MI8130 Statistical and Geostatistical Data Analysis
Professor: Xavier Emery
Assistant professor: Fabián Soto F.

Complementary class 1: Fundamentals of statistics

P1 Typical error $\frac{\sigma}{\sqrt{n}}$. We're looking for n with $\frac{\sigma}{\sqrt{n}} \leqslant 0{,}01$

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

Thus, at least 278 samples are needed in order to estimate the mean grade of stock pile with a typical error less than 0.01 % (277 would also be right answer, 278 it's more strict) .

P2 Some properties of the exponential function and for Gaussian distribution:

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

Be X a random variable with Gaussian distribution with mean μ and variance σ^2 we will write:

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

(2) The probability density function (pdf) of X is:

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

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

(4) Let be $\{X_i\}_{i=1}^{n}$ independent and identically distributed random variables, with Gaussian distribution $\mathcal{N}(\mu, \sigma^2) \implies Y = \sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, n\sigma^2)$

. . . .

Mean

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

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

In order to solve a let's make a variable change $u = t - s$

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

Multiply by $-1/2$

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

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

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

Median

The median of X is m. The exponential function is increasing (the order of values is kept), m is the mid value for $X \implies e^m$ is the mid value for T.

Expected value of geometric mean

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

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

Due to (4),

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

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

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

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

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

Summary

Mean	$e^{-m+\frac{s^2}{2}}$
Median	e^m
Expected value of geometric mean	$e^{m+\frac{s^2}{2n}}$