TMA4245 Oppsummering Blok 1

$\boxed{\text{Deskriptiv Statistikk}}$ kap 1

$\boxed{\text{Sannsynlighetsteori}}$ kap 2-7 + notater

$\boxed{\text{Inferens}}$

Deskriptiv Statistikk

Hvordan beskrive en samling av observasjoner $x_1, x_2, \ldots x_n$ gjennom informative figurer og tall?

- Histogramm
- Boxplot
- Normal-kvantil plot
- Gjennomsnitt
- Median
- Varians og standard avvik

Eksam (Eks1)

Sannsynlighet (kap 2)

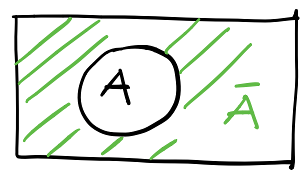
- Stokastisk forsøk : experiment der resultatet er underlagt tilfeldigheter

- Utfallsrommet : alle mulige utfall fra et forsøk S

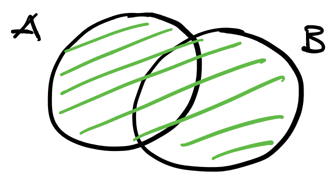
- Hendelse : $A \subseteq S$ delmengde av S

Operasjoner om hendelser

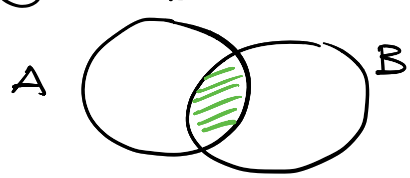
komplement $\bar{A} = \{e \in S \mid e \notin A\}$



Union $A \cup B = \{e \in S \mid e \in A \text{ eller } e \in B\}$



Snitt $A \cap B = \{e \in S \mid e \in A \text{ og } e \in B\}$



Definisjon av sannsynlighet

$P(A)$ er en funksjon fra alle mulige hendelser A i S til $\mathbb{R}$ slik at

1) $0 \leq P(A) \leq 1 \quad \forall A$

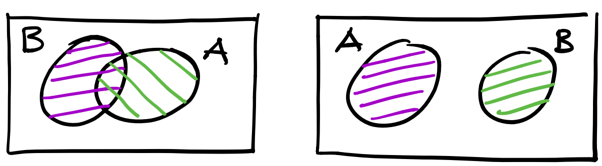
2) $P(S) = 1 \quad [\text{eller } P(\emptyset) = 0]$

3) $A_1, A_2, \ldots$ parvis disjunkte $[A_i \cap A_j = \emptyset \ \forall\, i \neq j]$

$$P(\cup A_i) = \sum P(A_i)$$

Regneregler for sannsynlighet

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



Komplement $\quad P(A) = 1 - P(\bar{A})$

Betinget sanns $\quad P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Uavhengighet $\quad P(A \cap B) = P(A)P(B)$

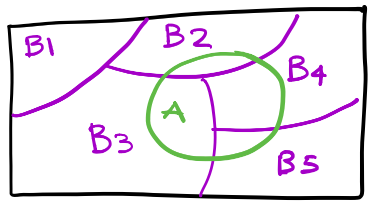
$P(A|B) = P(A)$

Bayes regel $\quad P(A|B) = \dfrac{P(B|A)\,P(A)}{P(B)}$

Lov om total sanns

$B_1\ B_2 \ldots\ B_n$ partisjon av S

$$P(A) = \sum P(A \cap B_i) = \sum P(A|B_i)\,P(B_i)$$



Eksam (Eks 2 - Eks 3)

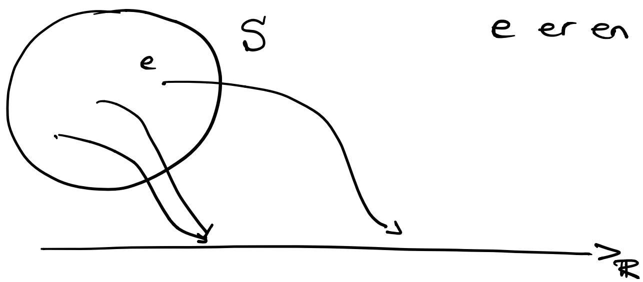
Telleregler – kombinatorikk

$n! = 1 \cdot 2 \cdot 3 \ldots n$ antall måter å ordne n elementer

$_nP_r = \frac{n!}{(n-r)!} = n(n-1)\ldots(n-r+1)$ antall ordnende utvalg når r elem velges blant n uten tilbakelegging

$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{_nP_r}{r!}$ antall ikke-ordnende utvalg når r elem velges blant n uten tilbakelegsing

STOKASTISK VARIABEL



e er en hendelse i S

En stokastisk variabel $X(e)$ på utfallsrommet S er en funksjon $X(e): S \to \mathbb{R}$.

$$\underbrace{X(e)}_{\text{Stok. Variabel (funksjon)}} = \underbrace{x}_{\text{realisasjon (tall)}}$$

Vi har
- diskret SV : X kan ta endelig (eller) tellbart antall verdier
- kontinuerlig SV : X kan ta uendelig antall verdier

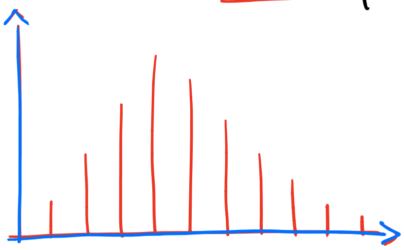
Hvordan beskriver vi en SV?

$$f(x) \begin{cases} \text{diskret SV : punktsannsynlighet } f(x) = P(X = x) \\ \text{kontinuerlig SV : sannsynlighetstetthet} \end{cases}$$

