distribution, defined in (3.9), taking $c = 1$. From (3.11), the probability of δ-record is explicitly computed as

$$\begin{aligned} p_{n,\delta} &= \int_{\max\{1,\delta\}}^{\infty} \frac{x-\delta}{x^2(x+n-1-\delta)} dx \\ &= \frac{1}{(n-1-\delta)^2} \left((n-1) \log(\tfrac{n-\min\{1,\delta\}}{\max\{1,\delta\}}) - \min\{1,\delta\}(n-1-\delta) \right), \end{aligned} \tag{3.16}$$

if $\delta \neq n-1$ and $p_{n,\delta} = \frac{1}{2(n-1)}$, if $\delta = n-1$. Figure 3.2 shows the behaviour of $p_{n,\delta}$ as a function of n and δ.

3.4 First order approximations for the δ-record probability

The quantity $p_{n,0}$ for the non-asymptotic setting and usual record probabilities in the LDM was studied in [33] by means of first order Taylor approximations. Also, it was suggested in [120] to apply this methodology to δ-records in the CRM (Classical Record Model). Here, we will develop this idea for the CRM and the LDM for the case $\delta \neq 0$.

First, we study the behaviour of the δ-record probability, $p_{n,\delta}$ when the r.v. are (X_n) are drawn from the CRM. We recall that in this setting

$$p_{n,\delta} = \mathbb{P}\left(X_n > \bigvee_{i=1}^{n-1} X_i + \delta\right) = \int_{-\infty}^{\infty} F(x-\delta)^{n-1} F(dx). \tag{3.17}$$

This last integral is usually not analytically solvable, so in order to compute it for a wider range of distributions, we will explore the possibility of a first order approximation for $F(x-\delta)$ at the point δ. Taking $\mid \delta \mid \ll 1$, we have

$$F(x-\delta) = F(x) - \delta f(x) + \mathcal{O}(\delta^2),$$

and thus

$$F(x-\delta)^{n-1} = F(x)^{n-1} - (n-1)F(x)^{n-2}\delta f(x) + \mathcal{O}(\delta^2).$$

This expression, together with (3.17), yields

$$p_{n,\delta} = \frac{1}{n} - \delta(n-1)I_n + \mathcal{O}(\delta^2), \tag{3.18}$$

where

$$I_n := \int_{-\infty}^{\infty} F(x)^{n-2} f(x)^2 dx.$$

Summarizing, $\delta(n-1)I_n$ plays the role of a *first order correction term* in the δ-record probability for small δ.

Moreover, according to the definition of near-record in (1.2.4), and since the first term $1/n$ is the usual record probability as pointed out in (1.1), we get the following approximation for the near record probability for $a \ll 1$ taking $a = -\delta$

$$p_{n,a}^{near} = a(n-1)I_n + \mathcal{O}(a^2).$$

Let us now consider the case where the r.v. of interest (Y_n) follow the LDM. Under this model the δ-record probability $p_{n,\delta}$ was computed in (3.1). In this setting, in addition to $\mid \delta \mid \ll 1$, we take $c \approx \mid \delta \mid$ for simplicity, so that we can compute a first order approximation around the points $(cj - \delta)$. Note that this condition also implies that the approximation will only be valid for small n. For the terms in the product in (3.1) we have

$$F(x + cj - \delta) = F(x) + f(x)(cj - \delta) + \mathcal{O}(\delta^2),$$

and so, for the whole product in (3.1) and after some algebra, we get

$$\prod_{j=1}^{n-1} F(x + cj - \delta) = F(x)^{n-1} + F(x)^{n-2} f(x) \left(c\frac{n(n-1)}{2} - \delta(n-1)\right) + \mathcal{O}(\delta^2),$$

Finally, substituting this last expression in (3.1), we obtain, for fixed n,

$$p_{n,\delta}(c) = \frac{1}{n} + (n-1)\left(\frac{cn}{2} - \delta\right) I_n + \mathcal{O}(\delta^2), \tag{3.19}$$

where the term I_n is the same as in the case of the CRM.

In order to study the performance of the approximations in the LDM, we define the excess probability.

Definition 3.4.1. Let denote the excess probability $E_n(c, \delta)$ as $p_{n,\delta}(c) - 1/n$, that is, the difference between the probability of a δ-record in the LDM and the probability of a usual record in the CRM.

Note that, unlike in the CRM case, in the LDM we do not have that $(n-1)(cn/2 - \delta) I_n$ is an approximation of the near-record probability, since part of it corresponds to the contribution of the trend c to the usual record probability.

Some notes about the approximations.

The accuracy of these δ-record probability approximations has been assessed via Montecarlo simulation. Nevertheless, the analytical expressions of the approximations illustrate some of the features of δ-record probabilities.

For instance, the approximation in both cases is consistent with the influence of the parameters δ and c. Indeed, for increasing (decreasing) δ, the occurrence of a δ-record is more difficult (easier), and the approximate probability will be lower (higher). For the LDM, the influence of δ is the same as in the CRM, while for c the behaviour is also consistent since, for a higher (lower) trend parameter c, the occurrence of a δ-record will be easier (more difficult), and then the approximate probability will be higher (lower).

Also, for the LDM, taking $c \approx \mid \delta \mid$, there is not an interaction term between c and δ, being the influence of δ of the same magnitude as in the CRM. Nevertheless, the influence of c is of a higher order of magnitude, revealing that in order to facilitate the appearance of δ-records it is better to increase the underlying trend than to widen the δ-record condition varying the value of δ.

I_n terms

It is important to note that, although the computation of I_n cannot be guaranteed to be analitically solvable, it avoids the problem that arises in the integrals (3.1) and (3.17), where there is a delay between the points appearing in the argument of the cdfs F with respect to the pdf. Some I_n terms have been computed previously in the literature, showing a relationship between these terms and the domains of attraction of extreme values (see Section 1.3.4). In particular, the authors in [33], find the following patterns:

- In the Fréchet class of heavy-tailed distributions they consider r.v. of the Pareto family with pdf $f(x) = \mu x^{-\mu-1} 1_{\{x>1\}}$. Computations yield

  $$I_n \propto n^{-2-1/\mu}$$

  and thus $I_n \propto n^{-\alpha}$, for a parameter $\alpha > 2$.

- $I_n \propto a_n n^{-2}$ with a_n growing at a slow logarithmic rate in the Gumbel class of exponential-like tailed distributions.

  For this chapter, we have computed the term I_n for another distribution in the Gumbel class. More specifically, for the Gumbel distribution as defined in Section 3.3.1. For this distribution we find that the term I_n is exactly n^{-2}, consistent with the results in [33] for other distributions in this family.

- In the Weibull class of distributions with a right endpoint they consider a $Beta(1, b)$ r.v. with $b > 1/2$, finding

  $$I_n \propto b\Gamma(2 - 1/b)n^{-2+1/b},$$

  and, in particular, $I_n \propto n^{\alpha}$, $\alpha < 2$. In this family we can include the case where X_n is the Type III max-stable distribution as in example 3.3.2, where we find $I_n = n^{-1}$. This is not surprising since it is well know that the right tail of this distribution behaves similarly to a uniform random variable in $[0, 1]$ which has $I_n = (n - 1)^{-1}$.

3.4.1 Correction terms for the LDM and qualitative classification

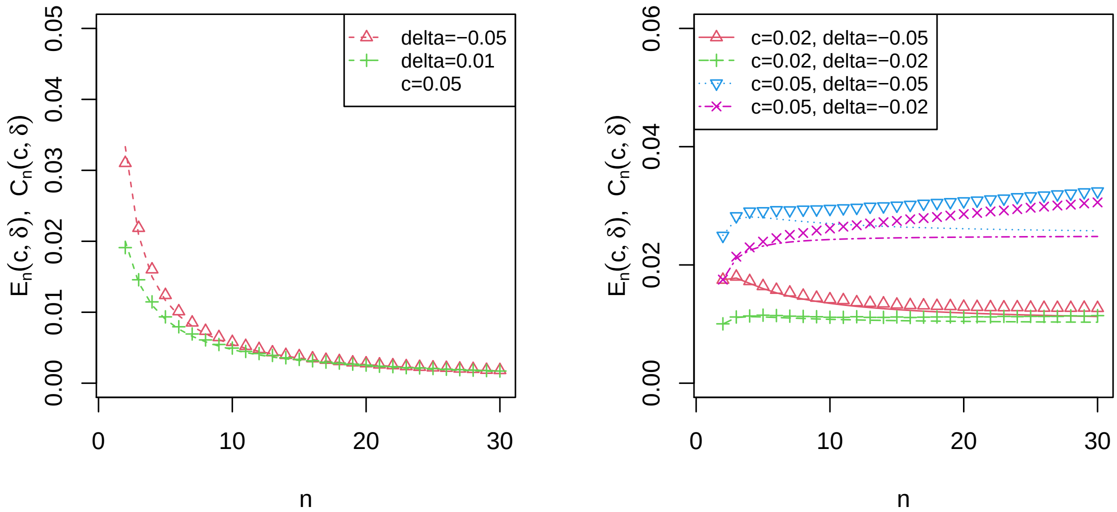
In order to study the influence of c and δ on $p_{n,\delta}(c)$, we define a *correction term* as

$$C_n(c, \delta) = (n - 1)\left(\frac{cn}{2} - \delta\right) I_n.$$

This corresponds to the main term of the excess probability $E_n(c, \delta)$ added by the influence of c and δ to the usual record probability $1/n$; see (3.19) and Definition 3.4.1.

We classify the behaviour of $C_n(c, \delta)$ depending on the growth rate of I_n, which reveals a dependence on the extreme-value family as pointed above. The predicted

Figure 3.3
Points: Estimations of the excess probability $E_n(c,\delta)$ via simulation with 10^8 iterations. Lines: $C_n(c,\delta)$. Left: Results for the Pareto distribution in the LDM. Right: Results for the Gumbel distribution in the LDM.



behaviour of $E_n(c,\delta)$ arising from the classification below has been assessed via Montecarlo simulations for several values of c and δ, finding a good agreement between the approximations and the simulations in all cases except for some distributions with unbounded pdf in the Weibull class.

1. $I_n \propto n^{-\alpha}$, $\alpha > 2$ (Fréchet family). The correction term is

$$C_n(c,\delta) \propto c\frac{1}{2n^{\alpha-2}} - \delta\frac{1}{n^{\alpha-1}},$$

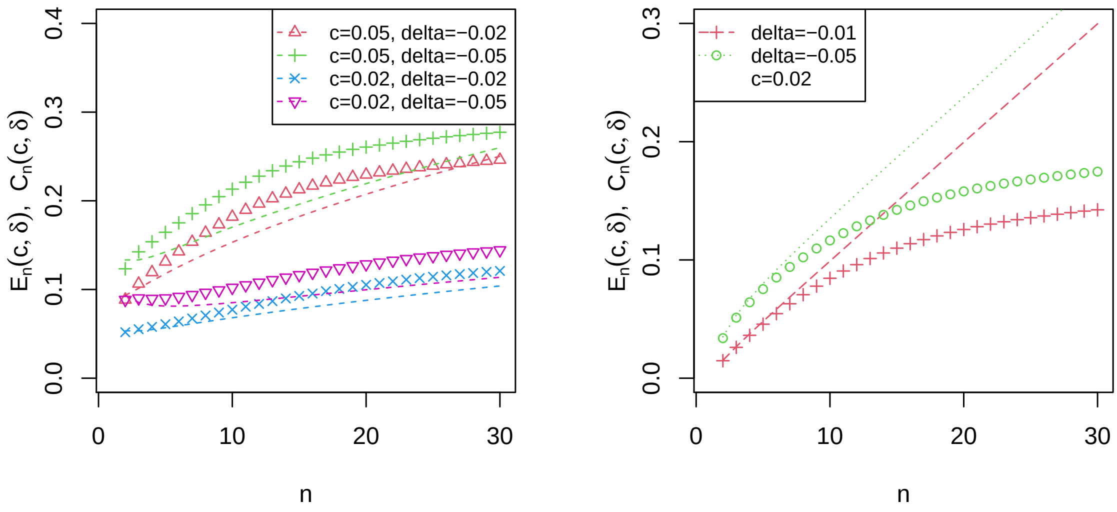
and thus the influence of c and δ vanishes quickly as n increases as it can be seen in Figure 3.3 (left), where the exact value of $C_n(c,\delta)$ and the estimated (by simulation) value of $E_n(c,\delta)$ have been displayed for the Pareto distribution.

2. $I_n \propto n^{-2}$, (Gumbel family). The correction term is

$$C_n(c,\delta) \propto c\frac{1}{2} - \delta\frac{1}{n}.$$

This means that the dependence on δ is weak, while for two values c_1, c_2, the difference between the correction terms should be proportional to $(c_2 - c_1)/2$. We can observe this phenomenon in Figure 3.3 (right) for the Gumbel distribution. For small parameter values, approximations are good, and the expected difference induced by $(c_2 - c_1)/2$ is predicted fairly well, since the proportionality constant is 1 in this case because $I_n = n^{-2}$.

Figure 3.4
Points: Estimations of the excess probability $E_n(c,\delta)$ via simulation with 10^8 iterations. Lines: $C_n(c,\delta)$. Left: Results for the $Beta(1,2)$ distribution in the LDM. Right: Results for the Type III max-stable distribution in the LDM.



3. $I_n \propto n^{-\alpha}$, $1 < \alpha < 2$ (Weibull family). The correction term is

$$C_n(c,\delta) \propto c\frac{n^{2-\alpha}}{2} - \delta\frac{1}{n^{\alpha-1}},$$

expecting an increasing influence of c and decreasing influence of δ as n increases. Figure 3.4 (left) shows the results for the $Beta(1,2)$ distribution. It can be observed that the influence of the δ parameter is negligible while the approximations tend to be grouped by the value of the trend as expected.

4. $I_n \propto n^{-1}$, (Weibull family). The correction term is

$$C_n(c,\delta) \propto c\frac{n}{2} - \delta,$$

revealing that while specially in the short term the influence of c is strong, the effect of δ is a translation proportional to δ units. This is exactly the phenomena shown in Figure 3.4 (right) for the Type III max-stable distribution, for which $I_n = n^{-1}$. The dependence on c is observed for the first observations. The influence of δ induces a constant difference for two different values of δ. This phenomena seems to be valid even in the limit.

5. $I_n \propto n^{-\alpha}$, $0 < \alpha < 1$, (Weibull family). The correction term is

$$C_n(c,\delta) \propto c\frac{n^{2-\alpha}}{2} - \delta n^{1-\alpha},$$

which shows an increasing influence of both parameters as n grows. However, simulations show that (3.19) is not accurate, at least for the distributions that we have considered. The reason is that those distributions have a finite right-endpoint and an unbounded pdf.

3.4.2 Conclusions about the approximations

In the CRM, our first order approximations seem to capture well the influence of δ in $p_{n,\delta}$ for small δ, even for not too small values of n. Moreover, we have found an approximation for the near-record probability. We have chosen not to show any plot for the CRM since estimations are very close to the approximations.

In the LDM, for moderate c and δ, estimations seems to have more variability due to the existence of two sources of error. Also, while c has a greater influence than δ, we can still find the effect of δ on the δ-record probability. This is confirmed both via approximations and simulations.

We find that the qualitative behaviour predicted by the first order approximations fits reasonably well the simulation results in the small n regime. This behaviour is still critically related to the different domains of attraction of extreme-values, and among the considered distributions, we do not find two distributions from two different families with a similar behaviour. In particular, we find that heavy-tailed distributions (Fréchet class) are the least influenced by the parameters c and δ, which it is not surprising since these distributions tend to take large values more often. The Gumbel class is an intermediate case between the two other families, being the Weibull the most dependent on the value of δ, except in the case where there is an asymptote in the right-endpoint, which makes the approximation inaccurate.

3.5 Correlations

The indicators of δ-records are in general not independent in the case of i.i.d. random variables, see [53]. In [123] the authors study the dependence of record events in the LDM, by means of the following dependence index ($\delta = 0$ in their case)

$$l_n(c,\delta) := \frac{\mathbb{P}(\text{obs. } n \text{ and } n+1 \text{ are } \delta\text{-records})}{\mathbb{P}(\text{obs. } n \text{ is } \delta\text{-record})\mathbb{P}(\text{obs. } n+1 \text{ is } \delta\text{-record})} = \frac{\mathbb{E}(1_{n,\delta}1_{n+1,\delta})}{\mathbb{E}(1_{n,\delta})\mathbb{E}(1_{n+1,\delta})}.$$

If the events are independent, then $l_n(c,\delta) = 1$. Otherwise, values greater or smaller than 1 indicate positive or negative correlation, respectively. That is, neighbouring δ-records tend to attract or repel each other, if $l_n > 1$ or $l_n < 1$.

In order to manipulate $\mathbb{E}(1_{n,\delta}1_{n+1,\delta})$ we consider the decomposition

$$\mathbb{E}\left(1_{n,\delta}1_{n+1,\delta}\right) = \mathbb{E}\left(1_{n,\delta}1_{n+1,\delta}1_{\{Y_n<Y_{n+1}\}}\right) + \mathbb{E}\left(1_{n,\delta}1_{n+1,\delta}1_{\{Y_n>Y_{n+1}\}}\right), \tag{3.20}$$

which, for $\delta < 0$, can be written as

$$\begin{aligned}\mathbb{E}(1_{n,\delta}1_{n+1,\delta}) &= \int_{-\infty}^{\infty}\Bigg(\int_{s-c}^{\infty}\prod_{j=1}^{n-1}F(s+cj-\delta)f(t)dt + \int_{s-c+\delta}^{s-c}\prod_{j=2}^{n}F(t+cj-\delta)f(t)dt\Bigg)f(s)ds\\ &= \int_{-\infty}^{\infty}\Bigg((1-F(s-c))\prod_{j=1}^{n-1}F(s+cj-\delta) + \int_{s-c+\delta}^{s-c}\prod_{j=2}^{n}F(t+cj-\delta)f(t)dt\Bigg)f(s)ds,\end{aligned} \tag{3.21}$$

and, for $\delta \geq 0$,

$$\begin{aligned}\mathbb{E}(1_{n,\delta}1_{n+1,\delta}) &= \int_{-\infty}^{\infty}\int_{s-c+\delta}^{\infty}\prod_{j=1}^{n-1}F(s+cj-\delta)f(t)dtf(s)ds\\ &= \int_{-\infty}^{\infty}(1-F(s-c+\delta))\prod_{j=1}^{n-1}F(s+cj-\delta)f(s)ds,\end{aligned} \tag{3.22}$$

since the second term in (3.20) vanishes.

As for $\mathbb{E}(1_{n,\delta})$, it is not possible to explicitly compute $\mathbb{E}(1_{n,\delta}1_{n+1,\delta})$, in general. Nevertheless, it is still possible to describe the behaviour of the dependence index in some particular and illustrative cases.

3.5.1 The Gumbel distribution

Example 3.5.1. Let $c > 0$ and F the Gumbel distribution, as in Section 3.3.1. When $\delta < 0$, elementary but lengthy computations yield

$$\lim_{n\to\infty}\mathbb{E}(1_{n,\delta}1_{n+1,\delta}) = \frac{(e^c-1)^2(e^c-e^\delta+1)}{(e^c+e^\delta-1)(e^{2c}+e^\delta-1)}$$

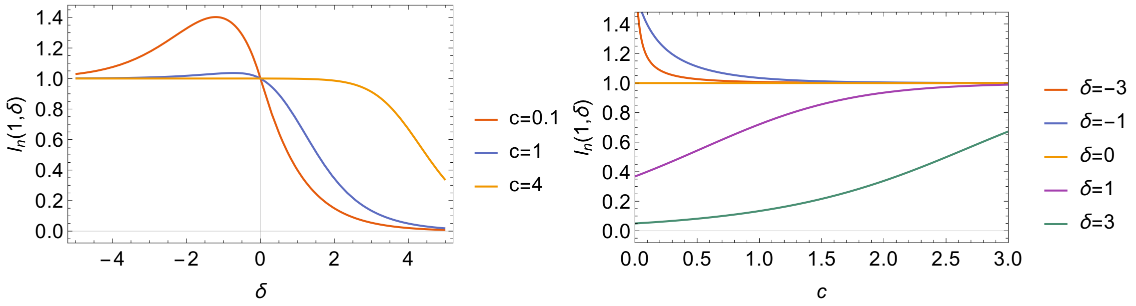
and

$$l_\infty(c,\delta) := \lim_{n\to\infty} l_n(c,\delta) = \frac{(e^c+e^\delta-1)(e^c-e^\delta+1)}{(e^{2c}+e^\delta-1)}.$$

By differentiating with respect to c, we see that $l_\infty(c,\delta)$ is decreasing in c and bounded below by 1, since $\lim_{c\to\infty} l_\infty(c,\delta) = 1$. With respect to δ we find that the derivative $\frac{\partial l_\infty}{\partial\delta}$ vanishes at

$$\delta = \log(1-e^{2c}+\sqrt{e^{4c}-e^{2c}}),$$

Figure 3.5
Dependence index $l_\infty(c,\delta)$ for the Gumbel distribution.



and then, for any c,

$$\max_{\delta<0} l_\infty(c,\delta) = \frac{2e^{2c}\left(\sqrt{e^{2c}(e^{2c}-1)} - e^{2c} + 1\right)}{\sqrt{e^{2c}(e^{2c}-1)}} = 2\left(e^{2c} - \sqrt{2e^{3c}\sinh(c)}\right).$$

Note also that $\lim_{\delta\to-\infty} l_\infty(c,\delta) = 1$.

For $\delta \geq 0$,

$$\lim_{n\to\infty} \mathbb{E}(1_{n,\delta}1_{n+1,\delta}) = \frac{e^c(e^c-1)^2}{(e^c+e^\delta-1)(e^{c+\delta}-e^c+e^{2c}-e^\delta+e^{2\delta})}$$

and

$$l_\infty(c,\delta) = \frac{e^c(e^c+e^\delta-1)}{e^{c+\delta}-e^c+e^{2c}-e^\delta+e^{2\delta}}.$$

We note that $l_\infty(c,\delta) = 1$, $\forall c > 0$, if $\delta = 0$, which results in the asymptotic independence of consecutive record indicators in the LDM. Also, there are no critical points for the index when $\delta \geq 0$. So, in this case $l_\infty(c,\delta)$ is increasing in c with $\lim_{c\to\infty} l_\infty(c,\delta) = 1$, and decreasing in δ, with $\lim_{\delta\to\infty} l_\infty(c,\delta) = 0$, as can be seen in Figure 3.5. Gathering these results, we conclude that $l_\infty(c,\delta) > 1$ if and only if $\delta < 0$. The asymptotic independence for records ($\delta = 0$) was proved in [14]; we have shown here that δ-records attract each other for $\delta < 0$ and repel each other for $\delta > 0$.

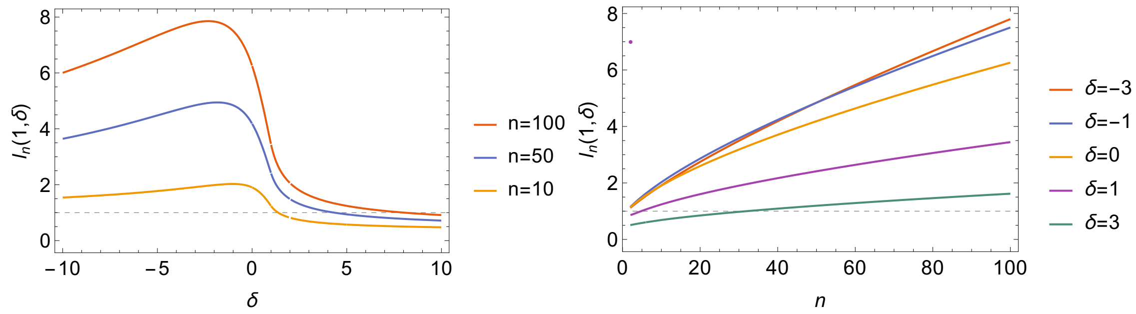
3.5.2 The Pareto distribution

Example 3.5.2. Let $c = 1$ and F be as in (3.9). The probability of δ-record is given in (3.16). For $\mathbb{E}(1_{n,\delta}1_{n+1,\delta})$ and $n > 2$, we use (3.21) for $\delta < 0$ and (3.22) for $\delta \geq 0$. Computations of $l_n(c,\delta)$ are cumbersome so we omit the details. The explicit expression of $l_n(1,\delta)$ is shown below as a function of δ.

