Aplicación del modelo Weibull en el análisis de eventos críticos en precios bursátiles

Juan de la Cruz Mejía Téllez*

Fecha de recepción: 29 de junio de 2012
Fecha de aceptación: 18 de diciembre de 2012

* Universidad Autónoma Metropolitana, Unidad Azcapotzalco.
Departamento de Sistemas.
jmt@nechikali.azc.uam.mx

Resumen

Es bien sabido que la distribución de probabilidad Weibull encuentra aplicación en la modelación de procesos o fenómenos que involucren riesgo, por ejemplo sismos, falla en la operación de un equipo o sistema, tiempo de vida, etc. Para inversionistas financieros es importante contar, para la toma de decisiones, con un método o metodología que estime cuándo será más probable que ocurra un evento crítico, como pudiera ser la caída o alza considerable en el precio de cotización de una determinada acción bursátil. En este trabajo se aplica el modelo Weibull en la distribución del tiempo al cual ocurren dos eventos críticos consecutivos. Caso de análisis: acción bursátil CEMEX.

Clasificación JEL: C22, C46 y C58

Palabras clave: distribución Weibull, ajuste de funciones, series de tiempo financieras, altas y bajas en precio.

Weibull Model Application for the Analysis of Critical Events in Stock Prices

Abstract

It is well known that the Weibull probability distribution finds application in phenomenon or processes modeling involving risk, such as earthquakes, break down of equipment or system, life-time, etc. Financial investors wish to count on methods or methodologies to support decisions making, estimating when a critical event is more likely to occur, such as downs or ups of a certain share price. In this paper we present the Weibull model application for the occurrence time random variable between two consecutive critical events. Study case: CEMEX equity shares.

JEL classification: C22, C46 y C58

Key words: *Weibull distribution, function adjustment, financial time series, market price movements.*

1. Introducción

Tradicionalmente la distribución de probabilidad Weibull se ha empleado en la teoría de confiabilidad para modelar fenómenos de falla y degradación (Kalbfleish,(1980)), encontrando aplicación en el mantenimiento preventivo de elementos, ensambles, equipos, sistemas, etc. (Lyonnet, (1991). A la variable aleatoria involucrada se le ha denominado de muy diversas formas como por ejemplo, tiempo de vida, tiempo de servicio, tiempo de falla, entre otras. Conceptualmente la variable aleatoria se entiende como el tiempo o espacio que ocurre entre dos eventos considerados caóticos o críticos en la operación o funcionamiento de un equipo o sistema. El analista o administrador del equipo o sistema desearía estimar la fecha de ocurrencia del próximo evento para tomar la decisión más adecuada antes de que se presente el evento, como pudiera ser darle servicio de mantenimiento, remplazarlo, e incluso puede dar pauta para cancelar el servicio que brinda tal equipo.

La modelación de datos financieros en las operaciones de transacción es un tópico en desarrollo en la econometría. Se ha creado una nueva literatura comúnmente referida como "la econometría de las finanzas de alta (ultra) frecuencia" Bauwens-Hautsch (2007). Las variables de interés más comunes en este tipo de operaciones son el tiempo entre dos transacciones consecutivas (trade- duration) y el tiempo entre dos cotizaciones consecutivas (quote–duration). El análisis de datos financieros con alta frecuencia teniendo la característica de que el espaciamiento entre transacciones consecutivas es irregular, fue tratado originalmente por Engle and Rusell (1998), quienes introdujeron el modelo de Duración Condicional Autoregresivo (ACD). Muchos otros investigadores han seguido la pauta marcada por Engle and Rusell desarrollando modelos en el campo de la econometría financiera: Zang *et al,* (2001), Pathmanathan D. *et al,* (2009) por citar algunos de éstos. Peiris M.S. *et al,* (2007), hacen una revisión de las publicaciones realizadas sobre series de tiempo financieras con el enfoque ACD. En todos estos trabajos se ha aplicado la distribución de probabilidad Weibull como una herramienta de modelación exitosa.

Cabe mencionar que en ninguno de los trabajos existentes en la literatura se ha aplicado la distribución Weibull para la modelación del tiempo en que ocurren dos eventos consecutivos considerados críticos en la baja del precio de cotización de una acción bursátil, o bien de alza en el precio, que es el caso de estudio del presente trabajo.

