

Análisis estadístico de la serie anual del número de récords de temperatura media diaria en España

JOSÉ ANTONIO LÓPEZ DÍAZ AEMET

Propiedades probabilísticas de la serie de récords

Precisemos en primer lugar lo que se entiende por récord: un récord es un valor al menos igual que el máximo de todos los anteriores de la serie. También se define un récord por debajo como un valor menor que el mínimo de todos los anteriores.

El resultado fundamental es que, para una serie de variables aleatorias independientes e idénticamente distribuidas (aleatoria simple), la probabilidad de ocurrencia de un récord en la posición N-ésima de la serie vale $1/N$. También se cumple que la ocurrencia de un récord en la posición N-ésima es un suceso independiente de la ocurrencia de récord en cualquier otro lugar (Resnick 1987). La prueba de que la probabilidad de récord en la posición N es $1/N$ es fácil si se repara en que para una serie aleatoria, como la considerada aquí, las distintas ordenaciones relativas en cuanto a magnitud de los N primeros valores tienen todas la misma probabilidad (está claro que al tener todas las variables la misma distribución, y ser independientes, ninguna ordenación relativa puede tener preferencia).

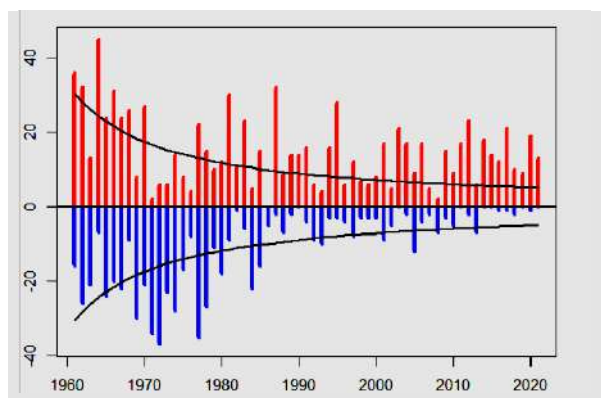


Figura 1: Número de récords cálidos y fríos de temperatura media diaria para cada año de 1961 a 2021. Líneas negras: valor esperado del número de récords anual (comienzo en 1951).

Si representamos cada ordenación relativa asignando un número entero entre 1 y N a cada uno de los N primeros valores de la serie según su posición en la serie ordenada (su rango), se deduce que el rango en la posición N puede ser cualquiera de los N valores 1, 2, ..., N, todos con la misma probabilidad. Y para que se tenga récord se precisa que el rango sea precisamente N, por tanto su probabilidad es $1/N$.¹

Datos

La serie de partida es la formada por las temperaturas medias diarias de España desde 1950 hasta 2021. Esta serie ha sido elaborada por José Ángel Núñez Mora, meteorólogo de la AEMET,

basándose en todos los datos disponibles cada día de temperatura en la Base Nacional de Datos de la AEMET, excluyendo aquellos datos que exceden 2.5 veces el error estadístico diario, utilizando una interpolación espacial. En la figura 1 se representan el número de récords, cálidos y fríos, de temperatura media diaria cada año del periodo 1961-2021 (ver el Informe del Clima 2021 de la AEMET, disponible en su página web). La figura muestra con claridad que el número de récords cálidos excede con claridad el valor esperado. También es evidente que, por contra, el número de récords fríos en años recientes está por debajo de lo esperado. Pero también salta a la vista que hay numerosos casos en años anteriores de valores bien claramente por encima, bien por debajo de lo esperado. Para valorar estadísticamente estas desviaciones sería útil disponer de un test que combinara, para cada año, ambos tipos de récords.

Test de significación estadística global

Si queremos valorar de una forma precisa el grado de significación estadística de las desviaciones respecto de los valores esperados apreciables a simple vista en la figura 1, es preciso hacer hipótesis simplificadoras naturales, denominadas en estadística hipótesis nula, que permitan un modelo probabilístico lo suficientemente sencillo como para calcular la distribución de una cantidad derivada de los datos bajo esas hipótesis y de esa forma evaluar la probabilidad de las desviaciones observadas. Dos tipos de hipótesis se pueden de forma natural introducir en este caso, unas referidas al número de récords para un día de calendario fijo a lo largo de los distintos años, y otras referidas al número de récords en los diferentes días de un mismo año.

