Varians og normalfordeling - repetition

Rasmus Waagepetersen
Institut for Matematiske Fag Aalborg Universitet

February 20, 2022

Basal statistik - empirisk middelværdi og varians

På baggrund af stikprøve $X_1, X_2, \ldots, X_n$ kan vi udregne empirisk middelværdi og empirisk varians:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Til disse svarer teoretisk middelværdi og varians:

$$\mu = \mathbb{E}X \quad \text{og} \quad \sigma^2 = \mathbb{E}(X - \mu)^2$$

Empirisk middelværdi og varians vil tilnærme sig de teoretiske værdier når antal observationer (n) går mod uendelig.

Udregning af varians for diskret stokastisk variabel

Varians er vægtet sum af kvadrerede afvigelser mellem X's mulige værdier og X's middelværdi.

Vægtene er sandsynlighederne for de forskellige mulige værdier.

Regneregler

X og Y stokastiske variable.

$$\mathbb{E}[X+Y] = \mathbb{E}X + \mathbb{E}Y \quad \mathbb{E}aX = a\mathbb{E}X$$

$$\mathbb{V}\mathrm{ar}aX = a^2\mathbb{V}\mathrm{ar}X$$

Hvis X og Y uafhængige:

$$\mathbb{V}\mathrm{ar}[X+Y] = \mathbb{V}\mathrm{ar}X + \mathbb{V}\mathrm{ar}Y$$

Spredning:

$$\sigma = \sqrt{\mathbb{V}\mathrm{ar}X}$$

Binomialfordelingen og Poisson-fordelingen

Når $X \sim b(n, p)$ gælder

$$\mu = np$$

$$\sigma^2 = np(1-p)$$

Hvis X er Poisson-fordelt med middelværdi μ kan X antage alle værdier $x = 0, 1, 2, \dots$ og

$$P(X = x) = \exp(-\mu)\frac{\mu^x}{x!}$$

Varians for Poisson-fordelt X:

$$\mathbb{V}\mathrm{ar}X = \mathbb{E}X = \mu$$

Normalfordeling

Normalfordelingen er en model for målinger som kan antage alle talværdier.

Klokkeformet tæthedsfunktion.

Normalfordelingen specificeres vha. μ og σ^2: symmetrisk omkring μ

Spredning $\sigma = \sqrt{\sigma^2}$ karakteriserer "bredde" af fordelingen:
$\mu \pm 2\sigma$: 95.4%
$\mu \pm 3\sigma$: 99.7%