	Emne	Kompenedium
1	Repetisjon	
2	Kontinuerlige stokastiske variabler	[I] 1.1
3	Forventning og varians	[I] 1.2 - 1.4
4	Median og andre prosentiler	[I] 1.5 - 1.6

	Oppgaver for Forelesning 7	
	Oppgaver fra arbeidsboken	[DA] 7.1 - 7.2
	Oppgaver fra kompendiet	[I] 1.15, 1.19, 1.22 - 1.24

① Repetisjon:

$$Q(\underline{x}) = \underline{x}^T A \underline{x} = \underline{u}^T D \underline{u} = \lambda_1 u_1^2 + \lambda_2 u_2^2 + \dots + \lambda_n u_n^2$$

kvadr. form

$P^T A P = D$ ortogonal diag.

$b_1 x_1 + b_2 x_2 + \dots$ = "

$$f(\underline{x}) = \underline{x}^T A \underline{x} + B\underline{x} + C$$

$$f'(\underline{x}) = \begin{pmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \vdots \\ \partial f/\partial x_n \end{pmatrix} = 2A\underline{x} + B^T$$

A pos. semidefin. $\Leftrightarrow \lambda_i \geq 0 \Leftrightarrow Q$ konvels

A neg. semidefin. $\Leftrightarrow \lambda_i \leq 0 \Leftrightarrow Q$ konkav

Stasjonære pkt:

$$2A\underline{x} + B^T = \underline{0}$$

$$2A\underline{x} = -B^T \quad \text{lin. system}$$

A pos. semidefin: alle stasjonære pkt globale min

A neg. semidefin: alle stasjonære pkt er globale max

[DAJ 6.4] $Q(\underline{x}) = \underline{x}^T A \underline{x}$ (A symmetrisk) $\Rightarrow Q'(\underline{x}) = 2A\underline{x}$

$= \begin{pmatrix} \partial Q/\partial x_1 \\ \vdots \\ \vdots \end{pmatrix}$

Alt: $\frac{\partial Q}{\partial x_1} = \frac{\partial}{\partial x_1}\left(\underline{x}^T A \underline{x}\right) = \frac{\partial}{\partial x_1}\left(\underline{x}^T\right) \cdot A\underline{x}$ ($\underline{x}^T = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}^T$)

$+ \underline{x}^T \frac{\partial}{\partial x_1}(A\underline{x})$

$= (1\ 0\ \dots\ 0) \cdot A\underline{x} + \underline{x}^T \cdot A \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \underline{e}_1^T A\underline{x} + \underline{x}^T A \underline{e}_1$

$= e_1^T A\underline{x} + (\underline{x}^T A e_1)^T = e_1^T A\underline{x} + e_1^T A^T \underline{x} = 2 \cdot e_1^T A\underline{x}$

$\frac{\partial Q}{\partial x_i} = 2\, e_i^T A\underline{x} = e_i^T \cdot (2A\underline{x})$

$= (0\ 0 \dots 1 \dots 0) \cdot 2A\underline{x} = (2A\underline{x})_i$ (1 på i'te plass)

$\underline{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ ← i'te plass

$Q'(\underline{x}) = \begin{pmatrix} \partial Q/\partial x_1 \\ \vdots \\ \partial Q/\partial x_n \end{pmatrix} = 2A\underline{x}$ ☺

Alt: $Q(\underline{x}) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \dots + a_{1n}x_1x_n + a_{21}x_1x_2 + \dots$

$\frac{\partial Q}{\partial x_1} = a_{11} \cdot 2x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{21}x_2 + a_{31}x_3$

$Q'_{x_1} = 2a_{11}x_1 + 2a_{12}x_2 + 2a_{13}x_3 + \dots$

$Q'_{x_2} = 2a_{21}x_1 + 2a_{22}x_2 + \dots$

$\}\ Q'(\underline{x}) = 2A\underline{x}$

[DA] 6.8 $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$

i	x_1	x_2	$\dots$	x_n	Y
1	x_{11}	x_{21}		x_{n1}	y_1
2	x_{12}	x_{22}		x_{n2}	y_2
$\vdots$					
N	x_{1N}	x_{2N}		x_{nN}	y_N

