Plan:

1. Simultant fordelte stokastiske variabler
2. Uavhengighet
3. Store talls lov

utsatt til neste uke

Pensum:

[L] 4.4

Merk: Alle eksempler på simultant fordelte variabler er diskret (sitt i tabell), men alle formler gjelder også for kontinuerlige variabler!

Repetisjon:

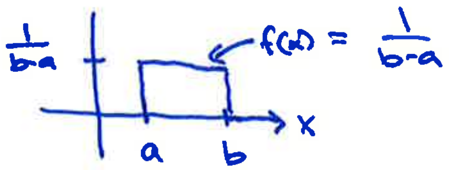
Krav: i) $f(x) \geq 0$ for alle x
ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

① Kontinuerlige stokastiske variabler

X kontinuerlig stokastisk variabel ------> $f(x)$ sannsynlighets-fordeling eller tetthetsfunksjon

$X(S)$ kont. mengde (intervall)

Eks: Uniform fordeling på $[a, b]$:



"alle verdier mellom a og b like sannsynlige"

Defn.: $\boxed{P(a \leq X \leq b) = \int_a^b f(x)\,dx}$

"$f(x)$ = sannsynlighet for x-verdier i området omkring x"

Husk: $\boxed{P(X = a) = 0}$

alle pkt. sannsynligheter er null

Viktige defn:

- Kummulativ fordelingsfunksjon:

$$\boxed{F(x) = \int_{-\infty}^{x} f(t)\,dt = P(X \leq x)}$$

- Forventningsverdi $\mu = E(X)$:

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

diskret: $E(X) = \sum_x x \cdot p(x)$

- Varians og std. avvik: $\sigma^2 = Var(X)$

$$\boxed{Var(X) = E[(X - \mu)^2] = E(X^2) - \mu^2}$$

$\mu = E(X)$

Regneregler:

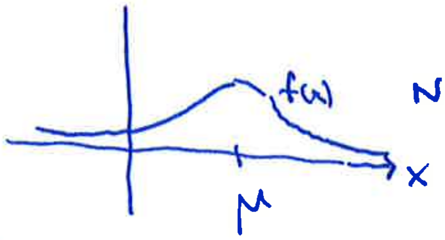
i) $E(aX + b) = a \cdot E(X) + b$

ii) $Var(aX + b) = a^2 \cdot Var(X)$

iii) $Var(X) = E(X^2) - \mu^2$

② **Normalfordeling:** $f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$ for μ, σ med $\sigma > 0$

Teoretisk viktig fordeling:



$N(\mu,\sigma)$

$E(X) = \mu$
$Var(X) = \sigma^2$

Viktig egenskap:

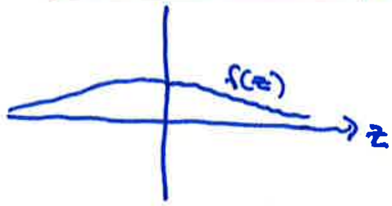
" $\frac{1}{\sigma}(x-\mu) = \frac{1}{\sigma}\cdot X + (-\frac{\mu}{\sigma})$ "

X normalfordelt $N(\mu,\sigma)$ $\Rightarrow$ $Z = \frac{X-\mu}{\sigma}$ normalfordelt $N(0,1)$ = std. normalfordelt

Z = antall std. avvik fra μ

Alternativt:

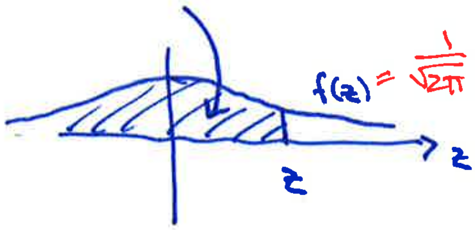
X normal-fordelt $\Rightarrow$ $aX+b$ normal-fordelt



Std. normalfordeling $N(0,1)$ med $\mu=0$, $\sigma=1$

Kummulativ fordelingsfunksjon for $N(0,1)$:

