

POTENCIA DE PRUEBAS DE RACHAS PARA ALTERNATIVA DE TENDENCIA

por

Emilse Gómez Torres¹, Raydonal Ospina² & Jimmy Corzo³

Dedicado al profesor Don Jairo Charris Castañeda

Resumen

Gómez Torres, E., Raydonal Ospina & Jimmy Corzo: Potencia de pruebas de rachas para alternativa de tendencia, Rev. Acad. Colomb. Cienc. **28** (106): 27–38, 2004. ISSN 0370-3908.

Se presenta un estudio de simulación de la potencia de varias pruebas de rachas para una alternativa de tendencia monótona, en muestras obtenidas de distribuciones de Laplace, logística y normal, y para funciones de tendencia monótona con comportamiento de crecimiento lento a moderado.

Palabras Clave. Simulación, tendencia monótona en localización, pruebas de rangos, hipótesis de aleatoriedad

Abstract

A simulation study with observations coming from Laplace, logistic and normal distributions is presented to estimate the power of run tests for trend alternative and to compare them with Daniels test.

Key Words and Phrases. Simulation, location trend, rank tests, hypothesis of randomness

1. Introducción

Sea Y_1, Y_2, \dots, Y_T una sucesión de variables aleatorias independientes que representan T observaciones en el tiempo de una variable Y , con funciones de distribución continuas $F_t(y \pm a(t))$ para $t = 1, \dots, T$ con varianza

constante, donde $a(t)$ es una función monótona en t , continua en el intervalo $[1, T]$ y derivable en cada punto del intervalo $(1, T)$. La hipótesis de aleatoriedad se formula como sigue

$$H_1 : F_1(x) = \dots = F_T(x) = F(x) \text{ para todo } x,$$

¹Profesora Asistente del Departamento de Estadística de la Universidad Nacional de Colombia (egomez@matematicas.unal.edu.co)

²Estadístico de la Universidad Nacional de Colombia (raydonal@cox.de.ufpe.br).

³Profesor Asociado del Departamento de Estadística de la Universidad Nacional de Colombia (jcorzo@matematicas.unal.edu.co).

lo cual equivale en términos de $a(t)$ a la hipótesis:

$$H_1 : a'(t) = 0, \text{ para todo } t \in (1, T) \quad (1)$$

Se consideran las siguientes dos alternativas de tendencia:

$$K_1 : a'(t) > 0 \text{ para todo } t \in (1, T) \quad (2)$$

tendencia monótona creciente

$$K_2 : a'(t) < 0 \text{ para todo } t \in (1, T)$$

tendencia monótona decreciente

Obsérvese que la función $a'(t)$ es responsable de la tendencia en las alternativas consideradas.⁴

2. Estadístico basado en rachas

Para la construcción de un estadístico de rachas⁵ es necesario transformar la sucesión Y_1, Y_2, \dots, Y_T es una sucesión dicótoma⁶, que contenga información sobre la tendencia de la sucesión. Para esto se define la siguiente función de cambio en la cual se compara una observación en el tiempo i con otra que se encuentra m unidades más adelante:

$$\Delta_{i,m} := Y_{i+m} - Y_i$$

donde: $i = 1, \dots, T-1$, y m es un número que depende del número máximo de comparaciones entre una observación y cada una de las posteriores a ella; este número máximo de comparaciones se denota por k y se fija de acuerdo con la cantidad de información sobre la tendencia que se quiere recuperar. Así $m = 1, \dots, \min\{k, T-i\}$, esto es $m = 1, \dots, k$ cuando $i = 1, \dots, T-k$ y $m = 1, \dots, T-i$ cuando $i = T-k+1, \dots, T-1$.

Por ejemplo, si se toma $k = 3$, es porque se quiere comparar cada observación con las tres subsiguientes a ella. En tal caso cuando $i = 1, \dots, T-3$, $m = 1, 2, 3$; cuando $i = T-2$, $m = 1, 2$ y cuando $i = T-1$, $m = 1$; a partir de los cuales se generan las siguientes $3(T-2)$ cambios: $\Delta_{1,1}, \Delta_{1,2}, \Delta_{1,3}, \Delta_{2,1}, \Delta_{2,2}, \Delta_{2,3}, \dots, \Delta_{T-3,1}, \Delta_{T-3,2}, \Delta_{T-3,3}, \Delta_{T-2,1}, \Delta_{T-2,2}, \Delta_{T-1,1}$.

⁴Aiyar, Guillier y Albers (1979) utilizan el siguiente modelo más dinámico $Y_t = \beta_T a(t) + b + \varepsilon_t$ donde el parámetro β_T depende del número de observaciones en la sucesión y b es una constante real que representa un cambio en localización, para estudiar la eficiencia relativa asintótica de ciertas pruebas de rangos para alternativas de tendencia monótona.

⁵Corresponde al término inglés *run*, el cual según Gibbons se puede definir así: Dada una secuencia de dos o más tipos de símbolos, una racha es definida como una sucesión de uno o más símbolos idénticos que son seguidos o precedidos por un símbolo diferente o por ningún símbolo (Gibbons (1971, pág. 50))

⁶Una sucesión dicótoma es un arreglo de dos símbolos distinguibles. La transformación de una sucesión de variables aleatorias a una sucesión de variables aleatorias binarias se denomina dicotomización. La correspondiente sucesión de realizaciones de las variables binarias también se denominará sucesión dicótoma.

