Introduktion til Sandsynlighedsteori og Statistik
FORMELSAMLING

Théodore Doubinsky og Gustav Krogh

October-December 2021

Contents

Forord

Formelsamlingen indeholder alle relevante formler fra undervisningen i Introduktion til sansynlighedsteori og statistik. Vi forbeholder os retten til, at det ikke er vores ansvar hvis formler, eksempler osv. indeholder fejl, som kunne lede til fejl i eksempelvis svar til eksamen. Dog vil vi pointere at formlerne er kopieret direkte fra tekstkilden på probabilitycourse.com, og disse inderholder derfor ikke fejl, medmindre hjemmesiden gør. Eksemplerne kan dog indeholde fejl, idet de er skrevet af fra bogen. Formelsamlingen er desuden rettet igennem, af folk som har meldt sig frivilligt.
I definitionsboxene, vil henvisningen til *Introduction to Probability, Statistics and Random Processes*, være i en parentes på formen: (s. 123). Hvis flere sider er brugt, vil de være listet i rækkefølgen svarende til placering i definitionsboxen. ((s. 123, 456, 789)).

Chapter 1

Diskrete Stokastiske Variable

1.1 Marginale Diskrete Stokastiske Variable

Definition 1.1.1: PMF af Stok. Var (Side 109 og 112)

Lad X være en diskret stokastisk variabel. PMF'en, $P_X(x)$ af X er defineret som følgende (s. 109):

$$P_X(x) = P(X = x)$$

PMF'en har følgende vigtige egenskaber (s. 112):

- $0 \leq P_X(x) \leq 1$
- $\sum_{x \in R_X} P_X(x) = 1 \qquad$ (Lov om total SS)
- Lad $A \subset R_X, \quad P(X \in A) = \sum_{x \in A} P_X(x)$

Leder du efter hvordan man får en Marginal PMF givet den Simultane, så se 1.2.1 (side 7)

Definition 1.1.2: CDF af Stok. Var (Side 135 og 137)

CDF af en stokastisk variabel X er givet ved (s. 135):

$$F_X(x) = P(X \leq x), \text{ for all } x \in \mathbb{R}.$$

En vigtig egenskab for CDF'en er (s. 137):

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

Definition 1.1.3: Diskret Middelværdi (Side 139 142 og 144)

Middelværdien af en stokastisk variabel X er givet ved (s. 139):

$$EX = \sum_{x_k \in R_X} x_k P(X = x_k) = \sum_{x_k \in R_X} x_k P_X(x_k)$$

Linearitet af middelværdi siger, at følgende gælder (uafhængighed er **ikke** nødvendig) (s. 142):

$$E[aX + b] = aEX + b$$

$$E[X_1 + X_2 + \cdots + X_n] = EX_1 + EX_2 + \cdots + EX_n$$

Tager man en funktion på en stokastisk variabel kan LOTUS bruges (s. 144):

$$E[g(X)] = \sum_{x_k \in R_X} g(x_k) P_X(x_k)$$

Definition 1.1.4: Diskret Varians (Side 146-148)

Variansen af en stokastisk variabel X er givet ved:

$$\mathrm{Var}(X) = E\left[(X - \mu_X)^2\right] = E\left[X^2\right] - \left[EX\right]^2$$

Følgende regneregler gælder for variansen:

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$$

Ved uahf:

$$\mathrm{Var}(X_1 + \ldots + X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n)$$

Standardafvigelsen, σ er defineret således:

$$\mathrm{SD}(X) = \sigma_X = \sqrt{\mathrm{Var}(X)}$$

Problemtype 1.1.5: Funktioner af diskrete stok. var (side 143)

Lad X være en disk. stokastisk variabel, og lad Y være en stok. var. hvor $Y = g(X)$. Find PMF af Y.
Step 1: Find R_Y.
R_Y er defineret ved:

$$R_Y = \{g(x) : x \in R_X\}$$

Step 2: Find $P_Y(y)$.
For alle $y \in R_Y$:

$$\begin{aligned} P_Y(y) &= P(Y = y) \\ &= P(g(X) = y) \\ &= \sum_{x:g(x)=y} P_X(x) \end{aligned}$$

Dette skal gentages for alle $y \in R_Y$.

Eksempel på et problem af type 1.1.5 (side 144)
Lad X være en stokastisk variabel med $R_X = \{-1, 0, 1, 2, 3\}$ med $f_X(x) = \frac{1}{5}$ for alle x. Lad nu $Y = 2|X|$. Find PMF af Y.
Først findes R_Y:

$$\begin{aligned} R_Y &= \{g(x) : x \in R_X\} \\ &= \{g(-1), g(0), g(1), g(2), g(3)\} = \{2, 0, 2, 4, 6\} \\ &= \{0, 2, 4, 6\} \end{aligned}$$

Nu findes $f_Y(y)$:

$$\begin{aligned} P_Y(0) &= P(Y = 0) = P(2|X| = 0) = \frac{1}{5} \\ P_Y(2) &= P(2|X| = 2) = P(X = 1 \vee X = 2) \\ &= P(X = 1) + P(X = -1) = \frac{2}{5} \\ P_Y(4) &= P(2|X| = 4) = P(X = 2) + P(X = -2) = \frac{1}{5} + 0 = \frac{1}{5} \\ P_Y(6) &= \text{Det samme som for } P_Y(4) \end{aligned}$$

1.2 Simultane Diskrete Stokastiske Variable

Definition 1.2.1: Simultan PMF af Stok. Var (Side 220-221)

Den simultane PMF for to stokastiske variable, X og Y er givet ved:

$$P_{XY}(x, y) = P(X = x, Y = y)$$

Ligesom for marginal PMF, gælder følgende for den simultane:

$$\sum_{(x_i, y_j) \in R_{XY}} P_{XY}(x_i, y_j) = 1$$

$$P\big((X, Y) \in A\big) = \sum_{(x_i, y_j) \in (A \cap R_{XY})} P_{XY}(x_i, y_j)$$

Endvidere, kan den **marginale PMF** findes ud fra den simultane således:

$$P_X(x) = \sum_{y_j \in R_Y} P_{XY}(x, y_j), \qquad \text{for any } x \in R_X$$
$$P_Y(y) = \sum_{x_i \in R_X} P_{XY}(x_i, y), \qquad \text{for any } y \in R_Y$$

Definition 1.2.2: CDF af Stok. Var (Side 223-234)

Den simultane CDF for to diskrete stok. var. er givet ved:

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

På samme måde som for PMF'en, kan de marginale CDF'er findes ud fra den simultane:

$$F_X(x) = F_{XY}(x, \infty) = \lim_{y \to \infty} F_{XY}(x, y), \qquad \text{for any } x,$$
$$F_Y(y) = F_{XY}(\infty, y) = \lim_{x \to \infty} F_{XY}(x, y), \qquad \text{for any } y$$

Definition 1.2.3: Diskret simultan LOTUS (Side 235)

Lad X, Y være to diskrete stokastiske variable. Lad $g(X, Y)$ være en funktion. Så gælder der:

$$E[g(x, y)] = \sum_{(x_i, y_j) \in R_{XY}} g(x_i, y_j) P_{XY}(x_i, y_j)$$

1.2.1 Betingede Diskrete Stok. Var.

Definition 1.2.4: Betinget PMF (Side 227-228)

Lad X, Y være stokastiske variable og A være en hændelse. PMF'en for $X|A$ er:

$$P_{X|A}(x_i) = P(X = x_i|A) = \frac{P(X = x_i \text{ og } A)}{P(A)}$$

PMF'en for $X|Y$ og pr. generalitet $Y|X$ er:

$$P_{X|Y}(x_i|y_j) = \frac{P_{XY}(x_i, y_j)}{P_Y(y_j)},$$
$$P_{Y|X}(y_j|x_i) = \frac{P_{XY}(x_i, y_j)}{P_X(x_i)}$$

For enhver $x_i \in R_X, y_j \in R_Y$.

Definition 1.2.5: Betinget middelværdi (Side 291)

Lad X, Y være diskrete stokastiske variable, og A være en hændelse, så vil:

$$E[X|A] = \sum_{x_i \in R_X} x_i P_{X|A}(x_i),$$
$$E[X|Y = y_j] = \sum_{x_i \in R_X} x_i P_{X|Y}(x_i|y_j)$$

Definition 1.2.6: Lov om total (...) ved betingelser (Side 233 og 244)

Loven om total sandsynlighed:

$$P(X \in A) = \sum_{y_j \in R_Y} P(X \in A|Y = y_j)P_Y(y_j), \quad \text{for any set } A.$$

Loven om total middelværdi:
Hvis $B_1, B_2, \ldots, B_n$ udfører en partition af S:

$$EX = \sum_i E[X|B_i]P(B_i)$$

For to stokastiske variable X, Y, hvoraf Y skal være diskret:

$$EX = \sum_{y_j \in R_Y} E[X|Y = y_j]P_Y(y_j)$$

Loven om total varians (s. 244)

$$\mathrm{Var}(X) = E[\mathrm{Var}(X|Y)] + \mathrm{Var}(E[X|Y])$$

Definition 1.2.7: Diskret Uafhængighed (Side 228)

To diskrete stokastiske variable X, Y er uafhængige hvis:

$$P_{XY}(x, y) = P_X(x)P_Y(y), \quad \text{for all } x, y$$

Ækvivalent:

$$F_{XY}(x, y) = F_X(x)F_Y(y), \quad \text{for all } x, y$$

Og generelt for n uafhængige diskrete stokastiske variable $X_1, X_2, \ldots, X_n$

$$P(X_1 = x_1, X_2 = x_2, ..., X_n) = P(X_1 = x_1)P(X_2 = x_2) \cdot ... \cdot P(X_n)$$
$$\text{for alle } x_1, x_2, ..., x_n$$

Definition 1.2.8: Middelværdier ved uafhængighed (Side 242)

Lad X, Y være uafhængige stokastiske variable. Følgende formler gælder:

$$\begin{aligned} E[X|Y] &= EX \\ E[g(X)|Y] &= E[g(X)] \\ E[XY] &= EXEY \\ E[g(X)h(Y)] &= E[g(X)]E[h(Y)] \end{aligned}$$

Chapter 2

Kontinuerte Stokastiske Variable

2.1 Marginale Kontinuerte Stokastiske Variable

Definition 2.1.1: PDF af stok. var. (Side 164-165)

En kontinuert stokastisk variabel, X med en kontinuert CDF, $F_X(x)$, har en PDF, $f_X(x)$:

$$f_X(x) = \frac{dF_X(x)}{dx} = F'_X(x)$$

For en PDF gælder følgende:

1. $f_X(x) \geq 0$ for alle $x \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f_X(u)du = 1$
3. $P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(u)du$
4. Mere generelt for enhver mængde A, $P(X \in A) = \int_A f_X(u)du$.

