

# ¿ Son útiles para la detección del Cambio Climático los cambios en la frecuencia de los sucesos climáticos extremos ?

José Antonio López Díaz. AEMET - Director del Programa de Técnicas Climatológicas

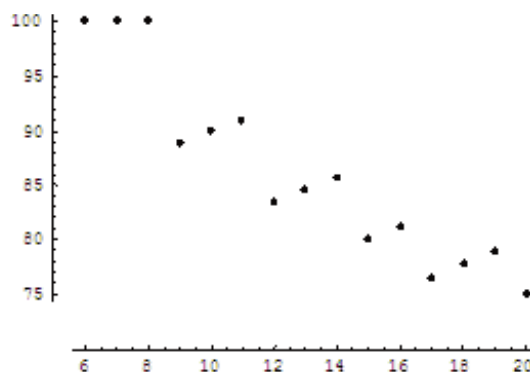
## Introducción

No cabe duda que los sucesos climáticos extremos, como olas de calor o frío, sequías, etc., tienen un gran impacto en la opinión pública y en la sociedad en general. Por esa razón, estos sucesos son objeto de especial atención por los servicios meteorológicos, que suelen tener programas especiales para su predicción, seguimiento, ayuda en la prevención y otros. Nada que objetar a todo ello, pero en este trabajo me voy a centrar en la cuestión de hasta qué punto podemos utilizarlos como heraldos del cambio climático, pues está claro que existe la tentación de relacionarlos con el cambio climático dado que actúan como poderosos imanes sobre la atención social, y parecen ideales para concienciar a la sociedad sobre el problema. Lo que voy a argumentar en este artículo es que para esa tarea son bastante poco útiles, sobre todo comparados con variables como la temperatura media, si hemos de atenernos al necesario rigor científico y no nos dejamos llevar por la impresión del momento o la aparente fuerza de algunos titulares.

## Cambio en la frecuencia de sucesos extremos

Me voy a referir en lo sucesivo a la frecuencia de ocurrencia de eventos extremos, que cambia lógicamente con un cambio del clima. Para simplificar la discusión consideremos dos estados climáticos A y B, ambos con el mismo número de años, de tal forma que en B la probabilidad de ocurrencia de un suceso extremo (por ejemplo la superación de un umbral alto de temperatura) es superior al estado A. Para esta situación un sencillo test estadístico basado en la ley binomial nos permite calcular el número mínimo de ocurrencias del evento en el estado B, para un número total de ocurrencias en A y B dado, que es necesario para poder afirmar con rigor, o sea, con un nivel de significación estadística como el habitual 5% bilateral, que la probabilidad de ocurrencia del evento en B es superior a la misma en A.

Los resultados se muestran en la Fig. 1, que contiene en abscisas el número total de ocurrencias del fenómeno (suma en A y en B), y en ordenadas el porcentaje mínimo de ocurrencias en B sobre el total, para que el cambio de

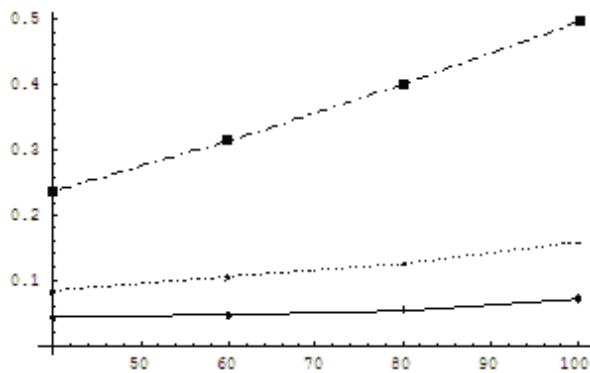


**Figura 1** - Porcentaje mínimo de ocurrencias en una muestra sobre el total en dos muestras del mismo tamaño para detectar al nivel de significación 5% bilateral el cambio en la frecuencia de ocurrencia. En abscisas el tamaño total de las dos muestras.