También se define la siguiente variable que indica si el cambio $\Delta_{i,m}$ corresponde a un **estado de incremento** en el valor de Y en el tiempo $i+m$ con respecto al tiempo i o no:

$$\begin{aligned} \eta_{i,m} &:= \begin{cases} 1 & \text{si } \Delta_{i,m} > 0 \\ 0 & \text{si } \Delta_{i,m} \leq 0 \end{cases} \\ &:= \frac{1}{2} \{1 + \text{sign}(\Delta_{i,m})\} \end{aligned} \quad (3)$$

para $i = 1, \dots, T-1$ y $1 \leq m \leq T-i$, donde

$$\text{sign}(x) := \begin{cases} 1 & \text{si } x > 0 \\ -1 & \text{si } x \leq 0 \end{cases} \quad \text{para } x \in R$$

El número de elementos incluidos en la sucesión dicótoma depende de k , por esta razón se denota por $N(k)$ el cual se obtiene sumando las $k(T-k)$ comparaciones entre k elementos sucesivos a todas las comparaciones posibles de los últimos k elementos tomados por pares, así: $N(k) = k(T-k) + k(k-1)/2 = k(T - (k+1)/2)$. Retomando el ejemplo anterior, cuando $k = 3$, $N(k)$ se obtiene sumando los $3(T-3)$ cambios de los $\Delta_{i,m}$ con $i = 1, \dots, T-3$ y $m = 1, 2, 3$ con los 3 cambios de las últimas tres observaciones tomadas en pares, por tanto $N(k) = 3(T-3) + 3(3-1)/2 = 3(T - (3+1)/2) = 3(T-2)$

La forma de interpretar la sucesión dicótoma se ilustra en el siguiente ejemplo:

En caso de que $\eta_{i,m} = 1$ para todo i y para todo m es porque $Y_1 < Y_2 < \dots < Y_T$, lo cual significa que la sucesión de los Y_t es estrictamente creciente.

Si el caso es $\eta_{i,m} = 1$ para $i = 1, \dots, T-2$ y $\eta_{T-1,1} = 0$, es porque $Y_1 < Y_2 < \dots < Y_{T-1}$, $Y_i < Y_T$ para $i = 1, \dots, T-2$ y $Y_{T-1} > Y_T$, esto significa que la sucesión de los Y_t es creciente hasta la penúltima observación.

Con este mismo razonamiento se puede intuir que cualquier alteración en el orden estricto de incremento de la sucesión de los Y_t se verá reflejada en la presencia de ceros en la sucesión de los $\eta_{i,m}$. De esta manera cuando el número de unos presentes en la sucesión de los $\eta_{i,m}$ es grande se podrá interpretar que la mayor

o decremento hasta el tiempo $t(i, m)$, en el arreglo de las comparaciones sucesivas $\Delta_{i,m}$ de los elementos de la sucesión $\{Y_1, Y_2, \dots, Y_T\}$. Por ejemplo en la Tabla 1, el ${}^{-}r_{t(2,3)} - 1 = {}^{-}r_8 - 1$ representa el número de veces que la subsucesión $\eta_{1,1}, \eta_{1,2}, \eta_{1,3}, \eta_{2,1}, \eta_{2,2}, \eta_{2,3}, \eta_{3,1}, \eta_{3,2}$ ha permanecido en unos o en ceros; así si esta subsucesión fuera 1, 1, 0, 1, 0, 0, 0, 1 entonces ${}^{-}r_8 - 1 = 3$ indicando tres permanencias en algún estado, dos marcadas con líneas debajo y la última corresponde a los dos ceros en negrilla. Se define la función que indica con un valor positivo un incremento y con un valor negativo un decremento en un instante dado, mediante la función:

$$\delta_{t(i,m)} := g(i, m) \text{sign}(\Delta_{i,m}) = \delta_t \quad (7)$$

cuyo signo está dado por la función de cambio. Dos posibles formas de la función $g(i, m)$ que son independientes de m se utilizarán más adelante: $g(i, m) = 1$ y $g(i, m) = i$. Finalmente como el número de antirrachas en la sucesión de η_t es una variable aleatoria, se utiliza la función $\varsigma({}^{-}r_{N(k)})$ para normar el estadístico de rachas, que se define como sigue:

$$G_k := \frac{1}{\varsigma({}^{-}r_{N(k)})} \sum_{t=1}^{N(k)} \delta_t {}^{-}r_t \quad (8)$$

Valores extremos positivos de G_k se obtienen cuando hay muchas permanencias (${}^{-}r_t - 1$ grande) y son en

su mayoría incrementos (δ_t positivo), éstos reflejan una tendencia de la sucesión Y_1, \dots, Y_T a permanecer en estado creciente. Análogamente, valores extremos negativos de G_k se obtienen cuando hay muchas permanencias (${}^{-}r_t - 1$ grande) y son en su mayoría decrementos (δ_t negativo), éstos reflejan una tendencia de la sucesión Y_1, \dots, Y_T a permanecer en estado decreciente. En consecuencia, la hipótesis nula se rechazará en favor de la alternativa de tendencia a permanecer en algún estado cuando el estadístico toma valores extremos, tanto positivos como negativos.

El estudio de potencia se hará para varios casos particulares de G_k que se distinguen de dos maneras: por el tipo de función que indica el cambio $g(i, m)$ y por la función que normaliza el valor de la estadística $\varsigma({}^{-}r_{N(k)})$. Para facilitar la interpretación estos casos se distinguen con las letras ${}^{-}C_k$ y γ_k y se muestran en la Tabla 2.

Nótese que en los casos estudiados, los estadísticos de la forma ${}^{-}C_k$ tienen como factor de normalización ${}^{-}r_{N(k)}$ y los de la forma γ_k no están normalizados. Además la función $g(i, m)$ se mantiene constante en ${}^{-}C_k$, mientras que γ_k es una función del primero de los dos instantes de observación comparados en la función de cambio.