Definition 2.1.2: Kontinuert Middelværdi (Side 167-168)

Lad X være en kontinuert stok. var. $E[X]$ er givet ved:

$$EX = \int_{-\infty}^{\infty} x f_X(x)dx$$

Igen gælder LOTUS:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx$$

Definition 2.1.3: Kontinuert varians (Side 169)

For en kontinuert stokastisk variabel kan varians findes på følgende måde:

$$\begin{aligned}\mathrm{Var}(X) = E[(X-\mu_x)^2] &= \int_{-\infty}^{\infty}(x-\mu_X)^2 f_X(x)dx \\ = EX^2 - (EX)^2 &= \int_{-\infty}^{\infty} x^2 f_X(x)dx - \mu_X^2\end{aligned}$$

Problemtype 2.1.4: Transformationer

Lad X være en kontinuert stokastisk variabel, og lad endvidere $Y = g(X)$ (**$g(x)$ skal være strengt monoton og differentiabel for $x \in R_X$!**). Problemet er, at finde CDF og PDF.

Sætning 4.1* (Ugeseddel 6)

Lad X være en stokastisk variabel og $Y = g(X)$. Lad $R_Y = \{g(x) : x \in R_X\}$. For at bruge denne sætning er det vigtigt at erkende, at:

1. $g(x)$ er strengt aftagende i R_Y og dvs. at den er strengt monoton.
2. $g(X)$ er kontinuert og differentiabel i R_Y

Når dette er gjort kan man definere:

$$f_Y(y) = \begin{cases} |h'(y)|\, f_X(h(y)) & \text{for } y \in R_Y, \\ 0 & \text{ellers,}\end{cases}$$

Hvor $h(y)$ er den inverse funktion til $g(x)$. CDF kan findes ved at integrere.
Eksempel: (4.9 i bogen) Lad X have PDF givet ved:

$$f_X(x) = \begin{cases} 4x^3 & 0 < x \le 1 \\ 0 & \text{ellers}\end{cases}$$

Lad endvidere Y være givet ved $\frac{1}{X}$ Find $f_Y(y)$.
Man kalder $Y = g(X) = \frac{1}{X}$. Man ser, at ved at proppe $x \in (0,1]$ in i funktionen, får vi at $R_Y = [1,\infty)$ (Range af Y, definitionsmængde). Det anbefales at man forestiller sig grafen få at få dette resultat.

$$\begin{aligned} g(x) = y = \frac{1}{x} &\Rightarrow x = \frac{1}{y} = h(y) \\ h'(y) = (\frac{1}{y})' = -\frac{1}{y^2} &\Rightarrow |h'(y)| = \frac{1}{y^2}\end{aligned}$$

Altså for $y \in R_Y$:

$$\begin{aligned} f_Y(y) &= \frac{1}{y^2}\cdot 4(\frac{1}{y})^3 \\ &= \frac{4}{y^5}\end{aligned}$$

Og altså kan hele PDF'en skrives som:

$$f_Y(y) = \begin{cases} \frac{4}{y^5} & \text{y} \in R_y \\ 0 & \text{Ellers} \end{cases}$$

2.2 Simultane Kontinuerte Stokastiske Variable

Definition 2.2.1: Simultan PDF (Side 257-259)

To kontinuerte stokastiske variable kan have en simultan PDF, hvor der gælder at:

$$P\big((X,Y) \in A\big) = \iint\limits_A f_{XY}(x,y)dxdy \tag{2.1}$$

At finde bestemte sandsynligheder (f.eks. når y kun er veldefineret når y<x) kan være tricky, og her henvises til side 283-284 (Første solved problem)
Vil man finde de marginale PDF'er givet den simultane kan dette gøres ved følgende formler:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy, \qquad \text{for all } x,$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx, \qquad \text{for all } y.$$

Definition 2.2.2: Simultan kontinuert CDF (Side 263)

Lad X og Y være kontinuerte stokastiske variable. De to funktioner siges at have en simultan PDF noteret ved $f_{XY}(x,y)$. Hvis den simultane PDF er kontinuert (Hvilket den nok er med 99% sikkerhed), kan den simultane CDF opskrives således:

$$F_{XY}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(u,v)dudv$$

Definition 2.2.3: Simultan Middelværdi (Side 279)

Middelværdi kan ikke spytte et enkelt tal ud for to forskellige variable, og altså findes $E[X,Y]$ ikke. Hvis der står $E[XY]$ menes der altså $E[g(X,Y)]$, hvor $g(x,y) = x \cdot y$. Dette kan man til gengæld godt finde ved LOTUS:

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{XY}(x,y)dxdy$$

Definition 2.2.4: Betinget PDF og CDF (Side 266-269)

Lad X være en kontinuert stokastisk variabel. Lad A være det område, hvor $a < X < b$ ($a, b \in \mathbb{R} \cup \{-\infty, \infty\}$) så:

$$F_{X|A}(x) = \begin{cases} 1 & x > b \\ \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)} & a \leq x < b \\ 0 & x < a \end{cases}$$

$$f_{X|A}(x) = \begin{cases} \frac{f_X(x)}{P(A)} & a \leq x < b \\ 0 & \text{otherwise} \end{cases}$$

Lad nu Y være en kontinuert stokastisk variabel. Lad det være givet, at $Y = y$. Så:

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$
$$P(X \in A|Y = y) = \int_A f_{X|Y}(x|y)dx$$
$$F_{X|Y}(x|y) = P(X \leq x|Y = y) = \int_{-\infty}^{x} f_{X|Y}(x|y)dx$$

Definition 2.2.5: Betinget Middel. og var. (Side 267 og 271)

Lad ting være defineret som i boksen ovenfor. Så gælder:

$$E[X|A] = \int_{-\infty}^{\infty} x f_{X|A}(x)dx,$$
$$E[g(X)|A] = \int_{-\infty}^{\infty} g(x) f_{X|A}(x)dx,$$
$$\mathrm{Var}(X|A) = E[X^2|A] - (E[X|A])^2$$

Igen når $Y = y$ er givet

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx$$
$$E[g(X)|Y = y] = \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y)dx$$
$$Var(X|Y = y) = E[X^2|Y = y] - (E[X|Y = y])^2$$

Definition 2.2.6: Kontinuert uafhængighed (Side 272)

X, Y er uafhængige stokastiske variable hvis og kun hvis:

$$\begin{aligned} f_{XY}(x,y) &= f_X(x)f_Y(y), \qquad \text{for all } x, y. \\ F_{XY}(x,y) &= F_X(x)F_Y(y), \qquad \text{for all } x, y. \end{aligned}$$

Når X, Y er uafhængige så gælder at:

$$\begin{aligned} &E[XY] = EXEY, \\ &E[g(X)h(Y)] = E[g(X)]E[h(Y)]. \end{aligned}$$

Definition 2.2.7: Lov om total (...) ved betingelser (Side 275)

Loven om total sandsynlighed:

$$P(A) = \int_{-\infty}^{\infty} P(A|X=x)f_X(x) \; dx$$

Loven om total middelværdi:

$$\begin{aligned} E[Y] &= \int_{-\infty}^{\infty} E[Y|X=x]f_X(x) \; dx \\ &= E[E[Y|X]] \end{aligned}$$

Loven om total varians:

$$\mathrm{Var}(Y) = E[\mathrm{Var}(Y|X)] + \mathrm{Var}(E[Y|X])$$

2.2.1 Covarians og korrelation

Definition 2.2.8: Covarians (Side 291)

Lad X og Y være stokastiske variable. Deres kovarians er defineret ved:

$$\mathrm{Cov}(X,Y) = E\big[(X-EX)(Y-EY)\big] = E[XY] - (EX)(EY)$$

En positiv covarians betyder at der er en positiv korrelation, en negativ covarians betyder at der er en negativ korrelation, og en covarians på 0 betyder, at der er tale om uafhængighed.

Definition 2.2.9: Regneregler for Covarians (Side 292-293)

Covariansen har følgende egenskaber:

1. $\text{Cov}(X, X) = \text{Var}(X)$
2. X og Y uafhængige $\Leftrightarrow \text{Cov}(X, Y) = 0$
3. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
4. $\text{Cov}(aX, Y) = a\text{Cov}(X, Y)$
5. $\text{Cov}(X + c, Y) = \text{Cov}(X, Y)$
6. $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$
7. Helt generelt:

$$\text{Cov}\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \text{Cov}(X_i, Y_j)$$

Et fedt resultat fra disse regler er:

$$\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X, Y)$$

Definition 2.2.10: Korrelationskoefficient (Side 294-295)

Korrelationskoefficienten, ρ_{XY} (rho) er defineret ved:

$$\rho_{XY} = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var(X) Var(Y)}}} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Korrelationskoefficienten er altid mellem -1 og 1, så det hvis du ikke har det, har du nok lavet en fejl i dine udregninger.
Hvis $\rho_{XY} = 0$:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

Chapter 3

Fordelinger

3.1 Fordelingstyper

3.1.1 Diskrete stokastiske variable

Bernoulli

Definition 3.1.1: Bernoulli fordeling (s. 115)

En stokastisk variabel X er en *Bernoulli* stokastisk variabel med parameter p, skrevet $X \sim Bernoulli(p)$, hvis dens PMF og CDF er givet:

$$P_X(x) = \begin{cases} p & \text{for } x = 1 \\ 1 - p & \text{for } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 - p & \text{for } 0 \leq x < 1 \\ 1 & \text{for} 1 \leq x \end{cases}$$

hvor $0 < p < 1$.
Hvis $X \sim Bernoulli(p)$, så er $EX = p$ og $Var(X) = p(1 - p)$

Bernoulli fordeling er en enten-eller fordeling. Enten har man succes eller fiasko.