En primer lugar, para cada día de calendario del año, podemos suponer que las temperaturas medias de España en sucesivos años son realizaciones de variables aleatorias independientes (esto no es problemático, dado que transcurren 365 días entre cada dos sucesivas) e igualmente distribuidas, denominémoslo hipótesis H1. Esto último supone clima estacionario, sin variabilidad ni natural ni antropogénica, lo cual no se cumple, pero podemos aceptarlo como hipótesis de partida más simple. Entonces, por lo visto antes, la probabilidad de que en un año la temperatura media diaria de un día cualquiera sea récord vale $1/A$, siendo A el número de años transcurridos desde el inicio de la serie (en este caso A es el número de años desde 1951).

En segundo lugar, se puede también postular H2, que las temperaturas medias en días sucesivos del mismo año también son independientes, lo cual sabemos que no es cierto en rigor debido a fenómenos como situaciones sinópticas persistentes. Bajo este supuesto, el número total de récords cálidos (fríos) en un año tiene una distribución binomial $\text{Bin}(365, 1/A)$, con índice $N = 365$ y parámetro $p = 1/A$.

¹ La demostración de la independencia probabilística de los sucesos "récord en posición M" y "récord en posición N" para $M < N$ arbitrarios es más compleja. Un argumento se basa en condicionar al máximo de los primeros M elementos, llamémoslo $\max(M)$. El suceso "récord en N" solo depende, en cuanto a los primeros M valores, de $\max(M)$, no de los detalles de su ordenación. Por tanto, si condicionamos a que $\max(M)$ tome un valor determinado, los sucesos "récord en N" y "récord en M" son independientes, pues el segundo solo depende de la ordenación relativa de los M primeros valores, y el primero del máximo al que condicionamos. Además la probabilidad de "récord en M" condicionada a un valor de $\max(M)$ es constante, igual a $1/M$, por el mismo argumento de simetría empleado para probar la probabilidad de récord sin condicionar. Así pues, tenemos $\text{pr}(\text{récord en N y en M} \mid \max(M)) = \text{pr}(\text{récord en N} \mid \max(M)) * 1/M$, y ahora integrando respecto a la pdf de $\max(M)$ hallamos $\text{pr}(\text{récord en N y M}) = \text{pr}(\text{récord en N}) * 1/M = \text{pr}(\text{récord en N}) * \text{pr}(\text{récord en M})$ c.q.d.

Utilizando la distribución binomial anterior del número de récords podemos calcular el p-valor de un test bilateral sobre el número de récords observado, tanto para los cálidos como para los fríos. Pero sin duda sería deseable poder combinar ambos p-valores para formar un único test global para cada año. Esto es factible con tal de asumir también H3, que las dos variables aleatorias (número de récords cálidos y fríos) son independientes, lo cual parece bastante razonable. En este caso, cada p-valor se distribuye aproximadamente como una variable uniforme en $[0, 1]^2$, y ambos p-valores son independientes. Entonces $-2 \cdot \log(p_1 p_2)$, siendo p_1 y p_2 los respectivos p-valores, se distribuye según una distribución ji-cuadrado con 4 grados de libertad. Para este test global la hipótesis nula H_0 consiste en la conjunción de H_1 , H_2 y H_3 . El p-valor de este test global se obtendrá calculando la probabilidad de la cola derecha de la ji-cuadrado con 4 grados de libertad.

En la figura 2 se representan los p-valores del test global anterior cada año, indicando con colores si los p-valores de número de récords cálidos y fríos son congruentes en su sentido, bien hacia año más cálido, bien hacia año más frío, o no son congruentes. Se observa una acumulación de años con desviación global hacia cálido en los últimos 10 años, como era de prever a la vista de figura 1. La línea de p-valor 1 % permite apreciar que hay un número relativamente grande de años significativos para este test al 1 % (si se cumplieran la hipótesis nula del test todos los años esperaríamos tan solo unos $61/100 \sim 0.6$ años en promedio significativos a ese nivel, y lo son 34). Es claro que el calentamiento global causa un muy significativo aumento del número de récords cálidos, pero además la figura permite apreciar que ya en la década de los 60 o 70 del pasado siglo no se cumplía la hipótesis nula de este test global. Esto último parece razonable atribuirlo sobre todo a la falta de independencia estocástica en los sucesos "récord en el día D" para días sucesivos, debido a las situaciones sinópticas persistentes con temperaturas extremas, como las de bloqueo.