1. If $\delta < 0$,

$$l_n(1,\delta) = \frac{B+C}{A},$$

Figure 3.6
Dependence index $l_n(1,\delta)$ for the Pareto distribution as a function of δ and n.



where $a = n - \delta$, $A = (\delta - 2)(\delta(1-a) + (n-1)\log a)(n\log(a+1) - \delta a)$,

$$B = -\big(\delta^3(n-2) + \delta - 2n^3 - 2\delta^2(n^2-2) + \delta(n-1)(n+5)n + n + 1\big)\log(a+1),$$

$$C = (a-1)\log(a+1-\delta) - (\delta-2)a\left(\delta(a-1)^2 - (n-1)a\log(4a)\right) + (1-a)\log((a-\delta+1)(a+1)).$$

2. If $0 < \delta < 1$,

$$l_n(1,\delta) = \frac{a^2(B+C)}{A},$$

where $a = n - \delta$, $A = (\delta-1)^2(\delta-a)(\delta(1-a) + (n-1)\log a)(-\delta a + n\log(a+1))$,

$$B = (a-\delta)\left((\delta-1)(\delta^2(a-1) + (\delta-1)(n-1)\log\left(\frac{a-\delta+1}{(2-\delta)a}\right)\right),$$

$$C = -\log(2-\delta)(\delta(\delta+2) - 2\delta n + n - 1)\big) + (\delta-1)^2(n-1)\log(a-\delta+1).$$

3. If $\delta = 1$,

$$l_n(1,1) = \frac{(n-1)^2((n-2)n - 2(n-1)\log(n-1))}{2(n-2)(-n+(n-1)\log(n-1)+2)(-n+n\log(n)+1)}.$$

4. Finally, if $\delta > 1$, and $\delta \notin \{n/2, n, n+1\}$ (otherwise, the values of the index are given by the continuous extension at these points),

$$l_n(1,\delta) = \frac{a^2(B+C)}{A_1 A_2},$$

where $a = n - \delta$, $A_1 = (\delta + \log\delta - n\log\delta - n + (n-1)\log(n-1) + 1)(\delta-1)^2(\delta-a)$,

$$A_2 = (\delta - n\log\delta - n + n\log n),$$

$$B = (\log\delta)\big(2\delta(\delta^2 + 2\delta - 1) + (2\delta-1)n^2 - 5\delta^2 n + n\big),$$

$$C = (\delta-1)^2(n-1)\log(n-1) - (a-1)\big((\delta-1)(\delta-a) + (2\delta-1)\log(2\delta-1)(a-1)\big).$$

As a consequence of these results, we have

$$\lim_{\delta \to -\infty} l_n(1, \delta) = 1$$

and

$$\lim_{\delta \to \infty} l_n(1, \delta) = 1 - \log(2) \approx 0.3069,$$

for every $n > 1$.

Also, $\lim_{n \to \infty} l_n(1, \delta) = \infty$ for all $\delta \in \mathbb{R}$, that is, δ-record-attraction grows unboundedly, as n increases. Moreover, it can be proved that $l_n(1, \delta) \sim C \frac{n}{(\log n)^2}$ as $n \to \infty$, where C is a constant depending on δ.

The sublinear growth of $l_n(1, \delta)$ as n increases can be observed in the right panel of Figure 3.6, for different values of δ, as well as the decrease in δ. Also, for fixed n (left panel of Figure 3.6), there is a negative value of δ where the correlation reaches a maximum, as in the Gumbel case. Note that, for negative and small positive values of δ, $l_n(1, \delta) > 1$, while, for large values of δ, $l_n(1, \delta) < 1$.

3.6 Asymptotic behaviour of $N_{n,\delta}$

In Sections 3.2 and 3.3 we have presented properties of the probability that observation n is a δ-record. In this section we analyze the random variable $N_{n,\delta}$, defined as the number of δ-records among the first n observations, and study its behaviour as $n \to \infty$.

Depending on F, c and δ, it might be the case that only finitely many δ-records are observed. We give necessary and sufficient conditions for this to happen. On the other hand, if $N_{n,\delta}$ grows to infinity, we investigate if the ratio $N_{n,\delta}/n$ converges (in a certain stochastic sense) to p_δ and, in that case, how the fluctuations of $N_{n,\delta}/n$ around p_δ are distributed.

Recall that, in the Classical Record Model ($c = 0$), the number of records $N_{n,0}$ grows to infinity, and there are universal results ensuring that, for any continuous F, $N_{n,0}/\log n$ converges to 1, almost surely and $(N_{n,0} - \log n)/(\log n)^{1/2}$ has, asymptotically, a standard Gaussian distribution. However, when $\delta \neq 0$, results in [53] and [54] for the model with $c = 0$, show that $N_{n,\delta}$ may grow to a finite limit and, when it diverges, the corresponding limit laws depend both on δ and F. We begin by analyzing the situation where $N_{n,\delta}$ has a finite limit.

3.6.1 Finiteness of the total number of δ-records

Let $N_{\infty,\delta} = \lim_{n\to\infty} N_{n,\delta}$ be the total number of δ-records along the sequence $(Y_n)_{n\geq 1}$. In this section we find necessary and sufficient conditions for the finiteness of $N_{\infty,\delta}$ and $\mathbb{E}(N_{\infty,\delta})$. Clearly, these questions are related to the asymptotic behaviour of $p_{n,\delta}$. If $p_\delta > 0$, then we can expect $N_{\infty,\delta} = \infty$. On the other hand, if $p_\delta = 0$, it may happen that $N_{n,\delta}$ grows sublinearly to ∞ or $N_{\infty,\delta} < \infty$.

In Theorem 3.6.4, we give a complete characterization of the (almost sure) finiteness of the number of δ-records depending on c, δ and the cdf F. To that end, we first prove the following results.

Proposition 3.6.1. *Let $c = 0$ and $\delta > 0$. The following conditions are equivalent:*

(a) $N_{\infty,\delta} < \infty$,

(b) $\mathbb{E}(N_{\infty,\delta}) < \infty$,

(c)

$$\int_0^\infty \frac{1 - F(x+\delta)}{(1 - F(x))^2} f(x) dx < \infty.$$

Proof. It is clear that Y_n is a δ-record if and only if $e^{X_n} > e^\delta \max\{e^{X_1}, \dots, e^{X_{n-1}}\}$. That is, if the n-th observation in the sequence $(e^{X_n})_{n\geq 1}$ is a geometric record, with parameter $k = e^\delta$, according to [54]. In Section 2.1.1 of that paper, it is shown that the total number of geometric records, in a sequence of i.i.d. random variables, with cdf G, is finite if and only if

$$\int_1^\infty \frac{1 - G(kx)}{(1 - G(x))^2} dG(x) < \infty. \tag{3.23}$$

Moreover, in Section 2.3.4 of that paper, it is shown that (3.23) is equivalent to the finiteness of the expectation of the total number of geometric records. Since $G(x) = F(\log(x))$, the result is proved. □

Lemma 3.6.2. *1. If $c < 0$, $x_- > -\infty$ and $\mu^+ < \infty$, then $\mathbb{E}(N_{\infty,\delta}) < \infty$, $\forall \delta \in \mathbb{R}$.*

2. Let $\tilde{X}_1$ be a random variable with cdf G, and $(\tilde{X}_n)_{n\geq 2}$ an i.i.d. sequence, independent of $\tilde{X}_1$, with common cdf F, such that $G(x) \leq F(x)$, $\forall x$. Let $\tilde{Y}_n = \tilde{X}_n + cn, n \geq 1$. Then, if $c < 0$,

$$\mathbb{E}\left(\sum_{j=1}^{\infty} 1_{\{\tilde{Y}_j > \vee_{i=1}^{j-1} \tilde{Y}_i + \delta\}}\right) \leq \mathbb{E}(N_{\infty,\delta}).$$

Proof. (i) First we bound $p_{n,\delta}(c)$ as follows

$$\begin{aligned} p_{n,\delta}(c) &= \int_{-\infty}^{\infty} \prod_{j=1}^{n-1} F(x+cj-\delta) f(x)dx \\ &= \int_{-\infty}^{\infty} \prod_{j=1}^{n-1} F(x+cj-\delta) 1_{\{x+c(n-1)-\delta > x_-\}} f(x)dx \\ &= \int_{x_- - c(n-1)+\delta}^{\infty} \prod_{j=1}^{n-1} F(x+cj-\delta) f(x)dx \\ &\leq 1 - F(x_- - c(n-1)+\delta). \end{aligned}$$

So, $\sum_{j=1}^{n} p_{j,\delta} \leq \sum_{j=1}^{n} (1-F(x_- - c(j-1)+\delta))$ yielding

$$\mathbb{E}(N_{\infty,\delta}) \leq \sum_{j=1}^{\infty} (1-F(x_- - c(j-1)+\delta)) < \infty,$$

since $\mu^+ < \infty$.

(ii) It suffices to check that the δ-record probability for the $\tilde{Y}_n$ fulfills

$$\begin{aligned} \mathbb{E}\left(1_{\{\tilde{Y}_j > \vee_{i=1}^{j-1} \tilde{Y}_i + \delta\}}\right) &= \int_{-\infty}^{\infty} G(x+c-\delta) \prod_{i=2}^{j-1} F(x+ci-\delta) f(x)dx \\ &\leq \int_{-\infty}^{\infty} \prod_{i=1}^{j-1} F(x+ci-\delta) f(x)dx = p_{j,\delta}(c). \quad \square \end{aligned}$$

□

Proposition 3.6.3. *If $c < 0$ and $\mu^+ < \infty$, then $\mathbb{E}(N_{\infty,\delta}) < \infty$.*

Proof. It suffices to consider $\delta < 0$, since the number of δ-records is decreasing with δ. Also, we take $x_- = -\infty$ as, otherwise, the result follows from Lemma 3.6.2 (i). Moreover, since there exists $c_1 \in \mathbb{R}$ such that $\mathbb{P}(X_n + c_1 > 0) > 0$, and the number of δ-records is the same for the sequences $Y_n = X_n + cn$ and $\tilde{Y}_n = X_n + cn + c_1$, we assume without loss of generality that $\mathbb{P}(X_n > -\delta) > 0$.

Let $N = \inf\{n \in \mathbb{N} \mid X_n > -\delta\}$, then N is a geometric random variable and

$$N_{\infty,\delta} = \sum_{j=1}^{N} 1_{j,\delta} + \sum_{j=N+1}^{\infty} 1_{j,\delta} = \sum_{j=1}^{N} 1_{j,\delta} + \sum_{j=N+1}^{\infty} 1_{j,\delta} 1_{\{X_j > 0\}}.$$

For $j > N$, let $\tilde{1}_{j,\delta} = 1_{\{X_j > \vee_{i=N}^{j-1}(X_i + c(i-j)+\delta)\}} 1_{\{X_j > 0\}}$, then

$$1_{j,\delta} 1_{\{X_j > 0\}} = 1_{\{X_j > \vee_{i=1}^{j-1}(X_i + c(i-j)+\delta)\}} 1_{\{X_j > 0\}} \leq \tilde{1}_{j,\delta}.$$

Note that the $\tilde{1}_{j,\delta}$, defined for $j > N$, are the δ-record indicators of the sequence $\{X_N, X_{N+1}1_{\{X_{N+1}>0\}} + c, X_{N+2}1_{\{X_{N+2}>0\}} + 2c, \dots\}$. Now, taking expectations we have

$$\mathbb{E}(N_{\infty,\delta}) \leq \frac{1}{\mathbb{P}(X_1 > -\delta)} + \sum_{i=1}^{\infty} \mathbb{E}(\tilde{1}_{i,\delta}) < \infty,$$

since the last sum is bounded by Lemma 3.6.2 (*ii*). □

Theorem 3.6.4. *$N_{\infty,\delta} < \infty$ a.s. if and only if one of the following conditions holds*

1. *$c < 0$ and $\mu^+ < \infty$,*
2. *$c = 0$, $\delta > 0$ and $\int_0^\infty \frac{1-F(x+\delta)}{(1-F(x))^2} f(x)dx < \infty$,*
3. *$c > 0$ and $x_+ - x_- \leq \delta - c$.*

Moreover, $N_{\infty,\delta} < \infty$ a.s. if and only if $\mathbb{E}(N_{\infty,\delta}) < \infty$.

Proof. Since, by Theorem 3.2.2, the positivity of p_δ is linked to the finiteness of μ^+, we split the analysis into two cases:

1. $\mu^+ = \infty$. In this situation, $N_{\infty,\delta} = \infty$ *a.s.* for any $c, \delta \in \mathbb{R}$.

To check this assertion, we first prove that $M_n := \max\{Y_1, \dots, Y_n\} \to \infty$. Observe that $\mu^+ = \infty$ implies $x_+ = \infty$ and

$$\sum_{n=1}^{\infty} \mathbb{P}(Y_n > a) = \sum_{n=1}^{\infty} \mathbb{P}(X_n > a - cn) = \sum_{n=1}^{\infty} (1 - F(a - cn)) = \infty, \ \forall a \in \mathbb{R}. \quad (3.24)$$

From (3.24) and the second Borel-Cantelli lemma, we conclude that $Y_n > a$ infinitely often (i.o.), for any a, and so, $M_n \to \infty$, with probability one. This fact clearly implies $N_{\infty,0} = \infty$. Now, since, for $\delta < 0$, $N_{\infty,\delta} \geq N_{\infty,0}$, we get $N_{\infty,\delta} = \infty$. On the other hand, for $\delta > 0$, the event

$$\{X_n + (c-\delta)n > \max_{1 \leq j \leq n-1} \{X_j + (c-\delta)j\}\} \text{ implies } \{X_n + cn > \max_{1 \leq j \leq n-1} \{X_j + cj\} + \delta\},$$

that is, $1_{n,0}(c - \delta) \leq 1_{n,\delta}(c)$. Therefore, $N_{\infty,\delta}(c) \geq N_{\infty,0}(c - \delta) = \infty$.

2. $\mu^+ < \infty$. We distinguish three scenarios depending on the sign of c.

If $c > 0$, we first assume $x_+ - x_- > \delta - c$. In this case, we have $p_\delta > 0$ and $N_{\infty,\delta} = \infty$ is an immediate consequence of the law of large numbers in Theorem 3.6.7 below. If $x_+ - x_- \leq \delta - c$, only the first observation will be a δ-record as shown in the proof of Theorem 3.2.2, so $N_{\infty,\delta} = 1$.

If $c = 0$ and $\delta \leq 0$, then $N_{\infty,\delta} = \infty$, since $N_{\infty,\delta} \geq N_{\infty,0} = \infty$. If $c = 0$ and $\delta > 0$, the situation is more complicated. In fact $N_{\infty,\delta} < \infty$ if and only if

$$\int_0^\infty \frac{1 - F(x+\delta)}{(1 - F(x))^2} f(x)dx < \infty,$$

which is also equivalent to $\mathbb{E}(N_{\infty,\delta}) < \infty$, as it is shown in Proposition 3.6.1 by relating this question to the counting process of geometric records.

If $c < 0$, we proceed as in (3.24) to obtain

$$\sum_{n=1}^{\infty} \mathbb{P}(Y_n > a) = \sum_{n=1}^{\infty} \mathbb{P}(X_1 > a - cn) < \infty, \quad \forall a \in \mathbb{R},$$

where the last inequality follows from $\mu^+ < \infty$. Thus, the first Borel-Cantelli lemma ensures that $\mathbb{P}(Y_n > a \text{ i.o.}) = 0$, for all $a \in \mathbb{R}$, so $Y_n \to -\infty$. Then, there exists a random variable $N < \infty$ such that $\lim_{n\to\infty} M_n = M_N$ and, consequently, $N_{\infty,\delta} < \infty$. In this case, we can also prove that $\mathbb{E}(N_{\infty,\delta}) < \infty$; see Proposition 3.6.3. □

Remark 3.6.5. Theorem 3.6.4 answers a conjecture posed in [33], stating that the expected number of records ($\delta = 0$) in the LDM, with negative trend, remains finite, based on the observed exponential decay of p_n, in a particular case. We have shown that the conjecture holds if and only if $\mu^+ < \infty$.

3.6.2 Growth of $N_{n,\delta}$ to infinity

Now that we have completely characterized the asymptotic finiteness of the number of δ-records in the LDM, we turn our attention to the case $N_{\infty,\delta} = \infty$. More precisely, we are interested in the convergence of the proportion of δ-records to p_δ.

For records ($\delta = 0$) it was shown in [8] and [9] that $N_{n,0}/n \to p_0$ and that fluctuations of $N_{n,0}$ around p_0 are asymptotically Gaussian. We show here that these results carry over to the case of $\delta \neq 0$. As in the aforementioned works, we assume $\mu^+ < \infty$ and $c > 0$ and, additionally, that $x_+ - x_- > \delta - c$. Note that, by Theorems 3.2.2 and 3.6.4, we have $p_\delta > 0$ and $N_{\infty,\delta} = \infty$.

Bilateral version of the LDM.

In order to work with a stationary process, we consider a bilateral version of the LDM defined as in (1.2), but letting $n \in \mathbb{Z}$ instead of $n \in \mathbb{N}$.

Associated to this model, we will write M_n for the maximum up to the n-th observation starting at an observation indexed by 1, in contraposition with the starred M_n^* where the maximum is taken up to the n-th observation in the double-ended sequence:

$$M_n := \bigvee_{i=1}^{n} Y_i, \qquad M_n^* := \bigvee_{i \le n} Y_i.$$

In the same way, given $n \in \mathbb{N}$, the random variables $1_{n,\delta}$, will represent the occurrence of a δ-record over the single-ended sequence

$$1_{n,\delta} = \begin{cases} 1 \text{ if } Y_n > M_{n-1} + \delta, \\ 0 \text{ otherwise.} \end{cases} \tag{3.25}$$

while, for $n \in \mathbb{Z}$, $1^*_{n,\delta}$ will be the analogous over the double-ended one

$$1^*_{n,\delta} = \begin{cases} 1 \text{ if } Y_n > M^*_{n-1} + \delta, \\ 0 \text{ otherwise.} \end{cases} \tag{3.26}$$

The sum of these indicators will be the total number of δ-records at time n, $N_{n,\delta}$ and $N^*_{n,\delta}$, both quantities starting at point 1 just changing the definition of δ-records

$$N_{n,\delta} = \sum_{j=1}^{n} 1_{j,\delta}, \qquad N^*_{n,\delta} = \sum_{j=1}^{n} 1^*_{j,\delta}. \tag{3.27}$$

We first prove the following Law of Large Numbers for the number of δ-records up to time n in the LDM. In order to prove this result, we will require the Birkhoff's Ergodic Theorem which we state for completeness.

Lemma 3.6.6. *(Birkhoff's Ergodic Theorem) Let $X = (X_1, X_2, \dots)$ be a stationary (strict sense) ergodic random sequence with $\mathbb{E}(\mid X_1 \mid) < \infty$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mathbb{E}(X_1) \ \textit{a.s. and in } L_1.$$

Proof. See [103] page 385, Theorem 3.

Theorem 3.6.7. *Assume $\mu^+ < \infty$, $c > 0$ and $x_+ - x_- > \delta - c$. Then, as $n \to \infty$,*

$$N_{n,\delta}/n \to p_\delta \quad a.s.$$

and

$$\mathbb{E}(N_{n,\delta}/n) \to p_\delta.$$

Proof. It is clear that

$$\lim_{n\to\infty} \mathbb{P}(Y_n > a) = \lim_{n\to\infty} \mathbb{P}(X_n > a - cn) = 1, \ \forall a \in \mathbb{R},$$

thus $Y_n \to \infty$ and $M_n \to \infty$ a.s. Also, since $\mu^+ < \infty$, it is known by a Borel-Cantelli argument that $M_0^* < \infty$ a.s. Gathering these facts, we know that $\exists\, 0 < N < \infty$ a.s. such that $1^*_{N,0} = 1$ almost surely. From the definition of $1^*_{n,0}$, given $n \in \mathbb{N}$ we have $1_{n,0} \geq 1^*_{n,0}$, and so $1_{N,0} = 1$ a.s., entailing $M_n^* = M_n$ and $1_{n,\delta} = 1^*_{n,\delta}$ *a.s.* $\forall n > N$. So,

$$\sum_{k=N+1}^{\infty} 1_{k,\delta} = \sum_{k=N+1}^{\infty} 1^*_{k,\delta} \ a.s.$$

Also, we know that $1^*_{n,\delta}$ is a strictly stationary and ergodic sequence. Applying Birkhoff's Ergodic Theorem we have

$$\lim_{n\to\infty} \frac{N^*_{n,\delta}}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1^*_{k,\delta} = \mathbb{E}(1^*_{0,\delta}) \ a.s.$$

Now, let $(a_n)_{n\geq 1}$ be a real sequence diverging to ∞. Then

$$\left| \frac{N_{n,\delta} - N^*_{n,\delta}}{a_n} \right| \leq \left| \frac{N}{a_n} \right| \to 0 \ a.s.$$

since N does not depend on n. Finally, since $\left| \frac{N_{n,\delta} - N^*_{n,\delta}}{n} \right| \to 0$ *a.s.* and $\frac{N^*_{n,\delta}}{n} \to \mathbb{E}(1^*_{0,\delta})$ *a.s.*, we have $\frac{N_{n,\delta}}{n} \to \mathbb{E}(1^*_{0,\delta})$ *a.s.* Finally, $\mathbb{E}(1^*_{0,\delta})$ can be written as the rhs in (3.2), yielding $\mathbb{E}(1^*_{0,\delta}) = p_\delta(c)$.

Convergence in L_1 follows from the Dominated Convergence Theorem (see for example [5], result 1.6.9). □

Some related results can be obtained from the proof of Theorem 3.6.7. Let us now consider δ-record times in the LDM as defined in Definition 1.2.7 for the sequence (Y_n).

Theorem 3.6.8. *Under the hypothesis of Theorem 3.6.7, we have* $n^{-1}L_n(\delta) \to p_\delta^{-1}$ *as* $n \to \infty$ *a.s. and in* L_1.

Proof. Since by definition we have $N_{L_n(\delta),\delta} = n$, and $N_{L_n(\delta),\delta}/L_n(\delta) = p_\delta$ by Theorem 3.6.7 a.s. and in L_1, the result holds since $p_\delta > 0$. □

Remark 3.6.9. As it can be seen in the proof of Theorem 3.6.7, the assumption on independence of the X_n can be relaxed to stationary and ergodic and prove that $N_{n,\delta}/n \to \mathbb{E}(1^*_{0,\delta})$. This is useful because it allows to deal with a wider range of scenarios, including stationary autoregressive-moving-average (ARMA) processes. Note, however, that $\mathbb{E}(1^*_{0,\delta})$ could differ from p_δ in (3.2).

A proof of Gaussian convergence for the number of δ-records, based on the ideas in [8], is not straightforward. The main problem arises when considering the joint probability of two observations being δ-records. While in the case of records this quantity can be explicitly written as follows

$$\mathbb{E}(1_{i,0}1_{i+m,0}) = \int_{-\infty}^{\infty} \prod_{k=1}^{i-1} F(y+ck) \int_{y-cm}^{\infty} \prod_{j=1}^{m-1} F(s+cj) f(s) ds f(y) dy,$$

in the setting $\delta \neq 0$ there is no such analytical expression. In order to solve this problem we introduce the following general bounds, which do not depend on the specification of the model for the sequence $(Y_n)_{n\geq 1}$.

Proposition 3.6.10. *Let $(Y_k)_{k\in\mathbb{Z}}$ be a sequence of random variables and consider the events $A = \left\{ \bigvee_{k=-\infty}^{i-1} Y_k + \delta < Y_i \right\}$, $B = \left\{ \bigvee_{k=i+1}^{i+m-1} Y_k + \delta < Y_{i+m} \right\}$, $C = \{Y_i - \delta < Y_{i+m}\}$ and $E = \{Y_i + \delta < Y_{i+m}\}$. Then, if $\delta \leq 0$,*

a1) $\mathbb{P}(A \cap B \cap C) \leq \mathbb{E}(1^*_{i,\delta} 1^*_{i+m,\delta})$

and

a2) $\mathbb{P}(A \cap B \cap E) \geq \mathbb{E}(1^*_{i,\delta} 1^*_{i+m,\delta})$.

Also, if $\delta \geq 0$,

b) $\mathbb{P}(A \cap B \cap E) = \mathbb{E}(1^*_{i,\delta} 1^*_{i+m,\delta})$.

Proof. a1) Note that $1^*_{j,\delta}$ is the indicator of $D_j = \left\{ \bigvee_{k=-\infty}^{j-1} Y_k + \delta < Y_j \right\}, j = i, i+m$. Then we must show that $A \cap B \cap C \subseteq D_i \cap D_{i+m}$.

First, it is clear that $A = D_i$. Also, observe that $C \subseteq E$ and that $A \cap C \subseteq \left\{ \bigvee_{k=-\infty}^{i-1} Y_k + \delta < Y_{i+m} \right\}$, since $\delta \leq 0$. From the inclusions above we have

$$A \cap B \cap C \subseteq \left\{ \bigvee_{k=-\infty}^{i-1} Y_k + \delta < Y_{i+m} \right\} \cap E \cap B = D_{i+m}$$

and the conclusion follows.

a2) Trivial.

b) It is clear that $D_i \cap D_{i+m} \subseteq A \cap B \cap E$ and that $A \cap B \cap E \subseteq D_i$, because

$A = D_i$. Also, since $\delta \geq 0$, we have $A \cap E \subseteq \left\{ \bigvee_{k=-\infty}^{i-1} Y_k + \delta < Y_{i+m} \right\}$, so

$$A \cap B \cap E \subseteq \left\{ \bigvee_{k=-\infty}^{i-1} Y_k + \delta < Y_{i+m} \right\} \cap E \cap B = D_{i+m},$$

which completes the proof. □

Note that, although it is unnecessary in our setting, the reverse *a1)* inequality in Proposition 3.6.10 also holds for $\delta \geq 0$. Under the assumptions of the LDM, the lhs of the first two bounds in the previous proposition have analytical expressions. The strategy to prove Gaussian convergence is to work with the corresponding bounds of $\mathbb{E}(1_{i,\delta} 1_{i+m,\delta})$, which are shown to be tight enough to achieve our purpose. So, with this result we slightly modify the necessary bounds and rebuild the martingale approach in [8], to prove convergence to the Gaussian distribution.