Geometrisk fordeling

Definition 3.1.2: Geometrisk fordeling (s. 116)

En stokastisk variabel X er en *Geometrisk* stokastisk variabel med parameter p, skrevet $X \sim Geometric(p)$, hvis dens PMF er givet:

$$P_X(k) = \begin{cases} p(1-p)^{k-1} & \text{for } k = 1, 2, 3, ... \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x) = 1 - (1-p)^{\lfloor x \rfloor} \quad \text{for } x \geq 0$$

hvor $0 < p < 1$, og $\lfloor x \rfloor$ betyder rund ned (floor function).
Hvis $X \sim Geometric(p)$, så er $EX = \frac{1}{p}$ og $Var(X) = \frac{1-p}{p^2}$

Geometrisk fordeling er sandsynligheden for succes efter k forsøg.

Binomial fordeling

Definition 3.1.3: Binomial fordeling (s. 118)

En stokastisk variabel X er en *Binomial* stokastisk variabel med paramterne n og p, skrevet $X \sim Binomial(n, p)$, hvis dens PMF er givet:

$$P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{for } k = 0, 1, 2, \cdots, n \\ 0 & \text{otherwise} \end{cases}$$

hvor $0 < p < 1$.
Hvis $X \sim Binomial(n, p)$, så er $EX = np$ og $Var(X) = np(1-p)$

Binomial fordelingen er sandsynligheden for k succeser efter n forsøg.

Pascal fordeling

Definition 3.1.4: Pascal fordeling (s. 121)

En stokastisk variabel X er en *Pascal* stokastisk variabel med parametre m og p, skrevet $X \sim Pascal(m, p)$, hvis dens PMF er givet:

$$P_X(k) = \begin{cases} \binom{k-1}{m-1} p^m (1-p)^{k-m} & \text{for } k = m, m+1, m+2, m+3, ... \\ 0 & \text{otherwise} \end{cases}$$

hvor $0 < p < 1$.
Hvis $X \sim Pascal(m, p)$, så er $EX = \frac{m}{p}$ og $Var(X) = \frac{m(1-p)}{p^2}$. Mere klart, hvis det er givet at du fx har slået 6 seksere og sandsynligheden for at slå en sekser er $1/2$. Da finder pascal "hvor mange slag du skal slå for at få 6 6'ere".

Pascal fordelingen er sandsynligheden for m succeser ved k forsøg.

Poisson fordeling

Definition 3.1.5: Poisson fordeling (s. 123)

En stokastisk variabel X er en *Poisson* stokastisk variabel med parameter λ, skrevet $X \sim Poisson(\lambda)$, hvis dens range $R_X = \{0, 1, 2, 3, \ldots\}$, og dens PMF givet:

$$P_X(k) = \begin{cases} \frac{e^{-\lambda}\lambda^k}{k!} & \text{for } k \in R_X \\ 0 & \text{otherwise} \end{cases}$$

Hvis $X \sim Poisson(\lambda)$, så er $EX = \lambda$ og $Var(X) = \lambda$

Poisson fordelingen er sandsynligheden for k succeser, i det faste tidsrum givet af λ (husk $\lambda = \frac{succes}{tidsenhed}$)
Bemærk!: λ er succeser per tidsenhed. Denne kan skaleres op eller ned. For eksempel $\lambda = \frac{3\ succeser}{1time}$ er ækvivalent med $\lambda = \frac{6\ succeser}{2timer}$.

3.1.2 Kontiunerte stokastiske variable

Uniform fordeling

Definition 3.1.6: Uniform fordeling (s. 180)

En kontinuert stokastisk variabel X kaldes *Uniform* fordelt over intervallet $[a, b]$, skrevet $X \sim Uniform(a, b)$, hvis dens PDF er givet:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & x < a \text{ or } x > b \end{cases}$$

Hvis $X \sim Uniform(a, b)$ så er $EX = \frac{a+b}{2}$ og $Var(X) = \frac{(b-a)^2}{12}$.

Kan ses som den kontinuerte version af Bernoulli fordeling. (For intuitionens skyld).
Uniform fordeling beskriver sandsynligheden for at få et tilfældigt valgt tal i intervallet a,b.

Eksponential fordeling

Definition 3.1.7: Eksponential fordeling (s. 180)

En kontinuert stokastisk variabel X er *eksponential* fordelt, med parameter $\lambda > 0$, skrevet $X \sim Exponential(\lambda)$, hvis dens PDF og CDF er givet:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x) = 1 - e^{-\lambda x}, \quad x > 0$$

Hvis $X \sim Exponential(\lambda)$ så er $EX = \frac{1}{\lambda}$ og $Var(X) = \frac{1}{\lambda^2}$

Kan ses som den kontinuerte version af den Geometriske fordeling. (For intuitionens skyld).
Eksponentiel beskriver ventetid til første succes, hvor λ er gennemsnitlig succes per interval.

Normal fordeling

Definition 3.1.8: Normal Fordeling (s. 187)

Hvis X er en stokastisk variabel, med middelværdi μ og varians σ^2 skrevet $X \sim N(\mu, \sigma^2)$, er dens PDF og CDF givet:
$\mathrm{f}_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$,
$F_X(x) = P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$,
hvor $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{u^2}{2}\right\} du.$
Normalfordelingen har $EX = \mu$ og $Var(X) = \sigma^2$

Standard normal fordelingen er et specialtilfælde af Normal fordelingen, hvor $\mu = 0$ og $\sigma = 1$. Standard normalfordelingen har range: alle reelle tal.

Gamma fordeling

Definition 3.1.9: Gamma fordeling

En kontinuert stokastisk variabel X er *gamma* fordelt, med parametrene $\alpha > 0$ og $\lambda > 0$, skrevet $X \sim Gamma(\alpha, \lambda)$, hvis dens PDF er givet:

$$f_X(x) = \begin{cases} \frac{{}^{\alpha} x^{\alpha-1} e^{-}}{\Gamma(\alpha)} & x > 0 \\ 0 \quad ellers & \end{cases}$$

Hvis $X \sim Gamma(\alpha, \lambda)$ så er $EX = \frac{\alpha}{\lambda}$ og $\frac{\alpha}{\lambda^2}$

$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \mathrm{d}x, \qquad \text{for } \alpha > 0.$

Gamma fordelingen beskriver sandsynligheden for α ulykker(succeser) per tidsenhed x. Her er λ gennemsnitlige ulykker per tidsenhed.

3.1.3 Chi i anden fordelingen

Definition 3.1.10: Chi i anden fordelingen(s 462)

Hvis $Z_1, Z_2, ..., Z_n$ er uafhængige standard normal stokastiske variable, er den stokastiske varibel Y defineret

$$Y = Z_1^2 + Z_2^2 + \cdots + Z_n^2 \tag{3.1}$$

chi i anden fordelt med n grader af frihed (degrees of freedom). skrevet

$$Y \sim \chi^2(n). \tag{3.2}$$

Egenskaber:

- Chi i anden fordeling er en special case af gamma fordelingen. Mere specifikt:

$$Y \sim Gamma\left(\frac{n}{2}, \frac{1}{2}\right). \tag{3.3}$$

  så,

$$f_Y(y) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}}, \quad \text{for } y > 0.$$

- $EY = n$, $Var(Y) = 2n$
- For hvilken som helst $p \in [0, 1]$ og $n \in \mathbb{N}$, definerer vi $\chi^2_{p,n}$ som den reele værdi for

$$P(Y > \chi^2_{p,n}) = p, \tag{3.4}$$

  hvor $Y \sim \chi^2(n)$.

3.1.4 T-fordelingen

Definition 3.1.11: T-fordelingen (side 464)

Lad $Z \sim N(0,1)$, og $Y \sim \chi^2(n)$, hvor $n \in \mathbb{N}$. Antag at Z og Y er uafhængige. Den stokastiske variabel T, defineret som

$$T = \frac{Z}{\sqrt{Y/n}} \tag{3.5}$$

, er $t - fordelt$ med n frihedsgrader skrevet

$$T \sim T(n). \tag{3.6}$$

Egenskaber:

- t-fordelingen har en klokke formet PDF centreret i 0, men dens PDF er mere spredt ud end den normale PDF
- $ET = 0$, for $n > 0$. Men ET er udefineret for $n = 1$.
- $Var(T) = \frac{n}{n-2}$, for $n > 2$. Men, $Var(T)$ er udefineret for $n = 1, 2$.
- Når n bliver stor, går t densiteten mod standard normal PDF. Mere formelt kan det skrives

$$T(n) \xrightarrow{d} N(0,1). \tag{3.7}$$

- For hvilket som helst $p \in [0,1]$, og $n \in \mathbb{N}$, definerer vi $t_{p,n}$ som den reele værdi for hvor

$$P(T > t_{p,n}) = p. \tag{3.8}$$

Siden t-fordelingen har en symmetrisk PDF, haves det

$$t_{1-p,n} = -t_{p,n}. \tag{3.9}$$

3.2 Regneregler for fordelinger

3.2.1 Bernoulli

Hvis $X \sim Bernoulli(p)$ så kan

$$P_X(x) = \begin{cases} p & \text{x=1} \\ p-1 & \text{x=0} \\ 0 & \text{ellers} \end{cases}$$

Skrives $P_X(x) = (1-p)^{1-x} \cdot p^x$, hvor $x \in \{0,1\}$

Binomial

(side 118, Lemma 3.1) Hvis $X_1, X_2, ..., X_n$ er uafhængige Bernoulli(p) stokastiske variable, så er den stokastiske variabel X, defineret $X = X_1 + X_2 + ... + X_n$, $Binomial(n,p)$ fordelt.

Poisson

(side 125) Hvis $X \sim Binomial(n,p)$ og n, antal forsøg, er meget stor, og p, sandsynligheden for succes, er meget lille, så er $X \sim Binomial(n,p) \approx X \sim Poisson(n \cdot p)$.