$G(z) = P(Z \le z)$



$f(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-z^2/2}$

På kalkulator: $\boxed{z \rightleftarrows p}$

fra z-verdi (antall std. avvik) til p-verdi (sannsynlighet)

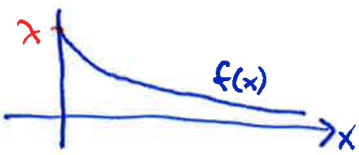
$\lambda > 0$

③ **Eksponentialfordeling:**

$f(x) = \lambda e^{-\lambda x}$, $x \ge 0$

$E(X) = 1/\lambda$
$Var(X) = 1/\lambda^2$

eksponentialfordeling med parameter λ



$X(S) = [0, \to)$

I en Poisson prosess med forventet antall ankomster λ per tidsenhet, er X = tid til første ankomst eksponentialfordelt

Oppgaverekk 6

6d) $X \sim N(20,5)$: X normalfordelt $\mu = 20,\ \sigma = 5$

$p(\underbrace{|X-20|} > 3)$

$|X-20|$ betyr avstand mellom X og 20

Evt: $|X-20| > 3$

$\Updownarrow$

$X-20 > 3$ eller $-(X-20) > 3$

$X > 23$ $\qquad$ $X-20 < -3$

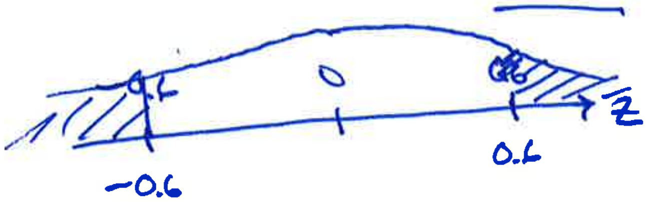
$X < 17$

$= p(X<17) + p(X>23)$

$= p(Z<-0.6) + p(Z>0.6)$

$= G(-0.6) + (1 - G(0.6))$

$= G(-0.6) \cdot 2 \simeq \underline{\underline{0.55}}$



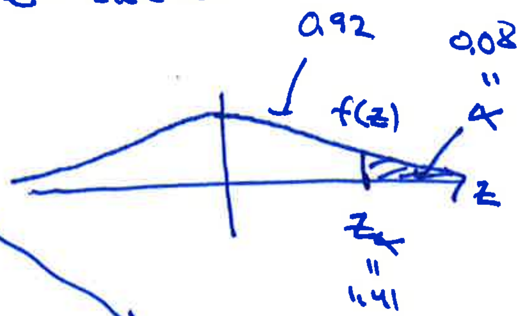
$Z = \dfrac{X-20}{5}$

Oppgave 9

Defn: z_α er den z-verdien som er slik at

$p(Z > z_\alpha) = \alpha$

når Z er $N(0,1)$.



$p(Z \le z_\alpha) = 1-\alpha$

$G(z_\alpha) = 1-\alpha$

$\alpha = 0.08$: $\quad G(z_\alpha) = 1-\alpha = 0.92$

$z_\alpha = \underline{\underline{1.41}}$

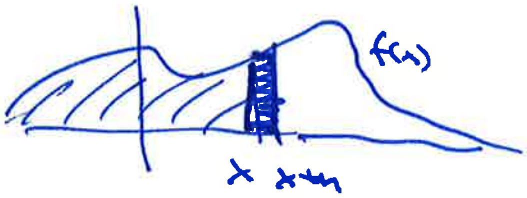
Kalk: $\boxed{\text{INV}}$ $\boxed{Z \geq p}$

$\alpha/2 = 0.04$: $\quad G(z_{\alpha/2}) = 1 - \frac{\alpha}{2} = 1 - \frac{0.08}{2} = 0.96$

$z_{\alpha/2} = \underline{\underline{1.75}}$

Opps. 11. Vis at $F'(x) = f(x)$.

Defn: $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$



$F'(x) = f(x)$ fordi:

$$\frac{F(x+h) - F(x)}{h} \simeq \frac{\cancel{h} \cdot f(x)}{\cancel{h}} = f(x)$$

når h er liten.