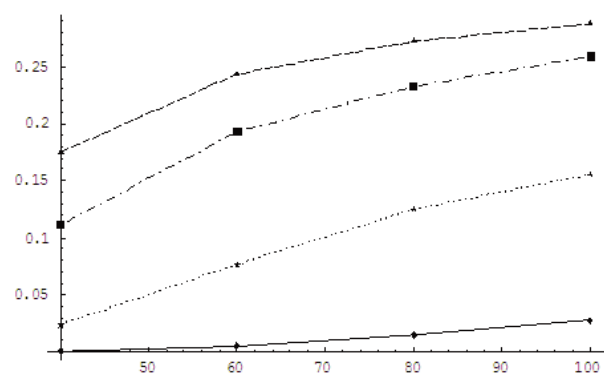
frecuencia sea significativo estadísticamente. Según esta figura, si leyéramos que en determinada región en los últimos 30 años el número de olas de calor (de acuerdo a alguna definición) ha sido de 8 mientras que en los 30 años anteriores ha sido de 2, deberíamos contener nuestro impulso inicial de considerarlo como “prueba” de cambio climático. Con los estándares habituales de rigor científico eso no es suficiente prueba, a pesar de la abultada razón de 1:4. Por tanto, los altos porcentajes de razón de frecuencias que muestra la Fig. 1, contrarios a la intuición, ya parecen indicar que la comparativa de conteos de fenómenos raros es poco prometedora. Pero para poner esto en perspectiva lo natural es compararlo con alguna alternativa.

## Comparación con el test de Mann-Whitney

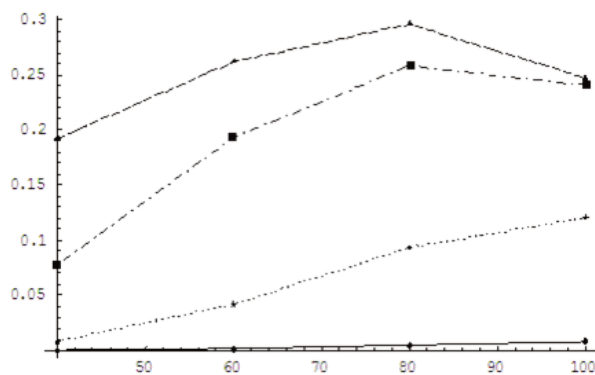
Pongamos algo más de estructura a nuestro par de estados climáticos y supongamos que en cada estado la variable se distribuye normalmente con la misma desviación típica en ambos, y que la diferencia de estados se debe a un cambio en la media. En esta situación se abren muchas posibilidades para estudiar estadísticamente si hay o no cambio de estado, con varios tests paramétricos y no paramétricos. Los más potentes suelen ser los paramétricos, pero como para las frecuencias de ocurrencia el test anterior es no paramétrico (y no hay buenas alternativas



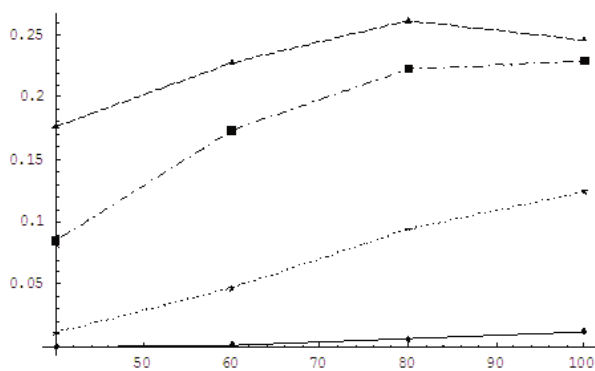
**Figura 2** - Probabilidad de detección significativa al 5% con el test de Mann-Whitney del cambio de media en dos muestras del mismo tamaño normalmente distribuidas. Cada línea en la figura corresponde a cambio de media igual a  $0.1\sigma$ ,  $0.2\sigma$  y  $0.4\sigma$  respectivamente, de abajo a arriba. Suma de tamaños de las dos muestras en abscisas.



**Figura 5** - Como en la figura anterior para un cambio en la media igual a  $0.4\sigma$ .



**Figura 3** - Razón entre la probabilidad de detección al 5% del cambio de media igual a  $0.1\sigma$  en dos muestras normales del mismo tamaño con el test de Mann-Whitney y con el test binomial para cambio de frecuencia de ocurrencia en la superación de un umbral extremo. Las tres líneas corresponden a umbral extremo el cuantil 0.99 (la inferior), 0.98, 0.90 y 0.80. Tamaño total en abscisas.