El objetivo principal de esta aplicación es el de poder contribuir con un elemento más, que consideramos relevante, en la toma de decisiones.

2.- Metodología.

Caracterización del modelo Weibull.

Función de densidad de probabilidad

donde α, β, son los parámetros de forma y escala, respectivamente.

Función de distribución acumulada

$$F(t, \alpha, \beta) = \int_0^t f(x)dx \ = 1 - e^{-(\frac{t}{\beta})^{\alpha}}$$

que es la probabilidad de que un componente no viva más allá del tiempo t.

Valor esperado:

$$E[T] = \beta \, \Gamma(\frac{1}{\alpha} + 1)$$

Varianza:

$$\sigma^2 = \beta^2 (\Gamma(\frac{2}{\alpha} + 1) - \Gamma^2(\frac{1}{\alpha} + 1))$$

donde Γ(α) es la función gama definida: $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$

con propiedades:

$$\Gamma(\alpha+1) = \alpha\, \Gamma(\alpha)$$

$$\Gamma(\alpha) = (\alpha-1)! \quad \text{si } \alpha \text{ es entero positivo}$$

Función de confiabilidad:

$$R(t) = \int_t^{\infty} f(x)dx = 1 - F(t) = e^{-(\frac{t}{\beta})^{\alpha}}$$

Función de riesgo o también llamada tasa de falla instantánea, obtenida por el cociente f(t)/R(t) y denotada por z(t):

$$z(t) = \frac{\alpha}{\beta}\left(\frac{t}{\beta}\right)^{\alpha-1}$$

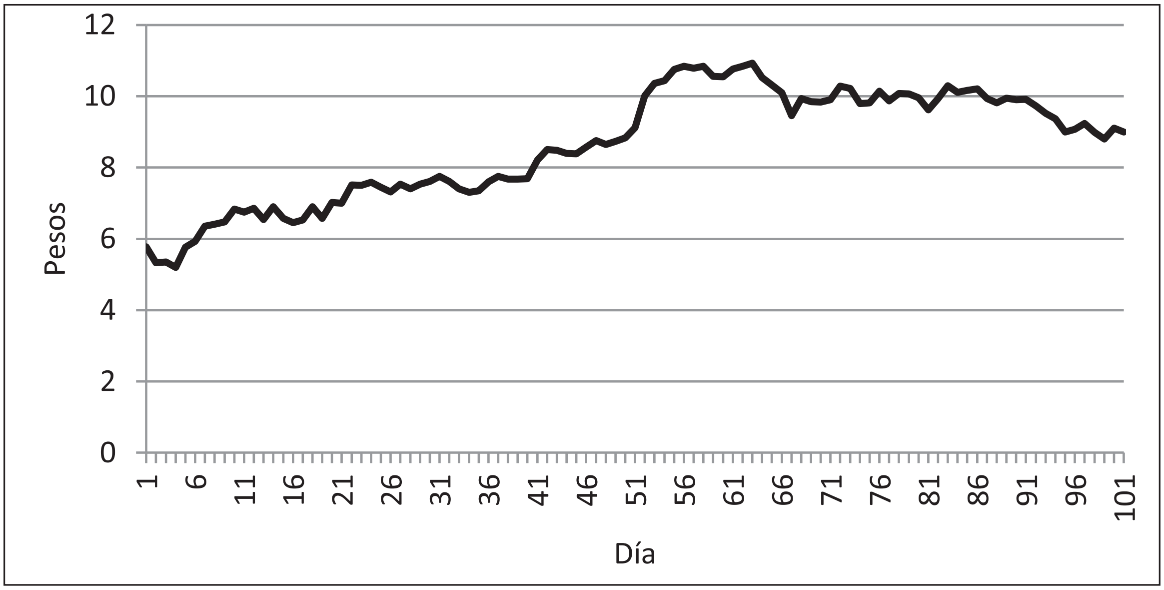
Para la identificación de los puntos críticos, se tomará en consideración el siguiente procedimiento:

Se evalúa dinámicamente la pendiente de la recta que une a dos puntos consecutivos en la serie de tiempo. Si la pendiente cambia de negativa a positiva a este punto se le denominará crítico, en particular punto mínimo, caso contrario punto máximo.