(1) Simultant fordelte stokastiske variabler

Vi tenker oss et stokastisk forsøk, og to variabler X og Y som avhenger av utfallet.

$$\boxed{p(x,y) = p(X=x, Y=y)}$$

simultan fordelingsfunksjon

Eks: Vi kaster to terninger, rød og blå.

X = rød terning

Y = største kast

x\Y	1	2	3	4	5	6
1	1/36	1/36	-	-	-	
2	0					
3						
4						
5						
6						

Kan finne de øvrige sannsynlighetene også, men gjorde det ikke på forelesning

$p(1,1) = 1/36 \ldots$

$p(1,2) = 1/36$

$p(2,1) = 0$

Eks:

x\Y	1	2	3	
1	0.1	0.2	0.2	0.5
2	0.1	0.1	0.3	0.5
	0.2	0.3	0.5	1

$p(1,1) = 0.1$

$p(1,2) = 0.2$

$\vdots$

$p(x,y) = p(X=x, Y=y)$

sannsynlighetsfordeling

Marginale sannsynlighetsfordelingene: rad/kolonnesummer

$$p_X(x) = \sum_{y=1}^{3} p(x,y) \qquad p_Y(y) = \sum_{x=1}^{2} p(x,y)$$

$p_X(1) = 0.5 \quad p_X(2) = 0.5$

$p_Y(1) = 0.2 \quad p_Y(2) = 0.3 \quad p_Y(3) = 0.5$

For X:

x	$p_X(x)$
1	0.5
2	0.5

$$E(X) = \sum_{x=1}^{2} x \cdot p_X(x) = 1 \cdot 0.5 + 2 \cdot 0.5 = \underline{\underline{1.5}}$$

$$Var(X) = \underline{E(X^2)} - 1.5^2 = 2.5 - 2.25 = \underline{\underline{0.25}}$$

$$E(X^2) = \sum_{x=1}^{2} x^2 \cdot p_X(x) = 1^2 \cdot 0.5 + 2^2 \cdot 0.5 = \underline{2.5}$$

$$\sigma_X = \sqrt{Var(X)} = \sqrt{0.25} = \underline{\underline{0.5}}$$

Kan finne $E(X)$, $Var(X)$, etc fra marginal fordeling $p_X(x)$.

For Y:

y	$p_Y(y)$
1	0.2
2	0.3
3	0.5

$$E(Y) = \sum_{y=1}^{3} y \cdot p_Y(y) = 1 \cdot 0.2 + 2 \cdot 0.3 + 3 \cdot 0.5 = \underline{\underline{2.3}}$$

$$Var(Y) = E(Y^2) - 2.3^2 = 5.9 - 5.29 = \underline{\underline{0.61}}$$

$$E(Y^2) = 1^2 \cdot 0.2 + 2^2 \cdot 0.3 + 3^2 \cdot 0.5 = 0.2 + 1.2 + 4.5 = 5.9$$

$$\sigma_Y = \sqrt{Var(Y)} = \sqrt{0.61} \simeq \underline{\underline{0.78}}$$

Forventning av $g(X,Y)$:

$$E[g(X,Y)] = \sum_{x=1}^{2} \sum_{y=1}^{3} g(x,y) \cdot p(x,y)$$

	1	2	3
1	0.1	0.2	0.2
2	0.1	0.1	0.3

$$E(XY) = \sum_{x,y} xy \cdot p(x,y)$$

$$= 1 \cdot 1 \cdot 0.1 + 1 \cdot 2 \cdot 0.2 + 1 \cdot 3 \cdot 0.2 + 2 \cdot 1 \cdot 0.1 + 2 \cdot 2 \cdot 0.1 + 2 \cdot 3 \cdot 0.3 = 1.1 + 2.4 = \underline{\underline{3.5}}$$

$$E(X) \cdot E(Y) = 1.5 \cdot 2.3 = \underline{3.45}$$

$\neq 0$

$Cov(X,Y) = 3.5 - 3.45 = \underline{\underline{0.05}}$

— Kan gjøres for kont. variabler også
— —"— for mer enn to variabler.

