REPASO DE ESTADÍSTICA

I. Conceptos de Probabilidad

A. Variables Discretas

1. En el mundo existen n estados posibles (eventos), en alguna fecha futura.

Ejemplo: un evento es el estado de la tecnología de IBM de aquí a un año.

2. La probabilidad de que el estado i se materializará es p_i, i=1,....,n.

Ejemplo: p_i = probabilidad que IBM tenga la tecnología i

- i=1: la misma tecnología, p_1 = $1/2$
- i=2: mejoramientos de la tecnología actual, p_2 = $1/4$
- i=3: un gran salto tecnológico, p_3 = $1/4$

3. Las probabilidades satisfacen las condiciones:

   a. $p_i \geq 0$ para todo i

   b. $\sum_{i=1}^{n} p_i = 1$

Un evento que es cierto corresponde a p_i = 1; un evento imposible, a p_i = 0.

4. Una variable aleatoria X es aquella que toma un valor dado x_i en cada evento i.

   Ejemplo: X = retorno proveniente de poseer una acción de IBM por un año, a partir de hoy.

5. X tomará el valor x_i cuando ocurre el evento i, con probabilidad: $P(X = x_i) = p_i$

   Ejemplo: si IBM no realiza un descubrimiento tecnológico (evento i=1), su retorno durante el próximo año será x_1 = -20%; en el evento i=2, IBM realiza modestos mejoramientos a su tecnología actual y el retorno es x_2 = +5%; en el evento i=3, IBM descubre una nueva tecnología y el retorno es x_3 = +25%.

6. Los siguientes momentos se utilizan para describir la distribución de la variable aleatoria X:

   a. **Valor esperado o esperanza**:

   $$M_1 = E[X] = \sum_{i=1}^{n} p_i x_i$$

   - Caracteriza el valor promedio de X

   b. **Varianza**:

   $$M_2 = V[X] = E\left[(X - E[X])^2\right] = \sum_{i=1}^{n} p_i \left(x_i - E[X]\right)^2$$

   - Caracteriza la dispersión de X alrededor de su valor promedio

   c. Desviación estándar: $\sigma(X) = \sqrt{V[X]}$

   - Lo mismo que la varianza, excepto que se mide en las mismas unidades que X

   d. Asimetría (*skewness*):

   - Tercer momento en torno a la media:

   $$M_3 = E\left[(X - E[X])^3\right] = \sum_{i=1}^{n} p_i (x_i - E[X])^3$$

   - Asimetría:

   $$A[X] = M_3 / (M_2)^{3/2}$$

   - Mide si la dispersión proviene de muchas desviaciones de la media positivas y pequeñas, y de unas pocas desviaciones negativas y grandes, o viceversa.

- Por ejemplo, caracteriza si las caídas en los precios del mercado accionario son más probables que sus recuperaciones (cola izquierda versus cola derecha) = asimetría negativa observada en los retornos accionarios.

e. Curtosis (*kurtosis*)

- Cuarto momento en torno a la media:

$$M_4 = E\left[(X - E[X])^4\right] = \sum_{i=1}^{n} p_i (x_i - E[X])^4$$

- Curtosis:

$$C[X] = M_4 / (M_2)^2$$

- Mide si la distribución es más apuntada, verticalmente, alrededor de la media, y si sus colas son más gruesas.

- Exceso de curtosis: $C[X] - 3$ (porque 3 es la curtosis de una distribución normal).

7. Si Y es otra variable aleatoria, que toma el valor y_i en el evento i, con probabilidad, $P(Y = y_i) = p_i$, entonces la relación entre las distribuciones de X e Y puede caracterizarse por:

a. Covarianza:

$$\begin{aligned} \operatorname{cov}(X, Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] - E[X]E[Y] \\ &= \sum_{i=1}^{n} p_i x_i y_i - \left(\sum_{i=1}^{n} p_i x_i\right)\left(\sum_{i=1}^{n} p_i y_i\right) \end{aligned}$$

- Caracteriza cuán cerca están las distribuciones de X e Y la una de la otra.

b. X e Y no están correlacionadas cuando $\operatorname{cov}(X, Y) = 0$. De lo contrario, se dice que están correlacionadas:

- positivamente correlacionadas cuando $\operatorname{cov}(X, Y) > 0$
- negativamente correlacionadas cuando $\operatorname{cov}(X, Y) < 0$

c. **Coeficiente de correlación**:

$$\rho(X,Y) = \mathrm{cor}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{V[X]V[Y]}}$$

(satisface que $-1 \leq \rho(X,Y) \leq +1$)

- El coeficiente de correlación mide el grado de asociación <u>lineal</u> entre dos variables aleatorias.