3. Análisis de la serie de tiempo

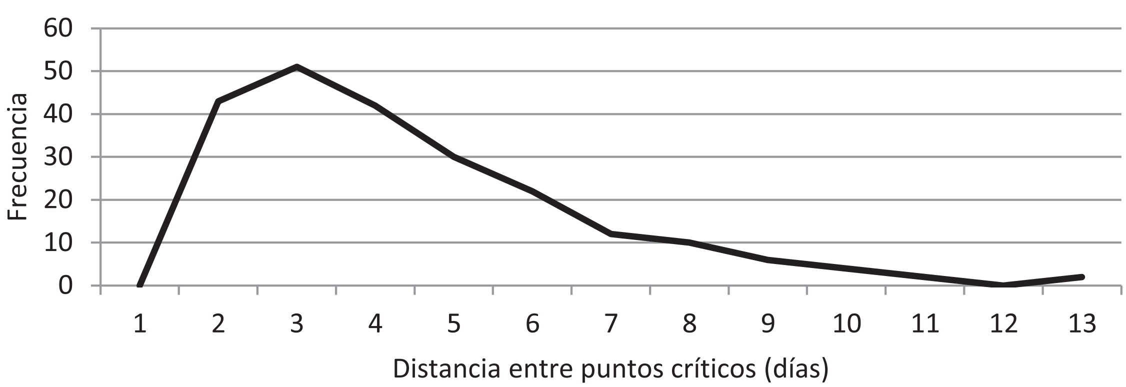
El objetivo en esta aplicación será el de modelar los tiempos a los cuales ocurre una baja en el precio (punto crítico) de la acción bursátil "CEMEX". Se tomó una muestra de 1000 datos históricos de esta serie de tiempo. En la Gráfica 1 se representan 100 datos recientes:

Gráfica 1. Cemex: Precios al cierre.



Identificados los puntos críticos, se determinan los tiempos de ocurrencia entre éstos, los cuales representarán los valores de la variable aleatoria. Para nuestro caso de estudio alcanzó un rango igual a 13 días. La Gráfica 2 ilustra el polígono de frecuencias.

Gráfica 2. Polígono de frecuencias.



En esta gráfica podemos observar que presumiblemente la variable aleatoria sigue la forma típica de la distribución Weibull, la cual será la hipótesis a considerar, contra la hipótesis alternativa de que no lo sea.

Considerando la función de confiabilidad:

$$R(t) = e^{-(\frac{t}{\beta})^{\alpha}}$$

tomando logaritmo natural

$$\ln R(t) = -\left(\frac{t}{\beta}\right)^{\alpha}$$

$$-\ln R(t) = \left(\frac{t}{\beta}\right)^{\alpha}$$

$$\ln(-\ln R(t)) = \alpha \ln\left(\frac{t}{\beta}\right) = \alpha \ln t - \alpha \ln \beta$$

obteniendo finalmente:

$$\ln \left(\ln\frac{1}{R(t)}\right) = \alpha \ln t - \alpha \ln \beta$$

Sean $x = \ln t$, $m = \alpha$, $b = \alpha \ln \beta$, $y = \ln \left(\ln\frac{1}{R(t)}\right)$,

resulta la forma: $y = mx - b$

que es la ecuación de una línea recta. Así entonces, si los datos corresponden al modelo Weibull, los puntos caerán sobre una línea recta, cuya pendiente y ordenada al origen, determinarán los valores de los parámetros α y β.

Cuadro 1. Valores de respuesta y de variable independiente para estimar la línea de regresión.

Días		
(t)	ln (t)	$\ln(\ln\frac{1}{R(t)})$
2	0.693147181	-1.525081273
3	1.098612289	-0.600447766
4	1.386294361	-0.063191602
5	1.609437912	0.317033912
6	1.791759469	0.605725609
7	1.945910149	0.805603419
8	2.079441542	1.021386719
9	2.197224577	1.169032176
10	2.302585093	1.393718246
11	2.397895273	1.552434286
12	2.48490665	1.552434286

Estimando la línea de regresión por el método de Mínimos Cuadrados se obtuvo:

$$y = 1.695397x - 2.51433$$

Pendiente	m=1.695397
Ordenada al origen	b= -2.51433

por tanto:

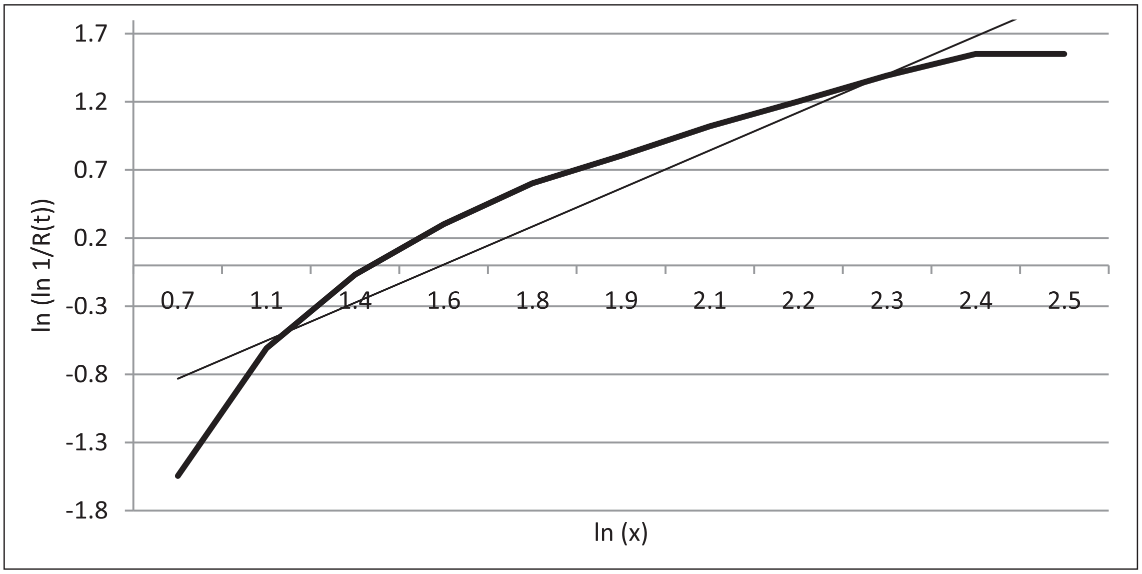
$$\alpha = m$$

$$\beta = e^{(-b/\alpha)}$$

$$\alpha = 1.6954$$

β= 4.4063

Gráfica 3. Contraste de datos ajustados con línea de regresión.



En lo siguiente se realizará la prueba de bondad de ajuste Ji-cuadrada para justificar o no que la variable aleatoria del tiempo entre ocurrencias de precios a la baja adopta el modelo Weibull.

$$X^2 = \sum_{i=1}^{k} \frac{(ei - fi)^2}{ei}$$

donde *ei* es el valor de la frecuencia teórica para el *i*-ésimo valor de tiempo:

ei= [F(t)-F(t-1)] * Frecuencia total

siendo F(t)-F(t-1) la probabilidad de que la variable aleatoria tome el valor t.
Procedemos a calcular los valores de la función de distribución acumulada F(t):

$$F(t) = 1 - e^{-(\frac{t}{\beta})^{\alpha}}$$

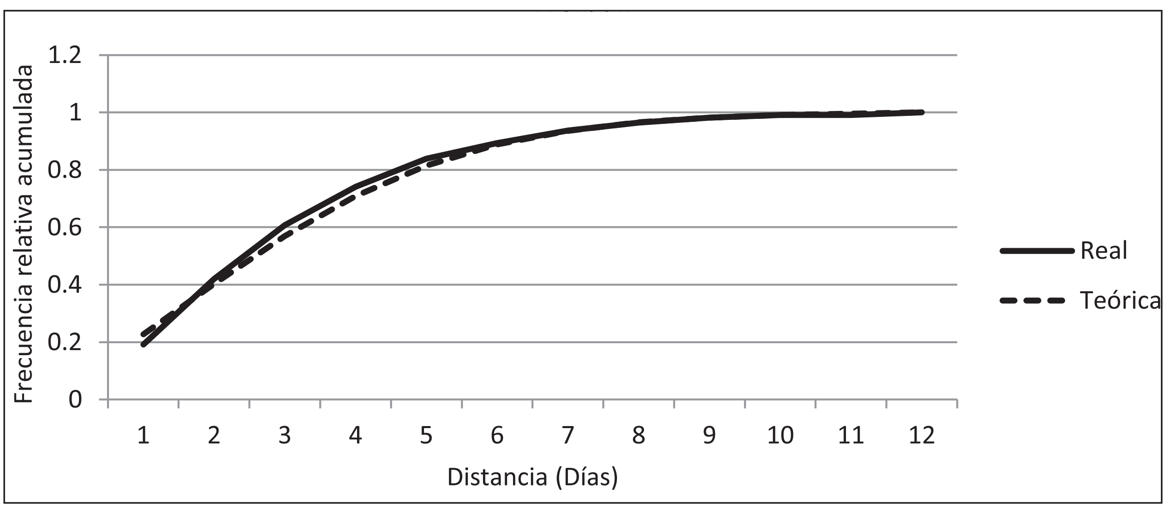
Cuadro 2. Función de distribución de probabilidad Weibull para