Se neste side for defn.

Defn: $\mathrm{Cov}(X,Y) = E[(X-\mu_X)\cdot(Y-\mu_Y)] = E(XY) - \mu_X \cdot \mu_Y$

Kovarians for X, Y

med $\mu_X = E(X)$

$\mu_Y = E(Y)$

$= E[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y]$

$= E(XY) - \mu_X \cdot E(Y) - \mu_Y E(X) + \mu_X \mu_Y$

$= E(XY) - \mu_X \cdot \mu_Y - \cancel{\mu_Y \cdot \mu_X} + \cancel{\mu_X \cdot \mu_Y}$

Regneregler:

i) $\mathrm{Cov}(X,Y) = E(XY) - \mu_X \cdot \mu_Y = E(XY) - E(X)\cdot E(Y)$

ii) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

iii) $E(aX + bY + c) = a \cdot E(X) + b\,E(Y) + c$

iv) $\mathrm{Var}(aX + bY + c) = a^2 \cdot \mathrm{Var}(X) + b^2 \cdot \mathrm{Var}(Y) + \underline{2ab \cdot \mathrm{Cov}(X,Y)}$

Eks: $\mathrm{Var}(X+Y) = \mathrm{Cov}(X+Y, X+Y)$

$= \underline{\mathrm{Cov}(X,X)} + \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,X) + \underline{\mathrm{Cov}(Y,Y)}$

$= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$

Defn: Korrelasjonskoeffisient til X og Y

$$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y}$$

(rho)

$= E\left[\frac{(X-\mu_X)}{\sigma_X} \cdot \frac{(Y-\mu_Y)}{\sigma_Y}\right]$

Merk:
- ρ er ~~ubenevnt~~ ubenevnt
- vi har alltid $-1 \le \rho \le 1$
- $\rho = 1 \Leftrightarrow Y = aX + b$ med $a > 0$

  $\rho = -1 \Leftrightarrow Y = aX + b$ med $a < 0$

Bevis: iv) $\mathrm{Var}(aX+bY+c) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X,Y)$

$$\mathrm{Var}(aX+bY+c) = \mathrm{Cov}(aX+bY+c,\ aX+bY+c)$$

$$= E[(aX+bY+c)^2] - E[aX+bY+c]^2$$

$$= E(a^2X^2 + b^2Y^2 + c^2 + 2abXY + 2acX + 2bcY) - (a\mu_x + b\mu_y + c)^2$$

$$= a^2E(X^2) + b^2E(Y^2) \cancel{+c^2} + 2abE(XY) \cancel{+2ac\mu_x} \cancel{+2bc\mu_y}$$
$$- a^2\mu_x^2 - b^2\mu_y^2 \cancel{-c^2} - 2ab\mu_x\mu_y \cancel{-2ac\mu_x} \cancel{-2bc\mu_y}$$

$$= a^2[E(X^2) - \mu_x^2] + b^2[E(Y^2) - \mu_y^2] + 2ab[E(XY) - \mu_x\mu_y]$$

$$= \underline{a^2\cdot \mathrm{Var}(X) + b^2\cdot \mathrm{Var}(Y) + 2ab\cdot \mathrm{Cov}(X,Y)}$$

Bevis: $-1 \le \rho \le 1$

$$\mathrm{Var}\left(\frac{X}{\sigma_x} + \frac{Y}{\sigma_y}\right) = \frac{1}{\sigma_x^2}\mathrm{Var}(X) + \frac{1}{\sigma_y^2}\mathrm{Var}(Y) + 2\cdot\frac{1}{\sigma_x}\cdot\frac{1}{\sigma_y}\mathrm{Cov}(X,Y)$$