3.2.2 Eksponentiel

(side 183) Hvis X er eksonential med parameter $\lambda > 0$ så er X en "memoryless" stokastisk variabel:

$$P(X > x + a | X > a) = P(X > x), \qquad \text{for } a, x \geq 0$$

3.2.3 Normal

(side 189, Theorem 4.3) Hvis $X \sim N(\mu_X, \sigma_X^2)$ og $Y = aX + b$, hvor $a, b \in \mathbb{R}$, så $Y \sim N(\mu_Y, \sigma_Y^2)$, hvor

$$\mu_Y = a\mu_X + b, \sigma_Y^2 = a^2\sigma_X^2$$

(side 283 Theorem 5.2) Hvis $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$ og $Z \sim X + Y$, så er

$$Z \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

3.2.4 Normalfordeling (Chi^2)(side 462)

Hvis $Z_i \sim N(0,1) : i = 1, 2, ..., n$, og $Y = Z_1^2 + Z_2^2 + ... + Z_n^2$, så er $Y \sim \chi^2(n)$. Chi i anden fordelingen er et specialtilfælde af gammafordleingen: $Y \sim Gamma(\frac{n}{2}, \frac{1}{2})$, med middelværdi $EY = n$ og $Var(Y) = 2n$.

3.2.5 Chi i anden

Theorem 8.3. Let X_1, X_2, $\cdots$, X_n be i.i.d. $N(\mu, \sigma^2)$ random variables. Also, let S^2 be the sample variance for this random sample. Then, the random variable Y defined as

$$Y = \frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

has a chi-squared distribution with $n-1$ degrees of freedom, i.e., $Y \sim \chi^2(n-1)$. Moreover, $\overline{X}$ and S^2 are independent random variables.

3.2.6 T-fordelingen

Theorem 8.4. Let X_1, X_2, $\cdots$, X_n be i.i.d. $N(\mu, \sigma^2)$ random variables. Also, let S^2 be the sample variance for this random sample. Then, the random variable T defined as

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

has a t-distribution with $n-1$ degrees of freedom, i.e., $T \sim T(n-1)$.

Chapter 4

Statistik

4.1 Basal statistik

Definition 4.1.1: Definition: stikprøve

Samlingen af stokastiske variable $X_1, X_2, ..., X_n$ er en *random sample/stikprøve* af størrelse n, hvis:

- $X_1, X_2, ..., X_n$ er uafhængige stokastiske variable, og
- De har samme fordeling, altså: $F_{X_1}(x) = F_{X_2}(x) = ... = F_{X_n}(x)$, for alle $x \in \mathbb{R}$

Definition 4.1.2: Egenskaber for stikprøve gennemsnit (sample mean) (s. 426)

- $E\overline{X} = \mu$
- $\text{Var}(\overline{X}) = \frac{\sigma^2}{n}$
- Den svage store tals love

$$\lim_{n\to\infty} P(|\overline{X} - \mu| \geq \epsilon) = 0 \tag{4.1}$$

.

- Den centrale grænseværdi sætning: Den stokastiske variabel

$$Z_n = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + X_2 + ... + X_n - n\mu}{\sqrt{n}\sigma} \tag{4.2}$$

konvergerer i fordeling til en standard normal stokastisk variabel når n går mod uendelig, altså:

$$\lim_{n\to\infty} P(Z_n \leq x) = \Phi(x), \qquad \text{for all } x \in \mathbb{R} \tag{4.3}$$

Hvor $\Phi(x)$ er standard normal CDF'en.

Definition 4.1.3: Bias (side 429 box 1 og box 2)

Lad $\hat{\Theta} = h(X_1, X_2, \cdots, X_n)$ være en punkt estimator for θ. **Bias** af punktestimatoren $\hat{\Theta}$ er defineret ved:

$$B(\hat{\Theta}) = E[\hat{\Theta}] - \theta. \tag{4.4}$$

Lad $\hat{\Theta} = h(X_1, X_2, \cdots, X_n)$ være en punkt estimator for θ. Det siges at $\hat{\Theta}$ er en **unbiased** estimator for θ hvis:

$$B(\hat{\Theta}) = 0, \qquad \text{for all possible values of } \theta. \tag{4.5}$$

$\bar{X}$ er altid unbiased (s. 429 eksempel 8.2)

Definition 4.1.4: MSE (mean squared error) (side 430 og side 431 box 1)

Mean squared error (MSE) af en punkt estimator $\hat{\Theta}$, skrevet $MSE(\hat{\Theta})$, er defineret som:

$$MSE(\hat{\Theta}) = E\big[(\hat{\Theta} - \theta)^2\big]. \tag{4.6}$$

Hvis $\hat{\Theta}$ er en punkt estimator for θ,

$$MSE(\hat{\Theta}) = \mathrm{Var}(\hat{\Theta}) + B(\hat{\Theta})^2, \tag{4.7}$$

hvor $B(\hat{\Theta}) = E[\hat{\Theta}] - \theta$ er **bias** af $\hat{\Theta}$
Per egenskaber for $\bar{X}$ er $MSE(\bar{X}) = Var(\bar{X}) = \frac{\sigma^2}{n}$ (s. 426)

Definition 4.1.5: Konsistent estimator (side 431 box 2

Lad $\hat{\Theta_1}, \hat{\Theta_2}, ..., \hat{\Theta_n}, ...,$ være en følge af punkt estimatorer for θ. Vi siger at $\hat{\Theta_n}$ er en **konsistent** estimator af θ, hvis

$$\lim_{n\to\infty} P\big(|\hat{\Theta}_n - \theta| \geq \epsilon\big) = 0, \text{ for all } \epsilon > 0. \tag{4.8}$$

Definition 4.1.6: Sample variansen og Sample standard afvigelse (side 434

Lad $X_1, X_2, ..., X_n$ være en stikprøve med middelværdi $EX_i = \mu < \infty$, og varians $0 < Var(X_i) = \sigma^2 < \infty$. **Sample variansen** af denne stokastiske stikprøve er defineret som:

$$S^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2 = \frac{1}{n-1}\left(\sum_{k=1}^{n} X_k^2 - n\overline{X}^2\right). \tag{4.9}$$

. Sample variansen er en unbiased estimator af σ^2. **Sample standard afvigelsen (Sample standard deviation)** er defineret som:

$$S = \sqrt{S^2}$$

, og er generelt brugt som en estimator for σ. S er en biased estimator af σ.

4.2 Likelihood funktioner og MLE

Definition 4.2.1: Likelihood og log likelihood funktionerne (side 437)

Lad $X_1, X_2, ..., X_n$ være en stikprøve fra en fordeling med parameter θ. Antag at vi har observeret $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$.

- Hvis X_i'erne er diskrete, så er **likelihood funktionen** defineret:

$$L(x_1, x_2, \cdots, x_n; \theta) = P_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n; \theta).$$

- Hvis X_i'erne er simultan (jointly) kontinuerte, så er likelihood funktionen defineret:

$$L(x_1, x_2, \cdots, x_n; \theta) = f_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n; \theta).$$

Ved nogle opgaver, er det nemmere at arbejde med **log likelihood funktionen** givet ved:

$$\ln L(x_1, x_2, \cdots, x_n; \theta).$$

HUSK! Når $x_1, \ldots, x_n$ er uafhængige (som de altid er når de er stikprøver) må PDF og PMF splittes op i produkt. ERGO:

$$P_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n; \theta) = P_{X_1}(x_1; \theta) P_{X_2}(x_2; \theta) \ldots P_{X_n}(x_n; \theta)$$
$$f_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n; \theta) = f_{X_1}(x_1; \theta) f_{X_2}(x_2; \theta) \ldots f_{X_n}(x_n; \theta)$$

Definition 4.2.2: Maksimum likelihood estimator (MLE) (side 440)

Maksimum likelihood estimator (MLE)

Lad $X_1, X_2, ..., X_n$ være en stikprøve fra en fordeling med parameter θ. Givet at vi har observeret $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$, en **maksimum likelihood estimator (MLE)** af parameteren θ, skrevet $\hat{\theta}_{ML}$ er en værdi af θ, som maksimerer likelihood funktionen:

$$L(x_1, x_2, \cdots, x_n; \theta).$$

En maksimum likelihood estimator (MLE) af parameteren θ, skrevet $\hat{\Theta}_{ML}$, er en stokastisk variabel $\hat{\Theta}_{ML} = \hat{\Theta}_{ML}(X_1, X_2, ..., X_n)$ hvis værdi når $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$ er givet ved $\hat{\theta}_{ML}$.

Metode

Når man ønsker at finde $\hat{\theta}_{ML}$, maksimerer man Likelihood funktionen eller log likelihood funktionen. Dette gøres ved at differentiere den mht θ, og sætte den lig nul. Herefter undersøges hvilke værdier er max eller min, max er $\hat{\theta}_{ML}$. Dette kan bl.a gøres ved hjælp af monotoniforhold.

Ved monotoniforhold dobbelt differentieres likelihood funktionen mht. θ. Her indsættes de punkter man fik fra at differentiere og sætte lig 0 ind. Hvis disse værdier er positive er værdien et minimum, og hvis negativ et max. altså:

$$\begin{array}{lll} \text{Lad} & f'(x_1) = 0 & \\ \text{hvis} & f''(x_1) > 0 & \text{så er } x_1 \text{ et minimum} \\ \text{hvis} & f''(x_1) < 0 & \text{så er } x_1 \text{ et max} \end{array}$$

Definition 4.2.3: Asymptotiske egenskaber for MLE (s 443)

Lad $X_1, X_2, ..., X_n$ være en stikprøve fra en fordeling med parameter θ. Lad $\hat{\Theta}_{ML}$ være maksimum likelihood estimator (MLE) for θ. Da, under nogle milde regulære betingelser,

- $\hat{\Theta}_{ML}$ er asymptotisk konsistent, i.e.,

$$\lim_{n\to\infty} P(|\hat{\Theta}_{ML} - \theta| > \epsilon) = 0. \tag{4.10}$$

- $\hat{\Theta}_{ML}$ er asymptotisk unbiased, i.e.,

$$\lim_{n\to\infty} E[\hat{\Theta}_{ML}] = \theta. \tag{4.11}$$

- Når n bliver stor er $\hat{\Theta}_{ML}$ approksimativt en normal stokastisk variabel. Mere præcist, den stokastiske variabel

$$\lim_{n\to\infty} E[\hat{\Theta}_{ML}] = \theta. \tag{4.12}$$

  konvergerer i fordeling mod $N(0, 1)$

4.3 Konfidens-intervaller

Definition 4.3.1: Interval estimation (s 449)

Interval Estimator Lad $X_1, X_2, ..., X_n$ være en stikprøve fra en fordeling med parameter θ som ønkses estimeret. En **interval estimator** med **konfidens-niveau** $1 - \alpha$ består af to estimatorer $\hat{\Theta}_l(X_1, X_2, ..., X_n)$ og $\hat{\Theta}_h(X_1, X_2, ..., X_n)$, så

$$P\left(\hat{\Theta}_l \leq \theta \, \text{and} \, \hat{\Theta}_h \geq \theta\right) \geq 1 - \alpha, \tag{4.13}$$

for hver mulig værdi af θ. Ækvivalent, siges det at $[\hat{\Theta}_l, \hat{\Theta}_h]$ er et $(1 - \alpha)100\%$ **konfidens-interval** for θ.