B. Variables continuas

1. Existe un continuo de estados posibles de la naturaleza (eventos), en una cierta fecha futura

   <u>Ejemplo</u>: El retorno accionario de IBM, durante el próximo año, puede tomar cualquier valor entre $-\infty$ y $+\infty$.

2. La función densidad f(x) de una variable aleatoria X es la probabilidad de que X tome un valor "cercano" a x.

   a. f(x)≥0 para todo x, con $\int_{-\infty}^{+\infty} f(x)dx = 1$

   b. Valor esperado o media: $E[X] = \int_{-\infty}^{+\infty} x f(x) dx$

   c. Varianza: $V[X] = E\left[(X - E[X])^2\right] = \int_{-\infty}^{+\infty} (x - E[X])^2 f(x) dx$

   d. Covarianza de dos variables aleatorias X e Y:

$$\mathrm{cov}(X,Y) = E[(X - E[X])(Y - E[Y])]$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - E[X])(y - E[Y]) f(x,y)dxdy$$

(f(x,y) es la densidad conjunta de X e Y)

C. Propiedades de los Momentos

- a, b, c, d denotan constantes, X, Y, Z variables aleatorias. Se tiene:

• $E[aX + b] = aE[X] + b$
• $E[X + Y] = E[X] + E[Y]$
• $E[aX + bY] = aE[X] + bE[Y]$
• $V[aX + b] = a^2 V[X]$

- $V[X+Y] = V[X] + V[Y] + 2\mathrm{cov}(X,Y)$
- $V[aX+bY] = a^2 V[X] + b^2 V[Y] + 2ab\mathrm{cov}(X,Y)$
- $\mathrm{cov}(X,X) = V[X]$
- $\mathrm{cov}(aX+b, cY+d) = ac\,\mathrm{cov}(X,Y)$
- $\mathrm{cov}(X, Y+Z) = \mathrm{cov}(X,Y) + \mathrm{cov}(X,Z)$

D. Ejemplo: La Distribución Normal o Gaussiana

1. La función densidad normal es:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < +\infty$$

2. Propiedades:

   a. La densidad es completamente caracterizada por sus dos primeros momentos. Se denota por $X \sim N(\mu, \sigma^2)$.

   b. $E[X] = \mu$ y $V[X] = \sigma^2$, Asimetría = 0, Curtosis = 3

   c. Si $X \sim N(\mu, \sigma^2)$ entonces $aX + b \sim N(a\mu + b, a^2\sigma^2)$.

   d. En particular, $\frac{X-\mu}{\sigma} \sim N(0,1)$ (normal estándar).

   e. Si $X \sim N(\mu, \sigma^2)$ y $Y \sim N(m, s^2)$, con $\mathrm{cor}(X,Y) = \rho$, entonces:

$$aX + bY \sim N(a\mu + bm, a^2\sigma^2 + b^2 s^2 + 2ab(\rho\sigma s))$$

3. Función distribución acumulada de una normal estándar $N(0,1)$:

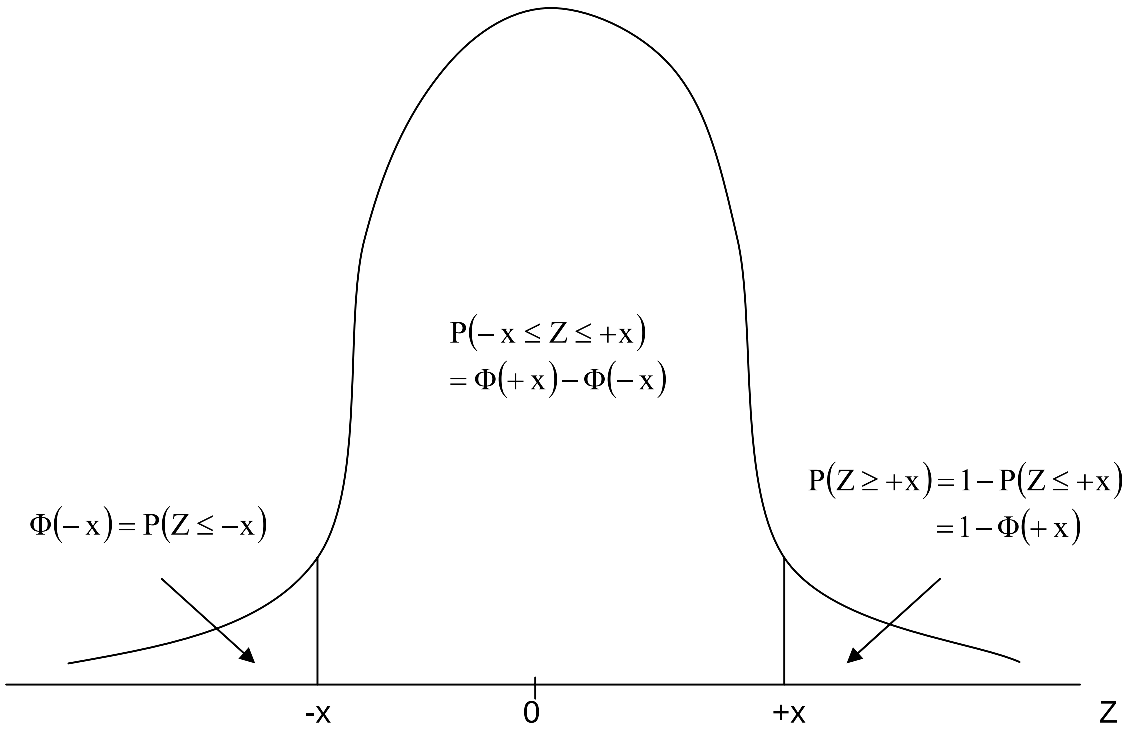
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{z^2}{2}\right\} dz, \quad -\infty < x < +\infty$$

(ver las tablas de la distribución normal estándar)

   a. Si $Z \sim N(0,1)$, entonces $\Phi(x)$ es la probabilidad de que Z sea menor que x, esto es: $\Phi(x) = P(Z \le x)$.

b. $\Phi(x) = P(Z \leq x)$ está dado por el área bajo la función densidad entre $-\infty$ y x.

c. Es útil para construir intervalos de confianza. Se tiene que $\Phi(-x)=1-\Phi(x)$:



d. Ejemplo: Suponga que el retorno mensual de la acción IBM, $R \sim N(\mu, \sigma^2)$ con μ = 1.5% y σ =2.7%. Un intervalo al 95% de confianza para R es un intervalo centrado en μ que contendrá el valor efectivo del retorno mensual de IBM el 95% de las veces:

- Construya $Z = \dfrac{R - \mu}{\sigma} \sim N(0,1)$. Sea α = 0.05 = 5%, 1-α = 0.95 = 95%. Deseamos encontrar un número $z_{1-\alpha/2}$ tal que:

  $P(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}) = 1 - \alpha$

  $\rightarrow$ Según las tablas de la distribución normal estándar, $z_{1-\alpha/2} = z_{0.975} = 1.96$

- Entonces:

$$
\begin{aligned}
95\% = 1-\alpha &= P(-1.96 \le Z \le 1.96) \\
&= P\left(-1.96 \le \frac{R-\mu}{\sigma} \le 1.96\right) \\
&= P(\mu - 1.96\sigma \le R \le \mu + 1.96\sigma)
\end{aligned}
$$

De esta manera, la probabilidad de que R esté entre μ - 1.96σ = -3.79% y μ + 1.96σ = 6.79%, es 95%. Por lo tanto, el intervalo de confianza es: [-3.79%, 6.79%].

II. Estadística de una Muestra

A. Media y Varianza de una Muestra

1. Ejemplo:

   a. Suponga que deseamos estimar el valor esperado μ y la varianza σ^2 del retorno mensual de IBM.
   b. Tenemos datos de los retornos accionarios de IBM para los últimos 24 meses: $R_1,...,R_T$, donde T=24.