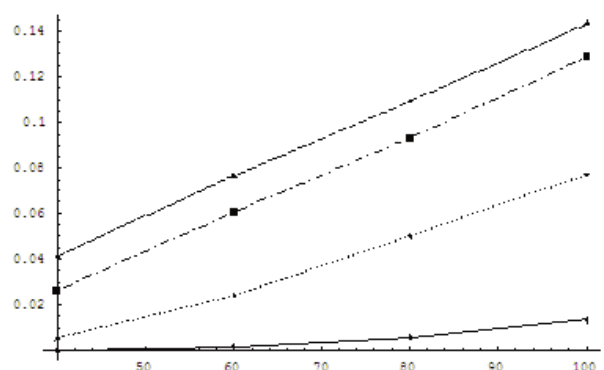
**Figura 4** - Como en la figura anterior para un cambio en la media igual a  $0.2\sigma$ .

paramétricas), consideraremos un buen test no paramétrico para mantener la equidad, el test de Mann-Whitney.

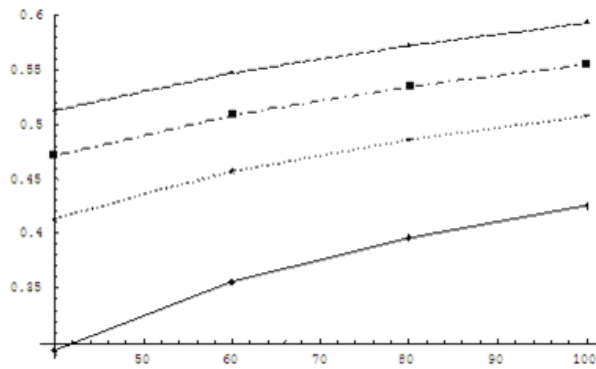
En la Fig. 2 se ha representado la probabilidad de que con este test detectemos el cambio de estado climático con

el habitual estándar de rigor científico para distintos cambios en la media. Frente a este patrón las figuras 3, 4 y 5 (correspondientes a cambios en la media de 0.1, 0.2 y 0.4 veces la desviación típica, respectivamente) contienen la razón de la probabilidad de detección con el test de la ocurrencia de sucesos extremos con relación al test de M-W. En cada figura las distintas líneas corresponden a distintas definiciones de suceso extremo, la inferior a superar el cuantil 0.99 de la normal, las siguientes el 0.98, 0.90 y 0.80. Las figuras muestran que el test basado en la frecuencia es mejor cuanto más bajo es el umbral, es decir, cuanto el suceso extremo es más frecuente. De hecho para los dos umbrales más extremos la eficacia relativa del test de frecuencias prácticamente está por debajo del 10%. Pero incluso para el umbral más bajo, el cuantil 0.80, con un periodo de retorno de 5 años (por tanto ya poco extremo) la eficacia relativa del test de las frecuencias prácticamente no llega al 25%. Por otra parte la eficacia relativa empeora algo con el aumento del cambio de media.

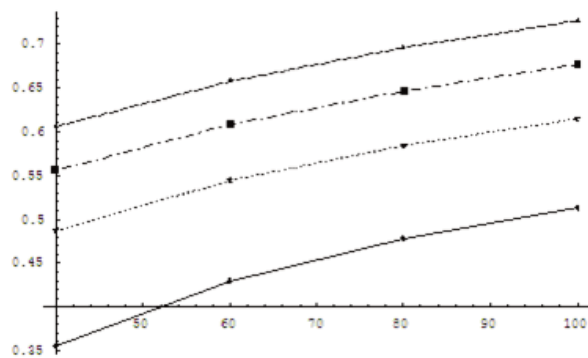
En la Fig. 6 se representan las probabilidades absolutas de detección con el test de las frecuencias para cambio de media  $0.4\sigma$ , que como se ve son inferiores al 15%. Con los cambios de media inferiores (no representadas) no llegan al 3%. En cambio la Fig. 7 recoge la probabilidad de que