Theorem 3.6.11. *Suppose that* $\int_0^\infty x^2 f(x)dx < \infty$ *and let* $c > 0,\ \delta \in \mathbb{R}$, *such that* $p_\delta > 0$. *Then, as* $n \to \infty$,

$$\sqrt{n}(n^{-1} N_{n,\delta}(c) - p_\delta(c)) \xrightarrow{\mathcal{D}} N(0, \sigma_\delta^2(c)),$$

where

$$\sigma_\delta^2(c) = p_\delta(c) - p_\delta^2(c) + 2\sum_{m=1}^{\infty} (\mathbb{E}(1^*_{i,\delta} 1^*_{i+m,\delta}) - p_\delta^2(c)). \tag{3.28}$$

Proof. For simplicity, we only consider the case $\delta \leq 0$ since the case $\delta > 0$ is analogous. We assume $-2\delta < x_+$ as, otherwise, we can define $X'_n = X_n + (-3\delta - x_+)$, $n \geq 1$; the number of δ-records in both models is the same and $-2\delta < x'_+$, where x'_+ is the right-end point of X'_n.

The proof is split into several steps.

1) We claim that

$$0 \leq p_{n,\delta} - p_\delta \leq c^{-1} \int_{c(n-1)/2-\delta}^{\infty} (1 - F(s))ds + F(-\delta)^{\lfloor (n-1)/2 \rfloor}. \tag{3.29}$$

The first inequality follows from

$$p_{n,\delta} - p_\delta = \int_{-\infty}^{\infty} \left(\prod_{j=1}^{n-1} F(y + cj - \delta) - \prod_{j=1}^{\infty} F(y + cj - \delta) \right) f(y)dy \geq 0.$$

For the second, let $u = \prod_{j=1}^{n-1} F(y + cj - \delta)$ and $v = \prod_{j=1}^{\infty} F(y + cj - \delta)$. Then, from the elementary inequality $u - v \leq u - uv$, we have

$$p_{n,\delta} - p_\delta \leq \int_{-\infty}^{\infty} u(1 - v) f(y) dy. \tag{3.30}$$

The integral in the rhs of (3.30) is split into two terms A, B, that we bound. Let $A = \int_{-\infty}^{-c(n-1)/2} u(1 - v) f(y) dy$ and $B = \int_{-c(n-1)/2}^{\infty} u(1 - v) f(y) dy$, then

$$\begin{aligned}
A &\leq \int_{-\infty}^{-c(n-1)/2} \prod_{j=1}^{n-1} F(-c(n-1)/2 + cj - \delta) f(y) dy \\
&\leq \prod_{j=1}^{n-1} F(c(j - (n-1)/2) - \delta) \\
&\leq \prod_{j=1}^{\lfloor (n-1)/2 \rfloor} F(c(j - (n-1)/2) - \delta) \\
&\leq \prod_{j=1}^{\lfloor (n-1)/2 \rfloor} F(-\delta) = F(-\delta)^{\lfloor (n-1)/2 \rfloor}.
\end{aligned} \tag{3.31}$$

For B we have

$$\begin{aligned}
B &\leq \int_{-c(n-1)/2}^{\infty} \left(1 - \prod_{j=n}^{\infty} F(y + cj - \delta)\right) f(y) dy \\
&\leq \int_{-c(n-1)/2}^{\infty} \sum_{j=n}^{\infty} (1 - F(y + cj - \delta)) f(y) dy \\
&\leq \int_{-c(n-1)/2}^{\infty} \left(\int_{z=n-1}^{\infty} (1 - F(y + cz - \delta)) dz\right) f(y) dy \\
&\leq \int_{-c(n-1)/2}^{\infty} \left(c^{-1} \int_{-c(n-1)/2 + c(n-1) - \delta}^{\infty} (1 - F(s)) ds\right) f(y) dy \\
&\leq c^{-1} \int_{c(n-1)/2 - \delta}^{\infty} (1 - F(s)) ds.
\end{aligned} \tag{3.32}$$

So, from (3.31) and (3.32), (3.29) holds.

2) Let $r_{m,\delta} = \mathbb{E}(1^*_{i,\delta} 1^*_{i+m,\delta})$, which is well defined since it does not depend on i. We bound $r_{m,\delta}$ by applying Proposition 3.6.10 as follows:

$$\begin{aligned}
r_{m,\delta} &= \mathbb{P}\left(Y_i, Y_{i+m} \text{ are } \delta\text{-records}\right) \\
&= \mathbb{P}\left(Y_i > \bigvee_{l<i} Y_l + \delta, Y_{i+m} > \bigvee_{l<i+m} Y_l + \delta\right) \\
&\leq \mathbb{P}\left(Y_i > \bigvee_{l<i} Y_l + \delta, Y_{i+m} > \bigvee_{l=1}^{m-1} Y_{i+l} + \delta, Y_{i+m} > Y_i + \delta\right) \\
&= \iint_{y<s+cm-\delta} \prod_{j=1}^{\infty} F(y+cj-\delta) \prod_{i=1}^{m-1} F(s+ci-\delta) f(s) ds f(y) dy.
\end{aligned}$$

If $r_{m,\delta} \geq p_\delta^2$, we apply the Fubini-Tonelli theorem, as well as the triangle inequality, to obtain

$$\begin{aligned}
|r_{m,\delta} - p_\delta^2| \leq & \left| \iint_{y<s+cm-\delta} \prod_{j=1}^{\infty} F(y+cj-\delta) \prod_{i=1}^{m-1} F(s+ci-\delta) f(s) ds f(y) dy - p_\delta^2 \right| \\
\leq & A + B,
\end{aligned}$$

where

$$A = \int_{-\infty}^{\infty} \prod_{j=1}^{\infty} F(y+cj-\delta) \left| \int_{-\infty}^{\infty} \prod_{i=1}^{m-1} F(s+ci-\delta) f(s) ds - p_\delta \right| f(y) dy$$

and

$$B = \int_{-\infty}^{\infty} \prod_{j=1}^{\infty} F(y+cj-\delta) \int_{-\infty}^{y-cm+\delta} \prod_{i=1}^{m-1} F(s+ci-\delta) f(s) ds f(y) dy.$$

Since variables are separated in A and applying the first step of this proof

$$\begin{aligned}
A &\leq \int_{-\infty}^{\infty} \prod_{j=1}^{m-1} F(s+cj-\delta) f(s) ds - p_\delta \\
&\leq c^{-1} \int_{c(n-1)/2-\delta}^{\infty} (1-F(s)) ds + F(-2\delta))^{\lfloor (m-1)/2 \rfloor}. \qquad (3.33)
\end{aligned}$$

While for B we have

$$
\begin{aligned}
B =& \int_{-\infty}^{cm/2} \prod_{j=1}^{\infty} F(y+cj-\delta) \int_{-\infty}^{y-cm+\delta} \prod_{i=1}^{m-1} F(s+ci-\delta) f(s) ds f(y) dy \\
&+ \int_{cm/2}^{\infty} \prod_{j=1}^{\infty} F(y+cj-\delta) \int_{-\infty}^{y-cm+\delta} \prod_{i=1}^{m-1} F(s+ci-\delta) f(s) ds f(y) dy \\
\leq& \int_{-\infty}^{-cm/2+\delta} \prod_{j=1}^{m-1} F(s+cj-\delta) f(s) ds \int_{-\infty}^{cm/2+} \prod_{j=1}^{\infty} F(y+cj-\delta) f(y) dy \\
&+ \int_{cm/2}^{\infty} \prod_{j=1}^{\infty} F(y+cj-\delta) f(y) dy \\
\leq& \prod_{j=1}^{m-1} F(-cm/2+cj) + 1 - F(cm/2) \\
\leq& F(-2\delta)^{\lfloor (m-1)/2 \rfloor} + 1 - F(cm/2). \qquad (3.34)
\end{aligned}
$$

Analogously, applying the corresponding bound in Proposition 3.6.10, we get the same conclusion if $r_{m,\delta} \leq p_\delta^2$ via (3.33) and (3.34), so

$$
|r_{m,\delta} - p_\delta^2| \leq c^{-1} \int_{c(m-1/2-\delta)}^{\infty} (1-F(s)) ds + 2F(-2\delta)^{\lfloor (m-1)/2 \rfloor} + 1 - F(cm/2). \qquad (3.35)
$$

3) Since $\int_0^\infty x^2 f(x) dx < \infty$, it is easy to check, from (3.35), that the series $\sum_{m=1}^{\infty} |r_{m,\delta} - p_\delta^2|$ converges; for $F(-2\delta)^{\lfloor (m-1)/2 \rfloor}$ convergence holds since $F(-2\delta) < 1$.

4) Using the strategy in the proof of Theorem 3.6.7, we get the following convergence in distribution

$$
\sqrt{n}(n^{-1} N_{n,\delta} - n^{-1} N_{n,\delta}^*) \xrightarrow{\mathcal{D}} 0. \qquad (3.36)
$$

5) Theorem 5.2 in [62] is applied to the $N_{n,\delta}^*$ in order to transfer the asymptotic normality to $N_{n,\delta}$, as a consequence of (3.36). This martingale result guarantees convergence to the Gaussian distribution, if the next two conditions hold:

1. $\sum_{k=1}^{\infty} \mathbb{E}(\xi_{k,\delta} \mathbb{E}(\xi_{l,\delta} | \mathcal{M}_0))$ converges $\forall l \geq 0$.
2. $\lim_{l \to \infty} \sum_{k=K}^{\infty} \mathbb{E}(\xi_{k,\delta} \mathbb{E}(\xi_{l,\delta} | \mathcal{M}_0)) = 0$ uniformly in $K \geq 1$.

where $\xi_{k,\delta} = 1_{k,\delta}^* - p_\delta$ and $\mathcal{M}_0$ is a certain sub-σ-algebra of events of the original probability space (see [62], page 128, for details). Moreover we have

$$\lim_{n\to\infty} n^{-1}\mathbb{E}\left(\left(\sum_{i=1}^{n}\xi_{i,\delta}\right)^2\right) = \sigma_\delta^2.$$

Given that the hypothesis $\delta \neq 0$ does not imply any extra difficulty in the application of this theorem, we omit the verification of these two conditions since the adaptation of this part of the proof is straightforward following the lines of [8]. □

3.7 A Law of Large Numbers for δ-records with Random Trend

In this section we consider a generalization of the LDM. The first model extending the LDM can be found in [9], where the i.i.d. assumption is relaxed to strict stationarity. In this regard Remark 3.6.9 already pointed that the Law of Large Numbers that it is shown in Theorem 3.6.7 can be generalized to this setting.

In 2015, the authors in [56] studied the asymptotic behaviour of record appearances from a Random Trend Model that can be written as

$$Y_n = X_n + T_n, \tag{3.37}$$

where $T_n := \sum_{k=1}^{n} \tau_k,\ n \geq 1$ is an stochastic drift process with ergodic stationary increments, and (X_n, τ_{n+1}) is a bivariate, strictly stationary and ergodic sequence and $0 < c := \mathbb{E}(\tau_1) < \infty$ appears as the mean increment or slope of the trend. Also, we assume $\mathbb{E}(X_1^+) < \infty$ where $x^+ := x \vee 0$, that is, we have $\mu^+ < \infty$.

It is easy to check that the expression in (3.37) represents a flexible model that comprises the LDM, its correspondent generalisation pointed in Remark 3.6.9 and studied in [9], and also other models like some drifted random walks.

Note that, while the decomposition $Y_n = X_n' + cn$ with $X_n' := X_n + T_n - cn$ could be done, it does not eliminate the random trend including it into the residuals of the LDM because X_n' is not stationary in general. Finally, we remark that (X_n) and (τ_n) are allowed to be dependent as long as they have a finite expectation.

The result presented in this section may be considered an extension of the analogous results in [53], devoted to the i.i.d. case, and in [56], where the asymptotic record rate for a model with random trend is proved. Moreover, for the particular case $\delta = 0$, our proof is much simpler than the one given in [56].

For theoretical purposes, we are interested in working in a similar framework to the bilateral version of the LDM in Section 3.6.2, which is legitimate since any stationary single-ended sequence can be extended to a double-ended one. That is, we can consider the indices as defined in $\mathbb{Z}$ yielding the model in (3.37) to

$$Y_n = X_n + T_n, \tag{3.38}$$

where $T_n := \sum_{k=1}^n \tau_k$, for $n \geq 1$ and, $T_n := \tau_1 - \sum_{k=n}^0 \tau_k$, given $n \leq 0$. Also, as in (3.37), (X_n, τ_{n+1}) is a bivariate, strictly stationary and ergodic sequence and $0 < c = \mathbb{E}(\tau_1) < \infty$.

Associated to this model and in the same way that for the bilateral version of the LDM, we define, for $n \in \mathbb{Z}$,

$$M_n^* = \max\{Y_i : \ i \leq n\}, \qquad 1_{n,\delta}^* = 1_{\{Y_n > M_{n-1}^* + \delta\}},$$

where we will see later that $M_n^* < \infty$ for all $n \in \mathbb{Z}$. Also, for $n \in \mathbb{N}$,

$$N_{n,\delta}^* = \sum_{k=1}^n 1_{k,\delta}^*,$$

being the non-starred version of these quantities as in the LDM but over the sequences defined in (3.38).

In the proof of the Law of Large Number, we will make use of the Dubins-Freedman Strong Law that we state next.

Lemma 3.7.1. *(Dubins-Freedman strong law) Let $(U_n)_{n\geq 1}$ be a sequence of nonnegative and bounded random variables, adapted to the increasing family of σ-algebras $(\mathcal{G}_n)_{n\geq 0}$. Then*

$$\left\{ \sum_{n\geq 1} U_n = \infty \right\} = \left\{ \sum_{n\geq 1} \mathbb{E}(U_n \mid \mathcal{G}_{n-1}) = \infty \right\} \ a.s.$$

and

$$\frac{\sum_{k=1}^n U_k}{\sum_{k=1}^n \mathbb{E}(U_k \mid \mathcal{G}_{k-1})} \to 1 \ on \ \left\{ \sum_{n\geq 1} \mathbb{E}(U_n \mid \mathcal{G}_{n-1}) = \infty \right\} \ a.s.$$

Proof. See [24].

Finally, we are able to prove the next Law of Large Numbers for the number of δ-records, or equivalently, for the asymptotic δ-record rate.

Theorem 3.7.2. *If $\delta \leq 0$, for the Random Trend Model it holds that*

$$\frac{N_{n,\delta}}{n} \to \mathbb{E}(1_{1,\delta}^*) = \mathbb{P}\left(X_1 > \bigvee_{k\geq 1} \left\{ X_{1-k} - \sum_{j=2-k}^{1} \tau_j \right\} + \delta \right) > 0$$

both a.s. and in L_1 as $n \to \infty$.

Proof. The proof is organized in six steps. The proof of $M_n \to \infty$ in 1 below and that of 2 are the same as in Lemma 1 and Proposition 1 of [56], but we include it here for completeness.

1. $M_n \to \infty$ and $N_{n,\delta} \to \infty$ a.s. as $n \to \infty$.

   Since M_n is an increasing sequence by construction, it either converges to a finite limit or diverges to ∞ a.s. Also, we have $\forall a \in \mathbb{R}$

$$\begin{aligned}\mathbb{P}(M_n > a) &\geq \mathbb{P}(X_n > a - T_n)\\ &\geq \mathbb{P}(X_n > a - nc/2, T_n \geq nc/2) \to 1,\end{aligned}$$

   in view of $\mathbb{P}(X_n > a - nc/2) = \mathbb{P}(X_0 > a - nc/2) \to 1$ and $\mathbb{P}(T_n \geq nc/2) \to 1$, by Birkhoff's Theorem. Thus, $M_n \to \infty$, and then $N_{n,0} \to \infty$.

   Since $1_{n,0} \leq 1_{n,\delta}$ $\forall n \in \mathbb{N}$ and $\delta \leq 0$, the result is straightforward.

2. $\mathbb{P}\left(\bigvee_{k\geq 1}\{X_{n-k} - \sum_{j=n+1-k}^{n} \tau_j\} \in \mathbb{R}\right) = 1$ and $M_n^* < \infty$ a.s. for all $n \in \mathbb{Z}$.

   Stationarity will guarantee the result if we prove that

$$\mathbb{P}\left(X_{-k} > \sum_{j=-k+1}^{0} \tau_j,\ i.o.\right) = 0.$$

   We know by Birkhoff's Theorem that $\mathbb{P}\left(\sum_{j=-k+1}^{0} \tau_j \leq kc/2,\ i.o.\right) = 0$, thus

$$\begin{aligned}\mathbb{P}\Big(X_{-k} > \sum_{j=-k+1}^{0} \tau_j\ i.o.\Big) &\leq \mathbb{P}\Big(X_{-k} > kc/2\ i.o.\Big)\\ &= \mathbb{P}\Big(X_0 > kc/2\ i.o.\Big) \text{ for } k \geq 1.\end{aligned}$$

   Also, since $\mathbb{E}(X_0^+) < \infty$ we know $\sum_{k=1}^{\infty} \mathbb{P}(X_0 > kc/2) < \infty$ and the result holds by the Borel-Cantelli lemma.

3. $\exists\ 0 < N < \infty$ a.s. such that $M_n^* = M_n$ and $1_{n,\delta} = 1_{n,\delta}^*$ *a.s.* $\forall n > N$.

   As a consequence of the previous result, we have

$$\mathbb{P}\left(\bigvee_{k\geq 1}\left\{X_{n-k} - \sum_{j=n+1-k}^{n} \tau_j\right\} \in \mathbb{R}\right) = 1\ \forall n \in \mathbb{Z},$$

   which together with $M_n \to \infty$ imply that $\exists\ 0 < N < \infty$ a.s. such that $1_{N,0}^* = 1$ almost surely. Also, given $n \in \mathbb{N}$ we have $1_{n,0} \geq 1_{n,0}^*$ by construction, and thus $1_{N,0} = 1$ a.s. Now, we trivially have $M_n^* = M_n$ and $1_{n,\delta} = 1_{n,\delta}^*$ *a.s.* $\forall n > N$.

4. $\mathbb{E}(1^*_{1,\delta}) > 0$ if $\delta \leq 0$.

   Knowing that $\mathbb{E}(1^*_{1,\delta}) \geq \mathbb{E}(1^*_{1,0})$ by definition, it suffices to check $\mathbb{E}(1^*_{1,0}) > 0$. Let us assume $\mathbb{E}(1^*_{1,0}) = \mathbb{P}\left(X_1 > \bigvee_{k\geq 1}\{X_{1-k} - \sum_{j=2-k}^{1} \tau_j\}\right) = 0$, then

$$\mathbb{P}\left(X_1 > \bigvee_{k\geq 1}\left\{X_{1-k} - \sum_{j=2-k}^{1} \tau_j\right\} \,\Big|\, \mathcal{F}_0\right) = 0 \text{ a.s.}$$

   and so stationarity implies

$$\mathbb{P}\left(X_n > \bigvee_{k\geq 1}\left\{X_{n-k} - \sum_{j=n+1-k}^{n} \tau_j\right\} \,\Big|\, \mathcal{F}_{n-1}\right) = 0 \text{ a.s. } \forall n \in \mathbb{N}.$$

   Since $\bigvee_{k\geq 1}\{X_{n-k} - \sum_{j=n+1-k}^{n} \tau_j\}$ and $\bigvee_{k=1}^{n-1}\{X_{n-k} - \sum_{j=n+1-k}^{n} \tau_j\}$ couple by the previous reasoning, then

$$\sum_{n=1}^{\infty}\mathbb{P}\left(X_n > \bigvee_{k\geq 1}\left\{X_{n-k} - \sum_{j=n+1-k}^{n} \tau_j\right\} \,\Big|\, \mathcal{F}_{n-1}\right) < \infty$$

   and the Dubins-Freedman strong law imply

$$\sum_{n=1}^{\infty}\mathbb{P}\left(X_n > \bigvee_{k\geq 1}\left\{X_{n-k} - \sum_{j=n+1-k}^{n} \tau_j\right\}\right) = \sum_{n=1}^{\infty} 1_{n,0} < \infty,$$

   resulting in a contradiction.

5. $N^*_{n,\delta}/n \to \mathbb{E}(1^*_{1,\delta}) > 0$ a.s. as $n \to \infty$.

   Choosing N such that the coupling of the part three has taken place, we have

$$\sum_{k=N+1}^{\infty} 1_{k,\delta} = \sum_{k=N+1}^{\infty} 1^*_{k,\delta} \; a.s. \tag{3.39}$$

   Since we have that $1^*_{n,\delta}$ is a strictly stationary and ergodic sequence by construction, we get the desired result by Birkhoff's Theorem:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} 1^*_{k,\delta} \to \mathbb{E}(1^*_{1,\delta}) \; a.s.$$

6. $N_{n,\delta}/n \to \mathbb{E}(1^*_{1,\delta}) > 0$ a.s. as $n \to \infty$.

   Since $\mathbb{E}(1^*_{1,\delta}) > 0$ then $\sum_{k=1}^{n} 1^*_{k,\delta} \to \infty$ $a.s.$ as $n \to \infty$. Because of step 3 above, we have $\sum_{k=1}^{\infty} 1_{k,\delta} = \infty$ $a.s.$ by (3.39). Now, for any $(a_n)_{n\in\mathbb{N}}$ real sequence such that $(a_n) \to \infty$ we have

$$\left|\frac{N_{n,\delta} - N^*_{n,\delta}}{a_n}\right| \leq \left|\frac{N}{a_n}\right| \to 0 \; a.s.$$

since N does not depend on n. Finally, if $n \to \infty$, we can conclude from $\left|\frac{N_{n,\delta}-N^*_{n,\delta}}{n}\right| \to 0$ $a.s.$ and $\frac{N^*_{n,\delta}}{n} \to \mathbb{E}(1^*_{1,\delta})$ $a.s.$ that $\frac{N_{n,\delta}}{n} \to \mathbb{E}(1^*_{1,\delta})$ $a.s.$

Finally, convergence in L_1 is straightforward by the dominated convergence theorem. □

Next, we show an example of an application of Theorem 3.7.2 that was analyzed in [56] for records.

Example 3.7.3. *Drifted Random Walk.* Consider a random walk, $S_n = \sum_{i=1}^n \tau_i$, where $(\tau_n, n \geq 1)$ is a stationary ergodic sequence with $E(\tau_1) > 0$. We want to get the asymptotic δ-record rate of that random walk. For this we need to fit this situation in our general framework in (3.37), for which we take $X_n = 0$, $\forall n \in \mathbb{N}$, and $T_n \equiv S_n$.

Theorem 3.7.2 guarantees that the asymptotic δ-record rate $N_{n,\delta}/n$ converges to $\mathbb{E}(1^*_{1,\delta})$ as $n \to \infty$, where

$$\begin{aligned}
\mathbb{E}(1^*_{1,\delta}) &= \mathbb{P}\left(X_1 > \bigvee_{k\geq 1}\left\{X_{1-k} - \sum_{j=2-k}^{1} \tau_j\right\} + \delta\right) \\
&= \mathbb{P}\left(0 > \bigvee_{k\geq 1}\left\{- \sum_{j=2-k}^{1} \tau_j\right\} + \delta\right) \\
&= \mathbb{P}\left(\bigwedge_{k\geq 1}\left\{\sum_{j=2-k}^{1} \tau_j\right\} > \delta\right).
\end{aligned}$$

4

Statistical Inference based on δ-records in the presence of a trend

In this chapter we study some statistical properties of δ-records in the LDM. In the first section we propose two estimators for the variance of the number of δ-records based on the ideas of [9], proving consistency. In the second section we develop a framework for Maximum Likelihood Estimation and we find analytic solutions for a family of distributions, analyzing its properties. We assess the performance of these estimators based on δ-records via Montecarlo simulation and we compare the results with those using records only. In the final section, we apply the results in this chapter, and also probabilistic results derived in Chapter 3, to a real data set of summer temperatures in Spain, where the LDM is consistent with the global-warming phenomenon.

4.1 Estimation of the variance of the number of δ-records

Making inference requires suitable estimators for the unknown quantities in order to apply the theoretical results studied in Chapter 3 to a time series. First, it is obvious that these properties about the δ-record rate, $n^{-1}\sum_{j=1}^{n} 1_{j,\delta}(c)$, make it a suitable estimator for $p_\delta(c)$.

