MI4040 Análisis Estadístico y Geoestadístico de datos
Profesor: Xavier Emery
Auxiliar: Fabián Soto F.

Auxiliar 1: Fundamentos de estadística

P1 Error típico $\frac{\sigma}{\sqrt{n}}$. Buscamos n tal que $\frac{\sigma}{\sqrt{n}} \leqslant 0{,}01$

$$\implies n \geqslant \frac{\sigma^2}{(0{,}01)^2} \underset{\substack{\uparrow \\ \text{Estimamos} \\ \sigma^2 \text{ con } s^2}}{\approx} \frac{s^2}{(0{,}01)^2} = 277$$

Así, se necesitan al menos 277 muestras para estimar la ley media del stock con un error típico menor que 0.01 %.

P3 Algunas propiedades de la exponencial y la distribución normal:

$$(1)\;\; e^{x+y} = e^x e^y \implies e^{\left(\sum_{i=1}^{n} X_i\right)} = \prod_{i=1}^{n} e^{X_i} \;,\; e^{xy} = (e^x)^y$$

Si X es una variable aleatoria con distribución normal con media μ y varianza σ^2 escribiremos

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

(2) La función densidad de probabilidad de X es:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

(3) Si $Y \sim \mathcal{N}(0,1) \implies T = \mu + \sigma Y$ con $T \sim \mathcal{N}(\mu, \sigma^2)$

(4) Sean $\{X_i\}_{i=1}^{n}$ variables aleatorias independientes e idénticamente distribuidas (i.i.d.), con distribución normal $\mathcal{N}(\mu, \sigma^2) \implies Y = \sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, n\sigma^2)$

. . . .

Media

$Y = e^x \overset{(3)}{=} e^{m+sT}$ con $T \sim \mathcal{N}(0,1)$

$$\implies \mathbb{E}(Y) = \mathbb{E}(e^{m+sT}) = \int_{-\infty}^{\infty} e^{m+sT} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \overset{(1)}{=} \int_{-\infty}^{\infty} \frac{e^m}{\sqrt{2\pi}} e^{st-\frac{t^2}{2}} dt = \frac{e^m}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} e^{st-\frac{t^2}{2}} dt}_{a}$$

Para resolver a hacemos un cambio de variable $u = t - s$

$$\implies u^2 - s^2 = (t-s)^2 - s^2 = t^2 - 2ts + s^2 - s^2 = t^2 - 2ts$$

Multiplicamos por $-1/2$

$$\implies -\frac{(u^2-s^2)}{2} = -\frac{t^2}{2} + ts = st - \frac{t^2}{2}$$

$$\implies \int_{-\infty}^{\infty} e^{st-\frac{t^2}{2}} dt = \int_{-\infty}^{\infty} e^{-\frac{(u^2-s^2)}{2}} du = \int_{-\infty}^{\infty} e^{\frac{-u^2}{2}} \cdot e^{\frac{s^2}{2}} du = e^{\frac{s^2}{2}} \int_{-\infty}^{\infty} e^{\frac{-u^2}{2}} du$$

$$\implies \mathbb{E}(Y) = \frac{e^m e^{\frac{s^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-u^2}{2}} du = \frac{e^{m+s^2}}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = e^{m+s^2}$$

Mediana

La mediana de X es m. La exponencial es una función creciente (conserva el orden de los valores), m es el valor central para $X \implies e^m$ es el valor central para T.

Esperanza de la media geométrica

$Y_i = e^{X_i}$, con $T_i \sim \mathcal{N}(0,1)$,

$$\implies \sqrt[n]{\prod_{i=1}^{n} Y_i} \overset{(1)}{=} \left(e^{\sum_{i=1}^{n} m+sT_i} \right)^{\frac{1}{n}} \overset{(1)}{=} e^{\frac{1}{n}\left(\sum_{i=1}^{n} m + \sum_{i=1}^{n} sT_i\right)} = e^{\frac{1}{n}(nm+s\sum_{i=1}^{n} T_i)} = e^m \cdot e^{\frac{s}{n}\sum_{i=1}^{n} T_i}$$

Pero por (4),

$$\sum_{i=1}^{n} T_i = Z \sim \mathcal{N}(0,n) \implies \mathbb{E}\left(\sqrt[n]{\prod_{i=1}^{n} Y_i}\right) = \mathbb{E}\left(e^m \cdot e^{\frac{s}{n}Z}\right) = \mathbb{E}\left(e^{m+\frac{s}{n}Z}\right)$$

$$m + \frac{s}{n}Z = \underbrace{m}_{\text{media}} + \underbrace{\frac{s}{n}\sqrt{n}}_{\text{dispersión}} U$$

Si $\sigma = \frac{s}{n}\sqrt{n} \implies \sigma^2 = s^2\frac{n}{n^2} = \frac{s^2}{n}$,

$$\implies m + \tfrac{s}{n}Z \iff V, \text{ con } V \sim \mathcal{N}(m, \tfrac{s^2}{n})$$

$$\mathbb{E}\left(\sqrt[n]{\prod_{i=1}^{n} Y_i}\right) = \mathbb{E}(e^V) \underset{\substack{\uparrow \\ \text{media de} \\ \text{lognormal}}}{=} e^{m+\frac{1}{2}\frac{s^2}{n}} = e^{m+\frac{s^2}{2n}}$$

....

Resumen

Media	$e^{-m+\frac{s^2}{2}}$
Mediana	e^m
Esperanza de la media geométrica	$e^{m+\frac{s^2}{2n}}$