      - Suponemos que los retornos son idénticamente distribuidos (pero no necesariamente normales), con valor esperado μ y varianza σ^2 comunes.
      - Esto significa que el retorno en el periodo t, con t=1, 2..., T, R_t, tiene la misma distribución, con media μ y varianza σ^2.

2. Para estimar el valor esperado y la varianza de los retornos de IBM, utilizamos:

   a. $\hat{\mu} = \overline{R} = \frac{1}{T}\sum_{t=1}^{T} R_t$: media de la muestra

   b. $\hat{\sigma}^2 = \frac{1}{T-1}\sum_{t=1}^{T}\left(R_t - \overline{R}\right)^2$: varianza de la muestra.

   c. Para estimar la desviación estándar σ, utilizamos $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$, la desviación estándar de la muestra.

   d. Estos estimadores son insesgados: $E[\hat{\mu}] = \mu$ y $E[\hat{\sigma}^2] = \sigma^2$

e. Para que este método funcione, tanto el valor esperado como la varianza de los retornos accionarios deben ser estables en un cierto periodo de tiempo.

B. Correlación de la Muestra

1. Supongamos que también tenemos datos de los retornos accionarios de Apple, $S_1,\ldots,S_T$, y deseamos estimar la correlación entre los movimientos de los precios accionarios de IBM y Apple.

2. Calcule primero la covarianza de la muestra:

$$\hat{c} = \frac{1}{T-1}\sum_{t=1}^{T}\left(R_t - \overline{R}\right)\left(S_t - \overline{S}\right)$$

3. Luego el coeficiente de correlación de la muestra es:

$$\hat{\rho} = \frac{\hat{c}}{\sqrt{\hat{\sigma}_{IBM}^{2}}\sqrt{\hat{\sigma}_{APL}^{2}}}$$

III. Análisis de Regresión

A. Especificación del Modelo

1. Ejemplo: El retorno R de la acción de IBM puede modelarse como $R_t = \alpha + \beta R_{M_t} + \varepsilon_t$ en cada periodo t = 1,...,T. R_M es el retorno del mercado, como un todo. (Este es el modelo CAPM).

2. Este modelo nos dice que, en cada periodo t, el retorno R_t de IBM puede explicarse (o predecirse) por el término lineal $\alpha + \beta R_{M_t}$.

a. El ruido ε_t es aleatorio (y no observable), y representa perturbaciones aleatorias a la relación entre el retorno del activo y el retorno del mercado.
b. α (el intercepto) y β (la pendiente) son desconocidos, pero son constantes o parámetros que pueden estimarse a partir de los datos históricos $R_1,\ldots, R_t$ y $R_{M_1},\ldots,R_{M_t}$.
c. El retorno R de IBM es la variable dependiente, el retorno del mercado R_M es la variable explicativa o regresor.

B. Estimación de los Parámetros

1. Para que el modelo esté bien especificado, el error no debe estar correlacionado con el regresor, y debe tener esperanza igual a cero.

2. En particular, la relación lineal postulada por el modelo se cumple, en forma exacta, sólo en promedio:

$$E[\varepsilon_t] = 0 \Rightarrow E[R_t] = \alpha + \beta E[R_{Mt}]$$

3. También:

$$\mathrm{cov}(R_t, R_{Mt}) = \mathrm{cov}(\alpha + \beta R_{Mt} + \varepsilon_t, R_{Mt}) = \beta \underbrace{\mathrm{cov}(R_{Mt}, R_{Mt})}_{=V[R_{Mt}]} + \underbrace{\mathrm{cov}(R_{Mt}, \varepsilon_t)}_{=0}.$$

Por lo tanto: $\beta = \dfrac{\mathrm{cov}(R_t, R_{Mt})}{V[R_{Mt}]}$ puede estimarse utilizando la varianza y covarianza muestrales, descritas anteriormente.

4. El valor estimado de α puede obtenerse de la siguiente relación:

$$\hat{\alpha} = \overline{R} - \hat{\beta}\,\overline{R}_M$$

5. $\hat{\beta}$ y $\hat{\alpha}$ obtenidos de esta manera se denominan estimadores de mínimos cuadrados ordinarios (OLS) de β y α.