With regard to the estimation of $\sigma_\delta^2(c)$ for a sample with size n, we propose the following two estimators following the ideas of [9] for records:

1.

$$\tilde{\sigma}_\delta^2 = \tilde{\gamma}_{n,\delta}(0) + 2\sum_{k=1}^{m} \tilde{\gamma}_{n,\delta}(k), \quad m \in \mathbb{N}, \tag{4.1}$$

where

$$\tilde{\gamma}_{n,\delta}(k) = n^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta} - n^{-1}N_{n,\delta}\right)\left(1_{j+k,\delta} - n^{-1}N_{n,\delta}\right). \tag{4.2}$$

2.

$$\hat{\sigma}_\delta^2 = \hat{\gamma}_{n,\delta}(0) + 2\sum_{k=1}^{m} \hat{\gamma}_{n,\delta}(k), \quad m \in \mathbb{N}, \tag{4.3}$$

where

$$\hat{\gamma}_{n,\delta}(k) = (n-k)^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}\right) - \left(n^{-1}N_{n,\delta}\right)^2. \tag{4.4}$$

In [9], the authors proved weak consistency for $\hat{\sigma}_0^2$ to get the same property for $\tilde{\sigma}_0^2$ under the condition $m(n) = o(n^{1/2})$. After that, they preferred the use of $\tilde{\sigma}_0^2$ over $\hat{\sigma}_0^2$ since the last one might yield negative estimations of σ_0^2. Nevertheless, the problem that the cited authors claim to avoid is not solved, since it is still present in $\tilde{\sigma}_0^2$ as well. For instance, considering the sequence $\{1, 1, 0, 1, 1\}$ representing the first 5 record indicators, we get $\tilde{\gamma}(0) = 4/25$, $\tilde{\gamma}(1) = -6/125$, $\tilde{\gamma}(2) = -7/125$ and hence $\tilde{\sigma}_0^2 = -6/125$. Nevertheless, both estimators will produce positive estimations asymptotically due to the weak convergence, but we do not see theoretical evidence of the benefits of using $\tilde{\sigma}_0^2$ over $\hat{\sigma}_0^2$, and so, we will use both estimators when applying this result to real data (Section 4.3).

In this section, we prove the weak consistency of both estimators of σ_δ^2 for any $\delta \in \mathbb{R}$. To that end we will make use of the notation introduced in Chapter 3, including the *starred* double-ended stationary sequence.

The following proposition establishes two useful results about sums of δ-record indicators.

Proposition 4.1.1. *In the LDM, and if $k \in \mathbb{N}$, then*

(a) $n^{-1}\sum_{j=1}^{n} 1_{j,\delta}1_{j+k,\delta} \to r_{k,\delta}$ *as $n \to \infty$ a.s. and in L_1.*

(b) If $\mathbb{E}(X_0^2) < \infty$*, the following series*

$$\sum_{h=1}^{\infty} \left| \mathbb{E}(1^*_{0,\delta} 1^*_{k,\delta} 1^*_{h,\delta} 1^*_{h+k,\delta}) - \mathbb{E}(1^*_{0,\delta} 1^*_{k,\delta})^2 \right|$$

converges.

Proof. **(a)** As it was proved in Theorem 3.6.7, we know that there exists $N \in \mathbb{N}$ such that $1^*_{N,0} = 1$ a.s. and that $1_{n,\delta} = 1^*_{n,\delta}$ for all $n > N$, which implies

$$\sum_{j=N+1}^{\infty} 1_{j,\delta} 1_{j+k,\delta} = \sum_{j=N+1}^{\infty} 1^*_{j,\delta} 1^*_{j+k,\delta} \quad a.s.$$

Since $(1^*_{j,\delta} 1^*_{j+k,\delta})$ is a strictly stationary and ergodic sequence, Birkhoff's Theorem yields

$$\lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} 1^*_{j,\delta} 1^*_{j+k,\delta} = \mathbb{E}\left(1^*_{0,\delta} 1^*_{k,\delta}\right) \quad a.s.$$

Finally, we have that

$$\left| \frac{\sum_{j=1}^{n} 1_{j,\delta} 1_{j+k,\delta} - \sum_{j=1}^{n} 1^*_{j,\delta} 1^*_{j+k,\delta}}{n} \right| \le \left| \frac{N}{n} \right| \to 0 \qquad a.s.$$

and therefore

$$\lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} 1_{j,\delta} 1_{j+k,\delta} = \mathbb{E}\left(1^*_{0,\delta} 1^*_{k,\delta}\right) = r_{k,\delta} \quad a.s.$$

Convergence in L_1 follows from the dominated convergence theorem.

(b) In order to prove convergence, it is enough to sum starting from $h > k$. Define the two following events

$$A_{1,h} := \Bigg\{ X_0 > \bigvee_{i=1}^{\infty} (X_{-i} - ci + \delta),\ X_k > \bigvee_{i=1}^{\infty} (X_{k-i} - ci + \delta),$$
$$X_h > \bigvee_{i=1}^{\infty} (X_{h-i} - ci + \delta),\ X_{h+k} > \bigvee_{i=1}^{\infty} (X_{h+k-i} - ci + \delta) \Bigg\},$$
$$A_{2,h} := \Bigg\{ X_0 > \bigvee_{i=1}^{\infty} (X_{-i} - ci + \delta),\ X_k > \bigvee_{i=1}^{\infty} (X_{k-i} - ci + \delta),$$
$$X_h > \bigvee_{i=1}^{h-k-1} (X_{h-i} - ci + \delta),\ X_{h+k} > \bigvee_{i=1}^{h-1} (X_{h+k-i} - ci + \delta) \Bigg\},$$

and thus we need to prove

$$\sum_{h=k+1}^{\infty} \left| \mathbb{P}(A_{1,h}) - \mathbb{E}(1^*_{0,\delta} 1^*_{k,\delta})^2 \right| < \infty.$$

Using the triangle inequality we have

$$\begin{aligned}\left|\mathbb{P}\left(A_{1,h}\right)-\left(\mathbb{E}(1^*_{0,\delta}1^*_{k,\delta})\right)^2\right| &\leq \left|\mathbb{P}\left(A_{1,h}\right)-\mathbb{P}\left(A_{2,h}\right)\right|+\left|\mathbb{P}\left(A_{2,h}\right)-\left(\mathbb{E}(1^*_{0,\delta}1^*_{k,\delta})\right)^2\right| \\ &= \mathbb{P}\left(A_{2,h}\setminus A_{1,h}\right)+\left|\mathbb{P}\left(A_{2,h}\right)-\left(\mathbb{E}(1^*_{0,\delta}1^*_{k,\delta})\right)^2\right| \\ &:= A_h+B_h,\end{aligned}$$

where the second equality holds since $A_{1,h}\subset A_{2,h}$.

We first see the convergence of the first part. Using basic set operations we bound A_h as

$$\begin{aligned}A_h \leq &\mathbb{P}\Bigg(\left\{X_h>\bigvee_{i=1}^{h-k-1}(X_{h-i}-ci+\delta),\ X_{h+k}>\bigvee_{i=1}^{h-1}(X_{h+k-i}-ci+\delta)\right\} \\ &\bigcap\left\{X_h>\bigvee_{i=1}^{\infty}(X_{h-i}-ci+\delta),\ X_{h+k}>\bigvee_{i=1}^{\infty}(X_{h+k-i}-ci+\delta)\right\}^{\mathsf{c}}\Bigg) \\ \leq &\mathbb{P}\Bigg(\left\{X_h>\bigvee_{i=1}^{h-k-1}(X_{h-i}-ci+\delta),\ X_{h+k}>\bigvee_{i=1}^{h-1}(X_{h+k-i}-ci+\delta)\right\} \\ &\bigcap\left\{X_h>\bigvee_{i=1}^{\infty}(X_{h-i}-ci+\delta)\right\}^{\mathsf{c}}\Bigg) \\ &+\mathbb{P}\Bigg(\left\{X_h>\bigvee_{i=1}^{h-k-1}(X_{h-i}-ci+\delta),\ X_{h+k}>\bigvee_{i=1}^{h-1}(X_{h+k-i}-ci+\delta)\right\} \\ &\bigcap\left\{X_{h+k}>\bigvee_{i=1}^{\infty}(X_{h+k-i}-ci+\delta)\right\}^{\mathsf{c}}\Bigg),\end{aligned}$$

which yields

$$\begin{aligned}A_h \leq &\mathbb{P}\left(\bigvee_{i=1}^{h-k-1}(X_{h-i}-ci+\delta)<X_h\leq\bigvee_{i=1}^{\infty}(X_{h-i}-ci+\delta)\right) \\ &+\mathbb{P}\left(\bigvee_{i=1}^{h-1}(X_{h+k-i}-ci+\delta)<X_{h+k}\leq\bigvee_{i=1}^{\infty}(X_{h+k-i}-ci+\delta)\right) \\ \leq &\mathbb{P}\left(X_h\leq\bigvee_{i=h-k}^{\infty}(X_{h-i}-ci+\delta)\right)+\mathbb{P}\left(X_{h+k}\leq\bigvee_{i=h}^{\infty}(X_{h+k-i}-ci+\delta)\right).\end{aligned}$$

For each term in the last sum we proceed as follows:

$$
\begin{aligned}
\mathbb{P}\left(X_h \leq \bigvee_{i=h-k}^{\infty} (X_{h-i} - ci + \delta) \right) &\leq \sum_{i=h-k}^{\infty} \mathbb{P}\left(X_h \leq X_{h-i} - ci + \delta\right) \\
&\leq \sum_{i=h-k}^{\infty} \mathbb{P}\left(\mid X_{h-i} - X_h \mid \geq ci - \delta\right) \\
&\leq \sum_{i=h-k}^{\infty} \mathbb{P}\left(\mid X_{h-i} \mid + \mid X_h \mid \geq ci - \delta\right) \\
&\leq 2 \sum_{i=h-k}^{\infty} \mathbb{P}\left(\mid X_0 \mid \geq (ci - \delta)/2\right).
\end{aligned}
$$

Gathering these facts, we have

$$
\sum_{h=k+1}^{\infty} A_h \leq 4 \sum_{h=k+1}^{\infty} \sum_{i=h-k}^{\infty} \mathbb{P}\left(\mid X_0 \mid \geq (ci - \delta)/2\right) < \infty,
$$

since $\mathbb{E}\left(X_0^2\right) < \infty$.

For the convergence of $\sum_{h=k+1}^{\infty} B_h$, we have

$$
\begin{aligned}
&\sum_{h=k+1}^{\infty} \left| \mathbb{P}\left(A_{2,h}\right) - \left(\mathbb{E}(1^*_{0,\delta} 1^*_{k,\delta})\right)^2 \right| \\
&= \sum_{h=k+1}^{\infty} \Bigg| \mathbb{E}(1^*_{0,\delta} 1^*_{k,\delta}) \mathbb{P}\Bigg(X_h > \bigvee_{i=1}^{h-k-1} (X_{h-i} - ci + \delta),\ X_{h+k} > \bigvee_{i=1}^{h-1} (X_{h+k-i} - ci + \delta) \Bigg) \\
&\quad - \left(\mathbb{E}(1^*_{0,\delta} 1^*_{k,\delta})\right)^2 \Bigg| \\
&\leq \sum_{h=k+1}^{\infty} \Bigg| \mathbb{P}\Bigg(X_h > \bigvee_{i=1}^{h-k-1} (X_{h-i} - ci + \delta),\ X_{h+k} > \bigvee_{i=1}^{h-1} (X_{h+k-i} - ci + \delta) \Bigg) - \mathbb{E}(1^*_{0,\delta} 1^*_{k,\delta}) \Bigg| \\
&\leq \sum_{h=k+1}^{\infty} \Bigg(\mathbb{P}\Bigg(X_0 > \bigvee_{i=1}^{h-k-1} (X_{-i} - ci + \delta),\ X_k > \bigvee_{i=1}^{h-k-1} (X_{k-i} - ci + \delta) \Bigg) - \mathbb{E}(1^*_{0,\delta} 1^*_{k,\delta}) \Bigg).
\end{aligned}
$$

Let us denote $1^k_{n,\delta} = 1$ if $Y_n > \bigvee_{i=1}^{k} Y_{n-i} + \delta$, and 0 otherwise. From the last inequality we get

$$\begin{aligned}
&\sum_{h=k+1}^{\infty}\left|\mathbb{P}\left(A_{2,h}\right)-\left(\mathbb{E}(1^*_{0,\delta}1^*_{k,\delta})\right)^2\right| \\
&\qquad\leq \sum_{h=k+1}^{\infty}\left(\mathbb{E}(1^{h-k-1}_{0,\delta}1^{h-k-1}_{k,\delta})-\mathbb{E}(1^*_{0,\delta}1^*_{k,\delta})\right) \\
&\qquad= \sum_{h=k+1}^{\infty}\left(\mathbb{E}(1^{h-k-1}_{0,\delta}1^{h-k-1}_{k,\delta}-1^*_{0,\delta}1^{h-k-1}_{k,\delta})-\mathbb{E}(1^*_{0,\delta}1^*_{k,\delta}-1^*_{0,\delta}1^{h-k-1}_{k,\delta})\right) \\
&\qquad= \sum_{h=k+1}^{\infty}\left(\mathbb{E}(1^{h-k-1}_{k,\delta}(1^{h-k-1}_{0,\delta}-1^*_{0,\delta}))+\mathbb{E}(1^*_{0,\delta}(1^{h-k-1}_{k,\delta}-1^*_{k,\delta}))\right) \\
&\qquad\leq \sum_{h=k+1}^{\infty}\left(\mathbb{E}(1^{h-k-1}_{0,\delta}-1^*_{0,\delta})+\mathbb{E}(1^{h-k-1}_{k,\delta}-1^*_{k,\delta})\right) \\
&\qquad= 2\sum_{h=k+1}^{\infty}\left(\mathbb{E}(1^{h-k-1}_{0,\delta}-1^*_{0,\delta})\right) \\
&\qquad= 2\sum_{h=k+1}^{\infty}\left(\mathbb{E}(1^{h-k-1}_{0,\delta})-p_\delta\right) \\
&\qquad= 2\sum_{n=0}^{\infty}(\mathbb{E}(1_{n,\delta})-p_\delta) \\
&\qquad< \infty,
\end{aligned}$$

because of the bound in equation (3.29), for which the convergence of the sum holds. □

We now establish the asymptotic behaviour of the estimators $\hat{\gamma}_{n,\delta}(k)$ and $\tilde{\gamma}_{n,\delta}(k)$.

Proposition 4.1.2. *In the LDM the following convergences hold:*

(a) $\hat{\gamma}_{n,\delta}(k)\to r_{k,\delta}-p_\delta^2$ *a.s. and in* L_1 *as* $n\to\infty$.

(b) $\tilde{\gamma}_{n,\delta}(k)\to r_{k,\delta}-p_\delta^2$ *a.s. and in* L_1 *as* $n\to\infty$.

Proof. **(a)** From Proposition 4.1.1(a) we have that for fixed k

$$\frac{1}{n-k}\sum_{j=1}^{n-k}1_{j,\delta}1_{j+k,\delta}\to r_{k,\delta}$$

as $n\to\infty$ a.s. and in L_1. Also, Theorem 3.6.7 guarantees that $n^{-1}N_{n,\delta}\to p_\delta$ as $n\to\infty$ in the a.s. and L_1 sense and therefore the result holds.

(b) Expanding the expression of $\tilde{\gamma}_{n,\delta}(k)$ and regrouping the terms conveniently we get

$$\begin{aligned}\tilde{\gamma}_{n,\delta}(k) &= \frac{1}{n}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta} - 1_{j,\delta}n^{-1}N_{n,\delta} - 1_{j+k,\delta}n^{-1}N_{n,\delta} + n^{-2}N_{n,\delta}^2\right)\\ &= \frac{1}{n}\sum_{j=1}^{n-k}(1_{j,\delta}1_{j+k,\delta}) - n^{-1}N_{n,\delta}\frac{1}{n}\sum_{j=1}^{n-k}1_{j,\delta} - n^{-1}N_{n,\delta}\frac{1}{n}\sum_{j=1}^{n-k}1_{j+k,\delta} + \frac{n-k}{n}n^{-2}N_{n,\delta}^2.\end{aligned}$$

Taking limits as $n \to \infty$, we have from the proof of Proposition 4.1.1(a) and Theorem 3.6.7 that

$$\tilde{\gamma}_{n,\delta}(k) \to r_{k,\delta} - p_\delta^2 - p_\delta^2 + p_\delta^2 = r_{k,\delta} - p_\delta^2,$$

both a.s. and in L_1. □

In the next theorem, the consistency of the estimators $\hat{\sigma}_{n,\delta}^2$ and $\tilde{\sigma}_{n,\delta}^2$ is proved.

Theorem 4.1.3. *In the LDM with $\mathbb{E}(X_0^2) < \infty$, and taking $m \in \mathbb{N}$ such that $m = m(n) = o(n^{1/2})$, the following convergences hold*

(a) $\hat{\sigma}_{n,\delta}^2 \to \sigma_\delta^2$, *in* L_1, *as* $n \to \infty$.

(b) $\tilde{\sigma}_{n,\delta}^2 \to \sigma_\delta^2$, *in* L_1, *as* $n \to \infty$.

Proof. **(1)** First, we recall the that $\sigma_\delta^2 = p_\delta - p_\delta^2 + 2\sum_{k=1}^{\infty}(r_{k,\delta} - p_\delta^2)$ as it has been proved in Theorem 3.6.11. Now, from the definition of $\hat{\gamma}_{n,\delta}(0)$ in (4.4), we observe that

$$\hat{\gamma}_{n,\delta}(0) = \frac{N_{n,\delta}}{n} - \left(\frac{N_{n,\delta}}{n}\right)^2 \to p_\delta - p_\delta^2$$

as $n \to \infty$ as a consequence of Theorem 3.6.7.

Now it remains to check that

$$\mathbb{E}\left(\left|\sum_{k=1}^{m}(\hat{\gamma}_{n,\delta}(k) - (r_{k,\delta} - p_\delta^2))\right|\right) \to 0, \quad \text{as } n \to \infty.$$

Interchanging the sum and the expectation and grouping terms we get the bound

$$
\begin{aligned}
&\mathbb{E}\left(\left|\sum_{k=1}^{m}\left(\hat{\gamma}_{n,\delta}(k)-(r_{k,\delta}-p_\delta^2)\right)\right|\right)\\
&\leq \sum_{k=1}^{m}\mathbb{E}\left(\left|\hat{\gamma}_{n,\delta}(k)-(r_{k,\delta}-p_\delta^2)\right|\right)\\
&=\sum_{k=1}^{m}\mathbb{E}\left(\left|(n-k)^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}-\left(n^{-1}N_{n,\delta}\right)^2\right)-r_{k,\delta}+p_\delta^2\right|\right)\\
&=\sum_{k=1}^{m}\mathbb{E}\left(\left|(n-k)^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}-\mathbb{E}\left(1_{j,\delta}^*1_{j+k,\delta}^*\right)\right)-\left(n^{-1}N_{n,\delta}\right)^2+p_\delta^2\right|\right).
\end{aligned}
$$

Using the triangle inequality we split the bound in three terms:

$$
\begin{aligned}
&\mathbb{E}\left(\left|\sum_{k=1}^{m}\left(\hat{\gamma}_{n,\delta}(k)-(r_{k,\delta}-p_\delta^2)\right)\right|\right)\\
&\leq\left(\sum_{k=1}^{m}\mathbb{E}\left(\left|(n-k)^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}-\mathbb{E}(1_{j,\delta}^*1_{j+k,\delta}^*)\right)\right|\right)\right)+m\mathbb{E}\left(\left|p_\delta^2-\left(n^{-1}N_{n,\delta}\right)^2\right|\right)\\
&\leq\sum_{k=1}^{m}\mathbb{E}\left(\left|(n-k)^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}-1_{j,\delta}^*1_{j+k,\delta}^*\right)\right|\right)\\
&+\sum_{k=1}^{m}\mathbb{E}\left(\left|(n-k)^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}^*1_{j+k,\delta}^*-\mathbb{E}\left(1_{j,\delta}^*1_{j+k,\delta}^*\right)\right)\right|\right)\\
&+2m\mathbb{E}\left(\left|n^{-1}N_{n,\delta}-p_\delta\right|\right)\\
&:=A_n+B_n+C_n,
\end{aligned}
$$

where the last inequality follows applying the triangle inequality after adding and subtracting the same terms, and from the fact that $p_\delta+n^{-1}N_{n,\delta}\leq 2$.

For the term A_n we note that

$$
\begin{aligned}
\mathbb{E}\left(\left|\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}-1_{j,\delta}^*1_{j+k,\delta}^*\right)\right|\right)&=\sum_{j=1}^{n-k}\mathbb{E}\left(1_{j,\delta}1_{j+k,\delta}-1_{j,\delta}^*1_{j+k,\delta}^*\right)\\
&\leq 2\sum_{j=1}^{n-k}\mathbb{E}\left(1_{j,\delta}-1_{j,\delta}^*\right)\\
&\leq 2\sum_{j=1}^{\infty}\left(\mathbb{E}(1_{j,\delta})-p_\delta\right)\\
&<\infty,
\end{aligned}
$$

where the first inequality and the convergence follow arguing as in the proof of Proposition 4.1.1(b). Finally, the convergence to 0 of A_n holds since

$$\sum_{k=1}^{m}(n-k)^{-1} \leq m(n-m)^{-1} \leq m(n-n^{1/2})^{-1} \leq mn^{-1/2} \to 0,$$

as $n \to \infty$.

To prove the convergence of B_n to 0 we first note

$$\begin{aligned} B_n &= \sum_{k=1}^{m}\left((n-k)^{-1}\mathbb{E}\left(\left|\sum_{j=1}^{n-k}\left(1^*_{j,\delta}1^*_{j+k,\delta} - \mathbb{E}\left(1^*_{j,\delta}1^*_{j+k,\delta}\right)\right)\right|\right)\right) \\ &\leq \sum_{k=1}^{m}\left((n-k)^{-1}\left(\mathbb{E}\left(\left[\sum_{j=1}^{n-k}\left(1^*_{j,\delta}1^*_{j+k,\delta} - \mathbb{E}\left(1^*_{j,\delta}1^*_{j+k,\delta}\right)\right)\right]^2\right)\right)^{1/2}\right) \end{aligned} \tag{4.5}$$

We now bound the expectation in the sum (4.5), obtaining

$$\begin{aligned} &\mathbb{E}\left(\left[\sum_{j=1}^{n-k}\left(1^*_{j,\delta}1^*_{j+k,\delta} - \mathbb{E}\left(1^*_{j,\delta}1^*_{j+k,\delta}\right)\right)\right]^2\right) \\ &= \mathbb{E}\left(\sum_{j=1}^{n-k}\left(1^*_{j,\delta}1^*_{j+k,\delta} - \mathbb{E}\left(1^*_{j,\delta}1^*_{j+k,\delta}\right)\right)^2\right) \\ &\quad + 2\mathbb{E}\left(\sum_{j=1}^{n-k}\sum_{i=1}^{j-1}\left(\left(1^*_{i,\delta}1^*_{i+k,\delta} - \mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}\right)\right)\left(1^*_{j,\delta}1^*_{j+k,\delta} - \mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}\right)\right)\right)\right) \\ &\leq n-k \\ &\quad + 2\mathbb{E}\Bigg(\sum_{j=1}^{n-k}\sum_{i=1}^{j-1}\Big(1^*_{i,\delta}1^*_{i+k,\delta}1^*_{j,\delta}1^*_{j+k,\delta} - \mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}\right)1^*_{j,\delta}1^*_{j+k,\delta} \\ &\qquad - \mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}\right)1^*_{i,\delta}1^*_{i+k,\delta} + \left(\mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}\right)\right)^2\Big)\Bigg). \end{aligned}$$

We now introduce the expectation into the double sum simplifying the expression and regrouping terms, yielding

$$
\begin{aligned}
&\mathbb{E}\left(\left[\sum_{j=1}^{n-k}\left(1^*_{j,\delta}1^*_{j+k,\delta}-\mathbb{E}\left(1^*_{j,\delta}1^*_{j+k,\delta}\right)\right)\right]^2\right)\\
&\leq n-k+2\sum_{j=1}^{n-k}\sum_{i=1}^{j-1}\left(\mathbb{E}\left(1^*_{i,\delta}1^*_{i+k,\delta}1^*_{j,\delta}1^*_{j+k,\delta}\right)-\left(\mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}\right)\right)^2\right)\\
&=n-k+2\sum_{j=1}^{n-k-1}(n-k-j)\left(\mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}1^*_{j,\delta}1^*_{j+k,\delta}\right)-\left(\mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}\right)\right)^2\right)\\
&\leq n-k+2(n-k)\sum_{j=1}^{\infty}\left|\mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}1^*_{j,\delta}1^*_{j+k,\delta}\right)-\left(\mathbb{E}\left(1^*_{0,\delta}1^*_{k,\delta}\right)\right)^2\right|\\
&\leq M(n-k),
\end{aligned}
$$

for a real constant M, because the infinite sum appearing in the second last line above is convergent by Proposition 4.1.1(b). Now, from (4.5) we have

$$
\begin{aligned}
B_n &\leq \sum_{k=1}^{m}\left((n-k)^{-1}\left(M(n-k)\right)^{1/2}\right)\\
&\leq M^{1/2}\sum_{k=1}^{m}(n-k)^{-1/2}\\
&\leq M^{1/2}m(n-m)^{-1/2}\\
&\leq M^{1/2}m(n-n^{1/2})^{-1/2}\\
&\to 0,
\end{aligned}
$$

since $m(n)=o(n^{1/2})$.