Datasett

$\varepsilon_1 = y_1 - (\beta_0 + \beta_1 x_{11} + \dots + \beta_n x_{n1})$

"

$X = \begin{pmatrix} 1 & \cdots \\ \vdots & \\ 1 & \cdots \end{pmatrix}$ $\underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{pmatrix}$ $\underline{\beta} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix}$

"

$\underline{\varepsilon} = \underline{y} - X \cdot \underline{\beta}$

$$\min_{\beta\text{-verdier}} \varepsilon_1^2 + \varepsilon_2^2 + \dots + \varepsilon_N^2 = \|\underline{\varepsilon}\|^2 = \underline{\varepsilon}^T \cdot \underline{\varepsilon} = (\underline{y} - X\underline{\beta})^T \cdot (\underline{y} - X\underline{\beta})$$

$$= (\underline{y}^T - \underline{\beta}^T X^T)(\underline{y} - X\underline{\beta}) = \underline{y}^T \cdot \underline{y} - \underline{\beta}^T X^T \underline{y} - \underline{y}^T X \underline{\beta} + \underline{\beta}^T (X^T X)\underline{\beta}$$

($\underline{y}^T X \underline{\beta}$ = "")

$$= \underbrace{\underline{\beta}^T (X^T X)\underline{\beta}}_{\text{kvadr. form}} - \underbrace{2(\underline{y}^T X)\underline{\beta}}_{\text{lin. form}} + \underbrace{\underline{y}^T \underline{y}}_{\text{konst.}}$$

Stasj. pkt: $2(X^T X)\underline{\beta} - 2(\underline{y}^T X)^T = \underline{0}$

$$(X^T X)\underline{\beta} = X^T \underline{y} \Rightarrow \underline{\beta} = (X^T X)^{-1}(X^T \underline{y})$$

($|X^T X| \neq 0$)

Fakta: i) $X^T X$ symm. positiv semidefin. matrise $\Rightarrow \underline{\varepsilon}^T \underline{\varepsilon}$ konvex $\Rightarrow$ Stasj. pkt min

ii) $|X^T X| \neq 0 \Leftrightarrow$ kolonnene i X er lineært uavhengige

(2) Kontinuerlige stokastiske variabler

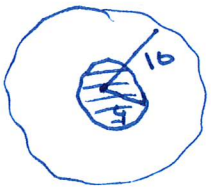
Stokastisk forsøk: mengden av mulige utfall (= utfallsrommet) er gitt, men utfallet er tilfeldig

Stokastisk variabel X: Variabel der verdien kommer an på utfallet av et stokastisk forsøk

Eks: Vi kaster to terninger X = summen, Y = max

Mulige verdier for X: $2, 3, \ldots, 12$ (diskret)

Eks: Vi kaster pil på en blink



X = avstand til sentrum

Mulige verdier for X: $[0, 10]$

$$P(X \leq 4) = \frac{\pi \cdot 4^2}{\pi \cdot 10^2} = \frac{16}{100} = 0.16$$

$$F(b) = P(X \leq b) = \frac{\pi \cdot b^2}{\pi \cdot 10^2} = \frac{b^2}{100} \qquad F(x) = \frac{x^2}{100}$$

$$f(x) = F'(x) = \frac{2x}{100} = x/50$$ sannsynlighetstetthet til X

Hvis X er en kontinuerlig stokastisk variabel, så kalles $f_X(x) = f(x)$ sannsynlighetstetthet til X hvis

$$p(a \le X \le b) = \int_a^b f(x)\,dx$$

En sannsynlighetstetthet oppfyller:

i) $f(x) \ge 0$ for alle x

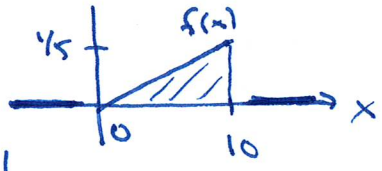
ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Den kumulative sannsynlighetstetthet er gitt ved:

$$F(b) = p(X \le b) = \int_{-\infty}^{b} f(x)\,dx$$

Vi ser at $F'(b) = f(b)$, dvs $F'(x) = f(x)$.