IV. Test de Hipótesis

A. Hipótesis nula y estadígrafos

1. Ejemplo: Suponga que deseamos testear la hipótesis de que la media μ del retorno mensual de la acción IBM, $R \sim N(\mu, \sigma^2)$, es igual a μ_0 = 1.5% (nula $H_0 : \mu = \mu_0$ versus alternativa $H_1 : \mu \neq \mu_0$).

    a. Como antes, tenemos datos de los retornos accionarios de IBM para los T meses pasados: $R_1,...,R_T$.
    b. Tanto μ como σ son desconocidos.
    c. Por lo tanto, primero deben estimarse.

2. Recordemos que:

$$\hat{\mu} = \overline{R} = \frac{1}{T}\sum_{t=1}^{T} R_t \qquad \text{y} \qquad \hat{\sigma}^2 = \frac{1}{T-1}\sum_{t=1}^{T}\left(R_t - \overline{R}\right)^2$$

3. Bajo la hipótesis nula, el estadígrafo $\sqrt{T}\left(\hat{\mu}-\mu_0\right)/\hat{\sigma}$ sigue una distribución t-Student con T-1 grados de libertad:

$$\frac{\sqrt{T}\left(\hat{\mu}-\mu_0\right)}{\hat{\sigma}} \sim t_{[T-1]}$$

a. Pregunta: ¿por qué multiplicamos por $\sqrt{T}$?
b. Porque el error estándar del estimador $\hat{\mu}$ es $\hat{\sigma}/\sqrt{T}$, y el estadígrafo t se define como la razón entre estimador y su error estándar:

$$\frac{\hat{\mu}-\mu_0}{\hat{\sigma}/\sqrt{T}} = \frac{\sqrt{T}\left(\hat{\mu}-\mu_0\right)}{\hat{\sigma}}$$

c. La mayoría de los paquetes estadísticos imprimirá directamente el error estándar del estimador.

B. Test de Hipótesis

1. Intuición:

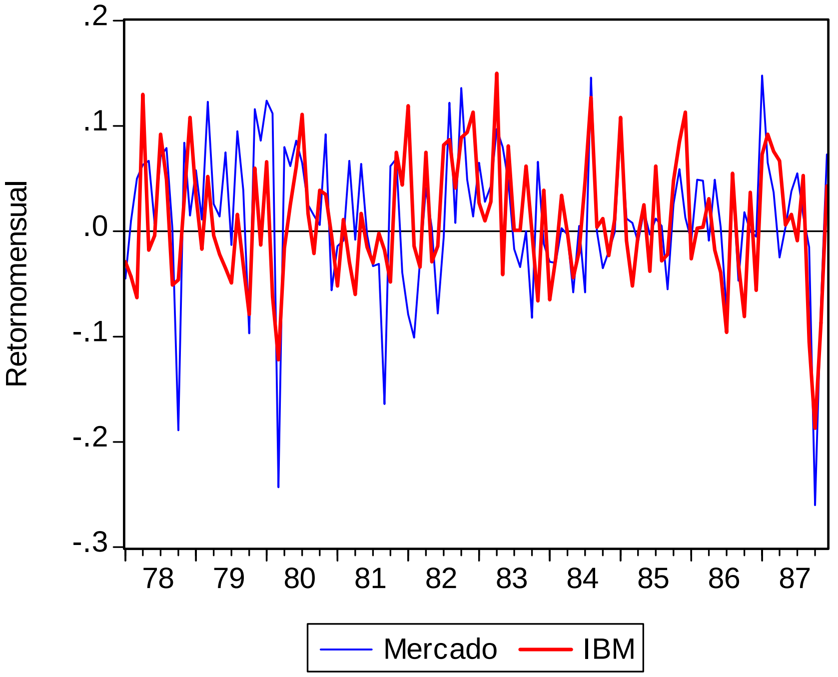
a. Si la nula es verdadera, $\hat{\mu}$ (el cual se aproxima a μ) debería estar en una vecindad cercana a μ_0.
b. Si realmente es así, entonces la nula será "aceptada" (= no rechazada). Si, en cambio, $\hat{\mu}$ se aleja de μ_0, la hipótesis nula será rechazada.