Finally we prove that $C_n\to 0$ as $n\to\infty$. To that end we apply the triangle inequality, obtaining

$$
\begin{aligned}
m\mathbb{E}\left(\left|n^{-1}N_{n,\delta}-p_\delta\right|\right) &= m\mathbb{E}\left(\left|n^{-1}N_{n,\delta}-n^{-1}\sum_{j=1}^{n}1^*_{j,\delta}+n^{-1}\sum_{j=1}^{n}1^*_{j,\delta}-p_\delta\right|\right)\\
&\leq m\mathbb{E}\left(\left|n^{-1}N_{n,\delta}-n^{-1}\sum_{j=1}^{n}1^*_{j,\delta}\right|\right)+m\mathbb{E}\left(\left|n^{-1}\sum_{j=1}^{n}1^*_{j,\delta}-p_\delta\right|\right)\\
&= mn^{-1}\sum_{j=1}^{n}\left(\mathbb{E}\left(1_{j,\delta}\right)-\mathbb{E}\left(1^*_{j,\delta}\right)\right)+mn^{-1}\mathbb{E}\left(\left|\sum_{j=1}^{n}\left(1^*_{j,\delta}-p_\delta\right)\right|\right).
\end{aligned}
$$

The convergence to 0 as $n\to\infty$ of the first term in the last expression is again guaranteed by equation (3.29) and the fact that $m(n)=o(n^{1/2})$. For the second part we proceed as with B_n, yielding

$$
\begin{aligned}
mn^{-1}\mathbb{E}\left(\left|\sum_{j=1}^{n}\left(1^*_{j,\delta}-p_\delta\right)\right|\right) &\leq mn^{-1}\left(\mathbb{E}\left(\left[\sum_{j=1}^{n}\left(1^*_{j,\delta}-p_\delta\right)\right]^2\right)\right)^{1/2} \\
&= mn^{-1}\left(\mathbb{E}\left(\sum_{j=1}^{n}\left(1^*_{j,\delta}-p_\delta\right)^2\right)\right. \\
&\quad\left.+2\sum_{j=1}^{n}\sum_{i=1}^{j-1}\mathbb{E}\left(\left(1^*_{i,\delta}-p_\delta\right)\left(1^*_{j,\delta}-p_\delta\right)\right)\right)^{1/2} \\
&\leq mn^{-1}\left(n+2\sum_{j=1}^{n}\sum_{i=1}^{j-1}\left(\mathbb{E}\left(1^*_{i,\delta}1^*_{j,\delta}\right)-p_\delta^2\right)\right)^{1/2} \\
&\leq mn^{-1}\left(n+2n\sum_{j=1}^{n-1}\left|r_{j,\delta}-p_\delta^2\right|\right)^{1/2}.
\end{aligned}
$$

We note that the sum involving the terms $r_{j,\delta}-p_\delta^2$ converges as $n\to\infty$ because of step 3 in the proof of Theorem 3.6.11. Finally, we notice that $mn^{-1}n^{1/2}\to 0$ as $n\to\infty$ because $m(n)=o(n^{1/2})$.

(2) From part (1) in this proof, it suffices to see that

$$\mathbb{E}\left(\left|\hat{\sigma}^2_{n,\delta}-\tilde{\sigma}^2_{n,\delta}\right|\right)\to 0 \quad \text{as } n\to\infty. \tag{4.6}$$

We first note that $\hat{\gamma}_{n,\delta}(0)=\tilde{\gamma}_{n,\delta}(0)$, so condition (4.6) boils down to proving

$$\mathbb{E}\left(\left|\sum_{k=1}^{m}\left(\hat{\gamma}_{n,\delta}(k)-\tilde{\gamma}_{n,\delta}(k)\right)\right|\right)\to 0, \quad \text{as } n\to\infty.$$

Writing the expressions of $\hat{\gamma}_{n,\delta}(k)$ and $\tilde{\gamma}_{n,\delta}(k)$, and regrouping conveniently, we get

$$
\begin{aligned}
&\mathbb{E}\left(\left|\sum_{k=1}^{m}\left(\hat{\gamma}_{n,\delta}(k)-\tilde{\gamma}_{n,\delta}(k)\right)\right|\right)\\
=&\mathbb{E}\Bigg(\Bigg|\sum_{k=1}^{m}\Bigg((n-k)^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}\right)-\left(n^{-1}N_{n,\delta}\right)^2\\
&-n^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}-n^{-1}N_{n,\delta}1_{j,\delta}-n^{-1}N_{n,\delta}1_{j+k,\delta}+\left(n^{-1}N_{n,\delta}\right)^2\right)\Bigg)\Bigg|\Bigg)\\
=&\mathbb{E}\Bigg(\Bigg|\sum_{k=1}^{m}\Bigg((n-k)^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}\right)-n^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}\right)\\
&-\left(n^{-1}N_{n,\delta}\right)^2+n^{-2}N_{n,\delta}\sum_{j=1}^{n-k}1_{j,\delta}+n^{-2}N_{n,\delta}\sum_{j=1}^{n-k}1_{j+k,\delta}-n^{-3}(n-k)N_{n,\delta}^2\Bigg)\Bigg|\Bigg)\\
\leq&\mathbb{E}\Bigg(\sum_{k=1}^{m}\left|(n-k)^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}\right)-n^{-1}\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}\right)\right|\\
&+n^{-1}N_{n,\delta}\sum_{k=1}^{m}\left|n^{-1}\sum_{j=1}^{n-k}1_{j,\delta}+n^{-1}\sum_{j=1}^{n-k}1_{j+k,\delta}-n^{-2}(2n-k)N_{n,\delta}\right|\Bigg), \qquad (4.7)
\end{aligned}
$$

where the last bound follows from the triangle inequality. Now, for the first sum in the expectation we have

$$
\begin{aligned}
\mathbb{E}\left(\sum_{k=1}^{m}\left(\left(\frac{1}{n-k}-\frac{1}{n}\right)\sum_{j=1}^{n-k}\left(1_{j,\delta}1_{j+k,\delta}\right)\right)\right)&=\sum_{k=1}^{m}\left(\frac{k}{n(n-k)}\sum_{j=1}^{n-k}\left(\mathbb{E}\left(1_{j,\delta}1_{j+k,\delta}\right)\right)\right)\\
&\leq n^{-1}\sum_{k=1}^{m}k\\
&=n^{-1}\frac{m(m+1)}{2}\\
&\to 0,
\end{aligned}
$$

as $n\to\infty$ since $m(n)=o(n^{1/2})$.

For the second sum in the expectation in (4.7) we use the triangle inequality and

$n^{-1}N_{n,\delta} \leq 1$, yielding

$$\begin{aligned}
&\mathbb{E}\left(n^{-1}N_{n,\delta}\sum_{k=1}^{m}\left|n^{-1}\sum_{j=1}^{n-k}1_{j,\delta}+n^{-1}\sum_{j=1}^{n-k}1_{j+k,\delta}-n^{-2}(2n-k)N_{n,\delta}\right|\right)\\
&\leq \mathbb{E}\left(\sum_{k=1}^{m}\left(\left|n^{-1}N_{n,\delta}+n^{-1}N_{n,\delta}-n^{-2}(2n-k)N_{n,\delta}\right|+n^{-1}\sum_{j=n-k+1}^{n}1_{j,\delta}+n^{-1}\sum_{j=1}^{k}1_{j,\delta}\right)\right)\\
&\leq \mathbb{E}\left(\sum_{k=1}^{m}\left(\frac{k}{n^2}N_{n,\delta}+\frac{2k}{n}\right)\right)\\
&\leq 3n^{-1}\sum_{k=1}^{m}k \to 0,
\end{aligned}$$

as $n \to \infty$ reasoning as in the previous case and the result is proved. □

4.2 Maximum Likelihood Estimation in the LDM

Once probabilistic properties for the LDM have been established in Chapter 3, the next step is addressing the question of making inference for such model based on δ-records. To that end, a natural approach is to consider maximum likelihood estimation.

Inference based on records in the LDM was studied for the first time by Smith in 1988 [110]. In that paper, maximum likelihood estimation was performed numerically for different trend models. Later, Feuerverger and Hall [28] considered a least-squares approach based on bootstrap techniques, while Hoayek et al. [64] proposed distribution-free estimators for the increasing variance model, which was shown to be equivalent to the LDM if the r.v. (X_n) have a Gumbel distribution, and thus losing the distribution-free property for the specific case of the LDM.

Here we study maximum likelihood estimation when using δ-record observations and assess the effect of the δ parameter in the estimation problem. While, in most instances, maximization must be carried out numerically, we obtain the explicit expression of the Maximum Likelihood Estimators (hereafter referred as MLE) for a family of distributions. Taking $\delta = 0$, we get the explicit expression of the MLE using usual records for that family. To the best of our knowledge, no explicit MLE based on records in the LDM have appeared in the literature.

The use of δ-records for estimation in the CRM has been considered in [45, 46, 53], both in the frequentist and the Bayesian framework, for estimation of the parameters and prediction of future records. Those papers show that the information provided by δ-records can be successfully used and estimations and predictions based on δ-records outperform those based on records only. Therefore, it is expectable that the use of δ-records is also advantageous in the LDM. As in those papers, we work only with $\delta \leq 0$; taking $\delta > 0$ would leave out the information provided by records, which are in fact necessary to compute δ-records, so it seems pointless to make inferences based on δ-records for $\delta > 0$.

We work with observations (Y_n) that follow a deterministic trend model, that is, (Y_n) is a sequence of r.v. such that

$$Y_j = X_j + t_j(c), \tag{4.8}$$

where (X_n) is a sequence of absolutely continuous i.i.d. r.v with cdf F and pdf f, and $t_n(c)$ is a deterministic sequence depending on a single parameter c.

We consider the following sampling scheme of the observations Y_j. First, a negative δ parameter is chosen to define the δ-record condition. Realizations of the trend model are drawn sequentially, and they are observed only if their value is a δ-record. That is, we observe realizations from the sequence

$$\mathbb{y} = (Y_1 1_{1,\delta}, Y_2 1_{2,\delta}, \dots, Y_n 1_{n,\delta}) \,. \tag{4.9}$$

Note that this is exactly the same sampling scheme considered in [110], taking δ-records instead of records (except that Smith considers lower records while we consider upper records). That is, the estimation problem in [110] can be seen as a particular case of our problem with $\delta = 0$.

Our aim is to estimate the parameters of the model using the information of the sample based only on δ-records. We remark here that the sampling scheme allows us to work with both the δ-record values as well as δ-record times.

We assume that the sequence of the underlying r.v. (X_n) depends on an unknown vector parameter θ in a parameter space Θ, that is, $X_n \sim f(x;\theta)$. Since the model has also the trend parameter c, which can be assumed to belong to a real set $\mathcal{C}$, the parameter space in the maximum likelihood estimation problem can be written as $\mathcal{C} \times \Theta$.

Given a realization of $\mathbb{y}$ of size n, $(y_1, \dots, y_n)$ and writing m_j for $\max\{y_1, \dots, y_j\}$, $j = 1, \dots, n$, we define the sequence (v_n) as:

$$v_j = \begin{cases} y_j & \text{if } y_j \text{ is a } \delta\text{-record,} \\ m_j + \delta & \text{otherwise.} \end{cases} \tag{4.10}$$

Example 4.2.1. In example 1.2.8, we considered the sequence

$$2, 4, 3, 6, 1, 6, 7, 1, 7, 8, 6, 7, 2, 4, 5, 8, 12.$$

Choosing $\delta = -1$, the observed sample of $\mathbb{y}$ is

$$\mathbb{y} = (2, 4, 0, 6, 0, 6, 7, 0, 7, 8, 0, 0, 0, 0, 0, 8, 12),$$

and the values v_j are

$$(v_j)_{j=1}^n = (2, 4, 3, 6, 5, 6, 7, 6, 7, 8, 7, 7, 7, 7, 7, 8, 12).$$

Note that, in the case of records ($\delta = 0$), the sample would be

$$\mathbb{y} = (2, 4, 0, 6, 0, 0, 7, 0, 0, 8, 0, 0, 0, 0, 0, 0, 12),$$

and the values v_j are the partial maxima m_j:

$$(v_j)_{j=1}^n = (2, 4, 4, 6, 6, 6, 7, 7, 7, 8, 8, 8, 8, 8, 8, 8, 12),$$

which is the type of sequence used in the aforementioned papers.

Since observations (X_n) in the model (4.8) are independent, the likelihood of the sample of δ-records (4.9), with size n, can be written explicitly as

$$\mathcal{L}(\mathbb{y}; c, \theta) = \prod_{j=1}^{n} \left(f(v_j - t_j(c); \theta)\right)^{1_{j,\delta}} F\left(v_j - t_j(c); \theta\right)^{1-1_{j,\delta}}. \tag{4.11}$$

Taking $\delta = 0$ we recover the likelihood for usual records obtained in [110] (written for upper instead of lower records).

Note that a location change in the deterministic trend model (4.8) can be due both to a change in the deterministic trend function $t_n(c)$ and to a change in the location parameter of the underlying r.v. (X_n).

We now assume that the observations (Y_n) are drawn from the LDM (1.2), which is a particular case of the deterministic trend model where $t_n(c) = cn$. Note that under the definition of the trend in the LDM we only consider location changes in the location parameter of the underlying variables (X_n) and not in the deterministic trend, and so the model is identifiable. The likelihood for the LDM is then

$$\mathcal{L}(\mathbb{y}; c, \theta) = \prod_{j=1}^{n} \left(f(v_j - cj; \theta)\right)^{1_{j,\delta}} F\left(v_j - cj; \theta\right)^{1-1_{j,\delta}}. \tag{4.12}$$

Note that taking $|\delta|$ large enough such that all the observations are δ-records, the expression (4.12) is equal to the classical likelihood function in a linear model.

The maximum likelihood estimator for the model parameters, the trend c and the distribution parameters θ, are

$$\left(\hat{c}_{mle}, \hat{\theta}_{mle}\right) = \underset{(c,\theta)\in\mathcal{C}\times\Theta}{\arg\max}\, \mathcal{L}(\mathbb{y}; c, \theta). \tag{4.13}$$

The estimators $\hat{c}_{mle}$ and $\hat{\theta}_{mle}$ cannot be computed analytically in general. In Subsection 4.2.2 we obtain numerically the MLE with a grid-search algorithm, and we assess its performance for different models via Montecarlo simulation.

4.2.1 Analytical solution for the MLE in a family of distributions

In this subsection we compute analytically the expression of the MLE based on δ-record data for the family of distributions with pdf

$$f(x) = \begin{cases} \lambda e^{\lambda(x-\alpha)} & x \leq \alpha, \\ 0 & \text{otherwise}, \end{cases} \tag{4.14}$$

and cdf $F(x) = e^{\lambda(x-\alpha)}$ if $x < \alpha$ and 1 otherwise. Note that in this family of distributions there is a location parameter α, and a shape parameter λ. Since these distributions are essentially equivalent to the shifted exponential distribution multiplied by -1, we will refer to this family as opposite-shifted exponential family. Additionally, this family is an extension of the Type III max-stable distribution, for which the probability of δ-record was analyzed in detail in example 3.3.2 if $\alpha = 0$ and $\lambda = 1$, which is particularly interesting since it is a limiting extreme value distribution, and thus it can be used to model the maxima arising from random variables in the Weibull class.

Writing $\Phi = \{j \in \{1, \ldots, n\} : 1_{j,\delta} = 1\}$, i.e., the set of indexes of δ-records in the sample, formula (4.11) for this family of distributions is

$$\mathcal{L}(\mathbb{y}; c, \alpha, \lambda) = \lambda^{N_{j,\delta}} e^{\lambda \sum_{j\in\Phi}(v_j - \alpha - cj)} \prod_{j\in\Phi} 1_{\{v_j \leq \alpha + cj\}} e^{\lambda \sum_{j\notin\Phi} \min\{v_j - \alpha - cj, 0\}}, \tag{4.15}$$

where $N_{n,\delta} = \sum_{j=1}^{n} 1_{j,\delta}$ is the number of δ-records in the sample.

In order to write (4.15) in a more manageable form, we prove the following property of the sequence (v_n).

Lemma 4.2.2. *Let $\alpha \in \mathbb{R}, c \geq 0$. Let (y_n) be a realization of $\mathbb{y}$ and (v_n) as defined in* (4.10). *Then, if*

$$y_j \leq \alpha + cj, \quad \textit{for all } j \textit{ such that } y_j \textit{ is a record}, \tag{4.16}$$

then

$$v_j \leq \alpha + cj, \tag{4.17}$$

for all $j = 1, \ldots, n$.

Proof. Since records are δ-records and $v_j = y_j$, the condition (4.17) trivially holds when observation j is a record. Suppose now that observation j is a δ-record but not a record and let $i < j$ be such that $y_i = m_j$ (the last record index before j). Then

$$v_j = y_j < m_j = m_i = y_i \leq \alpha + ci < \alpha + cj,$$

so (4.17) is proved when observation j is a δ-record.

Take now j such that observation j is not a δ-record, that is $y_j < m_j - a$. As above, let $i < j$ such that $y_i = m_j$, then

$$v_j = m_j - a = m_i - a = y_i - a \leq \alpha + ci - a < \alpha + cj$$

and the result is proved. □

We remark that Lemma 4.2.2 is a consequence of the definitions of record, δ-record and the sequence (v_n), so it can be applied for any distribution and not only for the family of distributions considered in this section.

By Lemma 4.2.2, the formula (4.15) can be written as:

$$\mathcal{L}(\mathbb{y}; c, \alpha, \lambda) = \lambda^{N_{n,\delta}} e^{\lambda \sum_{j=1}^{n}(v_j - \alpha - cj)} \prod_{j \in \Phi_R} 1_{\{y_j \leq \alpha + cj\}}, \tag{4.18}$$

where $\Phi_R = \{j \in \{1, \ldots, n\} : 1_{j,0} = 1\}$ is the set of indexes of records in the sample.

We can now give the explicit expression of the MLE in this family of distributions. Since there are three parameters to estimate (α, c, λ), we consider all the possibilities regarding which are known or unknown.

Theorem 4.2.3. *Let $\alpha \in \mathbb{R}$, $c \geq 0$, $\lambda > 0$. Let f be a density function as in* (4.14) *and consider a sample of size n as in* (4.9). *Let (v_n) be defined as in* (4.10). *Then:*

(a) The MLE of λ is:

$$\hat{\lambda}_{mle} = \frac{N_{n,\delta}}{\sum_{j=1}^{n} (\alpha + cj - v_j)}, \tag{4.19}$$

where α and c are either their values (if they are known) or their maximum likelihood estimations given in part (b).

(b) Both when λ is known and unknown, we have:

1. if α is unknown and c is known:

$$\hat{\alpha}_{mle} = \max_{j \in \Phi_R} \{y_j - cj\},$$

2. if α is known and c is unknown:

$$\hat{c}_{mle} = \max_{j \in \Phi_R} \left\{ \frac{y_j - \alpha}{j} \right\},$$

3. if α and c are unknown, the MLE of (α, c) is

$$(\hat{\alpha}_{mle}, \hat{c}_{mle}) = \left(\frac{jy_i - iy_j}{j-i}, \frac{y_j - y_i}{j-i} \right)$$

with $i, j \in \Phi_R$ or $(\hat{\alpha}_{mle}, \hat{c}_{mle}) = (y_l, 0)$, where $l = \max\{k \in \Phi_R\}$, such that $\hat{\alpha}_{mle} + \hat{c}_{mle} k \geq v_k$, for all $k \in \Phi_R$, which maximizes $2\hat{\alpha}_{mle} + (n+1)\hat{c}_{mle}$.

Proof. (a) From (4.18), we get that the likelihood is positive if and only if (4.16) holds. Thus, when α and c satisfy this condition, the logarithm of the likelihood is

$$\log \mathcal{L}(\mathrm{y}; c, \alpha, \lambda) = N_{n,\delta} \log \lambda - \lambda \sum_{j=1}^{n} (\alpha + cj - v_j).$$

Taking the derivative with respect to λ, we observe that, for every α, c such that (4.16) holds, the likelihood is maximized when λ is taken equal to (4.19).

(b) When λ is known, then, if α and c are such that (4.16) is satisfied, the likelihood is decreasing in $\sum_{j=1}^{n}(\alpha + cj - v_j)$. If λ is unknown, replacing it by its MLE in (a), the logarithm of the likelihood is equal to

$$N_{n,\delta} \left(\log N_{n,\delta} - \log \left(\sum_{j=1}^{n} (\alpha + cj - v_j) \right) - 1 \right),$$

which is decreasing in $\sum_{j=1}^{n} (\alpha + cj - v_j)$. So, in both cases, α and c must be taken such that $\sum_{j=1}^{n} (\alpha + cj - v_j)$ is minimized subject to condition (4.16). Therefore, the MLE of α and c are the solution of the following linear programming problem:

$$\begin{aligned} \text{Min. } & 2\alpha + (n+1)c \\ \text{s.t. } & \alpha + cj \geq v_j, \qquad j \in \Phi_R, \\ & c \geq 0. \end{aligned}$$

When either α or c are known, the solutions of (4.2.1) are readily shown to be those in the statement of the theorem. When both are unknown, then, the feasible set is clearly nonempty (by taking α and c large enough) and the problem is bounded (due to the signs of the coefficients in the objective function and the constraints). Therefore, the solution is an extremal point of the feasible set. The intersections of the lines defining the constraints are of the form

$$\left(\frac{jy_i - iy_j}{j-i}, \frac{y_j - y_i}{j-i} \right), \qquad i, j \in \Phi_R$$

and $(y_j, 0)$, $j \in \Phi_R$, which proves the result.

□

Remark 4.2.4. (a) By Theorem 4.2.3 the MLE of α and c, both when λ is known and unknown, depend only on the sequence of records, and not on near-records. Therefore, changing the value of δ does not change the estimation of α and c, but it does affect the estimation of λ.

(b) Taking $\delta = 0$ in Theorem 4.2.3 gives the explicit expression of the MLE of the parameters based on usual records. To the best of our knowledge, this is the first result giving explicit expressions of the MLE in the LDM, since all previous works searched for the estimators numerically.

(c) If we work with lower records instead of upper records and take a negative drift, as in [110], and consider that the distribution of X_n is the shifted exponential with parameter λ, that is $f(x; \alpha, \lambda) = \lambda e^{-\lambda(x-\alpha)} 1_{\{x \geq \alpha\}}$, then Theorem 4.2.3 can be modified in a straightforward way to yield the analytical expression of the MLE of the parameters.

4.2.2 Numerical results

In this section we assess the behaviour of the MLE for three families of distributions: shifted exponential, normal and opposite-shifted exponential. All of them have a location parameter $\alpha \in \mathbb{R}$ and a scale parameter $\lambda > 0$. Their respective pdf are:

$$\begin{aligned} f_1(x; \alpha, \lambda) &= \lambda e^{-\lambda(x-\alpha)} 1_{\{x \geq \alpha\}}, \\ f_2(x; \alpha, \lambda) &= \frac{\lambda}{\sqrt{2\pi}} e^{-(x-\alpha)^2 \lambda^2/2}, \\ f_3(x; \alpha, \lambda) &= \lambda e^{\lambda(x-\alpha)} 1_{\{x \leq \alpha\}}. \end{aligned}$$

Note that, in the three distributions, the parameter λ has been chosen to be the inverse of the standard deviation.

In order to analyze the behaviour of the MLE and the effect of different values of δ on the accuracy of the estimations, we use Montecarlo simulation. We have considered different settings: in the three distributions, we have kept the value of α equal to 5, and the value of c equal to 0.1. We have taken different values of λ =0.25, 0.5, 0.75 and 1, and different values of δ equal to 0 (usual records), -0.5, -1 and -1.5. These values have been chosen to represent situations in which the trend is small with respect to the variability of the underlying random variables. Note that if the trend is large, many observations will be record and δ-record, and therefore the estimates will be much more accurate. Furthermore, in practical applications it is common to work with weak trends, as in the case of temperatures in a climate change framework (see Section 4.3), or even to use these methods as a tool to detect the existence of such trends.

For each setting we have carried out 1000 simulation runs of length $n = 100$ observations. We have computed the MLE of α, c, λ in all 7 combinations defined by which parameters are unknown. In each case, the MSE (defined as the average of the square of the estimations minus the actual value) was computed. For the exponential and normal distributions, the MLEs of the parameters were found via a grid search; for the opposite-shifted exponential, the exact values of the MLEs were obtained with the formulas in Theorem 4.2.3. Tables 4.1, 4.2 and 4.3 show the results. Note that Table 4.3 includes only rows where λ needs to be estimated; this is because our primary interest is to assess the improvement in the estimations when using δ-records instead of records only, and we showed in Section 4.2.1 that, for this family of distributions, the MLEs for α and c using δ-records are equal to those obtained with records.

The tables are split by horizontal lines depending on the unknown parameters. Thus, for instance, the first line in Table 4.1 is the MSE of the estimation of α when c and λ are known; the fourth and fifth lines are the MSE of α and c when λ is known; the three next to last lines are the MSE of the three parameters when all are unknown. The last line of the tables is the mean number of δ-records in the sample. Also, some of the parameters have a $\times 10^{-k}$ attached; this means that all the figures in the line must be multiplied by that constant in order to get the MSE; for instance, in Table 4.1, the second line has a $\times 10^{-4}$, which means that all the line has to be multiplied by 10^{-4}, so the figure in the first column means that the MSE of the MLE of c for records when $\lambda = 0,25$ is equal to $21,18 \times 10^{-4}$.