Tabla 2. Estadísticos basados en rachas para pruebas de tendencia monótona

Estadístico	$g(i, m)$	$\varsigma({}^{-}r_{N(k)})$	Casos: Características
${}^{-}C_k = \frac{1}{{}^{-}r_{N(k)}} \sum_{t=1}^{N(k)} \delta_t {}^{-}r_t$	1	${}^{-}r_{N(k)}$	${}^{-}C_1 : \begin{cases} k = 1; m = 1; \\ N(k) = T - 1 \end{cases}$ ${}^{-}C_{T-1} : \begin{cases} k = T - i; m = 1, \dots, T - i; \\ N(k) = T(T - 1)/2 \end{cases}$
$\gamma_k = \sum_{i,m} \delta_{i,m} {}^{-}r_{i,m}$	i	1	$\gamma_1 : \begin{cases} k = 1; m = 1; \\ N(k) = T - 1 \end{cases}$ $\gamma_2 : \begin{cases} k = 2; m = 1, \text{mín} \{T - i, 2\}; \\ N(k) = 2T - 3 \end{cases}$ $\gamma_5 : \begin{cases} k = 5; m = 1, \dots, \text{mín} \{T - i, 5\}; \\ N(k) = 5T - 15 \end{cases}$ $\gamma_{T-1} : \begin{cases} k = T - i; m = 1, \dots, T - i; \\ N(k) = T(T - 1)/2 \end{cases}$

2.1 Regiones críticas

Tablas de los valores críticos de ${}^{-}C_1$ y ${}^{-}C_{T-1}$ y un procedimiento para obtener la distribución de los estadísticos, se encuentran en Corzo y Gómez (2000). El método que se utiliza para obtener la distribución de los estadísticos γ_k es el mismo que el empleado para el

cálculo de la distribución de ${}^{-}C_1$ y ${}^{-}C_{T-1}$ dado que la única diferencia entre las dos clases de estadísticos esta en el denominador⁸. Para tamaños de muestra entre 5 y 8 se utilizan los valores críticos exactos de las distribuciones de los seis estadísticos y para tamaños de muestra entre 9 y 50 se emplean las aproximaciones

⁸Detalles sobre el cálculo de las distribuciones y regiones críticas de los estadísticos γ_k se pueden consultar en Gómez (2002).

obtenidas mediante simulación. El mecanismo adoptado para aproximar los percentiles consiste en generar un número adecuado de permutaciones aleatorias de los enteros de 1 a T y posteriormente estimar los percentiles con la probabilidad asociada más cercana por debajo, al nivel de significancia deseado⁹.

3. Estudio de potencia

El estudio de la potencia se lleva a cabo simulando

muestras a partir del modelo $Y_t = \pm a(t) + \varepsilon_t$, en el cual $a^j(t)$ representa una función de tendencia monótona creciente (ocho casos de $a^j(t)$ se muestran en la Tabla 3), y ε_t es la parte aleatoria que se generó a partir de las distribuciones de Laplace, logística y normal. En este artículo sólo se presentan los resultados para $a(t)$ dado que los resultados para $-a(t)$ son muy similares (Gómez (2002)).

Tabla 3. Casos de tendencia monótona creciente utilizados

Caso j	$a^j(t)$	$a'^j(t)$: Rapidez de crecimiento
Caso 1	$a^1(t) = t^{1/4}$	$a'^1(t) = \frac{1}{4}t^{-3/4}$
Caso 2	$a^2(t) = t^{2/5}$	$a'^2(t) = \frac{2}{5}t^{-3/5}$
Caso 3	$a^3(t) = t^{1/2}$	$a'^3(t) = \frac{1}{2}t^{-1/2}$
Caso 4	$a^4(t) = t^{3/5}$	$a'^4(t) = \frac{3}{5}t^{-2/5}$
Caso 5	$a^5(t) = t^{7/10}$	$a'^5(t) = \frac{7}{10}t^{-3/10}$
Caso 6	$a^6(t) = t^{4/5}$	$a'^6(t) = \frac{4}{5}t^{-1/5}$
Caso 7	$a^7(t) = t^{9/10}$	$a'^7(t) = \frac{9}{10}t^{-1/10}$
Caso 8	$a^8(t) = t$	$a'^8(t) = 1$

Aproximación de las potencias (para cada caso de $a^j(t)$ de la Tabla 3)

Paso 1: Generar m muestras de tamaño T de la distribución seleccionada para ε_t , obtener $y_t = \pm a^j(t) + \varepsilon_t$ para $t = 1, \dots, T$, y calcular el número de rechazos de la hipótesis nula $\hat{\pi}_1$.

Paso 2: Generar el doble de muestras de tamaño T que en el paso anterior, obtener $y_t = \pm a^j(t) + \varepsilon_t$ para $t = 1, \dots, T$, calcular el número de rechazos de la hipótesis nula $\hat{\pi}_2$ y $|\hat{\pi}_2 - \hat{\pi}_1|$

Paso 3: Si $|\hat{\pi}_2 - \hat{\pi}_1| > 0.005$ ir a paso 2.

Paso 4: Terminar.

Una aproximación del nivel de significancia se obtiene con el anterior algoritmo para $a^j(t) = 0$

En este estudio, se comparan entre sí las potencias de las pruebas que utilizan los estadísticos propuestos ${}^-C_1$, ${}^-C_{T-1}$, γ_1 , γ_2 , γ_5 y γ_{T-1} , y con las potencias de la prueba de Daniels, que utiliza el estadístico:

$$\rho = \frac{12}{T(T^2 - 1)} \sum_{i=1}^T \left(i - \frac{T+1}{2} \right) \left(R_i - \frac{T+1}{2} \right) \quad (9)$$

donde R_i denota el rango de la i -ésima observación. Como ρ está basado en el coeficiente de correlación de Spearman, su rango de variación está entre -1 y 1; así que la prueba rechaza la hipótesis de aleatoriedad cuando ρ se acerca a sus valores extremos.

Como en los tamaños de muestra menores o iguales que ocho no era posible obtener el mismo nivel de significancia en todas las pruebas, la comparación se realizó sobre las potencias empíricas de las pruebas aleatorizadas. Para tamaños de muestra mayores, no fue necesario hacer la aleatorización de las pruebas porque los niveles de significancia resultaron muy parecidos para todas las pruebas.

3.1. Resultados

El estudio del comportamiento de las potencias de las pruebas se realizó de la siguiente manera:

⁹El número de permutaciones a generar es el tamaño de una muestra aleatoria simple del conjunto de las $T!$ permutaciones de los enteros de 1 a T . La varianza de los estadísticos se aproximó por el método Jackknife. Una descripción completa de la metodología de simulación se encuentra en Gómez (2002).