Eks: $f(x) = x/50$, $0 \le x \le 10$

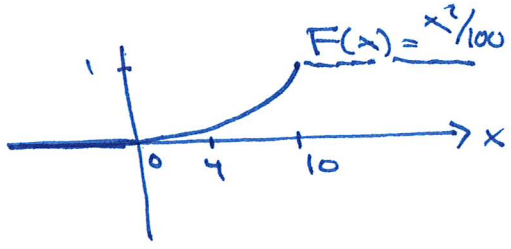


$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{10} x/50\,dx = \frac{10 \cdot 1/5}{2} = 1 \quad \text{ok}$$

$$F(b) = p(X \le b) = \int_{-\infty}^{b} f(x)\,dx$$

$$= \int_0^b x/50\,dx = \left[\tfrac{1}{50} \cdot \tfrac{1}{2}x^2\right]_0^b = \frac{1}{100}b^2 - 0$$

$$= \frac{1}{100}b^2 \Rightarrow F(x) = \frac{1}{100}x^2, \quad 0 \le x \le 10$$



- $F(x)$ er en voksende funksjon
- $\lim_{x \to \infty} F(x) = 1$
- $\lim_{x \to -\infty} F(x) = 0$

* Diskret:

$E(X) = x_1 \cdot p(x_1) + x_2 \cdot p(x_2) + \dots = \sum_{i=1}^{n} x_i \cdot p(x_i)$

(3) Forventning og varians:

Forventningsverdi: $E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$

Eks: $f(x) = x/50$, $0 \le x \le 10$ sannsynlighetstetthet for X

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_0^{10} x \cdot x/50\,dx = \left[\frac{1}{50} \cdot \frac{1}{3}x^3\right]_0^{10} = \frac{1000}{150} - 0 = \frac{100}{15} = \frac{20}{3} \simeq 6.67$$

Forventningsverdi til $h(X)$:

Eks: $E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx$

$$= \int_0^{10} x^2 \cdot x/50\,dx = \left[\frac{1}{50} \cdot \frac{1}{4}x^4\right]_0^{10} = \frac{10\,000}{200} - 0 = 50$$

$$E(h(X)) = \int_{-\infty}^{\infty} h(x)\, f(x)\,dx$$

Merk: $E(X^2) > E(X)^2$

$50 \neq E(X)^2 = \frac{20^2}{3^2} = \frac{400}{9}$

$Y = X^2$: stokastisk variabel

$$E(Y) = \int_0^{100} y \cdot f_Y(y)\,dy = \int_0^{100} y \cdot \frac{1}{100}\,dy = \left[\frac{1}{100} \cdot \frac{1}{2}y^2\right]_0^{100} = \frac{10.000}{200} - 0 = 50$$

$$F_Y(b) = P(Y \le b) = P(X^2 \le b) = P(X \le \sqrt{b}) = F_X(\sqrt{b}) = \frac{(\sqrt{b})^2}{100} = \frac{b}{100}$$

$$F_Y(y) = \frac{y}{100} \qquad f_Y(y) = \frac{1}{100}$$

Varians: X stokastisk variabel

$E(X) = \mu$

$$\mathrm{Var}(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

$$= \int_{-\infty}^{\infty} (x^2 - 2\mu x + \mu^2) f(x)\,dx = \int x^2 f(x)\,dx - 2\mu \int x f(x)\,dx + \mu^2 \cdot \int 1 \cdot f(x)\,dx$$

$$= E(X^2) - 2\mu \cdot \underbrace{E(X)}_{\mu} + \mu^2 \cdot 1 = E(X^2) - \mu^2 = \underline{E(X^2) - E(X)^2}$$

Eks: $f(x) = x/50$, $0 \le x \le 10$ $\quad \mu = E(X) = 20/3$

$$\mathrm{Var}(X) = \int_0^{10} (x - 20/3)^2 \cdot x/50\,dx$$

$$= \int_0^{10} \frac{1}{50} x \left(x^2 - \frac{40}{3}x + \frac{400}{9}\right) dx$$

$$= \int_0^{10} \frac{1}{50} \cdot \left(x^3 - \frac{40}{3}x^2 + \frac{400}{9}x\right) dx$$

$$= \left[\frac{1}{50}\left(\frac{1}{4}x^4 - \frac{40}{3}\cdot\frac{1}{3}x^3 + \frac{400}{9}\cdot\frac{1}{2}x^2\right)\right]_0^{10}$$

$$= \frac{1}{50}\left(\frac{10000}{4} - \frac{40\cdot 1000}{9} + \frac{400\cdot 100}{18}\right)$$

$$= \frac{10000}{50}\left(\frac{1}{4} - \frac{4}{9} + \frac{2}{9}\right) = 200\left(\frac{9-2\cdot 4}{36}\right) = \frac{200}{36} = \frac{50}{9}$$