We note that, both in the case of records and δ-records, the MSE is lower when there are more known parameters; of course this could be expected because all the information in the data is used to estimate a fewer number of parameters. Another common feature in the tables is that the estimation of α and c, both using records and δ-records is much better when there is less uncertainty in the model, that is, for greater values of λ. One source of this improvement is that there are more records and δ-records when there is less uncertainty in the model (see the last lines in the tables), but this does not account for all it. Indeed, note that, for instance, in Table 4.1 the MSE of the MLE of α using records when c and λ are known, with $\lambda = 0.25$, is equal to 1.89, and it is 0.06 when $\lambda = 1$, that is 30 times smaller, while the number of records only doubles from 6.67 to 12.82. A similar pattern is observed for c. For λ, even though the MSE increases as λ grows (for instance, when α and c are known, the MSE for λ using records ranges from 1.53×10^{-3} to 1.724×10^{-2}), this MSE should be taken relative to the value of λ. This is because multiplying λ by 4 (from 0.25 to 1) implies a factor of 16 in the MSE, so the value 1.724×10^{-2} for $\lambda = 1$ can be seen as smaller than 1.53×10^{-3} for $\lambda = 0.25$. When comparing Tables 4.1, 4.2 and 4.3, we observe that the MLE in the exponential distribution are the least accurate while those in the opposite-shifted exponential family are the most accurate. This can be explained by several factors: first, the exponential distribution generates less records and δ-records than the normal distribution, being the opposite-shifted exponential the one with the greatest number of records and δ-records; second, while the standard deviation of the three distributions is the same, the right-tail of the exponential distribution is heavier than the normal distribution

Table 4.1
Mean squared error of the MLE in the exponential distribution

Unknown	$\lambda = 0.25$				$\lambda = 0.50$				$\lambda = 0.75$				$\lambda = 1$			
		δ				δ				δ				δ		
parameter	records	-0.5	-1	-1.5	records	-0.5	-1	-1.5	records	-0.5	-1	-1.5	records	-0.5	-1	-1.5
α	1.89	1.75	1.55	1.45	0.44	0.36	0.26	0.21	0.15	0.10	0.06	0.04	0.06	0.03	0.02	0.01
$c\ (\times 10^{-4})$	21.18	19.47	18.06	14.04	3.21	2.61	2.02	1.41	0.95	0.65	0.41	0.23	0.38	0.19	0.08	0.01
$\lambda\ (\times 10^{-3})$	1.53	1.53	1.37	1.25	5.23	4.69	4.30	4.26	10.23	9.33	8.32	7.36	17.24	14.94	13.80	12.31
α	2.73	2.55	2.42	2.28	1.05	0.90	0.66	0.57	0.45	0.35	0.22	0.15	0.23	0.15	0.07	0.03
$c\ (\times 10^{-4})$	27.05	26.32	24.22	20.41	6.30	5.39	4.08	3.19	2.13	1.59	1.00	0.64	0.94	0.59	0.30	0.11
α	3.57	3.26	2.89	2.83	1.19	0.80	0.55	0.39	0.39	0.21	0.11	0.07	0.16	0.08	0.03	0.01
$\lambda\ (\times 10^{-3})$	4.88	4.76	4.19	4.06	27.61	18.91	12.27	9.89	45.10	24.53	16.13	12.00	59.53	31.32	21.11	15.21
$c\ (\times 10^{-4})$	36.12	33.31	29.40	28.49	8.78	8.02	4.91	4.70	2.53	1.61	1.40	0.70	0.91	0.48	0.21	0.04
$\lambda\ (\times 10^{-3})$	5.24	4.62	4.36	3.82	16.08	13.95	9.55	9.02	29.96	21.51	15.27	11.33	44.13	24.16	18.81	14.12
α	3.60	3.24	2.96	2.93	1.48	1.09	0.85	0.64	0.66	0.40	0.23	0.16	0.31	0.17	0.06	0.03
$c\ (\times 10^{-4})$	33.82	31.90	27.87	26.31	8.48	7.92	5.61	5.54	2.81	2.05	1.82	1.00	1.14	0.78	0.36	0.16
$\lambda\ (\times 10^{-3})$	11.49	9.17	8.42	7.17	39.75	26.31	15.68	12.58	64.20	33.96	20.92	13.70	79.10	35.50	21.80	14.77
$card(\Phi)$	6.67	7.39	8.27	9.18	8.53	10.59	13.17	16.21	10.67	15.08	20.50	28.81	12.82	19.94	30.83	45.30

Table 4.2
Mean squared error of the MLE in the normal distribution

	$\lambda = 0.25$				$\lambda = 0.50$				$\lambda = 0.75$				$\lambda = 1$			
Unknown		δ				δ				δ				δ		
parameter	records	-0.5	-1	-1.5	records	-0.5	-1	-1.5	records	-0.5	-1	-1.5	records	-0.5	-1	-1.5
α	0.64	0.51	0.50	0.47	0.12	0.09	0.08	0.07	0.05	0.03	0.03	0.02	0.02	0.02	0.01	0.01
c $(\times 10^{-5})$	23.83	19.13	18.17	16.54	4.67	3.83	3.19	2.86	1.78	1.16	0.88	0.57	0.59	0.27	0.07	0.05
λ $(\times 10^{-3})$	1.01	0.87	0.96	0.94	3.60	3.29	3.03	2.99	7.63	6.99	6.09	5.78	12.67	11.23	9.16	6.96
α	1.70	1.57	1.43	1.37	0.43	0.37	0.31	0.28	0.18	0.15	0.12	0.12	0.10	0.08	0.06	0.06
c $(\times 10^{-5})$	62.14	57.08	51.43	46.21	13.86	12.13	9.72	8.76	5.82	4.68	3.94	3.68	3.29	2.56	2.05	1.73
α	2.07	1.74	1.38	1.09	0.38	0.27	0.19	0.12	0.13	0.07	0.05	0.03	0.05	0.03	0.02	0.01
λ $(\times 10^{-3})$	6.00	4.81	3.45	2.63	15.29	8.97	6.88	5.12	25.19	13.88	9.37	7.31	35.94	19.43	11.51	7.92
c $(\times 10^{-5})$	72.03	59.24	53.05	38.92	10.06	7.74	5.72	4.25	3.62	2.22	1.71	0.91	1.43	0.79	0.38	0.13
λ $(\times 10^{-3})$	3.58	2.83	2.78	2.22	10.25	7.61	5.77	4.97	19.29	12.16	8.59	6.87	22.51	15.08	10.29	7.12
α	2.56	2.34	2.04	1.80	0.67	0.54	0.43	0.36	0.25	0.16	0.12	0.09	0.09	0.06	0.02	0.02
c $(\times 10^{-5})$	79.48	70.04	66.20	58.48	17.52	14.52	12.40	10.88	6.41	4.56	3.36	2.24	2.00	1.28	0.24	0.20
λ $(\times 10^{-3})$	8.11	6.06	4.58	3.32	18.58	10.34	7.55	5.41	26.63	14.18	9.32	7.44	36.23	19.62	11.58	7.92
$card(\Phi)$	8.55	10.19	12.12	14.18	12.16	16.78	22.58	29.48	15.64	24.44	34.99	47.76	18.68	32.07	47.97	64.69

Table 4.3
Mean squared error of the MLE in the opposite-shifted exponential distribution

Unknown	$\lambda = 0.25$				$\lambda = 0.50$				$\lambda = 0.75$				$\lambda = 1$			
		δ				δ				δ				δ		
parameter	records	-0.5	-1	-1.5	records	-0.5	-1	-1.5	records	-0.5	-1	-1.5	records	-0.5	-1	-1.5
$\lambda\ (\times 10^{-3})$	3.60	2.08	1.84	1.46	9.80	5.85	4.83	3.79	17.69	10.92	9.18	7.55	30.84	15.77	15.18	12.01
$\alpha\ (\times 10^{-3})$	3.58	3.58	3.58	3.58	0.77	0.77	0.77	0.77	0.37	0.37	0.37	0.37	0.21	0.21	0.21	0.21
$\lambda\ (\times 10^{-3})$	4.40	2.35	2.02	1.54	11.31	6.24	5.10	4.08	20.11	11.49	9.75	7.93	34.99	16.92	15.89	12.83
$c\ (\times 10^{-6})$	1.31	1.31	1.31	1.31	0.32	0.32	0.32	0.32	0.14	0.14	0.14	0.14	0.08	0.08	0.08	0.08
$\lambda\ (\times 10^{-3})$	4.39	2.34	2.03	1.56	11.30	6.27	5.10	4.06	20.14	11.48	9.73	7.96	34.60	16.79	15.81	12.81
$\alpha\ (\times 10^{-3})$	29.37	29.37	29.37	29.37	7.71	7.71	7.71	7.71	3.11	3.11	3.11	3.11	1.85	1.85	1.85	1.85
$c\ (\times 10^{-6})$	8.06	8.06	8.06	8.06	2.13	2.13	2.13	2.13	0.84	0.84	0.84	0.84	0.48	0.48	0.48	0.48
$\lambda\ (\times 10^{-3})$	5.43	2.76	2.23	1.68	13.42	6.91	5.48	4.38	23.37	12.40	10.51	8.39	39.79	18.22	16.83	13.76
$card(\Phi)$	20.17	29.54	37.86	45.18	26.79	42.98	55.52	65.54	31.78	52.75	68.03	77.73	35.83	61.03	76.25	85.80

(and, of course, than that of the opposite-shifted exponential family, which has a finite right endpoint in its support); third, the parameter α in the opposite-shifted exponential family is precisely the right endpoint of the support of the distribution, so it is expectable that a very good estimation is obtained since our sampling is based on upper records (and δ-records), that is, we are sampling the rightmost part of the distribution.

Regarding the effect of using δ-records we observe that, in most situations, the inclusion of δ-records in the sample improves estimations. This could be expected since δ-records include more information than usual records, but Theorem 4.2.3 shows that, for some distributions, this is not the case. Indeed, for the opposite-shifted exponential family of distributions, the MLE for α and c are based only on records, so taking δ-records does not improve their accuracy; the MLE of λ, however, does improve with the use of δ-records, as shown in Table 4.3. Except for those parameters in Table 4.3, in the rest of settings and parameters the MLE using δ-records outperform those using records only. Looking at the tables, we can see that the decrease in the MSE of the estimations using δ-records is roughly proportional to the number of extra observations included in the sample (near-records). For instance, in Table 4.1, we see that, for $\lambda = 0.25$, when c and λ are known, there is a reduction from 1.89 to 1.45 in the MSE of the MLE of α from using records to using δ-records with $\delta = -1.5$, while the number of observations increases from 6.67 to 9.18; that is, by multiplying the number of observations by 1.37 we divide the MSE by 1.31. This is consistent with what happens in i.i.d. samples, where the MSE of the MLE of the parameters is inversely proportional to the number of observations.

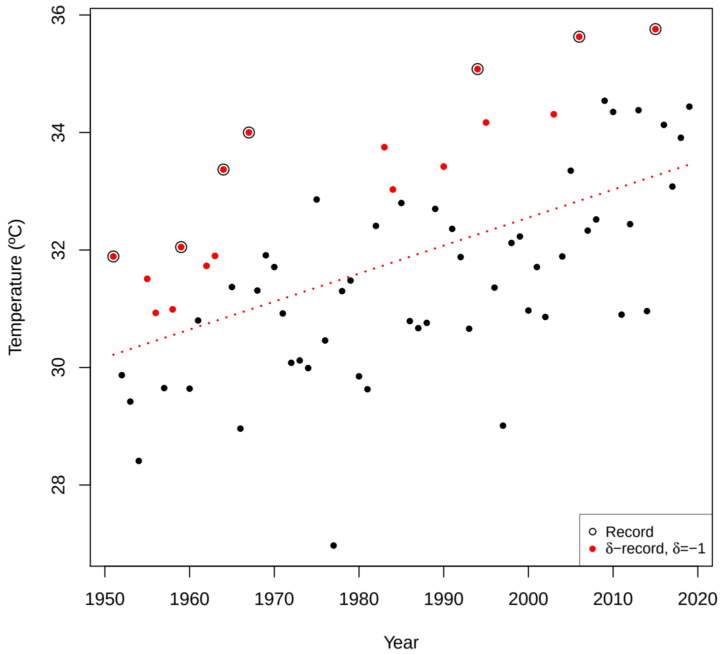
In Section 4.3 we show the MLE of the parameters, based on δ-records, in a real data set of temperatures, where the drift is induced by the global warming phenomenon.

4.3 Application

We present a practical application of the the theoretical results obtained in Chapters 3 and 4 to a real dataset of temperatures, where convergence to the stationary regime is seen for quite small values of n. As pointed out in the introduction, the LDM has been used by different authors to model temperature record data in the framework of climate-change.

Our dataset consists of means of daily maximum temperatures (in degrees Celsius), for every month of July, from 1951 to 2019, in the city of Saragossa, Spain. See Figure 4.1 for a data plot. For δ-records we choose the value $\delta = -1$, which is arbitrary and does not respond to any specific reason, other than interpretability of the example. Note that a year will have a δ-record temperature if the maximum

Figure 4.1
Monthly mean of maximum temperature in July, 1951-2019 in Saragossa (Spain).



average temperature in July is a record or if it is at a distance smaller than $1°C$ from the current maximum. In this framework, we find that 17 out of the 69 observations are δ-records (coloured in red), and 7 of them are records (with circle). The least squares line fitted to the data (in dotted red), reveals a linear increase of the maximum temperatures over time.

The simple linear model for the temperature takes the form

$$T_t = \beta_0 + \beta_1 t + \varepsilon_t, \tag{4.20}$$

where T_t is the temperature of year t (indexed by $\{1, 2, \ldots, 69\}$), and ε_t the error term. The results of the least-square estimators of the coefficients and their p-values (assuming Gaussian errors, with zero mean and standar deviation σ) are shown in Table 4.4.

In addition, we find an adjusted-R^2 of 0.2769 and an estimated standard deviation of $\sigma = 1.514$ degrees in the error terms. The hypothesis $\beta_1 = 0$ is clearly rejected, using the Student t-test. Moreover, the estimate of β_1, which represents the average increment of mean maximum temperatures by year, agrees well with previous estimates of the summer warming trend in Europe, see [124, 125]. Nevertheless, the intercept β_0 is irrelevant when counting δ-records.

In Figure 4.2 we show the classical diagnosis plots for linear regression. The top left panel indicates that a linear model is appropriate since no pattern in the

Table 4.4
Regression analysis estimations for the temperature data.

Coefficient	Estimate	Std.Error	p-value
β_0	30.1706	0.3686	< 2e-16
β_1	0.0476	0.00915	2.04e-06

residuals is observed. On the top right panel, the quantile-quantile plot of the residuals against the normal distribution, with all the values within the confidence lines, shows that the Gaussian assumption is adequate for errors; this is corroborated by a p-value of 0.58 in the Shapiro-Wilk test [102, Section 5.2.2] for normality of the standardized residuals. Moreover, the bottom panels show no significant autocorrelation or partial autocorrelation values, indicating the absence of serial correlation of the observations; this is confirmed by a p-value greater than 0.1 in the Kwiatkowski-Phillips-Schmidt-Shin test [75] for stationarity of a series around a deterministic trend. We conclude that the data are well fit by a linear regression in t, with Gaussian errors. Hence, an LDM with drift parameter $c = \hat{\beta}_1 = 0.0476$ and $(X_n)_{n\geq 1}$ independent normally distributed random variables, is adequate for the data. Note that, for applying Theorem 3.6.7, there is no need to assume any specific form of the distribution of the X_n.

Now, since 17 out of 69 observations were identified as δ-records, it is natural to estimate p_δ by the empirical record rate, that is,

$$\hat{p}_\delta = n^{-1}N_{n,\delta} = 17/69 \approx 0.2464.$$

Figure 4.3 illustrates how the empirical δ-record rate evolves with each extra observation and how it seems to stabilize around a constant value, as predicted by Theorem 3.6.7.

Concerning the asymptotic normality (Theorem 3.6.11), we need to estimate the variance σ_δ^2, defined in (3.28). In order to apply the estimators of the variance of the number of δ-records studied in Section 4.1 to our data, we must choose m, of order $o(\sqrt{n})$. In our case, $n = 69$ so we take $m = 8$ to obtain the estimates $\tilde{\sigma}_\delta^2 = 0.337$ and $\hat{\sigma}_\delta^2 = 0.331$, which are indeed very similar. Therefore, from Theorem 3.6.11, $N_{n,\delta}$ is approximately Gaussian, with mean 17 and variance 23.25 (0.337×69).

Now, we analyze the performance of these estimators for different values of $m = 5, 6, 7, 8$. In Tables 4.5 and 4.6 we show the results of applying estimators (4.1) and (4.3) respectively to compute the 95% confidence intervals for p_δ and the expected number of δ-records, $\mathbb{E}\left(N_{n,\delta}\right)$, with endpoints:

$$\hat{p}_\delta \pm 1.96\tilde{\sigma}_\delta/\sqrt{n}$$
$$n\hat{p}_\delta \pm 1.96\tilde{\sigma}_\delta\sqrt{n},$$

and the corresponding ones in the case $\hat{\sigma}_\delta^2$.

Figure 4.2
Diagnostic plots of the regression model. Top left: residuals vs year. Top right: quantile-quantile plot of the residual with the normal distribution. Bottom left: autocorrelation function. Bottom Right: partial autocorrelation function.

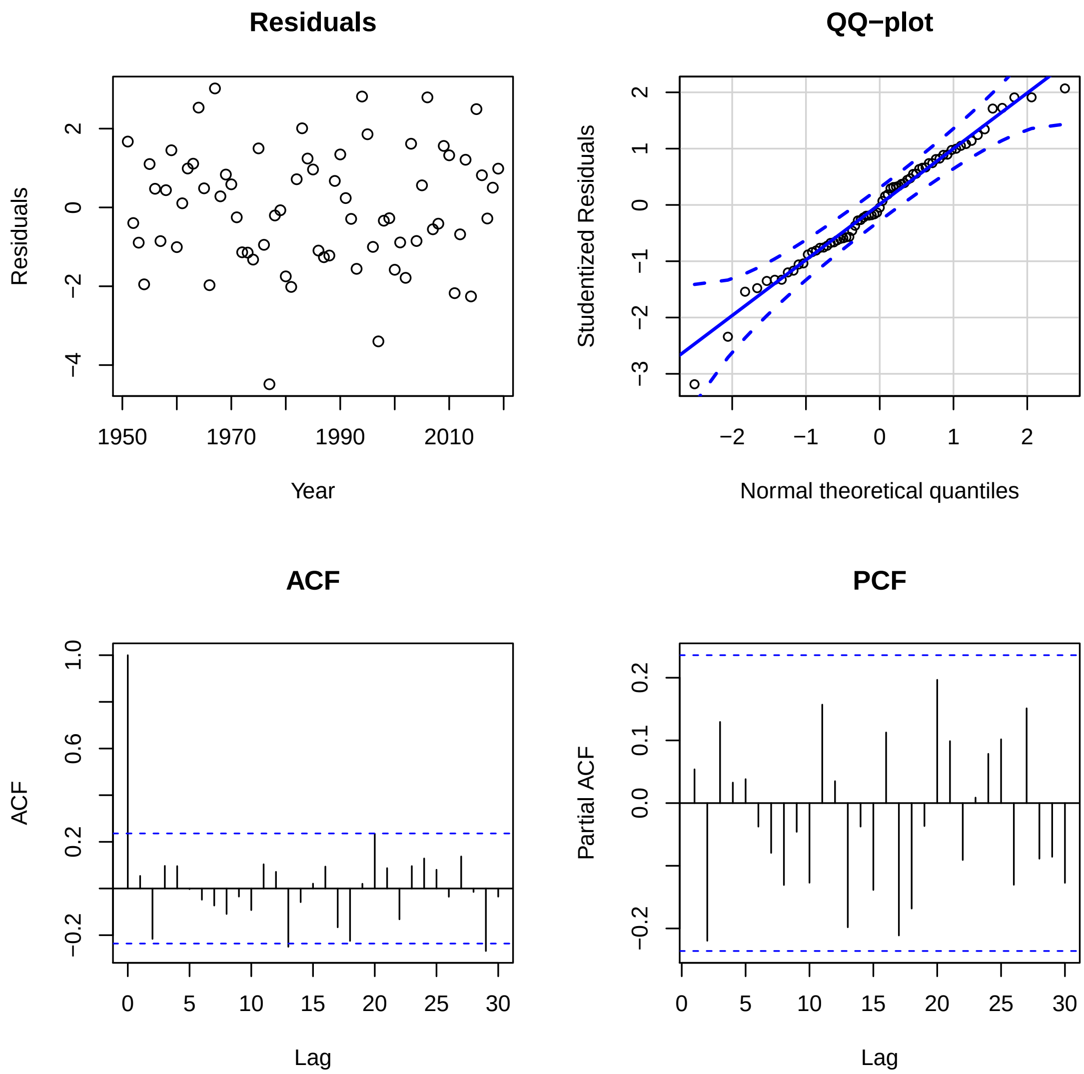


Figure 4.3
Evolution of the δ-record rate for the temperature data.

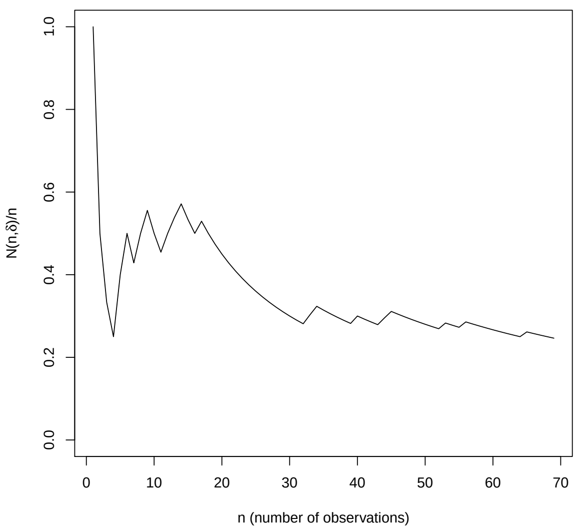


Table 4.5
Confidence intervals for the asymptotic δ-record rate and expected number of δ-records using $\tilde{\sigma}_\delta^2$ and different values of m.

	$m = 5$	$m = 6$	$m = 7$	$m = 8$
$\hat{p}_\delta - 1.96\tilde{\sigma}_\delta/\sqrt{n}$	0.118	0.121	0.119	0.109
$\hat{p}_\delta + 1.96\tilde{\sigma}_\delta/\sqrt{n}$	0.375	0.372	0.374	0.383
$n\hat{p}_\delta - 1.96\tilde{\sigma}_\delta\sqrt{n}$	8.114	8.362	8.198	7.543
$n\hat{p}_\delta + 1.96\tilde{\sigma}_\delta\sqrt{n}$	25.886	25.638	25.802	26.457

We can see that the results found are similar for both estimators and for every value of m considered. Also, we note that the intervals for both cases are not narrow, but this is clearly expected since we just have a small sample of 69 observations.

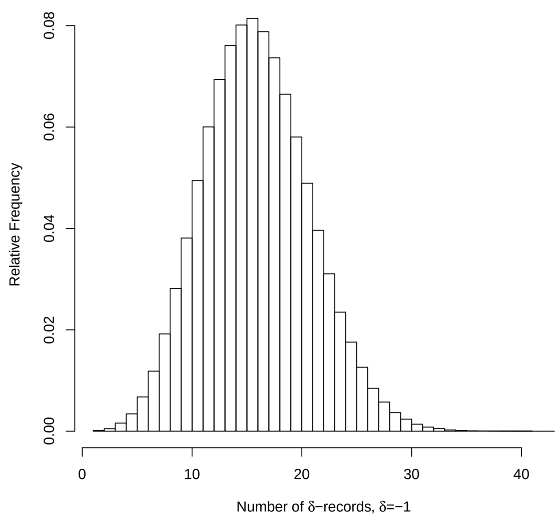
For assessing the goodness of fit, we simulate the adjusted model (4.20) 10^6 times, and compute the value of $N_{69,\delta}$. Figure 4.4 summarizes the total number of δ-records obtained at each of the 10^6 simulations. The histogram has a Gaussian shape, so the convergence in Theorem 3.6.11 to the Gaussian distribution seems to be fast. Moreover, the 0.025 and 0.975 quantiles of the normal distribution $N(17, 23.25)$ are, respectively, 7.54 and 26.45. The 0.025 and 0.975 empirical quantiles from the simulated data are 8 and 26, showing an excellent fit to the theoretical (asymptotic) distribution.

As a conclusion, we see that empirical results and theory are in very close agreement. This means that, even with a small sample, the approximations in Theorems 3.6.7 and 3.6.11 are good, at least for the model considered. Theoretical results take advantage over simulations since they do not need as many assumptions as the simulated results. In particular, simulations need the hypothesis of the distribution

Table 4.6
Confidence intervals for the asymptotic δ-record rate and expected number of δ-records using $\hat{\sigma}_\delta^2$ and different values of m.