- **Consistencia**¹⁰: Cambiando el tamaño de sucesión y fijando un caso de tendencia y una distribución. Los resultados de esta parte se presentan en la Tabla resumen 1 en la cual las filas muestran los tamaños de muestra que resaltan las diferencias entre las potencias aproximadas de las pruebas comparadas.
- **Potencia**: Cambiando el caso de tendencia y fijando un tamaño de la sucesión y una distribución. Los resultados de esta parte se presentan en la Tabla resumen 2 que contiene en sus cuatro filas las aproximaciones numéricas de las potencias de las siete pruebas para los cuatro casos de tendencia que más diferencias establecen entre las potencias aproximadas de las pruebas comparadas: $t^{1/4}$, $t^{1/2}$, $t^{4/5}$ y t (casos 1, 3, 6 y 8)¹¹.
- **Sensibilidad**¹²: Cambiando la distribución y fijando un tamaño de sucesión y un caso de tendencia.

Las Tablas Resumen presentan las aproximaciones de las potencias de las pruebas con un nivel de significancia de 5%¹³, simulando 50000 muestras¹⁴ de tamaño T que varía de 5 a 50 (incrementando de 1 en 1), para muestras de distribución de Laplace¹⁵, utilizando la siguiente notación:

T : tamaño de la muestra, aumenta de 5 en 5 desde 5 hasta 50

C1: potencia estimada de la prueba con $^{-}C_1$

C(T-1): potencia estimada de la prueba con $^{-}C_{T-1}$

gama1: potencia estimada de la prueba con γ_1

gama2: potencia estimada de la prueba con γ_2

gama5: potencia estimada de la prueba con γ_5

gama(T-1): potencia estimada de la prueba con γ_{T-1}

Daniels: potencia estimada de la prueba de Daniels (con ρ)

3.1.1. Consistencia

¹⁰Se realizó para muestras hasta de tamaño 50.

¹¹En Gómez (2002) se encuentran la totalidad de resultados enunciados en este artículo de manera que el lector interesado en profundizar en el tema pueda consultar un documento más detallado.

¹²Se refiere a la sensibilidad de las pruebas al grado de apuntamiento de las distribuciones muestreadas.

¹³Los resultados para un nivel de significancia del 1% son muy similares a los presentados aquí y se pueden consultar en Gómez (2002).

¹⁴Una muestra corresponde a una sucesión de variables de un tamaño dado con un comportamiento específico.

¹⁵Los resultados de las potencias con las distribuciones logística y normal son muy similares, razón por la cual no se hacen comentarios sobre ellos y sólo se presentan las tablas correspondientes en el anexo.

En la Tabla resumen 1 se evidencia el crecimiento de las potencias, en algunos casos muy lento y en otros casos bastante rápido, dependiendo de la rapidez de crecimiento de las funciones de tendencia y del número de comparaciones entre pares de observaciones consideradas en el estadístico de prueba. En general, el crecimiento más rápido lo presentan las potencias de la prueba de Daniels, seguido de cerca por las potencias de la prueba con $^{-}C_{T-1}$ y, un poco más despacio, por las potencias de la prueba con γ_{T-1} ; las potencias más bajas son las de las pruebas cuyos estadísticos son γ_5 , γ_2 , $^{-}C_1$ y γ_1 , en su orden.

Caso 1: $a^1(t) = t^{1/4}$. Las potencias de las pruebas basadas en γ_5 , γ_2 , γ_1 y $^{-}C_1$ son muy bajas (no superan el 12%) y muestran un comportamiento casi constante. Las potencias de las pruebas basadas en γ_{T-1} y $^{-}C_{T-1}$ presentan un crecimiento moderado llegando sólo al 41,68% y 71,05% respectivamente, mientras que la prueba de Daniels alcanza el 91,63%.

Caso 3: $a^3(t) = t^{1/2}$. Se refleja con claridad la mayor bondad de las pruebas que emplean todas las comparaciones (basadas en ρ , $^{-}C_{T-1}$ y γ_{T-1}), pues rápidamente obtienen la máxima potencia, aunque cada prueba lo hace a partir de un tamaño de muestra diferente ($T = 28$ para la prueba basada en ρ , $T = 37$ para la prueba basada en $^{-}C_{T-1}$ y $T = 46$ para la prueba basada en γ_{T-1}). Las potencias de las pruebas que utilizan pocas comparaciones (basadas en los estadísticos $^{-}C_1$, γ_1 y γ_2) presentan un crecimiento muy lento (sin llegar al 30%). Las potencias de la prueba con γ_5 se distinguen por acercarse con rapidez a las potencias de las pruebas basadas en ρ , $^{-}C_{T-1}$ y γ_{T-1} ; si bien en ese punto parece ser equidistante a los dos grupos descritos, al menos entre tamaños de sucesión 15 a 35.

Una tipología similar a la del caso 3 se observa en los casos 2 y 4, resaltando los siguientes aspectos:

Tabla resumen 1: Consistencia contra algunas alternativas de las pruebas con distribución de Laplace y $\alpha = 5\%$

