

MI4040 Análisis Estadístico y Geoestadístico de Datos

Profesor: Xavier Emery

Profesor Auxiliar: Fabián Soto F.



Clase Auxiliar 1: Fundamentos de estadística

P1 Error típico $\frac{\sigma}{\sqrt{n}}$. Nosotros estamos buscando un n tal que $\frac{\sigma}{\sqrt{n}} \leq 0,01$

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

Por lo tanto, se necesitan al menos 278 muestras para estimar la ley media de la reserva con un error típico inferior al 0,01 % (277 también sería la respuesta correcta, 278 es más estricto).

P2 Algunas propiedades de la función exponencial y para la distribución gaussiana:

(1)

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

Si X es una variable aleatoria con distribución gaussiana con media μ y varianza σ^2 , es posible representar dicha v.a como:

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

(2) La función de densidad de probabilidad (probability density function (pdf)) de X es:

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

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

(4) Sea $\{X_i\}_{i=1}^n$ variables aleatorias independientes e idénticamente distribuidas (iid), con distribución gaussiana. De esta forma se tiene que:

$$\mathcal{N}(\mu, \sigma^2) \Rightarrow Y = \sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2)$$

Media

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

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

En orden de resolver a realizamos 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 $-1/2$

$$\begin{aligned} \implies -\frac{(u^2 - s^2)}{2} &= -\frac{t^2}{2} + ts = st - \frac{t^2}{2} \\ \implies \int_{-\infty}^{\infty} e^{st - \frac{t^2}{2}} dt &= \int_{-\infty}^{\infty} e^{-\frac{(u^2 - s^2)}{2}} du = \int_{-\infty}^{\infty} e^{\frac{-u^2}{2}} \cdot e^{\frac{s^2}{2}} du = e^{\frac{s^2}{2}} \int_{-\infty}^{\infty} e^{\frac{-u^2}{2}} du \\ \implies \mathbb{E}(Y) &= \frac{e^m e^{\frac{s^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-u^2}{2}} du = \frac{e^{m+s^2}}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = e^{m+s^2} \end{aligned}$$

Mediana

La mediana de X es m . La función exponencial está aumentando (se mantiene el orden de los valores), m es el valor medio para $X \implies e^m$, de esta forma resulta ser el valor medio para T .

Valor esperado de la media geométrica

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

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

Debido a (4),

$$\begin{aligned} \sum_{i=1}^n T_i = Z \sim \mathcal{N}(0, n) &\implies \mathbb{E} \left(\sqrt[n]{\prod_{i=1}^n Y_i} \right) = \mathbb{E} \left(e^m \cdot e^{\frac{s}{n} Z} \right) = \mathbb{E} \left(e^{m + \frac{s}{n} Z} \right) \\ m + \frac{s}{n} Z &= \underbrace{m}_{\text{media}} + \underbrace{\frac{s}{n} \sqrt{n}}_{\text{desviación}} U \end{aligned}$$

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

$$\begin{aligned} \implies m + \frac{s}{n} Z &\iff V, \text{ con } V \sim \mathcal{N}(m, \frac{s^2}{n}) \\ \mathbb{E} \left(\sqrt[n]{\prod_{i=1}^n Y_i} \right) &= \mathbb{E}(e^V) \stackrel{\substack{\uparrow \\ \text{media de} \\ \text{f. lognormal}}}{=} e^{m + \frac{1}{2} \frac{s^2}{n}} = e^{m + \frac{s^2}{2n}} \end{aligned}$$

Resumen

Media	$e^{-m + \frac{s^2}{2}}$
Mediana	e^m
Valor esperado de una media geométrica	$e^{m + \frac{s^2}{2n}}$