α= 1.6954, β = 4.4063

Días T	F(t)
≤2	0.230537823
3	0.406154401
4	0.572043716
5	0.710327337
6	0.81507067
7	0.888293914
8	0.935992863
9	0.965135974
≥10	1

En la Gráfica 4 se muestra el contraste entre la distribución de frecuencias relativas acumuladas con la función de distribución de probabilidad Weibull:

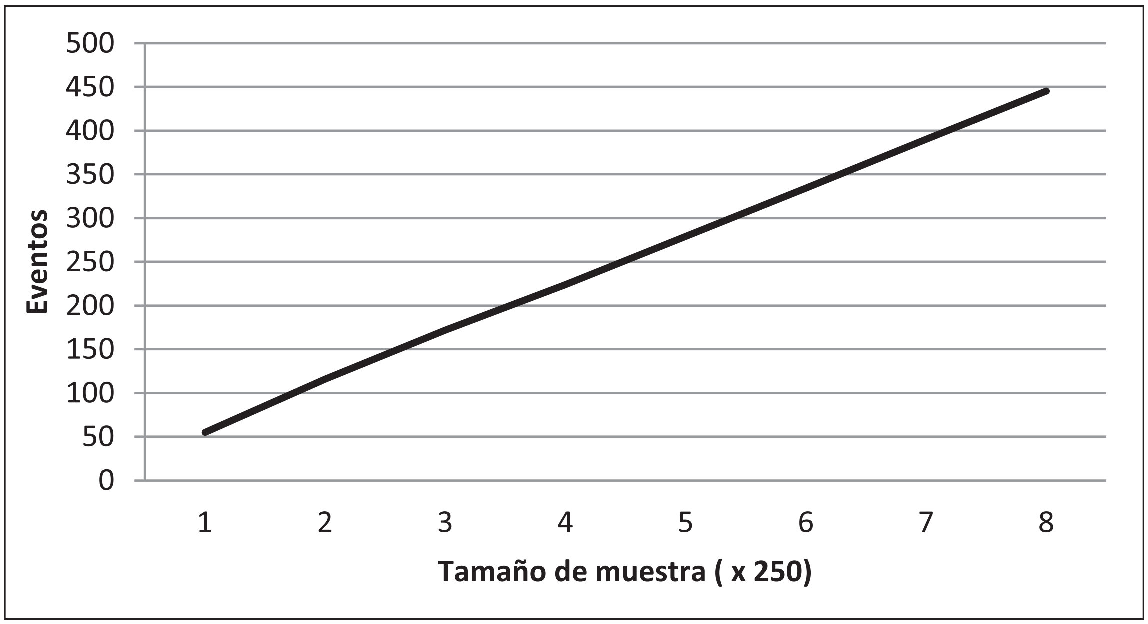
Gráfica 4. Contraste entre distribuciones acumuladas real con teórica Weibull.



Se observa en esta última figura que ambas curvas siguen la misma trayectoria, es decir, tienen la misma forma. Sin embargo, presentan una sensible discrepancia.

La variable Ji-cuadrada alcanza el valor de 7.2045. Refiriéndonos a valores tabulados para esta variable con 8 grados de libertad y nivel de significancia del 1%, se tiene el valor de 20.09. Con esto se interpreta que el valor calculado para la variable Ji-cuadrada es estadísticamente significativo como para no rechazar la hipótesis de que ocurra el modelo Weibull (E[T]= 4.43 días, σ^2 =5.77).