Caso 1							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	7.47%	7.08%	6.89%	8.03%	8.03%	9.43%	10.27%
10	6.85%	7.04%	7.50%	9.41%	10.47%	13.43%	20.42%
15	6.34%	6.83%	7.83%	10.50%	13.25%	20.99%	30.32%
20	7.09%	6.48%	7.47%	10.48%	17.11%	26.18%	42.48%
25	7.15%	6.69%	7.73%	11.48%	21.39%	34.77%	54.39%
30	7.81%	6.54%	7.47%	11.50%	26.70%	44.38%	65.29%
35	7.38%	6.29%	8.02%	11.61%	31.78%	51.20%	74.40%
40	7.38%	6.44%	7.89%	12.05%	37.22%	59.04%	81.98%
45	7.16%	6.45%	7.45%	12.18%	41.35%	65.93%	87.53%
50	6.64%	6.27%	8.35%	11.42%	41.68%	71.05%	91.63%
Caso 3							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	12.89%	11.99%	11.93%	15.00%	15.00%	19.86%	22.40%
10	15.47%	13.95%	17.33%	28.82%	34.83%	47.77%	65.19%
15	15.35%	13.76%	20.14%	39.78%	59.32%	79.85%	91.48%
20	17.08%	13.43%	21.40%	47.83%	80.95%	94.16%	99.13%
25	17.91%	13.89%	23.54%	56.34%	92.81%	99.07%	99.96%
30	19.18%	14.04%	23.52%	60.53%	98.01%	99.93%	99.99%
35	18.75%	13.84%	26.15%	65.08%	99.53%	99.99%	100.00%
40	18.87%	14.08%	26.11%	69.11%	99.94%	100.00%	100.00%
45	18.47%	14.02%	25.92%	72.61%	99.99%	100.00%	100.00%
50	17.55%	14.20%	28.61%	73.55%	99.95%	100.00%	100.00%
Caso 6							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	28.57%	27.19%	27.22%	34.47%	34.47%	45.71%	50.95%
10	47.97%	41.84%	56.92%	82.88%	88.72%	95.99%	99.15%
15	58.21%	50.17%	74.07%	97.31%	99.76%	100.00%	100.00%
20	67.32%	56.28%	84.52%	99.60%	100.00%	100.00%	100.00%
25	74.12%	62.32%	90.71%	99.95%	100.00%	100.00%	100.00%
30	79.45%	66.93%	94.02%	99.99%	100.00%	100.00%	100.00%
35	82.63%	70.65%	96.67%	100.00%	100.00%	100.00%	100.00%
40	85.41%	73.62%	97.75%	100.00%	100.00%	100.00%	100.00%
45	87.55%	76.89%	98.64%	100.00%	100.00%	100.00%	100.00%
50	88.60%	79.25%	99.16%	100.00%	100.00%	100.00%	100.00%
Caso 8							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	46.40%	45.03%	45.06%	54.19%	54.19%	68.14%	73.57%
10	80.22%	73.53%	88.68%	98.76%	99.50%	99.93%	100.00%
15	92.19%	87.66%	98.04%	100.00%	100.00%	100.00%	100.00%
20	97.48%	94.48%	99.74%	100.00%	100.00%	100.00%	100.00%
25	99.28%	97.74%	99.98%	100.00%	100.00%	100.00%	100.00%
30	99.78%	99.11%	100.00%	100.00%	100.00%	100.00%	100.00%
35	99.93%	99.62%	100.00%	100.00%	100.00%	100.00%	100.00%
40	99.99%	99.89%	100.00%	100.00%	100.00%	100.00%	100.00%
45	99.99%	99.96%	100.00%	100.00%	100.00%	100.00%	100.00%
50	100.00%	99.98%	100.00%	100.00%	100.00%	100.00%	100.00%

Caso 2: $a^2(t) = t^{2/5}$. Las potencias de la prueba con γ_5 empiezan a alejarse de las del grupo que emplea poca información, pero más lentamente que en el caso 3. Las pruebas con estadísticos que utilizan todas las comparaciones empiezan a alcanzar la potencia máxima.

Caso 4: $a^4(t) = t^{3/5}$. Hay mayor claridad en el acercamiento de las potencias de la prueba con γ_5 a las de las pruebas que utilizan todas las comparaciones, que cada vez logran el máximo en un tamaño de muestra más pequeño.

Caso 6: $a^6(t) = t^{4/5}$. Las potencias de las pruebas que emplean pocas comparaciones (basadas en los estadísticos ${}^{-}C_1$, γ_1 y γ_2) aumentan más rápido que en los casos 1 a 4; por ejemplo la prueba menos potente que es la basada en γ_1 alcanza una potencia de 62,3% para un tamaño de muestra relativamente pequeño, $T = 25$, potencia que nunca alcanza en los casos 1 a 4. Hay mayor cercanía de las potencias de las pruebas basadas en estadísticos que emplean todas las comparaciones y la prueba basada en γ_5 , que consigue la potencia máxima en $T = 31$, en tanto la prueba basada en γ_{T-1} la consigue para $T = 19$, la basada en ${}^{-}C_{T-1}$ la consigue para $T = 15$ y la prueba de Daniels la consigue para $T = 15$.

En el caso 5: $a^5(t) = t^{7/10}$, las curvas de las potencias estimadas son muy similares en su forma y en la velocidad de aumento, con la diferencia de que las potencias de las pruebas basadas en estadísticos que emplean pocas comparaciones exhiben un crecimiento más lento que en el caso 6.

Caso 8: $a^8(t) = t$. Con excepción de la prueba basada en γ_1 , todas las pruebas obtienen la potencia máxima, desde tamaños de muestra distintos: 10 para la prueba de Daniels, 12 para la prueba basada en ${}^{-}C_{T-1}$, 13 para la basada en γ_{T-1} , 15 para la basada en γ_5 , 30 para la basada en γ_2 y 46 para la basada en ${}^{-}C_1$. El mayor tamaño de muestra $T = 50$ se requiere para alcanzar la potencia máxima (99,98%) de la prueba basada en γ_1 .

En el caso 7: $a^7(t) = t^{9/10}$, las potencias crecen de manera similar al caso 8 aunque con más lentamente, empezando a verse que las potencias de las pruebas con ${}^{-}C_{T-1}$ y ρ se acercan bastante entre sí. De la misma forma las potencias de las pruebas con γ_5 y γ_{T-1} ; y las potencias de las pruebas que emplean pocas comparaciones se van acercando lentamente a las anteriores.