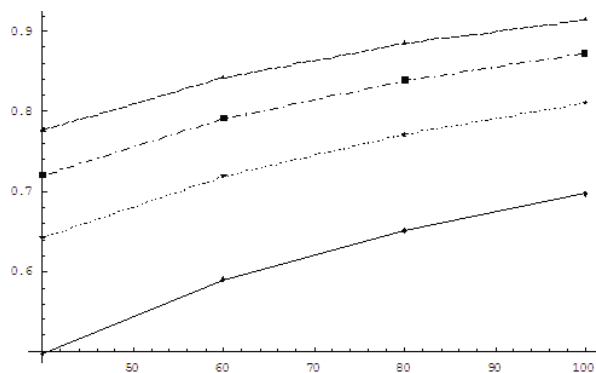
**Figura 6** - Probabilidades de detección con el test binomial basado en diferencia de ocurrencias para cambio en la media de  $0.4\sigma$ , umbrales como en las figuras anteriores



**Figura 7** - Probabilidad de que haya más ocurrencias en la segunda muestra que en la primera, para cambio en media igual a  $0.1\sigma$ .



**Figura 8** - Como en la figura anterior para cambio en media igual a  $0.2\sigma$ .



**Figura 9** - Como en la figura anterior para un cambio en media igual a  $0.4\sigma$ .

la frecuencia de ocurrencia en el estado B sea superior a la del estado A en cada caso, o sea, de que el cambio en frecuencias sea congruente con el cambio climático. Una comparación de las Figs. 6 y 9 muestra a las claras que existe una alta probabilidad (alrededor del 60%, aunque varía según el umbral) de que aquellos que no echen cuentas se dejen seducir por el cambio en la frecuencia de ocurrencia y lo confundan con una prueba con rigor de cambio climático. Es decir, la diferencia de probabilidades entre las dos figuras es la probabilidad de observar un cambio en la frecuencia de ocurrencia que sin embargo no llega a ser sig-

nificativo. Para cambios de media inferiores (Figs. 7 y 8) esta probabilidad de error disminuye algo pero sigue siendo peligrosamente alta.

### Conclusiones

La discusión anterior pone de manifiesto la relativamente poca eficacia de los cambios en la frecuencia de sucesos extremos a la hora de aducir pruebas con rigor de cambio de estado climático. Hemos visto que la alternativa consistente en utilizar la información de toda la población, y no sólo de los extremos, como con el test de Mann-Whitney, es sin duda más potente, por un factor de al menos 1:4. También es importante caer en la cuenta de que la probabilidad de que la frecuencia de ocurrencia observada apunte en el sentido del cambio de estado climático es mucho mayor que la probabilidad de que sea suficiente desde el punto de vista estadístico para demostrar ese cambio.

Bajo otra luz, los resultados anteriores constituyen un poderoso argumento mas para subrayar lo que esta en

juego con el cuidado o negligencia de la homogeneidad de las series climatológicas. Porque para aprovechar la potencia de tests como el de Mann-Whitney, es esencial que las series sean homogéneas. Una ruptura en la homogeneidad puede fácilmente hacer que este tipo de tests conduzca a conclusiones erróneas. Como pone de manifiesto la figura 3 un salto de  $0.4\sigma$  en el medio de una serie de longitud 100 tiene una probabilidad de cerca del 50% de ser significativo con el test de Mann-Whitney; por tanto si se debe a una inhomogeneidad esa es la probabilidad de ser conducidos a conclusiones erróneas sobre la tendencia de la serie.

En cambio hay que reconocer que si trabajamos en la cola de las distribuciones, donde viven los sucesos extremos, los problemas de homogeneidad se atenúan, al ser de todas formas mucho mayor la variabilidad intrínseca. La Fig. 5 muestra que para la misma serie y salto que el considerado antes, la probabilidad de detectar cambios significativos en la frecuencia de superación de umbrales extremos es como mínimo 4 veces menor que antes. En conclusión, aquellos servicios meteorológicos que no quieran ver cuarteada (o aun peor) su capacidad para hablar con fundamento de la evolución del cambio climático harían bien en esforzarse en preservar la homogeneidad de sus series climáticas. Los sucesos extremos serán más atractivos, tendrán mas impacto social y todo tipo de repercusiones, pero para este propósito no son la mejor opción.

### Bibliografía

**W.J. Conover, 1999**, *Practical nonparametric tests*. Wiley Series in Prob. and Statistics.