Definition 4.3.2: Pivotal Quantity (s 453)

Pivotal Quantity

Lad $X_1, X_2, ..., X_n$ være en stikprøve fra en fordeling med parameter θ, som ønskes estimeret. Den stokastiske vairabel Q er en pivot, hvis den har følgende egenskaber:

- Den er en funktion af de observerede data $X_1, X_2, ..., X_n$ og den ukendte parameter θ, men den afhænger ikke af nogle andre ukendte parametre:
$$Q = Q(X_1, X_2, \cdots, X_n, \theta). \tag{4.14}$$
- PDF'en af Q afhænger ikke af θ eller nogen anden ukendt parameter.

4.3.1 Vigtige cases af interval estimatorer

Her er indsat 4 billeder som findes på siderne 455, 458, 459, 460, 466, 467 og 469.

Assumptions: A random sample $X_1, X_2, X_3, \ldots, X_n$ is given from a distribution with known variance $\text{Var}(X_i) = \sigma^2 < \infty$; n is large.

Parameter to be Estimated: $\theta = EX_i$.

Confidence Interval: $\left[\overline{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$ is approximately a $(1-\alpha)100\%$ confidence interval for θ.

Assumptions: A random sample $X_1, X_2, X_3, \ldots, X_n$ is given from a $Bernoulli(\theta)$; n is large.

Parameter to be Estimated: θ

Confidence Interval: $\left[\overline{X} - \frac{z_{\frac{\alpha}{2}}}{2\sqrt{n}}, \overline{X} + \frac{z_{\frac{\alpha}{2}}}{2\sqrt{n}}\right]$ is approximately a $(1-\alpha)100\%$ confidence interval for θ. This is a conservative confidence interval as it is obtained using an upper bound for σ.

Figure 4.1: Anvendes næsten aldrig, og antager at $\sigma_{max} = \frac{1}{2}$

Assumptions: A random sample X_1, X_2, X_3, $\ldots$, X_n is given from a $Bernoulli(\theta)$; n is large.

Parameter to be Estimated: θ

Confidence Interval: $\left[\overline{X} - z_{\frac{\alpha}{2}} \sqrt{\frac{\overline{X}(1-\overline{X})}{n}}, \overline{X} + z_{\frac{\alpha}{2}} \sqrt{\frac{\overline{X}(1-\overline{X})}{n}}\right]$ is approximately a $(1-\alpha)100\%$ confidence interval for θ.

Figure 4.2: $\bar{X}(1-\bar{X}) = S^2$

Assumptions: A random sample X_1, X_2, X_3, $\ldots$, X_n is given from a distribution with unknown variance $\mathrm{Var}(X_i) = \sigma^2 < \infty$; n is large.

Parameter to be Estimated: $\theta = EX_i$.

Confidence Interval: If S is the sample standard deviation

$$S = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} (X_k - \overline{X})^2} = \sqrt{\frac{1}{n-1} \left(\sum_{k=1}^{n} X_k^2 - n\overline{X}^2 \right)},$$

then the interval

$$\left[\overline{X} - z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right]$$

is approximately a $(1-\alpha)100\%$ confidence interval for θ.

Assumptions: A random sample X_1, X_2, X_3, $\ldots$, X_n is given from a $N(\mu, \sigma^2)$ distribution, where $\mathrm{Var}(X_i) = \sigma^2$ <u>is known</u>.

Parameter to be Estimated: $\mu = EX_i$.

Confidence Interval: $\left[\overline{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$ is a $(1-\alpha)100\%$ confidence interval for μ.

Assumptions: A random sample X_1, X_2, X_3, $\ldots$, X_n is given from a $N(\mu, \sigma^2)$ distribution, where $\mu = EX_i$ and $\mathrm{Var}(X_i) = \sigma^2$ are unknown.

Parameter to be Estimated: $\mu = EX_i$.

Confidence Interval: $\left[\overline{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}, \overline{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right]$ is a $(1-\alpha)$ confidence interval for μ.

Assumptions: A random sample X_1, X_2, X_3, $\ldots$, X_n is given from a $N(\mu, \sigma^2)$ distribution, where $\mu = EX_i$ and $\mathrm{Var}(X_i) = \sigma^2$ are unknown.

Parameter to be Estimated: $\mathrm{Var}(X_i) = \sigma^2$.

Confidence Interval: $\left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right]$ is a $(1-\alpha)100\%$ confidence interval for σ^2.

4.4 Hypotese-tests

DET ANBEFALES KRAFTIGT AT MAN BRUGER EKSEMPLERNE I BOGEN KAPITEL 8.4 :)

Definition 4.4.1: Hypotese-test

Lad θ være en ukendt parameter. Der opstilles en hypotese, hvor man regner med at θ er i et bestemt accept-område S_0. Denne hypotese kaldes H_0. Den alternative hypotese kaldes H_1, hvor $\theta \in S_1$

Definition 4.4.2: Statistik

En statistik, W, er en funktion, der kun må afhænge af kendte parametre (eksempelvis stikprøve-værdier, varians [HVIS DEN ER KENDT!])

Definition 4.4.3: Type I fejl og Type II fejl (s. 478)

Type I fejl er den hændelse, hvor man forkaster en sand H_0. Sandsynligheden for en Type I fejl kaldes α (**Signifikans-niveau**).
Type II fejl er den hændelse, hvor man accepterer en falsk H_0. Sandsynligheden for Type II fejl hedder β. I relation til type II fejl kaldes **styrken** (π) for en hypotese-test, og er defineret ved $\pi = 1 - \beta$.

Definition 4.4.4: Accept-område (s. 477)

Et accept-område er det område, hvor man accepterer H_0-hypotesen. Dette afhænger ofte af en givet statistik W (**se evt. tabellerne nedenfor**) og en konstant c, som enten er givet, eller man selv skal finde. Et eksempel på et typisk accept-område kunne være $|W| < c$.
Lad S være et accept-område, så:

$$\text{Acceptér } H_0 \Leftrightarrow W \in S$$

Lad S være et accept-område. Når man har et accept-område, kan α, β og π findes med følgende formler (c **kan også findes via isolation!**):

$$\begin{aligned}
\alpha &= P(\text{Type I fejl}) = P(W \in S^{\complement} | H_0 \text{ Sand}) \\
\beta &= P(\text{Type II fejl}) = P(W \in S | H_1 \text{ Sand}) \\
\pi &= P(W \in S^{\complement} | H_1 \text{ Sand}) \qquad \textbf{*Pr. Thyge!!}
\end{aligned}$$

En mere visuel måde at se det på er med denne tabel:

	H_0 Sand	H_1 Sand
Acceptér H_0	$1 - \alpha$	β
Forkast H_0	α	π

4.4.1 Gode hypoteser for middelværdi

Definition 4.4.5: Middelværdi: To-sidet tests (s. 485)

Når man har en to-sidet test, ser hypoteserne således ud:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

Her er nogle cases (dog når vi ikke har en normalfordeling er bliver det ikke et eksakt.

Case	Test Statistic	Acceptance Region
$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ known	$W = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$\lvert W\rvert \leq z_{\frac{\alpha}{2}}$
n large, X_i non-normal	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$\lvert W\rvert \leq z_{\frac{\alpha}{2}}$
$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ unknown	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$\lvert W\rvert \leq t_{\frac{\alpha}{2},n-1}$

Definition 4.4.6: Middelværdi: En-sidet tests (s. 488)

En-sidet tests har to variationer. Første variation er når $H_0 : \mu \leq \mu_0$, $H_1 : \mu > \mu_0$. Her ser tabellen sådan ud:

Case	Test Statistic	Acceptance Region
$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ known	$W = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$W \leq z_\alpha$
n large, X_i non-normal	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$W \leq z_\alpha$
$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ unknown	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$W \leq t_{\alpha,n-1}$

Den anden case er når $H_0 : \mu \geq \mu_0$, $H_1 : \mu < \mu_0$

Case	Test Statistic	Acceptance Region
$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ known	$W = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$W \geq -z_\alpha$
n large, X_i non-normal	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$W \geq -z_\alpha$
$X_i \sim N\left(\mu, \sigma^2\right), \sigma$ unknown	$W = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$W \geq -t_{\alpha,n-1}$

Definition 4.4.7: P-værdier (s. 488)

P-værdien er det mindste signifikans-niveau, α, som resulterer i at man forkaster nulhypotesen.

