N 22.3.18 – 19

Tutorial # 10

\# of successes in n independent trials

$$X = \sum_{i=1}^{n} Y_i \ , \quad Y_i = \{\underset{\text{"failure"}}{0}, \underset{\text{"success"}}{1}\} \ , \ i = 1, \ldots, n$$

with prob. $q = 1 - p$, with prob. p

"a trial" (it's a random variable)

$$P(x) = \frac{n!}{x!\,(n-x)!}\, p^x \cdot \underbrace{q^{n-x}}_{=(1-p)^{n-x}} \longrightarrow \text{binomial distribution}$$

prob. density for X

(it involves only 1 parameter p, which deviation we are going to estimate)

For binomial distribution, $\mu = np$ – mean, $\sigma^2 = npq$ – variance

If n is large, can approximate binomial distribution by a normal distribution.

Formally, by the central limit th.: $\frac{1}{n}\sum_{i=1}^{n} Y_i \sim N(\frac{\mu}{n}, \sigma^2/n) \Rightarrow X = \sum_{i=1}^{n} Y_i \sim N(\mu, \sigma^2/n)$

Introduce (by shift and scaling) random variable $Z := \frac{X - \mu}{\sigma} = \frac{X - np}{(npq)^{1/2}}$.

Then, $Z \sim N(0, 1)$

normal distribution with zero mean and variance 1

This allows to measure deviation of X from its mean by measuring deviation of Z from 0.

Given a confidence level η, we can find a (e.g. from a table)

s.t. $\text{Prob}\{-a \le Z \le a\} = \eta$

$$= \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-u^2/2}\,du = \frac{2}{\sqrt{2\pi}} \int_0^a e^{-u^2/2}\,du = 2 \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-u^2/2} - 1$$

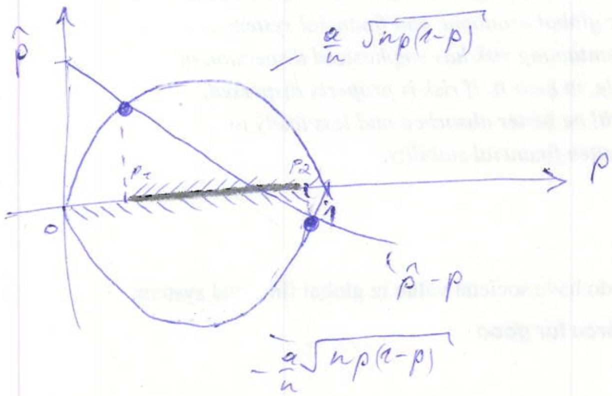
– tabulated

$$-a \le \underset{\text{"}\frac{X-np}{(np(1-p))^{1/2}}}{Z} \le a \iff -a\sqrt{np(1-p)} \le X - np \le a\sqrt{np(1-p)}$$

$$\Longleftrightarrow \quad -\frac{a}{n}\cdot\sqrt{np(1-p)} \le \frac{X}{n} - p \le \frac{a}{n}\cdot\sqrt{np(1-p)}$$

" $\hat{p}$ – can estimate from sample

$$\begin{cases} \hat{p} - p \le \frac{a}{n}\cdot\sqrt{np(1-p)} \\ \hat{p} - p \ge -\frac{a}{n}\sqrt{np(1-p)} \end{cases}$$



p_1, p_2 – points of intersections are roots of quadratic eq.:

$$\underbrace{(\hat{p} - p)^2}_{= \hat{p}^2 - 2\hat{p}p + p^2} = \frac{a^2}{n^2}\cdot n\cdot p(1-p)$$

$$\Rightarrow \quad \left(1 + \frac{a^2}{n}\right)p^2 - \left(2\hat{p} + \frac{a^2}{n}\right)p + \hat{p}^2 = 0$$

$$p_{1,2} = \frac{2\hat{p} + \frac{a^2}{n} \mp \sqrt{\cancel{4\hat{p}^2} + \frac{a^4}{n^2} + \frac{4\hat{p}a^2}{n} - \cancel{4\hat{p}^2} - \frac{4a^2\hat{p}^2}{n}}}{2\left(1 + a^2/n\right)}$$

$$= \frac{\hat{p} + \frac{a^2}{2n} \mp \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{a^2}{4n^2}}}{1 + a^2/n}$$

Hence, with given probab. γ p will be in $[p_1, p_2]$