Los diez años que más destacan por el p-valor del test global sobre el número de récords tanto cálidos como fríos se recogen en la tabla 1.

Año	1987	1971	1995	1964	2012	1977	2017	2003	1972	2020
Carácter	CC	FF	CC	CC	CC	CF	CC	CC	FF	CC

Tabla 1: Años con mayor significación estadística del test global sobre el número de récords cálidos/fríos, ordenados de más significativo estadísticamente a menos (orden creciente del p-valor). En el carácter se indica, con la primera letra el carácter cálido/frío según el número de récords cálidos, y análogamente con la segunda letra con los récords fríos.

Además de cuatro años del siglo XXI de carácter homogéneamente cálido, como era de esperar, llama la atención que hay tres años de la década de los 70, dos con carácter frío, y uno, 1977, con carácter contradictorio para ambos tipos de récords (en ese año hubo más récords cálidos de los esperable, pero también más récords fríos de lo normal). También es llamativo que 1964 se cuele en la lista en quinta posición con carácter cálido. El primer año del presente siglo, 2012, aparece en quinta posición.

Decenio	1961-70	1971-80	1981-90	1991-00	2001-10	2011-20
# Cálidos	3	0	5	2	4	8
# Fríos	1	4	1	0	0	0

Tabla 2: Para cada decenio, número de años con carácter global cálido al 1 % de nivel de significación, y análogamente para años fríos.

En la tabla 2 se resumen los resultados de este test global por decenios, al nivel de significación 1 %. Se confirma que el último decenio es el que más resultados significativos, y todos en dirección cálida, arroja (8 años), seguido del 1981-90, con 5 años cálidos y 1 frío. Por el lado de años fríos destaca el decenio 1971-80 con 4 años significativamente fríos y ninguno cálido. El decenio con menos años significativos es el 1991-2000, con tan solo 2 años cálidos.

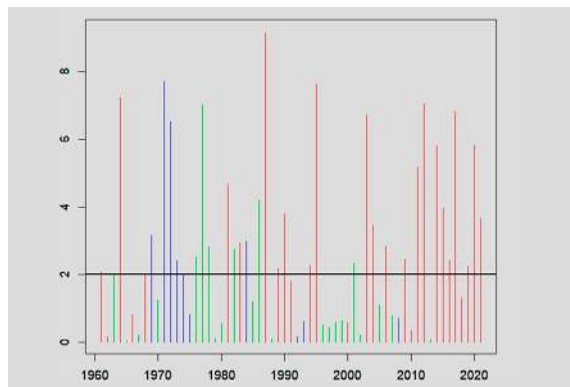


Figura 2: Logaritmos decimales cambiados de signo de los p-valores del test conjunto sobre el número de récords cálidos y fríos. En rojo, años en que tanto el número de récords cálidos como fríos indican cálido, en azul si ambos récords indican frío, en verde si las desviaciones de ambos récords son contradictorias en este sentido. La línea horizontal marca el límite de nivel de significación 1 %.

Estimación de la razón de riesgos de récords

Para cualquier suceso extremo, la razón de riesgos es el cociente entre la probabilidad de ocurrencia del suceso extremo en el clima actual y su probabilidad en un clima de referencia, que puede ser el pre-industrial. En nuestro caso podemos estimar esta razón de riesgos para récord cálido y frío estimando a partir de los datos un riesgo de récord promedio en cada año, y comparándolo con el teórico en serie estacionaria. Para ello podemos aplicar un filtro lineal a los datos, es decir, un promediado de los años contiguos con pesos adecuados, como por ejemplo un filtro gaussiano. Hay que decir que para un proceso Poisson, y este de ocurrencia de récords anuales lo es aproximadamente en cuanto el número esperado es relativamente bajo, el estimador de máxima verosimilitud del parámetro λ del proceso, que es igual al valor esperado del mismo, es precisamente la media aritmética de la serie. Por ello tiene mayor sentido aplicar un promedio ponderado como el filtro gaussiano, que sirva para suavizar la serie y evitar los saltos de año a año.