Gráfica 5. Número de eventos vs tamaño de muestra.



4. *Análisis de sensibilidad*

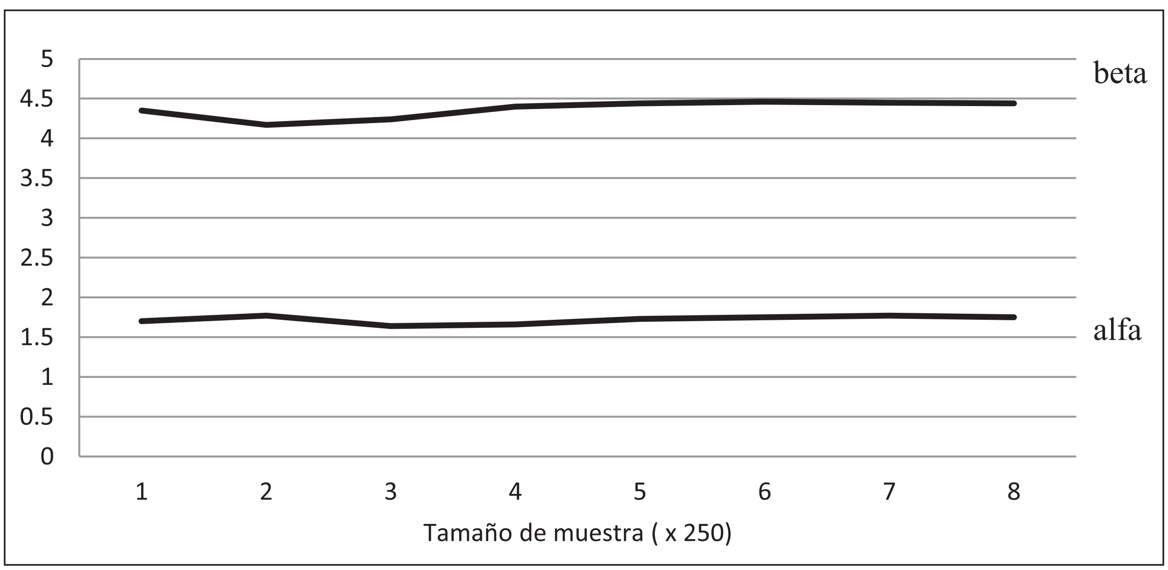
Con el propósito de observar los cambios en valor de los parámetros de escala y de forma, así como también en los dos primeros momentos de la distribución Weibull a medida que el tamaño de muestra (n) aumenta, se experimentó con varios niveles, 250, 500, 750, 1000 1250, 1500, 1750 y 2000, cubriendo un periodo de 1 a 8 años (aproximadamente) de información histórica. Se obtuvieron los siguientes resultados:

Cuadro 3. Resultados en el análisis de sensibilidad.

n	Eventos	α	β	Media	Varianza
250	55	1.70	4.35	4.38	5.59
500	116	1.77	4.17	4.21	4.77
750	172	1.64	4.24	4.29	5.70
1000	224	1.66	4.40	4.44	6.02
1250	279	1.73	4.44	4.46	5.64
1500	334	1.75	4.46	4.47	5.55
1750	390	1.77	4.45	4.46	5.42
2000	445	1.75	4.44	4.45	5.45

En la Gráfica 6 se observa que el número de eventos (mínimos locales) son directamente proporcionales al tamaño de muestra.