3.1.2. Potencia

Para tamaños de muestra pequeños hay una gran similitud entre las potencias de todas las pruebas. Para

$T = 5$, las pruebas basadas en estadísticos que usan el mismo número de comparaciones (γ_5 y γ_{T-1}) tienen potencias iguales; las pruebas basadas en estadísticos que usan pocas comparaciones (${}^{-}C_1$, γ_1 y γ_2) traslapan todas sus potencias.

Entre $T = 6$ y $T = 14$ se observan mayores diferencias entre las potencias para todos los casos, potencias que empiezan a acercarse nuevamente para $T = 15$, tamaño de muestra para el cual las pruebas basadas en estadísticos que utilizan más comparaciones alcanzan la potencia máxima en el caso 8. Las pruebas basadas en los dos estadísticos (${}^{-}C_1$ y γ_1) que emplean la comparación de pares sucesivos sólo alcanzan potencias de 87% y el 92%, respectivamente, las cuales alcanzan su máxima potencia para $T = 25$.

4. Conclusiones

Hay cinco puntos importantes para resaltar del estudio:

El comportamiento de las potencias exhibe una similitud entre pruebas que utilizan un número cercano de comparaciones que se deberá reflejar en la obtención de eficiencias relativas cercanas a uno, de acuerdo con la tipología descrita para cada caso de tendencia.

Si bien los estadísticos más sencillos presentan potencias bajas en los casos de tendencias más lentas, su bondad mejora hasta llegar a ser competitivos en los casos 6 a 8 (ver tabla 3), con la ventaja de que su cálculo es más rápido.

Las pruebas basadas en estadísticos que emplean mayor número de comparaciones entre las observaciones son más competitivos con la tradicional prueba de Daniels, manteniendo sus potencias un comportamiento análogo y muy cercano.

Es importante notar que para aumentar la potencia de la prueba basada en G_k se debe sospechar qué tipo de tendencia puede existir en la variable para definir así el mejor estadístico a emplear. En casos de tendencias muy lentas sería necesario considerar toda la información; cuando se sospecha una tendencia moderada se podría pensar en comparar únicamente 5 términos consecutivos y obtener resultados similares con mayor rapidez; si se puede suponer que hay una tendencia lineal, entonces con observar tres términos consecutivos sería suficiente y si la tendencia parece más rápida basta con examinar los pares sucesivos.

Tabla resumen 2: Potencias de las pruebas con distribución de Laplace y $\alpha=5$ para algunos tamaños de sucesión

T=5							
Caso	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
1	7.47%	7.08%	6.89%	8.03%	8.03%	9.43%	10.27%
2	10.19%	9.50%	9.40%	11.53%	11.53%	14.71%	16.39%
3	12.89%	11.99%	11.93%	15.00%	15.00%	19.86%	22.40%
4	16.69%	15.60%	15.58%	19.88%	19.88%	26.80%	30.37%
5	21.87%	20.61%	20.62%	26.32%	26.32%	35.33%	39.82%
6	28.57%	27.19%	27.22%	34.47%	34.47%	45.71%	50.95%
7	36.70%	35.28%	35.32%	43.70%	43.70%	56.98%	62.54%
8	46.40%	45.03%	45.06%	54.19%	54.19%	68.14%	73.57%
T=15							
Caso	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
1	6.34%	6.83%	7.83%	10.50%	13.25%	20.99%	30.32%
2	10.22%	9.80%	12.94%	22.65%	34.24%	53.37%	70.21%
3	15.35%	13.76%	20.14%	39.78%	59.32%	79.85%	91.48%
4	24.37%	20.81%	32.88%	64.37%	83.59%	95.67%	99.09%
5	38.88%	32.73%	52.46%	86.40%	96.75%	99.65%	99.99%
6	58.21%	50.17%	74.07%	97.31%	99.76%	100.00%	100.00%
7	78.18%	70.54%	90.72%	99.77%	99.99%	100.00%	100.00%
8	92.19%	87.66%	98.04%	100.00%	100.00%	100.00%	100.00%
T=25							
Caso	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
1	7.15%	6.69%	7.73%	11.48%	21.39%	34.77%	54.39%
2	11.59%	9.67%	13.72%	29.66%	64.87%	85.97%	96.61%
3	17.91%	13.89%	23.54%	56.34%	92.81%	99.07%	99.96%
4	29.61%	22.50%	42.09%	85.33%	99.69%	99.99%	100.00%
5	49.06%	38.17%	68.43%	98.26%	100.00%	100.00%	100.00%
6	74.12%	62.32%	90.71%	99.95%	100.00%	100.00%	100.00%
7	92.77%	85.81%	98.96%	100.00%	100.00%	100.00%	100.00%
8	99.28%	97.74%	99.98%	100.00%	100.00%	100.00%	100.00%

Finalmente, se encuentra evidencia para pensar que el denominador de \bar{C}_k incluye información importante al estadístico y mejora la prueba, pero los principales cambios están dados en términos del número de comparaciones entre pares de observaciones que se incluyan en la sucesión dicótoma, es decir del valor de k .¹⁶

Referencias

- [1] **Aiyar, R., Guillier C. & Albers, W.**, *Asymptotic Relative Efficiencies of Rank Tests for Trend Alternatives*. Journal of American Statistician Association **74** (179), 226–231.
- [2] **Corzo, J. A.**, *Una prueba de dispersión basada en secuencias*. Tesis de Maestría. Universidad Nacional de Colombia, 1984.
- [3] **Corzo, J. A.**, *Algunas propiedades distribucionales de estadísticas de rachas*. Reporte interno No. 25. Departamento de Matemáticas y Estadística. Universidad Nacional de Colombia, Santa Fe de Bogotá.
- [4] **Corzo, J. A.**, *Análisis de datos a través de rachas*. Coloquio Distrital de Matemáticas y Estadística, Santa Fe de Bogotá, 1996.
- [5] **Corzo, J. A. & E. Gómez**, *Una prueba de rachas para alternativa de tendencia con muestras pequeñas*. Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales **24** (2000), 417–426.
- [6] **Gómez, E.** *Estimación de la potencia de una prueba de rachas para alternativa de tendencia*. Trabajo de grado. Universidad Nacional de Colombia, 1999.
- [7] **Gómez, E.** *Estudio de eficiencia relativa de una prueba de rachas para alternativa de tendencia*. Tesis de maestría. Universidad Nacional de Colombia, (2002).