En la figura 3 se plotean los estimadores obtenidos para récords cálidos y fríos, con un filtro gaussiano de $T = 15$ años. Se ve que aunque la λ^* para récord frío disminuye monótonamente con los años, para los récords cálidos la λ^* fluctúa entre 12 y 15 desde 1980 hasta el presente. Esto se debe naturalmente al calentamiento que compensa con un aumento relativo de la probabilidad de récord respecto a la esperable al incremento de la longitud de las series. La razón de riesgos se representa en la figura 4, donde para los cálidos se ha representado el cociente λ^*/λ_0 entre el valor estimado de los datos y el teórico bajo la hipótesis de variables independientes e igualmente distribuidas, que se traduce en clima estacionario en nuestro caso, pero para los fríos

² No es del todo exacto esto debido al carácter discreto de la variable aleatoria, si fuese continua sería exacto el anterior resultado. Siendo B la variable binomial número de récords, he tomado como p-valor, si $B > \text{mediana}(B)$, $2^{1/2} \text{pr}(B=R) + \text{pr}(B > R)$ siendo R el número de récords observado, y análogamente de forma simétrica para la cola izquierda. Esto corresponde a repartir la masa de probabilidad del suceso $B=k$ de forma uniforme en el intervalo $[k-1/2, k+1/2]$.

Análisis estadístico de la serie anual del número de récords de temperatura media diaria en España

se ha representado el inverso de dicha razón de riesgos, para obtener valores comparables. Llama la atención que ambas razones siguen trayectorias similares hasta 2015 aproximadamente en que la razón de riesgos inversa para récord frío se dispara. Esto último tiene que ver con que el λ^* estimada es muy baja y hay errores muestrales altos. Las razones de riesgos (o sus inversas para fríos) toman valores próximos a la unidad hasta mediados de los ochenta, ascendiendo de ahí en adelante paulatinamente hasta valores del orden de 2 o superiores en la última década.

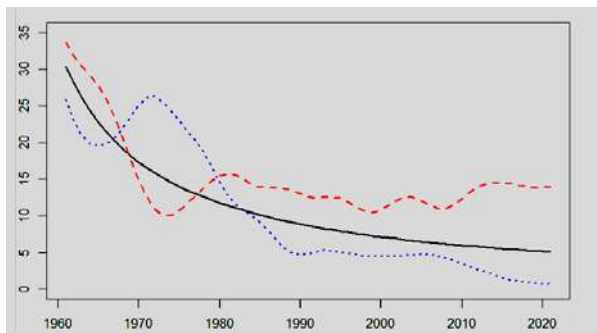


Figura 3: Valores esperados del número de récords cálidos y fríos (negro), estimación de ese valor esperado a partir de los datos con filtro gaussiano para récords cálidos (rojo) y fríos (azul)

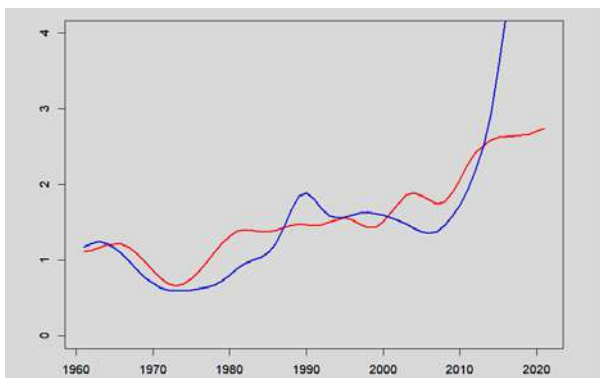


Figura 4: Para récords cálidos (rojo), cociente entre el valor esperado estimado y teórico λ^*/λ_0 , para récords fríos (azul), cociente λ_0/λ^* .

Estimación del agrupamiento de los récords

Ya se comentó a propósito del p-valor del test global, que combina los récords cálidos y fríos, los dos motivos principales para que no se cumpla su hipótesis nula: que el clima no es estacionario y que la ocurrencia de récords en días sucesivos del mismo año no es independiente. Por otra parte, si la única desviación de la hipótesis nula del test fuera debida a las tendencias del clima veríamos un aumento de los valores significativos en los decenios más recientes, pero como se ve en la figura 2 o en la tabla 2, en todos los decenios hay un número claramente mayor de años significativos que lo esperable bajo H_0 . Así pues debe haber una tendencia importante al *clustering* o agrupamiento de récords.