Definition 4.4.8: Likelihood-ratio test (s. 491)

Lad $X_1, X_2, \ldots, X_n$ være en stikprøve med de tilhørende værdier $X_1 = x_1, \ldots, X_n = x_n$. For at bestemme mellem $H_0 : \theta = \theta_0, H_1 : \theta = \theta_1$ bruges følgende formel:

$$\lambda(x_1, x_2, \ldots, x_n) = \frac{L(\theta_0, x_1, \ldots, x_n)}{L(x_1, \ldots, x_n, \theta_1)}$$

Likely-hood ratio testen giver et c (afhængig af et valgt α), hvor vi forkaster nulhypotesen hvis $\lambda < c$, og accepterer hvis $\lambda \geq c$

4.5 Lineær regression

Den **simple lineær regressions** model er givet ved $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$. Her er x_i deterministisk, dvs. et ikke stokastisk,ϵ_i stokastisk og derved er Y_i også stokastisk. $\epsilon_i \sim N(0, \sigma^2)$. (s 500)

Definition 4.5.1: Simpel lineær regression (side 501-502)

Simpel lineær regression

Givet observationerne $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, kan vi skrive regressionslinjen som

$$\hat{y} = \beta_0 + \beta_1 x. \tag{4.15}$$

Vi kan estimere β_0 og β_1 som

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}}, \tag{4.16}$$

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x}, \tag{4.17}$$

hvor

$$s_{xx} = \sum_{i=1}^{n} (x_i - \overline{x})^2, \tag{4.18}$$

$$s_{xy} = \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}). \tag{4.19}$$

For hver x_i, fås den **tilpassede værdi** $\hat{y}_i$ ved

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i. \tag{4.20}$$

. Kvantiteterne (The quantities)

$$e_i = y_i - \hat{y}_i \tag{4.21}$$

kaldes **residualerne**
I R kan $\hat{\beta}_1$ og $\hat{\beta}_0$ findes således:

$$\hat{\beta}_1 = \frac{cov(x, y)}{var(x)}$$

$$\hat{\beta}_0 = mean(y) - \hat{\beta}_1 \cdot mean(x)$$

Definition 4.5.2: Forklaringsgraden (Coefficient of determination) (side 505)

For de observerede data par, $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ vi definerer **Forklaringsgraden (Coefficient of determination)**, r^2 som

$$r^2 = \frac{s_{xy}^2}{s_{xx}s_{yy}}, \tag{4.22}$$

hvor

$$s_{xx} = \sum_{i=1}^{n}(x_i - \overline{x})^2, \quad s_{yy} = \sum_{i=1}^{n}(y_i - \overline{y})^2, \quad s_{xy} = \sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}). \tag{4.23}$$

Vi har $0 \leq r^2 \leq 1$. Større værdier af r^2 tyder typpisk på at vores lineære model

$$\hat{y_i} = \hat{\beta}_0 + \hat{\beta}_1 x_i \tag{4.24}$$

er et godt fit af på dataene.

Definition 4.5.3: Bias og Varians af $\hat{\beta}_1$ og $\hat{\beta}_0$(s. 518 opg 23 og 24

Per opgave 23 og 24 fås følgende

$$\hat{\beta}_1 \text{og} \hat{\beta}_0 \quad \text{er altid en unbiased estimator}$$

For $i = 1, 2, 3, ..., n$ er

$$cov(\hat{\beta}_1, Y_i) = \frac{x_i - \bar{x}}{s_{xx}}\sigma^2$$

Og

$$Cov(\hat{\beta}_1, \bar{Y}) = 0$$

Variansen for $\hat{\beta}_1$ og $\hat{\beta}_0$ er givet ved:

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{s_{xx}}$$
$$Var(\hat{\beta}_0) = \frac{\sum_{i=1}^{n} x_i^2}{n \cdot s_{xx}}\sigma^2$$

Og da $\hat{\beta}_1$ og $\hat{\beta}_0$ er unbiased, er MSE af disse lig variansen, per definition 4.1.3 og 4.1.4 (side 429 og 430 i lærebogen)

Chapter 5

Markov-kæder

Definition 5.0.1: Markov-kæde (Side 631)

$X_0, X_1, \ldots$ kaldes en markov-kæde (MK) med værdier i $S \subseteq \mathbb{N}_0$ hvis:

$$\begin{aligned} P(X_{m+1} = j | X_m = i, X_{m-1} = i_{m-1}, \cdots, X_0 = i_0) \\ = P(X_{m+1} = j | X_m = i), \end{aligned}$$

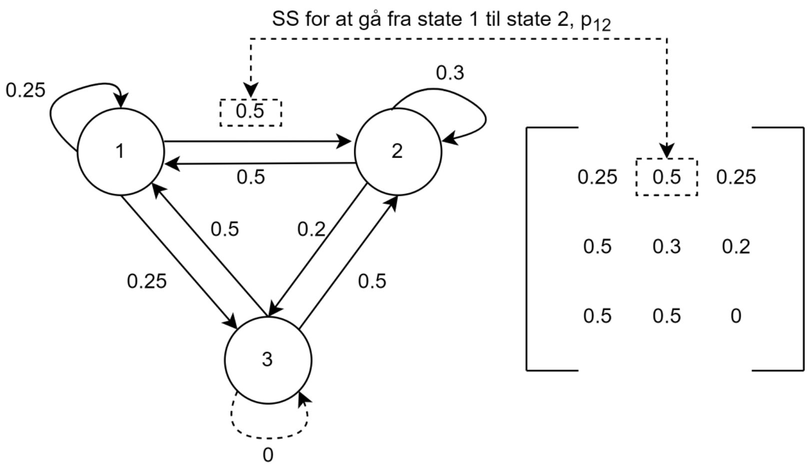
$\forall m, i, j, i_0, i_1, \ldots, i_{m-1}$. Er S en endelig mængde, har vi at gøre med en endelig markov-kæde.

Definition 5.0.2: Overgangsmatrix og -diagram (Side 631-632)

En markov-kæde vil næsten altid blive præsenteret enten med et overgangs-diagram eller en overgangs-matrix. Disse to modeller forklarer, hvordan en population har en tendens til at bevæge sig, givet at de er i en bestemt tilstand.

Overgangsmatricen bliver ofte betegnet med et P, og dens værdi på i'te række og j'te søjle bliver ofte betegnet med p_{ij}. Det smarte ved denne notation er dog også, at p_{ij} svarer til sandsynligheden for, at gå fra tilstand i til tilstand j.

Her er en figur der viser ligheden mellem et overgangsdiagram og -matrix.



Overgangs-diagram **Overgangs-matrix**

Definition 5.0.3: π-matricen (Side 634)

Hvor man intuitivt forstår P-matricen som noget, der beskriver en overgang, kan man forstå $\pi^{(n)}$-matrice, som en matrix/vektor der fortæller hvordan en befolkning/fordeling egentlig befinder sig til tid n. Man betegner også $i'te$ tilstand i $\pi^{(n)}$-vektoren som $\pi_i^{(n)}$. Følgende love gælder om $\pi^{(n)}$:

$$\pi^{(n+1)} = \pi^{(n)}P, \quad \text{for } n = 0, 1, 2, \cdots$$
$$\pi^{(n)} = \pi^{(0)}P^n, \quad \text{for } n = 0, 1, 2, \cdots .$$

Definition 5.0.4: Chapman-Kolmogorov ligningerne (Side 636)

$$p_{ij}^{(m+n)} = P(X_{m+n} = j | X_0 = i)$$
$$\sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}$$

Den n-trins overgangsmatrice er givet ved:

$$P^{(n)} = P^n, \qquad \forall n \in \mathbb{N}$$

Definition 5.0.5: Tilgængelighed (Side 637)

Lad $i, j \in S$ være tilstande. Man siger, at $i \to j$ (j **er tilgængelig for** i) hvis det er muligt at starte i tilstand i og så gå til j. Helt formelt er definitionen:

$$\exists n \in \mathbb{N}_0 : p_{ij}^{(n)} > 0 \Leftrightarrow i \to j$$

Hvis både $i \to j$ og $j \to i$ siger man, at i **og** j **kommunikerer**, noteret ved $i \leftrightarrow j$.

(Side 637) Det nævnes lige, at kommunikation er en ækvivalens-relation, dette betyder at følgende ting gælder:

- $i \leftrightarrow i, \quad \forall i \in S$
- Hvis $i \leftrightarrow j$, så $j \leftrightarrow i$
- Hvis $i \leftrightarrow j$ og $j \leftrightarrow k$, så $i \leftrightarrow k$

Definition 5.0.6: Klasser (Side 637-638)

Lad $i \in S$ være en tilstand. Man siger, at alle tilstande $j_1, j_2 \ldots \in S$ hvor $i \leftrightarrow j_k$ **er i samme klasse som** i. Dette betyder intuitivt, at dette er alle tilstande, hvor man kan gå fra i til j_k (i vilkårligt antal skridt!), samt at man kan gå fra j_k til i.

Definition 5.0.7: Irreducibel Markov-kæde (Side 638)

En markov-kæde siges værende irreducibel hvis og kun hvis:

1. Markov-kæden kun har en klasse
2. ELLER ÆKVIVALENT: Alle tilstandende i markov-kæden kommunikerer.

Matematisk er definitionen af begge punkter det samme, så det er i virkeligheden bare hvordan man vælger at bruge begreber der betyder noget her. Matematisk ser det sådan ud:

$$\forall i, j \in S : \quad i \leftrightarrow j$$

Definition 5.0.8: Rekurrente og transiente tilstande (Ugeseddel 13 Theorem 2)

Lad $\{X_n : n = 0, 1, 2, \ldots\}$ betegne en Markov-kæde med endeligt tilstandsrum S. Da er mængden af alle transiente tilstande præcis de tilstande $i \in S$ hvor der findes et $j \in S$ så $i \to j$, men $j \not\to i$. Specielt gælder der:

1. Hvis $i \in S$ opfylder at $i \to j$ medfører $j \to i$ for alle $j \in S$, så er i rekurrent.
2. Hvis Markov-kæde er irreducibel så er alle tilstande rekurrente.

Definition 5.0.9: Periode (Side 640)

Perioden for en tilstand $i \in S$ er defineret ved den største fælles divisor mellem antal skridt man kan tage fra tilstand i til sig selv. Med andre ord kan man lade $A = \{n \in \mathbb{N} : p_{ii}^{(n)} > 0\}$ betegne mængden af alle antal mulige skridt, og finde sfd(A). Hvis perioden er lig 1 kaldes tilstanden **aperiodisk**. En hurgtig måde at se, om en tilstand er aperiodisk er at se, om der er selvløkker ($p_{ii}^{(1)} > 0$), eller om andre tilstande i samme klasse har selvløkker.

Definition 5.0.10: Stationær fordeling (Side 651)

En stationær fordeling π er en vektor, som opfylder, at

$$\pi = P\pi$$

Definition 5.0.11: Grænsefordeling (Side 649)

En vektor $\pi = [\pi_1, \pi_2, \ldots]$ er den grænse fordeling hvis

$$\pi_j = \lim_{n\to\infty} P(X_n = j | X_0 = i)$$

For alle $i, j \in S$.
Altså hvis man får to værdier af π_j når man vælger forskellige i'er, er det ikke en grænsefordeling. En grænsefordeling er altid en stationær fordeling.