	$m=5$	$m=6$	$m=7$	$m=8$
$\hat{p}_\delta - 1.96\tilde{\sigma}_\delta/\sqrt{n}$	0.116	0.121	0.120	0.111
$\hat{p}_\delta + 1.96\tilde{\sigma}_\delta/\sqrt{n}$	0.377	0.371	0.373	0.382
$n\hat{p}_\delta - 1.96\tilde{\sigma}_\delta\sqrt{n}$	7.976	8.369	8.252	7.630
$n\hat{p}_\delta + 1.96\tilde{\sigma}_\delta\sqrt{n}$	26.024	25.631	25.748	26.370

Figure 4.4
Histogram of the total number of δ-records for the adjusted regression model (10^6 iterations of 69 observations).



of the residuals, forcing us to conjecture the entire law of the data.

We now consider the estimation of the parameters of the model based only on δ-records. In table 4.7 we give the MLE of the parameters for the temperature data. The values have been obtained via a grid search as explained in Section 4.2. We observe that, even when only the 7 records in the sample are used, the estimations for β_0 and β_1 are very close to the ones obtained with all the data (69 observations). However, the MLE for σ using only records is far from the value when all the observations are used. We observe an improvement in the estimations, especially in that of σ when δ-records are used.

Table 4.7
MLE for the temperature data

	Records	$\delta=-0.5$	$\delta=-1$	$\delta=-1.5$	$\delta=-2$	$\delta=-2.5$	**All data**
β_0	30.811	30.732	30.613	30.375	30.089	30.233	**30.171**
β_1	0.055	0.047	0.046	0.051	0.054	0.047	**0.048**
σ	0.955	1.088	1.150	1.207	1.350	1.428	**1.514**
$card(\Phi)$	7	11	17	25	31	38	**69**

"I en acabat, que cadascú es vesteixi com bonament li plagui, i via fora!, que tot està per fer i tot és possible".

Miquel Martí i Pol

5

Conclusions and future work

Throughout this monograph we have obtained novel results for δ-record and near-record observations.

In Chapter 2 we have studied the point process of near-record values when the observations are discrete, taking values in the integers. To that end, we start from the record value process, which is the so-called Shorrock's process. We use Shorrock's theorem to obtain the distribution of the number of near-record values in a set associated to a record by conditioning on its value, finding that this number is distributed as a geometric random variable starting at zero.

In addition, we obtain the distribution of the near-record values associated to a certain record value in a set, and show that these quantities and their number enjoy good independence properties. We use these results and the classical theory of point processes to combine these properties, finding that the process of near-records with values in a given set is a cluster process and we characterize it through its probability generating functional (p.g.fl.). From this p.g.fl. we derive the exact distribution of the number of near-records in a set in the whole sequence via its probability generating function (p.g.f.). For this number of near-records, we also obtain explicit expressions for some quantities of interest, such as the expected number, the variance and the covariance.

These results are applied to different distributions. For example, for the geometric distribution it is found that both the expected number and the variance of the number of near-records with value in $[0, n]$ are of the order of n. For the distribution with hazard rates $r_i = 1 - 1/(i+1)$, we find that the number of near-records with value exactly n is of the order of n^a, and its variance is of the order of n^{2a}, where a is the near-record parameter.

Once these properties are established, we analyze the asymptotic behaviour of the number of near-records in the interval $[0, n]$ as n grows. In the first place, it is obtained that the condition $\sum_{i=1}^{\infty} r_i^2 < \infty$ is sufficient to guarantee that the number of near-records in the whole sequence is finite almost surely. In this case, we characterize the distribution of the total number of near-records by giving the explicit expression of its p.g.f. As an example of a family with $\sum_{i=1}^{\infty} r_i^2 < \infty$, we analyze the zeta distribution, with $y_k = (k+1)^{-1}$. For this family, the expected number of near-records is exactly the near-record parameter a, and we give an explicit expression for the variance.

On the contrary, if $\sum_{i=1}^{\infty} r_i^2 = \infty$ holds, we obtain that the number of near-records in the whole sequence is infinite a.s. under mild conditions. Moreover, we find interesting asymptotic results for the number of near-records with values in $[0, n]$, denoted by $\eta([0, n])$, as n goes to infinity. Indeed, we find a strong law of large numbers and a central limt theorem for $\eta([0, n])$ both for for *heavy-* or *medium-tailed* distributions (those with $\limsup_{k\to\infty} r_k < 1$) and for light tailed distributions (those with $r_k \to 1$) under some additional assumptions on the speed of growth of $r_k \to 1$.

The end of this chapter is devoted to the characterization of all discrete distributions, taking values in the integers, satisfying that $M_n - cN_{n,\delta}$ is a martingale. This problem has been studied for records in [49], and we reached interesting results in the case of δ-records. If $\delta < -1$ we show that there is no distribution satisfying the martingale condition. If $\delta = -1$, that is, considering weak-records, we prove that only the geometric distribution is a solution. If $\delta > 0$, the problem of finding these distributions is shown to be equivalent to the problem of guaranteeing certain conditions for the solutions of linear recurrence relations. In this case, the solution to the linear recurrence is the sequence of the values of the survival function, and therefore we have to guarantee that this sequence lies in the interval $[0, 1]$, it is non-increasing and its limit is 0. The problem of the positivity of solutions in recurrence relations is an open problem in the literature, and therefore we have not obtained the solution in the general case. However, the problem has been solved for the case $\delta = 1$, obtaining that the there exist solutions if and only if $0 \leq c \leq 1/4$, and that these solutions are convex combinations of the Dirac delta distribution and two geometric distributions in general, except in the frontier case $c = 1/4$, where we find another solutions related with (although not equal to) the negative binomial distribution.

In Chapter 3 we have studied the behaviour of δ-records in the LDM. We have analyzed the asymptotic probability of δ-records, the dependence between δ-record events and the limiting distribution of the number of δ-records among the first n observations.

The behaviour of the asymptotic probability of δ-record shows similarities with the case of records ($\delta = 0$); for instance, for positive c, $p_\delta(c) > 0$ if and only if $\mu^+ < \infty$, regardless the value of δ (except for the trivial case $\delta \geq x_+ - x_- + c$, where no δ-records are observed). We also find that $p_\delta(c)$ is a continuous function of δ for every c, while, as a function of c, it is continuous for every $c \neq 0$, and a discontinuity arises at $c = 0$, if $x_+ < \infty$ and $\delta < 0$. This differs from records where $p_0(c)$ is a continuous function of c.

We have described in detail the probability of δ-record in different examples. For the Gumbel distribution, an explicit expression for $p_\delta(c)$ is found, showing that it decreases with δ, as a logistic function of $-\delta$. For the cases studied in the Dagum family of distributions, we have $p_\delta(c) = 0$, for every δ, c, since $\mu^+ = \infty$. For this family, we investigate whether or not the speed of convergence of $p_{n,\delta}(c)$ to 0, as $n \to \infty$, depends on δ. Since random variables X_n, with $\mu^+ = \infty$, may produce large values provoking abrupt changes in record values, we can expect that δ values close to 0 have a negligible impact and so, $p_{n,\delta}(c)/p_{n,0}(c) \to 1$. This happens in the case $c = 0$, where the number of δ-records grows at the same speed as the number of records, when the X_n are heavy-tailed. However, we find that, for some distributions in the Dagum family, $p_{n,\delta}(c)/p_{n,0}(c) \to a \neq 1$. The Uniform and the Type III max-stable distributions, both in the Weibull class of extreme values are analyzed, obtaining exact expressions for $p_{n,\delta}$. We also propose first order approximations for $p_{n,\delta}$ in the LDM, assessing positively its performance in the LDM.

The parameter δ has a clear impact in the qualitative behaviour of correlations of δ-record events. First, the expression of the limiting correlation is different for $\delta \geq 0$ and $\delta < 0$. For the Gumbel distribution, where record indicators are independent [14], dependence appears when $\delta \neq 0$; in fact, δ-records in this distribution attract each other for $\delta < 0$ and repel each other, for $\delta > 0$. For distributions with power law tails, it is known, for $c > 0$, that correlations between records are positive and increase with n; see [34]. We have studied the Pareto distribution with $c = 1$, and obtained that, while the correlations are positive (and increasing in n) for negative, zero and small positive values of δ, they are negative for big values of δ. In fact, for each n, the limiting correlation index, as $\delta \to \infty$, is 0.3069.

We also make a detailed analysis of the random variable $N_{n,\delta}(c)$. We completely solve the question of finiteness of $N_{\infty,\delta}(c)$, that is, the finitness of the number of δ-records along the infinite sequence of observations. We show that this cannot happen for $c > 0$, for any δ (except if the trivial condition $x_+ - x_- < \delta - c$ holds). It cannot happen either when $c < 0$ and the underlying random variables X_n have an infinite right-tail expectation. This last fact solves a problem posed in [33] by Franke et al., where the authors conjectured that, in the presence of a negative trend, the expected number of records in the whole sequence is finite. In the case $c > 0$ we analyze the asymptotic behaviour of the random variable $N_{n,\delta}$, which grows to infinity. We give a law of large numbers, showing that the ratio $N_{n,\delta}/n$ converges to $p_\delta(c)$ and that its asymptotic distribution is Gaussian, finding the explicit expression of its normalizing constants, which can be estimated from observed data. This result was already known for records and has been applied to different problems, such as athletic records [8, 9] and climate change [124, 125].

In the last section of this chapter a law of large numbers for a model with random trend, which is a generalization of the LDM, is proved.

Chapter 4 is devoted to the use of δ-records for statistical inference in the LDM. In the first section we propose two estimators of the variance of the number of δ-records. These two statistics depend only on the δ-record times, and not on their value, and so they only need the information of the sequence of the δ-record indices, $(1_j)_{j=1}^n$. Using techniques of ergodic theory as in Chapter 3, consistency for both estimators is proved.

In the second part of this chapter a framework for Maximum Likelihood Estimation is proposed. Based on the ideas of Smith [110], who analyzed the problem for usual records, we find the likelihood of the sample based on δ-records for models with deterministic trend and independent underlying random variables. That is, we consider observations drawn from the model $Y_j = X_j + t_j(c)$ such that (X_n) is a sequence of absolutely continuous i.i.d. r.v. and $t_n(c)$ is a deterministic sequence depending on a single parameter c. In this framework, we consider the sampling scheme in which only δ-records, with $\delta \leq 0$, are observed, i.e., we work both with δ-records values and times of occurrence.

The independence of the residuals in the deterministic model allows us to write the likelihood of the sample of e δ-records, and we find the analytical Maximum Likelihood Estimators (MLE) for a family of distributions, being this, to the best of our knowledge, the first analytical result for a MLE, not only for δ-records, but also for usual records in the LDM. Interesting properties are derived from the expressions of the MLEs for this family. For instance, the estimation of both the location and trend parameters depend only on the sequence of records and not on the δ-records which are not records. Nevertheless, this is a particular feature of this family which is not shared by many other distributions, such as the Gaussian or the shifted-exponential distributios, where all δ-records are used in the estimation of the parameters. By means of Montecarlo simulation, we have analyzed how the estimations of the unknown parameters improve with the use of δ-records. In particular, we see how in the Gaussian family and the shifted-exponential, the decrease in the mean squared error of the estimations using δ-records is roughly proportional to the number of extra observations included in the sample (near-records).

In the last section, we have illustrated the limiting results obtained in Chapter 3 for $N_{n,\delta}$ with a set of real data of temperatures of the city of Saragossa (Spain), showing a good agreement between the theoretical asymptotic results and the observed data in the example. Indeed, even for this relatively short series (69 data), the distribution of the number of δ-records is close to the theoretical limiting Gaussian distribution. Also, the estimators of the variance of the number of δ-records are applied to this dataset. These estimators depend on a number m which should be in the order of $o(n^{1/2})$ to enjoy consistency, as it is shown in the first part of this chapter. We show how both estimators yield similar estimations for different values of m, and we use these results to build confidence intervals for particular quantities of interest of the model, such as the asymptotic δ-record probability and the expected number of δ-records among the first n observations. Finally, we apply the MLE to the dataset of temperatures, showing a good performance and yielding accurate results even with a small sample of observations. In particular, we find good agreement between the results obtained and previous estimates of the summer warming trend in Europe [124, 125].

Ideas for future work

Below we list a series of ideas to develop in future work based on the results obtained in this monograph.

- The problem of characterizing the distributions satisfying that $M_n - cN_{n,\delta}$ is a martingale has been studied in Chapter 2 for discrete r.v. taking values in the integers. Preliminary results for the continuous case have been obtained by relating this problem with the theory of delay differential equations, which is a first step to solve the problem of finding all general distributions such that the martingale condition holds.

- The LDM has been extensively studied in Chapter 3. While the laws of large numbers have been shown to hold for more general models, like ARMA processes with an underlying trend, or the random trend model, it seems a natural continuation to look for central limit theorems for the number of δ-records as n increases.

- The explicit expression of the correlation between δ-record occurrences in the LDM was obtained in Chapter 3. Since this correlation indicator is the quantity appearing in the expression of the variance in the central limit theorem, exact expressions for the asymptotic distribution of the number of δ-records can be obtained. This result would allow us to develop asymptotic hypothesis testing methods for the trend parameter based on δ-records, and also to assess the power of such tests.

- The MLE methods analyzed in Chapter 4 could be complemented with additional results, such as consistency for particular families of distributions. Also, Bayesian or bootstrap techniques could be applied to propose alternative methods for making inference based on δ-records in the LDM.

- In addition to parametric inference, other statistical problems in the LDM can be addressed using δ-records. An example is the prediction of future records, which is very important in climatology. Another example is nonparametric estimation of the hazard function of the observations. These problems have been addressed successfully in the i.i.d. case, and we believe that the extension to the LDM will lead to results that outperform the inferences based on record values.

Conclusiones y trabajo futuro

A lo largo de esta memoria se ha obtenido una importante cantidad de nuevos resultados para las observaciones δ-récord y near-record (o récords cercanos).

En el Capítulo 2 se ha estudiado el proceso puntual de los valores near-record cuando las observaciones subyacentes son variables aleatorias discretas tomando valores en los números enteros. Para ello, nos hemos apoyado en el proceso de valores récord, que es conocido como proceso de Shorrock. Condicionando a la sucesión de valores récords, hemos obtenido la distribución del número de near-records que toman valores en un conjunto asociados a un determinado récord, encontrando que dicha distribución es en realidad una geométrica empezando en cero.

Además, hemos obtenido la distribución de esos valores near-records asociados a un determinado récord, demostrando asimismo que los valores, y el número de ellos, gozan de buenas propiedades de independencia. Combinando estos resultados con la teoría clásica de procesos puntuales, demostramos que el proceso de valores near-record que toman valores en un conjunto es un proceso de tipo cluster, caracterizándolo a través de su funcional generador de probabilidad. A partir de este funcional obtenemos la distribución exacta del número de near-records que toman valores en un conjunto a través de su función generatriz de probabilidad, así como expresiones explícitas para algunas cantidades de interés como la esperanza y la varianza de dicho número, o como la covarianza.

Estos resultados han sido aplicados a distintas distribuciones. Por ejemplo para la distribución geométrica obtenemos que tanto el número esperado de near-records en el intervalo $[0, n]$, como su varianza, son del orden de n. Para la distribución con tasa de riesgo $r_i = 1 - 1/(i + 1)$, se encuentra que el número de near-records con valor exactamente n es del orden de n^a, y su varianza del orden de n^{2a}, donde a es el parámetro de near-record.

A partir de estas propiedades, hemos analizado el comportamiento asintótico del número de near-records en el intervalo $[0, n]$ cuando n crece. En primer lugar se demuestra que la condición $\sum_{i=1}^{\infty} r_i^2 < \infty$ es suficiente para garantizar que el número de near-records a lo largo de toda la sucesión de observaciones es finito casi seguramente. En este caso, caracterizamos la distribución del número total de near-records dando su función generatriz de probabilidad. Como ejemplo de una familia con $\sum_{i=1}^{\infty} r_i^2 < \infty$, analizamos el caso de la distribución zeta, cuya función de supervivencia es $y_k = (k+1)^{-1}$. Para esta familia se demuestra que la esperanza del número de near-records coincide con el valor del parámetro de near-record a. Adicionalmente también se obtiene una expresión explícita para la varianza de este número.

Si por el contrario se tiene $\sum_{i=1}^{\infty} r_i^2 = \infty$, se demuestra que bajo condiciones débiles el número de near-records en toda la sucesión de observaciones es infinito casi seguramente. Además, demostramos interesantes resultados asintóticos para el número de near-records, $\eta([0,n])$, cuando n tiende a infinito. En efecto, demostramos una ley de grandes números y un teorema central del límite para $\eta([0,n])$, tanto para distribuciones con colas pesadas y medianas (aquellas que cumplen $\limsup_{k\to\infty} r_k < 1$) y para distribuciones de cola ligera (para las cuales $r_k \to 1$) bajo algunas condiciones adicionales sobre la velocidad de crecimiento de r_k cuando $r_k \to 1$.

El final de este capítulo se dedica al estudio de la caracterización de todas las distribuciones discretas tomando valores en los enteros que satisfacen que $M_n - cN_{n,\delta}$ es una martingala. Este problema fue estudiado en detalle en [49] en el caso de récords usuales. En el caso de los δ-récords, hemos encontrado interesantes resultados en esta memoria. Si $\delta < -1$, se demuestra que ninguna distribución puede satisfacer la condición de martingala propuesta. Si $\delta = -1$, que es el caso de los récords débiles, se demuestra que solo la distribución geométrica cumple dicha condición. Si $\delta > 0$, se demuestra que el problema de hallar las distribuciones que cumplen la condición de la martingala es equivalente al problema de garantizar ciertas condiciones para las sucesiones solución de unas ecuaciones lineales de recurrencia. En nuestro caso las soluciones de estas ecuaciones representan la función de supervivencia de las distribuciones, por lo que se ha de garantizar que dicha sucesión pertenece al intervalo $[0,1]$, y es no creciente con límite 0. El problema de garantizar la positividad de las soluciones en este tipo de ecuaciones de recurrencia es actualmente un problema abierto en la literatura, y por lo tanto no hemos obtenido una solución para el caso general. Sin embargo, sí que se ha resuelto el problema completamente para el caso $\delta = 1$, para el cual se ha obtenido que existe solución si y solo si $0 \le c \le 1/4$, y que además dichas soluciones son combinaciones convexas de la distribución Delta de Dirac y de dos distribuciones geométricas, excepto en el caso límite $c = 1/4$, donde se obtiene una solución relacionada (pero no igual) con la binomial negativa.

En el Capítulo 3 se ha estudiado el comportamiento de los δ-récords en el modelo con tendencia lineal (LDM). Se ha analizado el comportamiento asintótico de la probabilidad de δ-récord, la dependencia entre las ocurrencias δ-récord, y la distribución asintótica del número de δ-récords en las primeras n observaciones.

El comportamiento asintótico de los δ-récords muestra similitudes con el caso de los récords ($\delta = 0$). Por ejemplo, si la tendencia c es positiva, entonces $p_\delta(c) > 0$ si y solo si $\mu^+ < \infty$, independientemente del valor de δ (exceptuando el caso trivial en el cual se tiene $\delta \geq x_+ - x_- + c$, donde x_- y x_+ son los límites inferior y superior del soporte de las variables subyacentes respectivamente, situación en la cual no se observan δ-récords). También se demuestra que $p_\delta(c)$ es una función continua en δ para todo c. Como función de c es continua para todo $c \neq 0$, mientras que hay una discontinuidad en $c = 0$, si $x_+ < \infty$ y $\delta < 0$, lo cual difiere del caso de los récords usuales donde $p_0(c)$ siempre es una función continua en c.

Además, se ha estudiado en detalle la probabilidad de δ-récord en distintos ejemplos. Para la distribución Gumbel se halla una expresión explícita para la probabilidad asintótica $p_\delta(c)$, mostrando que decrece en δ como una función logística de $-\delta$. Para los casos estudiados en la familia de distribuciones Dagum se tiene que $p_\delta(c) = 0$, para todo c y δ, ya que $\mu^+ = \infty$. Para esta familia se estudia si la velocidad de la convergencia de $p_{n,\delta}(c)$ a 0 cuando $n \to \infty$ depende del parámetro δ. Como las variables aleatorias X_n con $\mu^+ = \infty$ pueden tomar valores muy grandes, provocando aumentos muy pronunciados en los valores récord, quizás sería esperable que valores de δ cercanos a 0 tuvieran un efecto despreciable, y que por tanto se tuviera $p_{n,\delta}(c)/p_{n,0}(c) \to 1$. Esto ocurre de hecho en el caso $c = 0$, donde el número de δ-récords crece con la misma velocidad que en el caso de los récords usuales cuando las variables aleatorias X_n son de cola pesada. Sin embargo, en este ejemplo encontramos que para ciertas distribuciones de la familia Dagum se tiene $p_{n,\delta}(c)/p_{n,0}(c) \to a \neq 1$. También se analizan las distribuciones uniforme, y la distribución límite de valores extremos de tipo III, ambas de la familia de extremos de la clase Weibull, para las que se obtienen expresiones exactas de $p_{n,\delta}$. Además, se proponen aproximaciones lineales de primer orden para la probabilidad $p_{n,\delta}$ en el LDM, evaluando su precisión numéricamente.

El parámetro δ tiene una clara importancia en el comportamiento cualitativo de las correlaciones de las ocurrencias δ-récord. En la distribución Gumbel, para la cual los indicadores de récord son independientes en el LDM [14], se demuestra que esto deja de ser así en el caso $\delta \neq 0$; de hecho, las ocurrencias de los δ-récords en esta distribución tienden a atraerse cuando $\delta < 0$, y a separarse si $\delta > 0$. Para distribuciones cuya cola decrece de manera polinómica ya era conocido que si $c > 0$, las correlaciones entre las ocurrencias de los récords son positivas y que incrementan con n [34]. En este marco, hemos estudiado la distribución Pareto con $c = 1$, obteniendo que si bien las correlaciones son positivas (y crecientes en n) para valores negativos o pequeños de δ, éstas son negativas para valores grandes de δ. De hecho, fijado n, el límite del índice de correlación cuando $\delta \to \infty$ es 0.3069, indicando una correlación negativa.

Realizamos además un estudio detallado del comportamiento de la variable aleatoria $N_{n,\delta}(c)$ y resolvemos completamente el problema de la finitud de $N_{\infty,\delta}(c)$, es decir, si el número de δ-récords en la sucesión completa es infinito o no. Demostramos que el número de δ-records es infinito si $c > 0$, para todo δ, excepto si se cumple la condición trivial $x_+ - x_- < \delta - c$. También es infinito si la tendencia es negativa, $c < 0$ y las variables subyacentes X_n no tienen una esperanza de su cola derecha finita, siendo el número finito si dicha esperanza es finita. Este resultado proporciona una solución al problema propuesto en [33] por Franke et al., donde los autores conjeturaron que en la presencia de una tendencia negativa, el número esperado de récords en la sucesión infinita es finito. En el caso $c > 0$ analizamos el comportamiento asintótico de la variable aleatoria $N_{n,\delta}(c)$, la cual ya habíamos demostrado que crece a infinito. Damos una ley de grandes números, probando que el cociente $N_{n,\delta}/n$ converge a $p_\delta(c)$, y que asintóticamente su distribución es gaussiana, encontrando una expresión explícita de las constantes normalizadoras, las cuales pueden ser estimadas de los datos. Este resultado ya era conocido para los récords usuales, y ha sido aplicado en diferentes problemas, como los récords en atletismo [8, 9] o el cambio climático [124, 125].

En la última sección de este capítulo se demuestra además un ley de grandes números para un modelo con tendencia aleatoria, el cual puede ser visto como una generalización del LDM.

El Capítulo 4 se dedica al uso de δ-récords en inferencia estadística en el LDM. En la primera sección proponemos dos estimadores para la varianza del número de δ-récords que se observan en el LDM. Estos dos estadísticos dependen únicamente de los tiempos de δ-récord, y no de sus valores, por lo que para su aplicación solo se necesita conocer el valor de las variables indicadoras $(1_j)_{j=1}^n$. Usando técnicas de teoría ergódica como en el Capítulo 3, se demuestra la consistencia en ambos estimadores.

En la segunda parte de este capítulo se considera el problema de la estimación máximo verosímil en el LDM. Basándonos en las ideas del trabajo de Smith [110], en el que se analizó el problema para récords, damos una expresión explícita para la verosimilitud de la muestra de δ-récords en modelos con una tendencia determinista con residuos independientes. En particular, consideramos que las observaciones siguen el modelo $Y_j = X_j + t_j(c)$, donde (X_n) es una sucesión de variables aleatorias absolutamente continuas e i.i.d., y la tendencia $t_n(c)$ es una sucesión determinista que depende de un único parámetro c. En este marco, consideramos el esquema de muestreo en el cual se observan únicamente los δ-récords, con $\delta \leq 0$, es decir, trabajamos únicamente con los valores δ-récord y sus tiempos de ocurrencia.