Para intentar valorar esta tendencia al agrupamiento podemos suponer que los récords siguen un modelo sencillo en que en cada ocurrencia del proceso binomial no se da un récord, sino un número m , pero de tal forma que su número esperado no varíe. El número de récords un año cualquiera no es ya

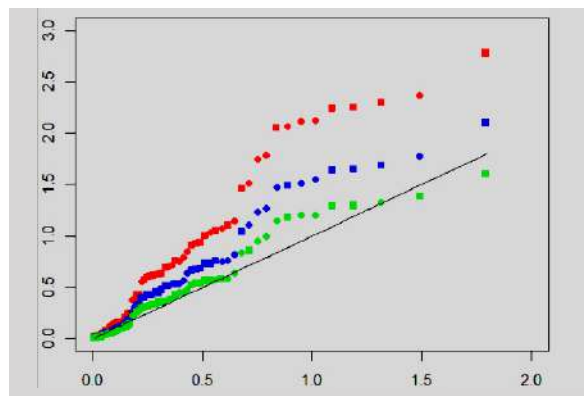


Figura 5: Diagrama p-p (en escala logarítmica decimal) para los p-valores del test global y para diferentes valores del parámetro de agrupamiento $m = 3, 4, 5$ (colores rojo, azul y verde, respectivamente). Cuadrados para los años 1961-1990, círculos para 1991-2021. Línea recta negra: ajuste perfecto.

una variable aleatoria $\text{Bin}(365, p)$ siendo p la probabilidad teórica, sino $m \text{Bin}(365, p/m)$. De esta forma el valor esperado de récords no varía, sigue siendo $365 p$ (pues la esperanza matemática del segundo proceso es $m \cdot 365 p/m = 365 p$), pero la varianza del proceso se multiplica por m aproximadamente³. Por tanto el segundo proceso dará p-valores más altos (menos significativos) al tener más varianza bajo H_0 , y podremos conseguir para un m adecuado unos p-valores más próximos a los que debería haber de cumplirse H_0 . En caso de cumplirse H_0 exactamente los p-valores deben ajustarse a una distribución uniforme, suponiendo realizaciones independientes del test, o sea, en este caso que los números de récords anuales son independientes en años sucesivos, lo cual podemos asumir como mucho menos problemático que la ocurrencia de récords en días sucesivos.

En la figura 5 se muestra de forma intuitiva el grado de ajuste a la distribución uniforme de los p-valores del test global para valores de $m = 3, 4$ y 5 . En dicha figura, que es un gráfico denominado p-p en escala doblemente logarítmica, el ajuste perfecto a la uniforme supondría que los puntos estarían alineados según la línea negra. Parece que el mejor ajuste se produce para $m = 5$, pero también es evidente que incluso en este caso los puntos no se ajustan bien a una recta. Esto indica que el modelo simple de agrupamiento considerado está lejos de reflejar bien los números de récords, de hecho sabemos que las tendencias climáticas no se han tenido en cuenta. No obstante esto último no parece ser determinante en la falta de ajuste porque la distribución de puntos de los primeros 30 años y de los últimos 31 (cuadrados y círculos en la figura respectivamente) es bastante uniforme sobre el tramo de los gráficos más disconformes con la linealidad (a partir de la abscisa 0.7 aproximadamente). Estos puntos de la derecha de la gráfica corresponden a los valores más alejados del valor esperado de número de récords. Pero también es evidente que un mayor grado de agrupamiento m del modelo propuesto acerca los datos al ajuste perfecto.

AGRADECIMIENTOS

A José María Sánchez-Laulhé Ollero por la cuidadosa revisión.

Referencias

- Cox, D.R., Hinkley, D.V., 1979: *Theoretical Statistics*, Chapman&Hall, ISBN 9780412161605
- Resnick, S., 1987: *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operational Research and Financial Engineering. ISBN 978-0-387-75953-1

³ Para el primer proceso la varianza es $V1 = 356 p (1-p)$ y para el segundo $V2 = m^2 365 p/m (1-p/m) = m 365 p (1-p/m)$, que al ser p es pequeña $V2 \sim m V1$.