Definition 5.0.12: Ved irreducibel og aperiodisk MK (Side 653

Lad en endelig Markov-kæde med tilstandsrum $S = \{1, 2, \ldots, r\}$ være irreducibel og aperiodisk. Så har:

1. Følgende ligning en entydig løsning $\pi = \pi \cdot P$ (Stat-fordeling!)
2. π fra ligningen ovenfor er også en grænsefordeling.
3. Vi har at mean return-time (se evt. 5.0.14 side 42) $r_j = \frac{1}{\pi_j}$.

Et meget vigtigt resultat der er relevant i forhold til stationær- og grænsefordelinger er, at hvis en markov-kæde er irreducibel og aperiodisk (Ugeseddel 14 samt side 653), kan man finde grænse/stat-fordeling ret nemt. Er man bedt om at finde denne fordeling kan man bruge en af de vedhæftede R-dokumenter. Helt formelt skal man for at finde grænse/stationær fordeling gøre således:

1. Hvis overgangsmatrice (O.M.) ikke er kendt, skal den findes.
2. Opstil A-matrice, ved at transponere O.M, trække 1 fra diagonalen samt erstatte hele sidste række med 1. (Søjler bliver til rækker og omvendt).

En visualisering af opsætningen af en A-matrix kan ses her:

$$\begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.3 & 0.5 & 0.2 \\ 0.4 & 0.6 & 0 \end{bmatrix} \xrightarrow{\text{Lav diagonal}} \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.3 & 0.5 & 0.2 \\ 0.4 & 0.6 & 0 \end{bmatrix} \xrightarrow{\text{Rotér/Spejl om linje}} \begin{bmatrix} 0.1 & 0.3 & 0.4 \\ 0.2 & 0.5 & 0.6 \\ 0.7 & 0.2 & 0 \end{bmatrix}$$

$$\xrightarrow{\text{træk 1 fra på diagonal}} \begin{bmatrix} 0.1 & 0.3 & 0.4 \\ 0.2 & 0.5 & 0.6 \\ 0.7 & 0.2 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.9 & 0.3 & 0.4 \\ 0.2 & -0.5 & 0.6 \\ 0.7 & 0.2 & -1 \end{bmatrix}$$

$$\xrightarrow{\text{erstat sidste række med 1}} \begin{bmatrix} -0.9 & 0.3 & 0.4 \\ 0.2 & -0.5 & 0.6 \\ 1 & 1 & 1 \end{bmatrix}$$

3. Definér b-matrix med en bredde på 1 og en højde på antal tilstande i MK. Altså i eksemplet ovenfor vil:

$$b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

4. Man kan nu finde π (grænsefordeling) ved at løse ligningen $A\pi = b$. (Her anbefales det at man bruger R-kommandoen `solve(A,b)`).

Definition 5.0.13: Absorberende tilstand (Side 644)

I en markovkæde er en tilstand $a \in S$ absorberende hvis $p_{aa} = 1$. Nogle opgaver beder om, at man skal finde sandsynligheden for, at hver tilstand bliver absorberet.
Lad $s \in S$ være en absorberende tilstand. Sandsynligheden, a_i, for at man bliver absorberet i s givet at man starter i en vilkårlig tilstand $i \in S$ er givet ved:

$$a_i = \sum_k a_k p_{ik}, \qquad \text{for } i \in S.$$

Her skal man evt opstille flere ligninger, og så huske, at absorberende tilstande $t \in S, t \neq s$, har $a_t = 0$, og $a_s = 1$.

Definition 5.0.14: Mean hitting time og Mean Return time (Side 645-646)

Lad S være mængden af tilstande i en markovkæde, og $A \subseteq S$. Lad T være tiden det tager at komme ind i A, her er hver indgang bestemt ved:

$$t_i = 1 + \sum_k t_k p_{ik}, \qquad \text{for } i \in S - A.$$

Lad $i \in S$ være en tilstand. Den gennemsnitlige ventetid for at komme tilbage til tilstand i er bestemt ved:

$$r_l = 1 + \sum_k t_k p_{lk}$$

Chapter 6

R-kommandoer og simuleringer

6.1 Fordelinger i R

Fordeling	CDF	PMF/PDF	Invers	Simulering
Normal	`pnorm(x, μ, σ)`	`dnorm(x, μ, σ)`	`qnorm(x, μ, σ)`	`rnorm(g, μ, σ)`
Eksponential	`pexp(x, λ)`	`dexp(x, λ)`	`qexp(x, λ)`	`rexp(g, λ)`
Poisson	`ppois(x, λ)`	`dpois(x, λ)`	`qpois(x, λ)`	`rpois(g, λ)`
χ^2	`pchisq(x, fg)`	`dchisq(x, fg)`	`qchisq(x, fg)`	`rchisq(g, fg)`
Gamma	`pgamma(x, α, λ)`	`dgamma(x, α, λ)`	`qgamma(x, α, λ)`	`rgamma(g, α, λ)`
t	`pt(x, fg)`	`dt(x, fg)`	`qt(x, fg)`	`rt(g, fg)`
Binomial	`pbinom(x, n, p)`	`dbinom(x, n, p)`	`qbinom(x, n, p)`	`rbinom(g, n, p)`
Geometrisk*	`pgeom(x-1,p )`	`dgeom(x-1, p)`	`qgeom(x, p)+1`	`rgeom(g, p)+1`
Pascal*	`pnbinom(x-m+1, m, p)`	`dnbinom(x-m+1, m, p)`	`qnbinom(x-m+1,m,p)`	`rnbinom(g, m, p)+m`

- x: Det x man smider i funktionen
- g: Antal gentagelser ved simulering
- fg: Frihedsgrader - det er bare parameteren, se evt. fordelingernes definition
- Alle andre bogstaver har samme betydning som bogen giver dem (kan afvige fra fordeling til fordeling)

Bemærkninger/Afvigelser

- Normalfordelingen i bogen bruger $N(\mu, \sigma^2)$ (Varians), hvor der i R bliver brugt $N(\mu, \sigma)$ (Standardafvigelse)
- **Den geometriske- og pascalfordelingen** har nogle meget anderledes parametre. Jeg er 95% sikker på, at den geometriske er rigtig - men kun 75% sikker på, at Pascal er rigtig - jeg anbefaler måske, at man bruger den teoretiske PMF i stedet... USE AT OWN RISK!

6.2 Gode kommandoer at kunne

```
X = c(14.4, 13.5, 21.9, 18.0, 24) #Nogle værdier
sd(X) #Stikprøve-standardafvigelse (S)
```

```
var(X) #Stikprøve-varians (S^2)
mean(X) #Gennemsnit (X-bar)
sum(X) #Summen af alle værdier af X
cumsum(X) #Kommulativ sum, altså spyttes
          #(14.4, 27.9, 49.8, ...) ud
sqrt(X) #Kvadratrod (af hvert element)
prod(X) #Produkt
cumprod(X) #Kommulativ produkt
```

6.3 Simuleringer

6.3.1 Eksempel på simulering (Aflevering 6)

Her bruges følgende fremgangsmåde:

1. Lav en x-vektor, som er defineret af simulerings-søjlen på tabellen på side 43.
2. Find de værdier, der bedes om. Ofte vil er der en smart kommando man kan bruge for at finde værdierne, der bedes om (se afsnit 6.2 - gode kommandoer).

Eksempel. Følgende opgave er givet:

Lad X betegne en Poissonfordelt stokastisk variabel med parameter 2, dvs. $X \sim \text{Poisson}(2)$.

1. Find Simulerede EX
2. Find simulerede $E\sqrt{X}$
3. Find simulerede σ_X

```
#Trin 1:
x = rpois(10000, 2) #10000 er et tilfældigt valgt, men stort, tal

#Trin 2:
mean(x)

## [1] 2.0063

mean(sqrt(x))

## [1] 1.268693

sd(x)

## [1] 1.428938
```

6.4 Matricer

Når man skriver en matrice i R, skrives det som man ville læse en bog. Altså når du kigger på matricen

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$

vil den i R skrives

```
A=matrix(c(2,3,4,5), nrow=2,byrow=TRUE)
```

Hvor `nrow` angiver antal rækker (R kan selv regne antal søjler ud). `byrow=` skal altid være `TRUE`.

Hvis man har to matricer

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} B = \begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix}$$

findes matrixproduktet ved anvendelse af `%*%`, hvis en matrix opløftet i x ønskes anvendes `%^%x`:

```
library(Matrix);
library(expm);

##
## Attaching package:  'expm'
## The following object is masked from 'package:Matrix':
##
##     expm

A=matrix(c(2,3,4,5), nrow=2,byrow=TRUE)
B=matrix(c(6,7,8,9), nrow=2,byrow=TRUE)
A %*% B

##      [,1] [,2]
## [1,]   36   41
## [2,]   64   73

A %^% B

##       [,1]   [,2]
## [1,] 44668  58905
## [2,] 78540 103573
```

En vigtig bemærkning er, at pakken `expm`, skal være installeret. Se afsnit 6.2 for installering.

6.5 Simmuleringer i R, anvendes når der skal plottes

Simulering af middelværdi terningekast:

$$a = sample(6, 100, replace = TRUE);$$
$$b = cumsum(a);$$
$$c = c(1 : 100);$$
$$d = b/c;$$
$$plot(d);$$

Simulering af middelværdi geometrisk fordeling:

$$a = rgeom(10000, 1/2) + 1;$$
$$b = cumsum(a);$$
$$c = c(1 : 10000);$$
$$d = b/c;$$
$$plot(d);$$

Law of rare events – PMF:

$$a = dbinom(0 : 25, 50, 1/50);$$
$$b = dpois(0 : 25, 1);$$
$$plot(a, col =' blue', ylim = c(0, 0.5));$$
$$par(new = TRUE);$$
$$plot(b, col =' red', ylim = c(0, 0.5));$$

Law of rare events – histogram:

$$x = rbinom(10000, 100, 1/100);$$
$$y = rpois(10000, 1);$$
$$layout(c(1, 2));$$
$$hist(x);$$
$$hist(y);$$

Simulering af sandsynligheder og middelværdier:

$$a = rbinom(100000, 50, 1/2);$$
$$b = a > 30;$$
$$mean(b);$$
$$c = exp(a);$$
$$mean(c);$$

Chapter 7

Appendix

7.1 Integralregning

Problemtype 7.1.1: Hvilken metode skal jeg bruge?