La independencia de los residuos en el modelo con tendencia lineal nos permite escribir la verosimilitud de la muestra de los δ-récords. A partir de ésta, encontramos los Estimadores Máximo Verosímiles (EMV) para una familia de distribuciones, siendo estos los primeros resultados analíticos en la literatura respecto a EMV, no solamente para δ-récords, sino también para los récords usuales. De estas expresiones analíticas obtenemos interesantes conclusiones, como por ejemplo, que en dicha familia la estimación de los parámetros de localización y de la tendencia dependen únicamente de los récords, y no de los δ-récords que no son récord. Sin embargo, esta propiedad no se cumple en otras familias de distribuciones, como la gaussiana o la exponencial, para las cuales todos los δ-récords se utilizan para la estimación de los parámetros. A través de simulación de tipo Montecarlo, se muestra como las estimaciones de los parámetros desconocidos mejoran con el uso de los δ-récords. En particular, vemos como en la familia gaussiana y exponencial, la reducción del error cuadrático medio usando δ-récords es, grosso modo, proporcional al número de observaciones adicionales incluidas en la muestra, es decir, al número de near-records respecto al de récords.

En la última sección se han ilustrado los resultados obtenidos en el Capítulo 3 para $N_{n,\delta}$ con un conjunto de datos reales en la ciudad de Zaragoza (España), para los cuales se muestra una excelente concordancia entre los resultados teóricos y los datos de la muestra. De hecho, para una serie de datos relativamente corta (69 datos), la distribución del número de δ-récords ya se aproxima bastante bien por la distribución normal tal y como predice la teórica asintótica desarrollada. Además, se aplican los estimadores de la varianza del número de δ-récords a este conjunto de datos. Estos estimadores dependen de un número, m, que ha de ser escogido de antemano y que debe ser del orden de $o(n^{1/2})$ para garantizar la consistencia de los estimadores, tal y como se ha demostrado en la primera parte de este capítulo. En la aplicación a los datos reales observamos como ambos estimadores, para distintos valores de m, arrojan estimaciones muy similares. Estas estimaciones se utilizan para construir intervalos de confianza para cantidades de interés del modelo, como la probabilidad asintótica de δ-récord o el número esperado de los mismos en las primeras n observaciones. Finalmente, utilizamos el EMV desarrollado en la sección anterior sobre este conjunto de temperaturas, para el cual se observa que arroja estimaciones precisas incluso con una muestra pequeña. En particular, los resultados encontrados son consistentes con estimaciones de otros autores sobre la tendencia del calentamiento en verano en Europa [124, 125].

Ideas de trabajo futuro

A continuación se ofrece una serie de ideas para desarrollar en trabajos futuros motivadas por los resultados obtenidos en esta monografía.

- En el Capítulo 2 abordamos el problema de caracterizar las distribuciones discretas con valores enteros que cumplen que $M_n - cN_{n,\delta}$ es una martingala. En el caso continuo ya se han obtenido resultados preliminares relacionando el problema con el de la resolución de ecuaciones diferenciales con retardo. Este es el siguiente paso hacia la caracterización de todas las distribuciones de cualquier tipo que cumplan la condición de martingala.
- En el Capítulo 3 se ha estudiado detalladamente la ocurrencia de δ-récords en el LDM. Nótese que las leyes de grandes números se han demostrado, no solo para el LDM, sino también para generalizaciones del mismo como procesos

ARMA con una tendencia subyacente y el modelo con tendencia aleatoria. Parece por tanto una continuación natural el tratar de demostrar resultados acerca de la normalidad asintótica de estas observaciones en estos modelos más generales.

- La expresión explícita de la correlación entre dos ocurrencias δ-récord se muestra en el Capítulo 3. Como este indicador de la correlación es la cantidad que aparece en la expresión de la varianza en el teorema central del límite, eso implica que se pueden obtener expresiones exactas para la distribución asintótica del número de δ-récords. Este resultado permitiría el desarrollo de contrastes de hipótesis basados en δ-récords sobre la tendencia subyacente de las observaciones, así como analizar la potencia de dichos contrastes.

- Los EMV analizados en el Capítulo 4 pueden ser complementados con resultados adicionales como la consistencia de los estimadores para familias de distribuciones particulares. También se está considerando el uso de técnicas de análisis bayesiano y bootstrap como métodos alternativos para hacer inferencia basada en δ-récords para el LDM.

- Además de la inferencia paramétrica, se pueden abordar otros problemas estadísticos en el LDM utilizando δ-récords. Por ejemplo para la predicción de récords futuros, cuyo interés en climatología es evidente. Otro ejemplo es la estimación no paramétrica de la función de riesgo de las observaciones. Estos problemas se han abordado con éxito en el caso i.i.d., y creemos que la extensión al LDM conducirá a resultados que mejoren la inferencia basada en valores récord.

Bibliography

[1] Ahsanullah, M. (1995). *Record statistics.* Nova Science, Commack, New York.

[2] Ahsanullah, M. and Nevzorov, V. B. (2015). *Records via probability theory.* Atlantis Press, Paris.

[3] Arnold, B. C. (2014). Univariate and multivariate Pareto models. *Journal of Statistical Distributions and Applications*, 1(1):1–16.

[4] Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1998). *Records.* Wiley, New York.

[5] Ash, R. B. (1972). *Real Analysis and Probability.* Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press, New York.

[6] Balakrishnan, N., Balasubramanian, K., and Panchapakesan, S. (1996). δ-exceedance records. *Journal of Applied Statistical Science*, 4(2-3):123–132.

[7] Balakrishnan, N., Pakes, A. G., and Stepanov, A. (2005). On the number and sum of near-record observations. *Advances in Applied Probability*, 37(3):765–780.

[8] Ballerini, R. and Resnick, S. I. (1985). Records from improving populations. *Journal of Applied Probability*, 22(3):487–502.

[9] Ballerini, R. and Resnick, S. I. (1987). Records in the presence of a linear trend. *Advances in Applied Probability*, 19(4):801–828.

[10] Basak, P. and Balakrishnan, N. (2003). Maximum likelihood prediction of future record statistics. In *Mathematical and Statistical Methods in Reliability*, pages 159–175.

[11] Benestad, R. (2003). How often can we expect a record event? *Climate Research*, 25:3–13.

[12] Benestad, R. (2004). Record-values, nonstationarity tests and extreme value distributions. *Global and Planetary Change*, 44(1):11–26.

[13] Bogachev, V. I. (2007). *Measure theory*, volume 1. Springer Berlin Heidelberg.

[14] Borovkov, K. (1999). On records and related processes for sequences with trends. *Journal of Applied Probability*, 36(03):668–681.

[15] Carlin, B. P. and Gelfand, A. E. (1993). Parametric likelihood inference for record breaking problems. *Biometrika*, 80(3):507–515.

[16] Cebrián, A. C., Castillo-Mateo, J., and Asín, J. (2021). Record tests to detect non-stationarity in the tails with an application to climate change. *Stochastic Environmental Research and Risk Assessment*, pages 1–18.

[17] Chandler, K. N. (1952). The Distribution and Frequency of Record Values. *Journal of the Royal Statistical Society: Series B (Methodological)*, 14(2):220–228.

[18] Coumou, D., Robinson, A., and Rahmstorf, S. (2013). Global increase in record-breaking monthly-mean temperatures. *Climatic Change*, 118(3-4):771–782.

[19] Daley, D. J. and Vere-Jones, D. (2003). *An introduction to the theory of point processes. Vol. I.* Probability and its Applications. Springer-Verlag, New York, second edition.

[20] De Haan, L. and Verkade, E. (1987). On extreme-value theory in the presence of a trend. *Journal of Applied Probability*, 24(1):62–76.

[21] Deheuvels, P. (1974). Valeurs extrémales d'échantillons croissants d'une variable aléatoire réelle. *Annales De L´ Institut Henri Poincaré. Probabilités et Statistiques*, 10:89–114.

[22] Dey, S., Dey, T., and Luckett, D. J. (2016). Statistical inference for the generalized inverted exponential distribution based on upper record values. *Mathematics and Computers in Simulation*, 120:64–78.

[23] Diersen, J. and Trenkler, G. (1996). Records tests for trend in location. *Statistics*, 28(1):1–12.

[24] Dubins, L. E. and Freedman, D. A. (1965). A Sharper Form of the Borel-Cantelli Lemma and the Strong Law. *The Annals of Mathematical Statistics*, 36(3):800–807.

[25] Einmahl, J. H. J. and Magnus, J. R. (2008). Records in Athletics Through Extreme-Value Theory. *Journal of the American Statistical Association*, 103(484):1382–1391.

[26] Eliazar, I. (2005). On geometric record times. *Physica A: Statistical Mechanics and its Applications*, 348:181–198.

[27] Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events*. Springer Berlin.

[28] Feuerverger, A. and Hall, P. (1998). On statistical inference based on record values. *Extremes*, 1(2):169–190.

[29] Fisher, R. A. and Tippett, L. H. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2):180–190.

[30] Foster, F. G. and Stuart, A. (1954). Distribution-Free Tests in Time-Series Based on the Breaking of Records. *Journal of the Royal Statistical Society: Series B (Methodological)*, 16(1):1–13.

[31] Foster, F. G. and Teichroew, D. (1955). A Sampling Experiment on the Powers of the Records Tests for Trend in a Time Series. *Journal of the Royal Statistical Society: Series B (Methodological)*, 17(1):115–121.

[32] Franke, J., Klözer, A., de Visser, J. A. G., and Krug, J. (2011). Evolutionary Accessibility of Mutational Pathways. *PLOS Computational Biology*, 7(8):e1002134.

[33] Franke, J., Wergen, G., and Krug, J. (2010). Records and sequences of records from random variables with a linear trend. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(10):P10013.

[34] Franke, J., Wergen, G., and Krug, J. (2012). Correlations of Record Events as a Test for Heavy-Tailed Distributions. *Physical Review Letters*, 108(6):64101.

[35] Gembris, D., Taylor, J. G., and Suter, D. (2002). Trends and random fluctuations in athletics: Sports statistics. *Nature*, 417(6888):506.

[36] Gembris, D., Taylor, J. G., and Suter, D. (2007). Evolution of Athletic Records: Statistical Effects versus Real Improvements. *Journal of Applied Statistics*, 34(5):529–545.

[37] Glick, N. (1978). Breaking records and breaking boards. *American Mathematical Monthly*, pages 2–26.

[38] Gnedenko, B. (1943). Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire. *The Annals of Mathematics*, 44(3):423.

[39] Godrèche, C. and Luck, J. M. (2020). Records for the moving average of a time series. *Journal of Statistical Mechanics: Theory and Experiment*, 2020(2):023201.

[40] Godrèche, C., Majumdar, S. N., and Schehr, G. (2015). Record statistics for random walk bridges. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(7):P07026.

[41] Godrèche, C., Majumdar, S. N., and Schehr, G. (2017). Record statistics of a strongly correlated time series: random walks and Lévy flights. *Journal of Physics A: Mathematical and Theoretical*, 50(33):333001.

[42] Gouet, R., Javier López, F., and Sanz, G. (2008). Laws of large numbers for the number of weak records. *Statistics & Probability Letters*, 78(14):2010–2017.

[43] Gouet, R., Lafuente, M., López, F. J., and Sanz, G. (2018). δ-Records Observations in Models with Random Trend. In *Studies in Systems, Decision and Control*, volume 142, pages 209–217. Springer, Cham.

[44] Gouet, R., Lafuente, M., López, F. J., and Sanz, G. (2020a). Exact and asymptotic properties of δ-records in the linear drift model. *Journal of Statistical Mechanics: Theory and Experiment*, 2020(10):103201.

[45] Gouet, R., López, F. J., Maldonado, L. P., and Sanz, G. (2014). Statistical inference for the geometric distribution based on δ-records. *Computational Statistics & Data Analysis*, 78:21–32.

[46] Gouet, R., López, F. J., Maldonado, L. P., and Sanz, G. (2020b). Statistical inference for the weibull distribution based on δ-record data. *Symmetry*, 12(1):20.

[47] Gouet, R., López, F. J., and San Miguel, M. (2001). A martingale approach to strong convergence of the number of records. *Advances in Applied Probability*, 33(4):864–873.

[48] Gouet, R., López, F. J., and Sanz, G. (2005). Central limit theorems for the number of records in discrete models. *Advances in Applied Probability*, 37(3):781–800.

[49] Gouet, R., López, F. J., and Sanz, G. (2007a). A characteristic martingale related to the counting process of records. *Journal of Theoretical Probability*, 20(3):443–455.

[50] Gouet, R., López, F. J., and Sanz, G. (2007b). Asymptotic normality for the counting process of weak records and δ -records in discrete models. *Bernoulli*, 13(3):754–781.

[51] Gouet, R., López, F. J., and Sanz, G. (2009). Limit laws for the cumulative number of ties for the maximum in a random sequence. *Journal of Statistical Planning and Inference*, 139(9):2988–3000.

[52] Gouet, R., López, F. J., and Sanz, G. (2012a). Central limit theorem for the number of near-records. *Communications in Statistics - Theory and Methods*, 41(2):309–324.

[53] Gouet, R., López, F. J., and Sanz, G. (2012b). On δ-record observations: asymptotic rates for the counting process and elements of maximum likelihood estimation. *TEST*, 21(1):188–214.

[54] Gouet, R., López, F. J., and Sanz, G. (2012c). On geometric records: rate of appearance and magnitude. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(01):P01005.

[55] Gouet, R., López, F. J., and Sanz, G. (2015a). On the point process of near-record values. *TEST*, 24(2):302–321.

[56] Gouet, R., López, F. J., and Sanz, G. (2015b). Records from stationary observations subject to a random trend. *Advances in Applied Probability*, 47(4):1175–1189.

[57] Gradshteyn, I. S. and Ryzhik, I. M. (2014). *Table of integrals, series, and products; 8th ed.* Academic Press, Amsterdam.

[58] Gulati, S., George, F., and Golam Kibria, B. M. (2019). Analysis of hurricane extremes and record values in the Atlantic. *Communications in Statistics Case Studies Data Analysis and Applications*, 5(2):101–110.

[59] Gulati, S. and Padgett, W. J. (2003). *Parametric and Nonparametric Inference from Record-Breaking Data*, volume 172. Springer New York.

[60] Haghighi-Talab, D. and Wright, C. (1973). On the distribution of records in a finite sequence of observations, with an application to a road traffic problem. *Journal of Applied Probability*, 10(3):556–571.

[61] Halava, V., Harju, T., and Hirvensalo, M. (2006). Positivity of second order linear recurrent sequences. *Discrete Applied Mathematics*, 154(3):447–451.

[62] Hall, P. and Heyde, C. C. (1980). *Martingale limit theory and its application.* Academic Press, New York.

[63] Hildebrandt, T. H. (1942). Remarks on the abel-dini theorem. *The American Mathematical Monthly*, 49(7):441–445.

[64] Hoayek, A. S., Ducharme, G. R., and Khraibani, Z. (2017). Distribution-free inference in record series. *Extremes*, 20(3):585–603.

[65] Jafari, A. A. and Zakerzadeh, H. (2015). Inference on the parameters of the Weibull distribution using records. *SORT*, 39(1):3–18.

[66] Jain, K. and Krug, J. (2005). Evolutionary trajectories in rugged fitness landscapes. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(04):P04008.

[67] Kearney, M. J. (2020). Record statistics for a discrete-time random walk with correlated steps. *Journal of Statistical Mechanics: Theory and Experiment*, 2020(2):023206.

[68] Khraibani, Z., Badran, H., and Khraibani, H. (2011). Records method for the natural disasters application to the storm events. *Journal of Environmental Sciences*, 5:643–651.

[69] Khraibani, Z., Jacob, C., Ducrot, C., Charras-Garrido, M., and Sala, C. (2015). A non parametric exact test based on the number of records for an early detection of emerging events: Illustration in epidemiology. *Communications in Statistics - Theory and Methods*, 44(4):726–749.

[70] Khraibani, Z., Khraibani, J., Kobeissi, M., Atoui, A., and Khraibani, Z. (2020). Application of records theory on the COVID-19 pandemic in Lebanon: Prediction and Prevention. *Epidemiology and Infection*, 148:e192.

[71] Kleiber, C. (2008). A Guide to the Dagum Distributions. In *Modeling Income Distributions and Lorenz Curves*, pages 97–117. Springer, New York, NY.

[72] Krug, J. (2007). Records in a changing world. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(7):P07001.

[73] Krug, J. and Karl, C. (2003). Punctuated evolution for the quasispecies model. *Physica A: Statistical Mechanics and its Applications*, 318(1-2):137–143.

[74] Krug, J. H. A. and Jain, K. (2005). Breaking records in the evolutionary race. *Physica A-statistical Mechanics and Its Applications*, 358:1–9.

[75] Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1-3):159–178.

[76] Lafuente, M., Ejea, D., Gouet, R., López, F. J., and Sanz, G. (2022). Approximations of δ-record probabilities in i.i.d. and trend models. In *Studies in Systems, Decision and Control*. Springer, Cham. To appear.

[77] Le Doussal, P. and Wiese, K. J. (2009). Driven particle in a random landscape: Disorder correlator, avalanche distribution, and extreme value statistics of records. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 79(5):051105.

[78] López-Blázquez, F. and Salamanca-Miño, B. (2013). Distribution theory of δ-record values. Case $\delta \leq 0$. *TEST*, 22(4):715–738.

[79] López-Blázquez, F. and Salamanca-Miño, B. (2015). Distribution theory of δ-record values: case $\delta \geq 0$. *TEST*, 24(3):558–582.

[80] López-Blázquez, F. and Salamanca-Miño, B. (2016). Geometric records from Pareto parents. *Journal of Applied Statistical Science*, 22(1-2):99–110.

[81] Majumdar, S. N., Schehr, G., and Wergen, G. (2012). Record statistics and persistence for a random walk with a drift. *Journal of Physics A: Mathematical and Theoretical*, 45(35):355002.

[82] Majumdar, S. N., von Bomhard, P., and Krug, J. (2019). Exactly Solvable Record Model for Rainfall. *Physical Review Letters*, 122(15):158702.

[83] Majumdar, S. N. and Ziff, R. M. (2008). Universal record statistics of random walks and Lévy flights. *Physical Review Letters*, 101(5):050601.

[84] Meehl, G. A., Tebaldi, C., Walton, G., Easterling, D., and McDaniel, L. (2009). Relative increase of record high maximum temperatures compared to record low minimum temperatures in the U.S. *Geophysical Research Letters*, 36(23):L23701.

[85] Mounaix, P., Majumdar, S. N., and Schehr, G. (2018). Asymptotics for the expected maximum of random walks and Lévy flights with a constant drift. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(8):083201.

[86] Nevzorov, V. and Stepanov, A. (2014). Records with confirmation. *Statistics & Probability Letters*, 95:39–47.

[87] Nevzorov, V. B. (2001). *Records: mathematical theory*, volume 194. American Mathematical Soc., Providence, Rhode Island.

[88] Oliveira, L. P., Jensen, H. J., Nicodemi, M., and Sibani, P. (2005). Record dynamics and the observed temperature plateau in the magnetic creep-rate of type-ii superconductors. *Phys. Rev. B*, 71:104526.

[89] Ouaknine, J. and Worrell, J. (2012). Decision problems for linear recurrence sequences. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 7550 LNCS, pages 21–28. Springer, Berlin, Heidelberg.

[90] Pakes, A. G. (2007). Limit theorems for numbers of near-records. *Extremes*, 10(4):207–224.

[91] Park, S.-C. and Krug, J. (2008). Evolution in random fitness landscapes: the infinite sites model. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(04):P04014.

[92] Park, S. C. and Krug, J. (2016). δ-Exceedance records and random adaptive walks. *Journal of Physics A: Mathematical and Theoretical*, 49(31).

[93] Rahmstorf, S. and Coumou, D. (2011). Increase of extreme events in a warming world. *Proceedings of the National Academy of Sciences*, 108(44):17905–17909.

[94] Redner, S. and Petersen, M. R. (2006). Role of global warming on the statistics of record-breaking temperatures. *Physical Review E*, 74(6):61114.

[95] Resnick, S. I. (1973). Limit laws for record values. *Stochastic Processes and their Applications*, 1(1):67–82.

[96] Richardson, T. O., Robinson, E. J., Christensen, K., Jensen, H. J., Franks, N. R., and Sendova-Franks, A. B. (2010). Record dynamics in ants. *PLoS ONE*, 5(3):e9621.

[97] Robinson, M. E. and Tawn, J. A. (1995). Statistics for Exceptional Athletics Records. *Applied Statistics*, 44(4):499.

[98] Rényi, A. (1962). Théorie des éléments saillants d'une suite d'observations. *Annales scientifiques de l'Université de Clermont. Mathématiques*, 8(2):7–13.

[99] Saddique, N., Khaliq, A., and Bernhofer, C. (2020). Trends in temperature and precipitation extremes in historical (1961–1990) and projected (2061–2090) periods in a data scarce mountain basin, northern Pakistan. *Stochastic Environmental Research and Risk Assessment*, 34(10):1441–1455.

[100] Samaniego, F. J. and Whitaker, L. R. (1986). On estimating population characteristics from record-breaking observations. i. parametric results. *Naval Research Logistics Quarterly*, 33(3):531–543.

[101] Samaniego, F. J. and Whitaker, L. R. (1988). On estimating population characteristics from record-breaking observations ii. nonparametric results. *Naval Research Logistics (NRL)*, 35(2):221–236.

[102] Sen, A. and Srivastava, M. (1990). *Regression Analysis : Theory, Methods, and Applications*. Springer New York.

[103] Shiryayev, A. N. (1984). *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer, New York.

[104] Shorrock, R. W. (1972). On record values and record times. *Journal of Applied Probability*, 9(2):316–326.

[105] Shoukri, M. M., Mian, I. U. H., and Tracy, D. S. (1988). Sampling properties of estimators of the log-logistic distribution with application to Canadian precipitation data. *Canadian Journal of Statistics*, 16(3):223–236.

[106] Sibani, P. (2007). Linear response in aging glassy systems, intermittency and the Poisson statistics of record fluctuations. *European Physical Journal B*, 58(4):483–491.

[107] Sibani, P., Brandt, M., and Alstrøm, P. (1998). Evolution and extinction dynamics in rugged fitness landscapes. *International Journal of Modern Physics B*, 12(4):361–391.

[108] Sibani, P., Rodriguez, G. F., and Kenning, G. G. (2006). Intermittent quakes and record dynamics in the thermoremanent magnetization of a spin-glass. *Physical Review B - Condensed Matter and Materials Physics*, 74(22):224407.

[109] Sire, C., Majumdar, S. N., and Dean, D. S. (2006). Exact solution of a model of time-dependent evolutionary dynamics in a rugged fitness landscape. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(07):L07001.

[110] Smith, R. L. (1988). Forecasting Records by Maximum Likelihood. *Journal of the American Statistical Association*, 83(402):331–338.

[111] Smith, R. L. and Miller, J. E. (1986). A Non-Gaussian State Space Model and Application to Prediction of Records. *Journal of the Royal Statistical Society: Series B (Methodological)*, 48(1):79–88.

[112] Soliman, A. A., Abd Ellah, A., and Sultan, K. (2006). Comparison of estimates using record statistics from weibull model: Bayesian and non-bayesian approaches. *Computational Statistics & Data Analysis*, 51(3):2065–2077.

[113] Stepanov, A. V. (1993). Limit Theorems for Weak Records. *Theory of Probability & Its Applications*, 37(3):570–574.

[114] Stepanov, A. V. (1994). A Characterization Theorem for Weak Records. *Theory of Probability & Its Applications*, 38(4):762–764.

[115] Tata, M. N. (1969). On outstanding values in a sequence of random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 12(1):9–20.

[116] Vervaat, W. (1973). Limit theorems for records from discrete distributions. *Stochastic Processes and their Applications*, 1:317–334.

[117] Wang, B. X. and Ye, Z.-S. (2015). Inference on the Weibull distribution based on record values. *Computational Statistics & Data Analysis*, 83:26–36.

[118] Weinreich, D. M., Delaney, N. F., DePristo, M. A., and Hartl, D. L. (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, 312(5770):111–114.

[119] Weinreich, D. M., Watson, R. A., and Chao, L. (2005). Perspective: sign Epistasis And Genetic Constraint On Evolutionary Trajectories. *Evolution*, 59(6):1165 – 1174.

[120] Wergen, G. (2013). Records in stochastic processes - Theory and applications. *Journal of Physics A: Mathematical and Theoretical*, 46(22):223001.

[121] Wergen, G. (2014). Modeling record-breaking stock prices. *Physica A: Statistical Mechanics and its Applications*, 396:114–133.

[122] Wergen, G., Bogner, M., and Krug, J. (2011a). Record statistics for biased random walks, with an application to financial data. *Physical Review E*, 83(5):51109.

[123] Wergen, G., Franke, J., and Krug, J. (2011b). Correlations Between Record Events in Sequences of Random Variables with a Linear Trend. *Journal of Statistical Physics*, 144(6):1206–1222.

[124] Wergen, G., Hense, A., and Krug, J. (2014). Record occurrence and record values in daily and monthly temperatures. *Climate Dynamics*, 42(5-6):1275–1289.

[125] Wergen, G. and Krug, J. (2010). Record-breaking temperatures reveal a warming climate. *EPL (Europhysics Letters)*, 92(3):30008.

[126] Yang, M. C. K. (1975). On the distribution of the inter-record times in an increasing population. *Journal of Applied Probability*, 12(1):148–154.