¹⁶Este artículo contiene resultados parciales de la tesis de maestría de Gómez (2002).

Apéndice

Tabla 3. Consistencia de las pruebas con distribución logística y $\alpha=5\%$

Caso 1							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	7.02%	6.68%	6.48%	7.45%	7.45%	8.65%	9.22%
10	6.46%	6.75%	7.04%	8.59%	9.37%	11.76%	17.58%
15	5.96%	6.53%	7.34%	9.52%	11.43%	17.93%	25.92%
20	6.63%	6.14%	6.99%	9.26%	14.47%	21.97%	36.45%
25	6.74%	6.40%	7.32%	10.31%	17.56%	28.84%	47.01%
30	7.40%	6.28%	6.98%	10.15%	21.98%	37.14%	57.35%
35	7.04%	6.05%	7.54%	10.33%	26.21%	43.04%	66.90%
40	7.00%	6.20%	7.45%	10.68%	30.62%	50.08%	75.19%
45	6.83%	6.22%	7.02%	10.68%	33.95%	57.01%	81.73%
50	6.30%	6.01%	7.86%	10.01%	34.65%	61.94%	86.98%
Caso 3							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	11.49%	10.69%	10.61%	13.31%	13.31%	17.38%	19.22%
10	13.19%	12.17%	14.76%	24.16%	29.62%	41.77%	60.72%
15	12.93%	11.92%	17.00%	32.98%	52.38%	75.08%	90.14%
20	14.69%	11.70%	17.84%	39.98%	75.43%	92.04%	99.08%
25	15.32%	12.03%	19.53%	47.35%	89.84%	98.59%	99.97%
30	16.37%	12.18%	19.64%	51.30%	97.10%	99.89%	100.00%
35	15.88%	11.90%	21.56%	55.29%	99.26%	99.99%	100.00%
40	16.04%	12.25%	21.62%	59.31%	99.88%	100.00%	100.00%
45	15.81%	12.27%	21.45%	62.92%	99.99%	100.00%	100.00%
50	14.89%	12.19%	23.62%	63.53%	99.93%	100.00%	100.00%
Caso 6							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	24.57%	23.21%	23.23%	30.30%	30.30%	41.12%	46.11%
10	41.17%	35.56%	50.48%	80.30%	87.25%	95.82%	99.46%
15	50.09%	42.45%	67.74%	96.76%	99.76%	100.00%	100.00%
20	58.91%	47.58%	78.88%	99.51%	100.00%	100.00%	100.00%
25	65.29%	52.49%	86.64%	99.94%	100.00%	100.00%	100.00%
30	71.37%	57.13%	90.54%	99.99%	100.00%	100.00%	100.00%
35	74.59%	60.89%	94.23%	100.00%	100.00%	100.00%	100.00%
40	77.64%	63.78%	96.01%	100.00%	100.00%	100.00%	100.00%
45	80.28%	66.80%	97.16%	100.00%	100.00%	100.00%	100.00%
50	81.06%	69.39%	98.14%	100.00%	100.00%	100.00%	100.00%
Caso 8							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	41.72%	40.26%	40.29%	50.31%	50.31%	65.14%	71.19%
10	75.30%	68.01%	86.71%	98.86%	99.59%	99.96%	100.00%
15	89.39%	83.57%	97.68%	100.00%	100.00%	100.00%	100.00%
20	96.17%	91.67%	99.70%	100.00%	100.00%	100.00%	100.00%
25	98.66%	96.10%	99.96%	100.00%	100.00%	100.00%	100.00%
30	99.61%	98.28%	100.00%	100.00%	100.00%	100.00%	100.00%
35	99.85%	99.21%	100.00%	100.00%	100.00%	100.00%	100.00%
40	99.96%	99.71%	100.00%	100.00%	100.00%	100.00%	100.00%
45	99.98%	99.85%	100.00%	100.00%	100.00%	100.00%	100.00%
50	100.00%	99.94%	100.00%	100.00%	100.00%	100.00%	100.00%