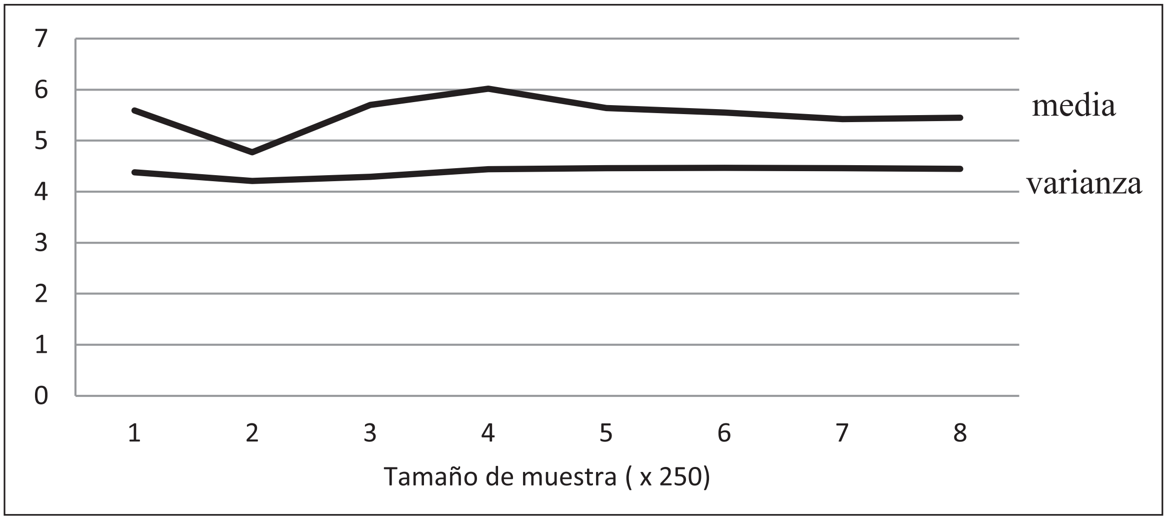
Gráfica 6. Valor de parámetros α y β vs tamaño de muestra.



Para un tamaño de muestra mayor a 1000 los parámetros de forma y de escala presentan estabilidad.

Al observar el comportamiento de media y varianza (Gráfica 7), se presume que la distribución se torna estacionaria para muestras de tamaño mayor a 1500 elementos.

Gráfica 7. Valores de media y varianza vs tamaño de muestra.



Se aplicó esta metodología de análisis a otras series de tiempo de emisoras bursátiles tales como América Móvil, Grupo Financiero Banorte, Grupo México, Walmart, obteniendo resultados exitosos, muy similares a los aquí mostrados.

5. Problema de aplicación

Un inversionista que posee una cantidad considerable de títulos de la empresa CEMEX. Desea contar con un esquema que le brinde de alguna manera información cuantitativa adicional para considerar si debe conservar los títulos o tomar la decisión de vender.

Solución:

Aprovechando el análisis elaborado en las secciones anteriores para la acción bursátil CEMEX, habiendo determinado que la distribución Weibull modela adecuadamente los eventos tipificados aquí como puntos críticos (baja, alza en precio), y teniendo en cuenta que el inversionista tiene títulos en su poder, será conveniente elaborar un cuadro informativo sobre las probabilidades de que ocurra baja en precio precisamente el día de mañana.

Los eventos a considerar para el cálculo de sus probabilidades son denominados como eventos condicionados, ya que conociendo el hecho de que han transcurrido d unidades de tiempo (días) desde que ocurrió la baja en precio, ésta se presente al día siguiente.

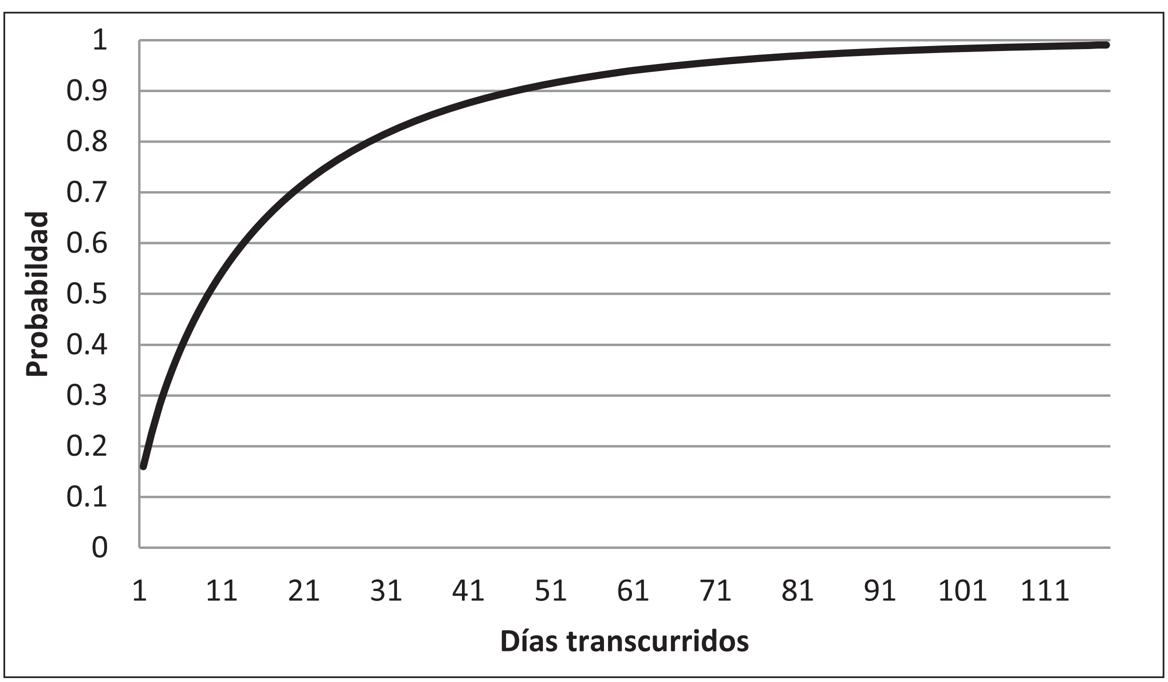
$$P\left(\frac{t < d+1}{t > d}\right) = \frac{F(d+1) - F(d)}{R(d)}$$