2. Significado estadístico de "cercano/lejano":

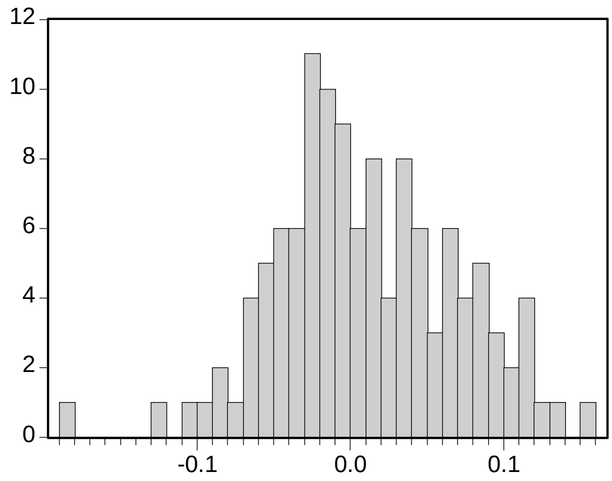
a. Si $\sqrt{T}\left(\hat{\mu}-\mu_0\right)/\hat{\sigma} > c_{[\alpha,T-1]} \vee \sqrt{T}\left(\hat{\mu}-\mu_0\right)/\hat{\sigma} < -c_{[\alpha,T-1]}$ entonces $\hat{\mu}$ y μ_0 están "lejos" el uno del otro, y la hipótesis nula es rechazada.
b. El valor crítico $c_{[\alpha,T-1]}$ se obtiene a partir de la tabla de la distribución t. α es el nivel de significancia del test, por ejemplo α = 0.05 para testear la hipótesis a un *nivel de confianza* (1-α) = 95%.
c. Para α = 0.05 y un T grande, $c_{[\alpha,T-1]} = 1.96$.

V Aplicaciones

El siguiente gráfico muestra los retornos mensuales de la acción de IBM y del portafolio de mercado para Estados Unidos, en el período enero 1978-diciembre 1987:



La serie de retornos de IBM puede describirse como sigue:



Serie: IBM	
Muestra 1978:01 1987:12	
Observaciones 120	
Media	0.0096
Mediana	0.0020
Máximo	0.150
Mínimo	-0.187
Dev. Std	0.059
Asimetría	-0.037
Curtosis	3.153
Jarque-Bera	0.144
Probabilidad	0.931

El gráfico anterior muestra un histograma de los retornos y algunos estadígrafos descriptivos. Por ejemplo, el máximo retorno observado en la muestra fue de 15% anual, mientras que el mínimo alcanzó a –18.7% anual (*crash* financiero de octubre de 1987). La asimetría y curtosis de los retornos se asemejan a los de una distribución normal, por lo cual no se rechaza la hipótesis de normalidad, según el test de Jarque-Bera[1]. Este se basa en el siguiente estadígrafo:

$$\boxed{JB = n\left[\frac{A^2}{6} + \frac{(C-3)^2}{24}\right]}$$

donde A = asimetría, C = curtosis y n= número de observaciones. La **hipótesis nula** es que los retornos se distribuyen **normal**. (Recordemos que para una distribución normal, A=0 y C=3). Cuando n es grande, el estadígrafo JB se distribuye χ^2 con 2 grados de libertad.