Alt: $\mathrm{Var}(X) = E(X^2) - E(X)^2 = 50 - \left(\frac{20}{3}\right)^2 = \frac{50\cdot 9}{9} - \frac{400}{9} = \frac{50}{9}$

Merk: $\mathrm{Var}(X) \ge 0$ for alle X

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

Eks: $\sigma = \sqrt{\frac{50}{9}} = \frac{\sqrt{25}\cdot\sqrt{2}}{\sqrt{9}} = \frac{5}{3}\sqrt{2}$

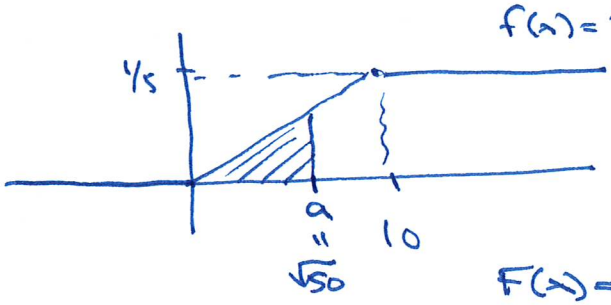
(4) Median:

Median $= a$

$F(a) = 1/2$

$\parallel$

$P(X \leq a)$



$f(x) = x/50, \quad 0 \leq x \leq 10$

$F(x) = \frac{x^2}{100}, \quad 0 \leq x \leq 10$

$\mu = 20/3 \simeq 6.67$

$\frac{a^2}{100} = \frac{1}{2}$

$a^2 = 100 \cdot \frac{1}{2} = 50$

$a = \sqrt{50} = 5\sqrt{2}$

$\simeq 7$

Prosentil:

$a =$ 90% prosentil

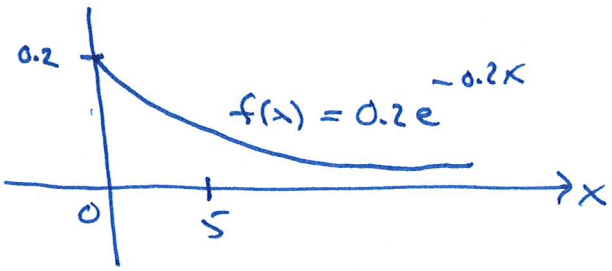
$F(a) = 0.90$

$\frac{a^2}{100} = 0.90$

$a^2 = 100 \cdot 0.9 = 90$

$a = \sqrt{90} = 3\sqrt{10}$

Eks: X har tetthetsfunksjon $f(x) = 0.2e^{-0.2x}$, $x \geq 0$



i) $f(x) \geq 0$ vok

ii) $\int_0^\infty 0.2e^{-0.2x}\,dx$

$= \lim_{b\to\infty} \left[-e^{-0.2x}\right]_0^b$

$= \lim_{b\to\infty} \left(-e^{-0.2b}\right) - \left(-e^0\right) = 1$ vok

$-e^{-0.2b} = -\frac{1}{e^{0.2b}} \to 0$

$E(x) = \int_0^\infty \underbrace{x}_{u} \cdot \underbrace{0.2e^{-0.2x}}_{v'}\,dx$

$u = -e^{-0.2x}$ $v = x$

$u' = 0.2e^{-0.2x}$ $v' = 1$

$= -xe^{-0.2x} - \int -e^{-0.2x} \cdot 1\,dx$

$= \left[-xe^{-0.2x} + \left(\frac{1}{-0.2}\right)e^{-0.2x}\right]_0^\infty$

$= \left[(-x-5)e^{-0.2x}\right]_0^\infty = \lim_{x\to\infty} \frac{-x-5}{e^{0.2x}} - (-5 \cdot 1)$

$\lim_{x\to\infty} \frac{-x-5}{e^{0.2x}} = 0$

$= 5$