$$= 1 + 1 + 2\rho = 2 + 2\rho$$

Siden $\mathrm{Var}(X/\sigma_x + Y/\sigma_y) \ge 0$, så er $2 + 2\rho \ge 0$

$$2\rho \ge -2$$
$$\underline{\rho \ge -1}$$

$$\mathrm{Var}\left(\frac{X}{\sigma_x} - \frac{Y}{\sigma_y}\right) = \frac{1}{\sigma_x^2}\mathrm{Var}(X) + \frac{1}{\sigma_y^2}\mathrm{Var}(Y) - 2\frac{1}{\sigma_x}\cdot\frac{1}{\sigma_y}\mathrm{Cov}(X,Y)$$

$$= 1 + 1 - 2\rho = 2 - 2\rho$$

Siden $\mathrm{Var}(X/\sigma_x - Y/\sigma_y) \ge 0$, så er $2 - 2\rho \ge 0$

$$2 \ge 2\rho$$
$$1 \ge \rho$$
$$\underline{\rho \le 1}$$

2 Uavhengighet

Defn: Simultant fordelte variabler X og Y er <u>uavhengige</u> hvis $p(x,y) = p_X(x) \cdot p_Y(y)$ for alle x, y.

$p(x,y)$: $p(X=x \text{ og } Y=y)$

$p_X(x)$: $p(X=x)$ $p_Y(y)$: $p(Y=y)$

A: $X=x$ B: $Y=y$

$$\boxed{p(A \cap B) = p(A) \cdot p(B)}$$

Eks:

X\Y	1	2	3	
1	0.1	0.2	0.2	0.5
2	0.1	0.1	0.3	0.5
	0.2	0.3	0.5	1

$(x,y)=(1,1)$: $p(1,1) = \underline{0.1}$

$p_X(1) \cdot p_Y(1) = 0.5 \cdot 0.2 = \underline{0.1}$

$(1,2)$: $p(1,2) = \underline{0.2} \neq$

$p_X(1) \cdot p_Y(2) = 0.5 \cdot 0.3 = \underline{0.15}$

X og Y <u>ikke uavhengige</u>

$\mathrm{Cov}(X,Y) = 0{,}05 \neq 0$

X\Y	1	2	3	
1	0.1	0.15	0.25	0.5
2	0.1	0.15	0.25	0.5
	0.2	0.3	0.5	1

X og Y uavhengige

Resultat:

X og Y uavhengige $\Longrightarrow$ $\mathrm{Cov}(X,Y) = 0$

Merk: Det motsatte gjelder ikke.

$\mathrm{Cov}(X,Y)=0$ kan gjelde selv om X,Y ikke er uavhengige

Hvorfor?

$$E[g(X)\cdot h(Y)] = E[g(X)]\cdot E[h(Y)]$$ når X og Y er uavhengige

Spesielt: $E(X\cdot Y) = E(X)\cdot E(Y)$

$$\Updownarrow$$

$$\mathrm{Cov}(X,Y) = E(XY) - \mu_X \cdot \mu_Y = 0$$

Viktig resultat:

X og Y uavhengige $\Longrightarrow$ $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Mer generelt:

$X_1, X_2, \ldots, X_n$ uavhengige :

$$\boxed{\begin{aligned}&\mathrm{Var}(a_1X_1 + a_2X_2 + \ldots + a_nX_n)\\ &= a_1^2\,\mathrm{Var}(X_1) + a_2^2\,\mathrm{Var}(X_2) + \ldots + a_n^2\,\mathrm{Var}(X_n)\end{aligned}}$$

Bevis: X og Y uavhengige $\Rightarrow$ Cov(X,Y) = 0

Anta X, Y uavhengige, dvs $p(x,y) = p_X(x) \cdot p_Y(y)$.

Da har vi:

$$E(XY) = \sum_{x,y} xy \cdot p(x,y) = \sum_{x,y} xy \cdot p_X(x) \cdot p_Y(y)$$

$$= \left[\sum_{x} x \cdot p_X(x)\right] \cdot \left[\sum_{y} y \cdot p_Y(y)\right] = E(X) \cdot E(Y)$$

$\Downarrow$

$$Cov(X,Y) = E(XY) - E(X)\,E(Y) = \underline{0}$$