Jeg har to funktioner med et plus imellem $(f(x) + g(x))$
$\int f(x) + g(x)dx = \int f(x)dx + \int g(x)dx$
Jeg har to funktioner, der er ganget sammen $(f(x) \cdot g(x))$
Se Partiel Integration: 7.1.3 på side 48
Jeg har en funktion af en funktion $(f(g(x)))$
Se Integration ved substitution 7.1.2 på side 47
Jeg har glemt de vigtiste integraler
Se på listen over integraler og afl. funk. 7.1.4 på side 48

Definition 7.1.2: Integration ved Substitution

Hvis man ønsker at integrere en sammensat funktion, anvendes intergation ved substitution. Valget af formlen afhænger af om man arbejder med grænser eller ej.

$$\int f(g(x)) \cdot g'(x)\, dx = \int f(t)\, dt$$
$$\int_a^b f(g(x)) \cdot g'(x)\, dx = \int_{g(a)}^{g(b)} f(t)\, dt, \qquad \text{hvor } t = g(x)$$

Eksempel: Vi ønsker at finde $\int x \cdot e^{x^2}\, dx$. Her ses det at vi har at gøre med en sammensat funktion.

- Vi finder den indre funktion, t, her $t = x^2$
- Vi differentierer den indre funktion: $\frac{dt}{dx} = 2x$
- Vi gør noget meget ulovligt og lader som om $\frac{dt}{dx}$ er en brøk og får: $dx = \frac{1}{2x}dt$
- Nu erstattes den indre funktion med t, og dx med $\frac{1}{2x}dt$: $\int x \cdot e^{x^2}\, dx = \int x \cdot e^t \cdot \frac{1}{2x}\, dt$
- Vi reducerer og får $\int x \cdot e^t \cdot \frac{1}{2x}\, dt = \int \frac{1}{2}e^t\, dt = \frac{1}{2}e^t + c$

- Nu sættes den indre funktion igen ind på t's plads, og integralet er fundet: $\int x \cdot e^{x^2}\, dx = \frac{1}{2}e^{x^2} + c$

Definition 7.1.3: Partiel Integration

Hvis man ønsker at integrere produktet af to funktioner, laves partiel integration. Her anvendes en af disse formler, afhængigt af om der er grænser eller ej.

$$\int f(x)g(x)\, dx = f(x)G(x) - \int f'(x)G(x)\, dx$$
$$\int_a^b f(x)g(x)\, dx = [f(x)G(x)]_a^b - \int_a^b f'(x)G(x)\, dx$$

Bemærk: valget af henholdsvis g og f er vigtigt. Sæt f lig den funktion som hurtigst bliver konstant når den differentieres nok gange.

Eksempel: Vi ønsker at finde $\int 2x \cdot \cos(x)\, dx$. Her husker vi at valget af f og g er vigtigt.

- Vi sættes f(x)=2x og g(x)=cos(x), da cos(x) aldrig vil blive en konstant ved at differentiere det!!!
- Nu skriver vi de fire størrelser op: $f(x) = 2x$, $g(x) = \cos(x)$, $f'(x) = 2$, $G(x) = \sin(x)$
- Indsat i formlen: $\int 2x\cos(x)\, dx = 2x \cdot \sin(x) - \int 2 \cdot \sin(x)\, dx$
- Nu ses det, at integralet på højre side sagtens kan udregnes:

$$\begin{aligned}\int 2x\cos(x)\, dx &= 2x \cdot \sin(x) - \int 2 \cdot \sin(x)\, dx \\ &= 2x \cdot \sin(x) - 2 \cdot (-\cos(x)) + c \\ &= 2x \cdot \sin(x) + 2\cos(x) + c\end{aligned}$$

Definition 7.1.4: Liste over integraler og afledte funktioner

$F(x)$	$f(x)$	$\frac{d}{dx}f(x)$
$a \cdot F(x)$	$a \cdot f(x)$	$a \cdot f'(x)$
$\frac{1}{n+1}x^{n+1}\ (n \neq -1)$	x^n	nx^{n-1}
$ln(\|x\|)$	$\frac{1}{x}$	
e^x	e^x	e^x
$\frac{1}{ln(x)}a^x$	a^x	$\ln(a)a^x$
$-cos(x)$	$sin(x)$	$cos(x)$
$sin(x)$	$cos(x)$	$-sin(x)$

170 - 170
Infinitesimalregning

Funktion $f(x)$	Differentialkvotienten $f'(x)$	Stamfunktionerne $F(x)$
$\sin(x)$	$\cos(x)$	$-\cos(x)+k,\ k\in\mathbb{R}$
$\cos(x)$	$-\sin(x)$	$\sin(x)+k,\ k\in\mathbb{R}$
$\tan(x)$	$1+\tan^2(x)=\dfrac{1}{\cos^2(x)}$	$-\ln\lvert\cos(x)\rvert+k, k\in\mathbb{R}$
$\cot(x)$	$-1-\cot^2(x)=-\dfrac{1}{\sin^2(x)}$	$\ln\lvert\sin(x)\rvert+k,\ k\in\mathbb{R}$
$\sin(c\cdot x)$	$c\cdot\cos(c\cdot x)$	$-\dfrac{\cos(c\cdot x)}{c}+k,\ k\in\mathbb{R}$
$\cos(c\cdot x)$	$-c\cdot\sin(c\cdot x)$	$\dfrac{\sin(c\cdot x)}{c}+k,\ k\in\mathbb{R}$
$\tan(c\cdot x)$	$c\cdot(1+\tan^2(c\cdot x))$	$-\dfrac{\ln\lvert\cos(c\cdot x)\rvert}{c}+k,\ k\in\mathbb{R}$
$\sin^2(x)$	$\sin(2\cdot x)$	$\dfrac{1}{2}\cdot x-\dfrac{1}{4}\cdot\sin(2\cdot x)+k,\ k\in\mathbb{R}$
$\cos^2(x)$	$-\sin(2\cdot x)$	$\dfrac{1}{2}\cdot x+\dfrac{1}{4}\cdot\sin(2\cdot x)+k,\ k\in\mathbb{R}$
$\tan^2(x)$	$\dfrac{2\cdot\tan(x)}{\cos^2(x)}$	$\tan(x)-x+k,\ k\in\mathbb{R}$
$1+\tan^2(x)$	$\dfrac{2\cdot\tan(x)}{\cos^2(x)}$	$\tan(x)+k,\ k\in\mathbb{R}$
$\dfrac{1}{\cos^2(x)}$	$\dfrac{2\cdot\tan(x)}{\cos^2(x)}$	$\tan(x)+k,\ k\in\mathbb{R}$
$\arctan(x)$	$\dfrac{1}{1+x^2}$	$x\cdot\arctan(x)-\dfrac{1}{2}\ln(1+x^2)+k,\ k\in\mathbb{R}$
$\sinh(x)$	$\cosh(x)$	$\cosh(x)+k,\ k\in\mathbb{R}$
$\cosh(x)$	$\sinh(x)$	$\sinh(x)+k,\ k\in\mathbb{R}$
$\tanh(x)$	$1-\tanh^2(x)=\dfrac{1}{\cosh^2(x)}$	$\ln(\cosh(x))+k,\ k\in\mathbb{R}$

Huskeregel:

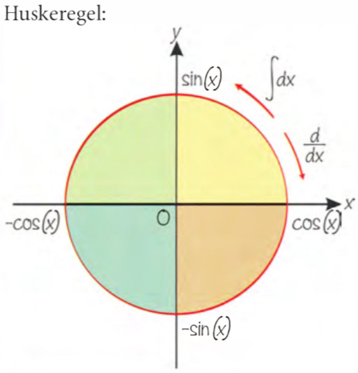


Figure 7.1: flere regler

7.2 Differential regneregler

Plus og minus

$$(f(x) \pm g(x))' = f'(x) \pm g'(x)$$

konstantreglen

$$(k \cdot f(x))' = k \cdot f'(x)$$

Kædereglen

$$(f(g(x)))' = f'(g(x)) \cdot g'(x)$$

Produktreglen

$$h(x) = f(x) \cdot g(x)$$
$$h'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

7.3 Love ved mængdeberegninger

7.3.1 Disse må du altid bruge

- $P(A^{\mathsf{C}}) = 1 - P(A)$
- $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- $P(A - B) = P(A) - P(A \cap B)$
- De Morgans lov:

$$(A_1 \cup A_2 \cup A_3 \cup \cdots A_n)^c = A_1^c \cap A_2^c \cap A_3^c \cdots \cap A_n^c$$

$$(A_1 \cap A_2 \cap A_3 \cap \cdots A_n)^c = A_1^c \cup A_2^c \cup A_3^c \cdots \cup A_n^c$$

- Inklusion/Eksklusionsprincip:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Bayes' lov:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

- Lov om total sandsynlighed ($B_1 \cup \ldots skaludgrepartition$):

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i)$$

- Distributive lov:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

7.3.2 Hvis der er uafhængighed

- $P(A|B) = P(A)$
- $P(A \cap B) = P(A)P(B)$
- A og B uaf $\Leftrightarrow$ A og $B^{\complement}$ uaf. $\Leftrightarrow$ $A^{\complement}$ og $B^{\complement}$ uaf.
- $P\big(A_1 \cup A_2 \cup \cdots \cup A_n\big) = 1 - \big(1 - P(A_1)\big)\big(1 - P(A_2)\big)\cdots\big(1 - P(A_n)\big)$.

7.3.3 Andre betingelser

Disjunkte mængder

- $A \cap B = \emptyset$
- $P(A \cup B) = P(A) + P(B)$

7.4 Uendelige summer/rækker

Den geometriske række:

$$\sum_{i=0}^{n} a^i = \frac{1 - a^{n+1}}{1 - a}$$
$$\sum_{i=0}^{\infty} a^i = \frac{1}{1 - a}$$

Eksponentialfunktionen kan også repræsenteres på følgende måde:

$$\sum_{i=0}^{\infty} \frac{a^i}{i!} = e^a$$

Index