

TMA4245 – Parallel 1

Geir-Arne Fuglstad

27. januar

Kovarians og korrelasjon

Definisjon:

For to stokastiske variabler X og Y er **kovariansen** gitt ved

$$\begin{aligned}\text{Cov}[X, Y] &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] - E[X]E[Y]\end{aligned}$$

Definisjon:

For to stokastiske variabler X og Y er **korrelasjonen** gitt ved

$$\text{Corr}[X, Y] = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X] \cdot \text{Var}[Y]}}$$

NB: Merk at kovariansen mellom X og X er det vi kaller **varians**,

$$\text{Cov}[X, X] = \text{Var}[X]$$

Regneregler for forventning og varians

Vi har tre konstanter $a, b, c \in \mathbb{R}$ og to stokastiske variabler X og Y .

Da gjelder **alltid**:

1. $E[a + bX + cY] = a + bE[X] + cE[Y]$
2. $\text{Var}[a + bX + cY] = b^2\text{Var}[X] + c^2\text{Var}[Y] + 2bc\text{Cov}[X, Y]$

Hvis X og Y er **uavhengige** gjelder:

1. $E[XY] = E[X]E[Y]$
2. $\text{Cov}[X, Y] = 0$

Binomisk fordeling

Definisjon:

La n være et positivt heltall og la $0 \leq p \leq 1$. En binomisk fordelt stokastisk variabel oppstår når vi teller antall suksess i n uavhengige forsøk hver med suksess/fiasko og suksessannsynlighet p . Vi skriver f.eks. $X \sim \text{Binom}(n, p)$ for å spesifisere en slik stokastisk variabel.

Definisjon:

En binomialfordeling med $n = 1$ kalles en bernoullifordeling. Vi skriver f.eks. $X \sim \text{Bernoulli}(p)$ for å spesifisere en slik stokastisk variabel.

Egenskaper

Punktsannsynlighet:

$$f(x; n, p) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, \dots, n,$$

Finnes i Python (f.eks. `scipy.stats.binom.pmf`), og oppgitt i **formelsamlingen**.

Kumulativ fordeling:

$$F(x; n, p) = \sum_{i \leq x} \binom{n}{i} p^i (1 - p)^{n-i}, \quad x \in \mathbb{R},$$

har ingen forenkling, men er tilgjengelig i **formelsamlingen gjennom tabell** og i Python (f.eks. `scipy.stats.binom.cdf`).

Forventning og varians:

$$E[X] = np \quad \text{og} \quad \text{Var}[X] = np(1 - p).$$

Disse er oppgitt i **formelsamlingen**.

Simulering:

Flere muligheter (`numpy.random.binomial` og `scipy.stats.binom.rvs`) hvis oppgaven ikke ber deg om å implementere en egen funksjon. Da må du bruke generell algoritme eller Bernoulli forsøksrekke.

Multinomisk fordeling

Se introduksjonsvideo, lærebok+temasider og oppgaver i plenumsregning.

Multinomisk fordeling har noen egenskaper listet i **formelsamlingen**.