sustituyendo valores para la distribución Weibull:

$$P\left(\frac{t \ < d+1}{t > d}\right) = 1 - e^{-\left(\frac{d+1}{\beta}\right)^{\alpha} + \left(\frac{d}{\beta}\right)^{\alpha}}$$

En la siguiente Gráfica se muestran esquemáticamente estas probabilidades condicionales, para α = 1.75, β = 4.44,

Gráfica 8. Probabilidades condicionales vs días transcurridos desde última baja en precio.



Por ejemplo, tomando el caso de que han transcurrido 10 días desde que se presentó la baja en precio, la probabilidad de que al día siguiente ocurra este evento será de 53% aproximadamente.

Es de esperarse que a mayor número de días transcurridos sin que se haya presentado la baja en precio, las probabilidades de que este evento ocurra al día siguiente son cada vez mayores.

6. Conclusiones

En este trabajo de investigación se ha aplicado la función de probabilidad Weibull, considerada en la literatura como una herramienta para la modelación de procesos que involucren riesgo. Trasladando la idea de "riesgo" al caso de pérdida cuando se tienen títulos financieros en propiedad y que se coticen a un precio menor al adquirido. Resulta valioso para el administrador financiero contar con un modelo que le brinde una medida más de apoyo en la toma de decisiones. Se encontró como resultado de esta aplicación que el modelo Weibull es robusto cuando el número de datos informativos sobre el valor de la acción financiera es grande (mayor a 1500). El enfoque aquí propuesto es válido en el análisis de alza en precio del título financiero.

Bibliografía

Bauwens, Luc. and Hautch Nicolas (2007). "Modelling Financial High Frecuency Data Using Point Processes". *Humboldt-Univesität Berlin*. SFB 649 discussion paper No. 2007, 066, Hdl.handle.net/104 19/25238.

Engle, R.F. and Rusell, J.R. (1998). "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data". *Econometric,* 66, pp. 1127-1162.

Lyonnet P. (2001). *Maintenance Planning Methods and Mathematics*. Chapman & Hall.

Kalbfleisch, J.D. and Prentice, R.L. (1980). *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, Inc.

Pathmanathan, D., Ng, K.H., Peiris, M.S. (2009). "On Estimation of Autoregressive Conditional Duration (ACD) Models Based on Different Error Distributions". *Sri Lankan Journal of Applied Statistics*, Vol. 10, pp.251-269.

Peiris, M.S., Ng, K.H. and Ibrahim, M. (2007). "A Review of Recent Development of Financial Time Series: ACD Modelling Using the Estimating Function Approach". *Sri Lankan Journal of Applied Statistics*, 8: pp. 1-17.

Zhang, M.Y., Rusell, J.R. and Tsay, R.S. (2001). "A Nonlinear Autoregressive Conditional Duration Model With Applications to Financial Duration Data". *Journal of Econometrics*, 104, pp. 179-207.