Advanced Statistics - Tutorial 18/11/2021

Variance: X random variable with mean μ

$$\Rightarrow \mathrm{Var}[X] = E[\underbrace{(X-\mu)^2}_{\geq 0}] = E[X^2] - \mu^2$$

$$\Rightarrow \mathrm{Var}[X] \geq 0 \text{ and } \mathrm{Var}[X] = 0 \text{ iff } P(X=\mu) = 1$$

Normal distribution

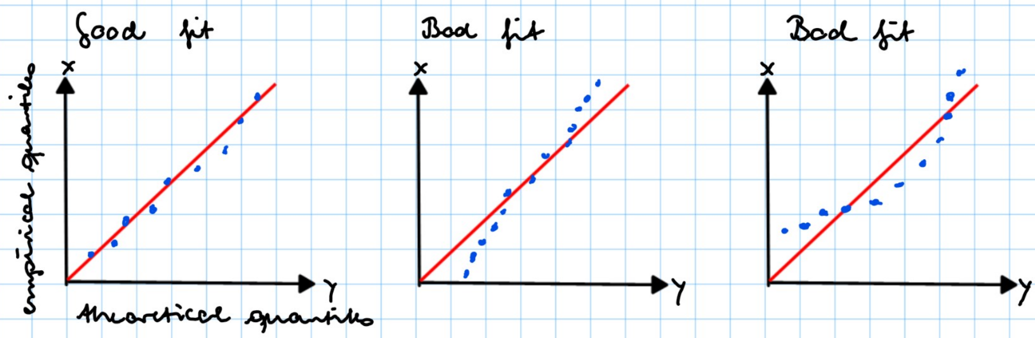
① Qualitative and quantitative tests for normal distribution

Q-Q plots

- F_t theoretical cdf (e.g. normal cdf)
- F_e empirical cdf - defined by

$$F_e(x) = \frac{\text{number of sample elements} \leq x}{\text{total number of sample elements}}$$

⇝ QQ plot shows $y = \underbrace{F_t^{-1}}_{\text{theoretical quantile}}(F_e(x))$ against x.



Skewness and kurtosis

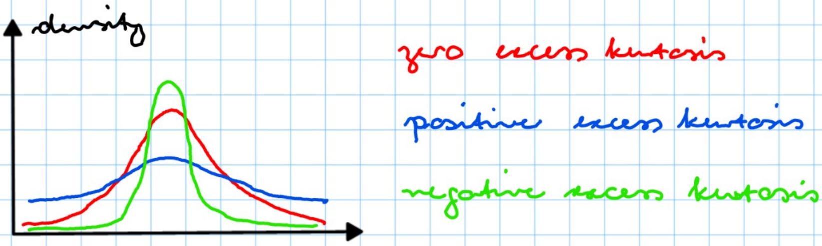
X random variable with mean μ and std σ

$\Rightarrow$ skewness $\quad E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$ - measure of asymmetry $(=0 \text{ if } X \sim \mathcal{N}(\mu,\sigma))$

kurtosis $\quad E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$ - measure of heaviness of tails $(=3 \text{ if } X \sim \mathcal{N}(\mu,\sigma))$

↓

excess kurtosis = kurtosis − 3



Kolmogorov - Smirnov test

Idea: Compare empirical and theoretical cdf

$\rightsquigarrow$ Calculate $\sup_x |F_e(x) - F_t(x)|$

② Central limit theorem

Let $n \in \mathbb{N}$ and $X_1, \dots, X_n$ be independent, identically distributed (i.i.d.) random variables with $\mu = E[X_i]$ and $\sigma = \sqrt{Var[X_i]}$

Let $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma}$.

Then $P\left(c \leq S_n \leq d\right) \xrightarrow[n \to \infty]{} \int_c^d \underbrace{\phi_{0,1}(x)}_{\text{density of } \mathcal{N}(0,1)} dx$