Tabla 4. Consistencia de las pruebas con distribución normal y $\alpha=5\%$

Caso 1							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	6.87%	6.56%	6.33%	7.35%	7.35%	8.36%	8.82%
10	6.40%	6.75%	6.76%	8.37%	9.00%	11.13%	16.54%
15	6.06%	6.60%	7.15%	9.29%	10.91%	17.03%	23.78%
20	6.82%	6.36%	7.21%	9.29%	13.72%	20.39%	33.45%
25	6.58%	6.21%	7.11%	9.74%	16.21%	26.61%	43.83%
30	7.36%	6.18%	7.05%	9.86%	20.41%	34.24%	53.62%
35	6.92%	6.08%	7.35%	9.75%	24.05%	39.29%	62.49%
40	6.96%	6.18%	7.46%	10.12%	28.39%	46.38%	70.85%
45	6.70%	6.08%	6.99%	10.31%	30.57%	52.58%	77.79%
50	6.15%	5.99%	7.66%	9.64%	31.69%	57.58%	83.77%
Caso 3							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	10.95%	10.20%	10.10%	12.68%	12.68%	16.30%	17.90%
10	12.26%	11.32%	13.53%	21.90%	27.02%	38.84%	57.77%
15	12.46%	11.64%	15.89%	30.68%	48.43%	71.76%	89.06%
20	13.88%	11.31%	16.87%	36.58%	71.64%	90.37%	98.94%
25	14.12%	11.40%	17.94%	42.99%	87.60%	98.29%	99.96%
30	15.46%	11.54%	18.33%	47.00%	96.31%	99.85%	100.00%
35	15.01%	11.46%	19.91%	50.95%	99.00%	99.99%	100.00%
40	14.91%	11.34%	19.95%	55.03%	99.85%	100.00%	100.00%
45	14.52%	11.34%	19.53%	58.03%	99.98%	100.00%	100.00%
50	13.72%	11.36%	21.82%	58.56%	99.88%	100.00%	100.00%
Caso 6							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	22.53%	21.16%	21.17%	28.10%	28.10%	38.66%	43.25%
10	37.45%	31.82%	46.49%	78.39%	86.18%	95.70%	99.67%
15	45.91%	38.48%	63.97%	96.39%	99.80%	100.00%	100.00%
20	54.60%	43.55%	75.32%	99.45%	100.00%	100.00%	100.00%
25	60.61%	47.96%	83.27%	99.92%	100.00%	100.00%	100.00%
30	66.84%	52.46%	88.09%	99.99%	100.00%	100.00%	100.00%
35	70.24%	56.05%	92.29%	100.00%	100.00%	100.00%	100.00%
40	73.43%	58.87%	94.56%	100.00%	100.00%	100.00%	100.00%
45	75.69%	61.86%	95.93%	100.00%	100.00%	100.00%	100.00%
50	76.86%	64.11%	97.48%	100.00%	100.00%	100.00%	100.00%
Caso 8							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	38.42%	36.87%	36.90%	47.54%	47.54%	62.98%	69.37%
10	71.97%	64.17%	85.16%	98.98%	99.66%	99.97%	100.00%
15	87.09%	80.16%	97.43%	99.99%	100.00%	100.00%	100.00%
20	95.05%	89.62%	99.64%	100.00%	100.00%	100.00%	100.00%
25	98.15%	94.74%	99.94%	100.00%	100.00%	100.00%	100.00%
30	99.32%	97.34%	99.99%	100.00%	100.00%	100.00%	100.00%
35	99.76%	98.78%	100.00%	100.00%	100.00%	100.00%	100.00%
40	99.91%	99.45%	100.00%	100.00%	100.00%	100.00%	100.00%
45	99.96%	99.74%	100.00%	100.00%	100.00%	100.00%	100.00%
50	99.99%	99.88%	100.00%	100.00%	100.00%	100.00%	100.00%

Tabla 5. Consistencia de las pruebas con distribución de Laplace y $\alpha=1\%$

Caso 1							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	2.11%	2.11%	2.22%	2.22%	2.22%	2.22%	2.23%
10	0.35%	2.81%	2.12%	2.73%	3.03%	2.46%	5.78%
15	1.83%	2.18%	2.17%	3.18%	4.04%	6.14%	10.42%
20	1.70%	2.08%	1.89%	3.06%	5.64%	8.34%	18.02%
25	1.65%	1.99%	2.04%	3.53%	7.54%	13.14%	27.10%
30	1.48%	1.86%	1.88%	3.39%	10.42%	20.24%	37.48%
35	1.64%	1.72%	2.07%	3.28%	13.45%	25.15%	47.70%
40	1.76%	1.77%	1.96%	3.50%	17.16%	32.20%	56.44%
45	1.82%	1.74%	1.80%	3.63%	19.62%	39.85%	66.16%
50	1.81%	1.67%	2.19%	3.12%	19.04%	44.86%	74.37%
Caso 3							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	5.41%	5.41%	5.75%	5.75%	5.75%	5.75%	5.84%
10	1.64%	6.53%	6.27%	12.10%	15.13%	17.12%	34.35%
15	6.31%	5.33%	7.46%	19.45%	33.72%	52.81%	71.72%
20	5.96%	5.46%	7.79%	25.01%	59.07%	78.98%	94.20%
25	5.89%	5.31%	8.70%	32.19%	79.66%	94.79%	99.42%
30	5.56%	5.21%	8.73%	35.78%	92.82%	99.32%	99.95%
35	5.97%	5.01%	10.02%	40.27%	97.84%	99.91%	100.00%
40	6.23%	4.87%	9.93%	44.70%	99.50%	100.00%	100.00%
45	6.44%	4.94%	9.72%	48.91%	99.91%	100.00%	100.00%
50	6.41%	4.74%	11.32%	49.16%	99.51%	100.00%	100.00%
Caso 6							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	17.72%	17.72%	18.60%	18.60%	18.60%	18.60%	18.78%
10	14.65%	26.64%	34.59%	62.46%	70.30%	77.93%	93.16%
15	38.46%	32.43%	52.76%	90.82%	98.27%	99.86%	99.98%
20	45.71%	38.78%	66.52%	98.21%	99.98%	100.00%	100.00%
25	52.29%	43.94%	77.07%	99.75%	100.00%	100.00%	100.00%
30	56.25%	48.32%	83.37%	99.96%	100.00%	100.00%	100.00%
35	63.01%	52.04%	89.53%	99.99%	100.00%	100.00%	100.00%
40	68.01%	55.58%	92.45%	100.00%	100.00%	100.00%	100.00%
45	72.03%	59.10%	94.60%	100.00%	100.00%	100.00%	100.00%
50	75.07%	62.04%	96.73%	100.00%	100.00%	100.00%	100.00%
Caso 8							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	33.92%	33.92%	35.23%	35.23%	35.23%	35.23%	35.26%
10	44.54%	58.59%	73.44%	94.11%	96.51%	98.42%	99.82%
15	82.01%	77.11%	93.47%	99.94%	100.00%	100.00%	100.00%
20	92.12%	88.09%	98.83%	100.00%	100.00%	100.00%	100.00%
25	96.90%	94.27%	99.81%	100.00%	100.00%	100.00%	100.00%
30	98.60%	97.23%	99.98%	100.00%	100.00%	100.00%	100.00%
35	99.60%	98.65%	99.99%	100.00%	100.00%	100.00%	100.00%
40	99.88%	99.48%	100.00%	100.00%	100.00%	100.00%	100.00%
45	99.96%	99.77%	100.00%	100.00%	100.00%	100.00%	100.00%
50	99.99%	99.91%	100.00%	100.00%	100.00%	100.00%	100.00%