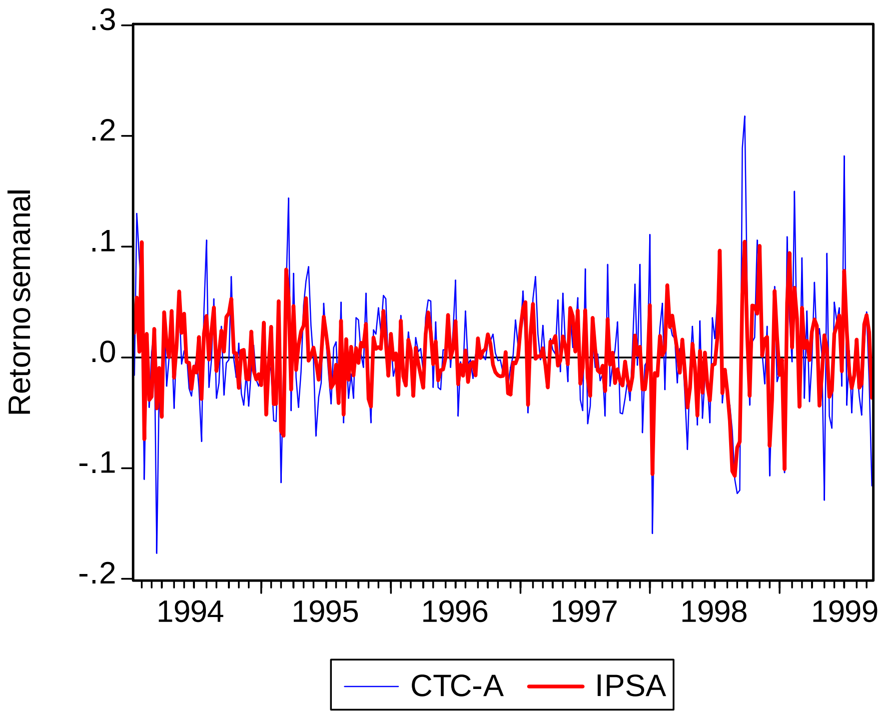
Por otra parte, se aprecia que la correlación entre el mercado e IBM es relativamente alta (0.52). Si estimamos una regresión lineal del retorno de IBM en una constante y el retorno del mercado, obtenemos lo siguiente:

Variable dependiente: IBM Muestra: 1978:01 1987:12				
Variable	Coeficiente	Error Std.	t-Statistic	Prob.
Constante	0.003	0.0047	0.697	0.4871
Mercado	0.453	0.0676	6.694	0.000
R^2	0.275			
R^2 ajustado	0.269			

Según lo anterior, el beta estimado de IBM es 0.45.

¿Qué hay de las empresas chilenas? Tomemos, por ejemplo, los retornos semanales de la acción CTC-A, para el período enero 1994-septiembre 1999:

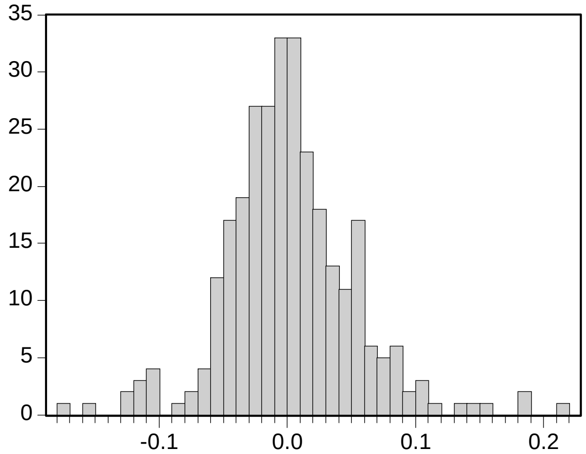
[1] El valor p (probabilidad), o nivel exacto de significancia, se define como el nivel más bajo de significancia al cual puede rechazarse la hipótesis nula. Como regla, se tiene que si el nivel de significancia escogido, α, es superior (inferior) al valor p, se rechaza (acepta) la hipótesis nula.



El gráfico anterior muestra los retornos para CTC-A y el IPSA (portafolio de mercado). La correlación entre ambas series es muy alta: 0.76. Ello indica que sólo el 24% de variación en el retorno de CTC-A corresponde a riesgo diversificable o específico a la empresa. La ecuación de CAPM para CTC-A viene dada por:

Variable dependiente: CTC-A				
Muestra: 1/14/1994-9/17/1999				
Variable	Coeficiente	Error Std.	t-Statistic	Prob.
Constante	-0.00048	0.0019	-0.250	0.803
IPSA	1.149	0.0568	20.205	0.000
R^2	0.581			
R^2 ajustado	0.579			

En este caso, la distribución de los retornos de CTC-A se aleja de la hipótesis de normalidad:



Serie: CTC-A	
Muestra: 1/14/1994 9/17/1999	
Observaciones: 297	
Media	0.0014
Mediana	-0.002
Máximo	0.218
Mínimo	-0.177
Std. Dev.	0.051
Asimetría	0.406
Curtosis	5.45
Jarque-Bera	82.78
Probabilidad	0.000