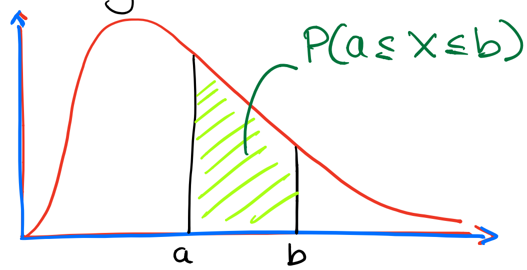
$F(x)$ – samme tålkning for diskret og kontinuerlig SV

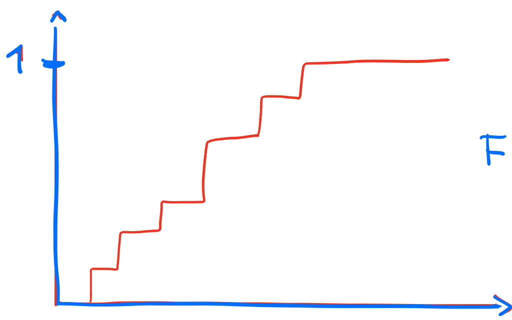
$$F(x) = P(X \leq x)$$

Sannsynlighetsfordelinger

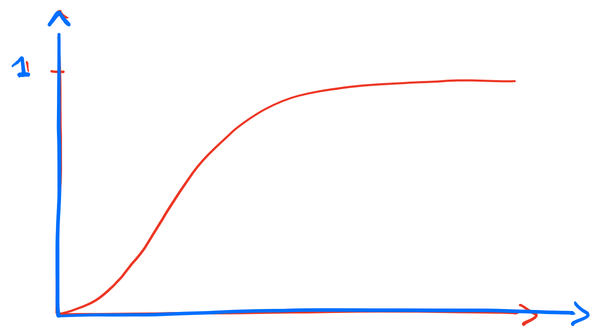


$f(x)$





$F(x)$



$$f(x) = P(X = x)$$

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$$

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx = F(b) - F(a)$$

$$F(x) = \int_{-\infty}^{x} f(u)\,du$$

To (eller flere) SV

- Simultan fordeling $\quad f(x,y)$
- Marginal fordeling
  - for X $\quad g(x) = \int_{\mathbb{R}} f(x,y)\,dy$
  - for Y $\quad h(y) = \int_{\mathbb{R}} f(x,y)\,dy$

- **Betinget fordeling** $f(x|y) = \frac{f(x,y)}{h(y)}$

$$f(y|x) = \frac{f(x,y)}{g(x)}$$

- **Uavhengighet** $f(x,y) = g(x) \cdot h(y)$ $\quad f(x|y) = g(x)$

Forventningsverdi og varians

Forventningsverdi $E(x) = \begin{cases} \sum x f(x) & \text{diskret} \\ \int x f(x)\,dx & \text{kontinuerlig} \end{cases}$

Regneregler: $E(x)$ er en linear operator

Tolkning: Gjennomsnitt av uendelig repetisjoner

Varians $Var(x) = E(x - E(x))^2 = \begin{cases} \sum (x - E(x))^2 f(x) \\ \int (x - E(x))^2 f(x)\,dx \end{cases}$

Regneregler

→ $Var(x) = E(x^2) - (E(x))^2$

→ $Var(aX+b) = a^2\,Var(x)$

→ $Var(X+Y) = Var(X) + Var(Y) + 2cov(X,Y)$

Tolkning Mål av spredning

NB $E(x)$ og $Var(x)$ er egenskaper av $f(x)$

$\bar{x}$ og s^2 er egenskaper av $(x_1, x_2, \dots x_n)$
og er også SV

Kovarians $\quad cov(X,Y) = E(X-E(X))(Y-E(Y))$

- Måler linear sammenheng mellom X og Y
- $X \perp Y \quad \Rightarrow \quad cov(X,Y) = 0$

$\quad\quad\quad \not\Leftarrow$

Eksam 2017 nov, opp 3 December 2009 opp1 Ek4 Eks 5

Viktige diskret sannsynlighetsfordelinger

Binomisk fordeling

: trekk fra urne med tilbakelegging
· kuler i 2 farge

n antall uavhengige forsøk

p sanns for suksess

X antall suksesser

$X \sim Binom(n, p)$

$$\boxed{f(x; p, n) = \binom{n}{k} p^x (1-p)^{n-x}}$$

$$\boxed{\begin{aligned} E(X) &= np \\ Var(X) &= np(1-p) \end{aligned}}$$

Multinomisk fordeling

k mulig utfall i hver forsøk

Hypergeometrisk fordeling

trekk fra urne uten tilbakelegging

N # kuler totalt

k # suksesser

n # kuler trukket

$X \sim HyperGeo(N, n, k)$

$$\boxed{f(x) = \frac{\binom{k}{x}\binom{N-k}{n-k}}{\binom{N}{n}}}$$

Negativ binomial

Bernoulli process, trekk til k suksesser

k – antall suksesser

p – sanns for suksess

X – # forsøk

$$\boxed{f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}} \quad x = k, k+1 \dots$$

Poisson :

λ : forventet antall suksesser i intervallet av lengde 1

X : # hendelser i $(0,t)$ i en Poisson prosess

$$\boxed{f(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}} \quad x = 0, 1, 2 \dots$$

$$\boxed{E(X) = \lambda t} \qquad \boxed{Var(X) = \lambda t}$$

- Eksam november 2017 – Oppgave 2
- Eksam mai 2009 – Oppgave 2

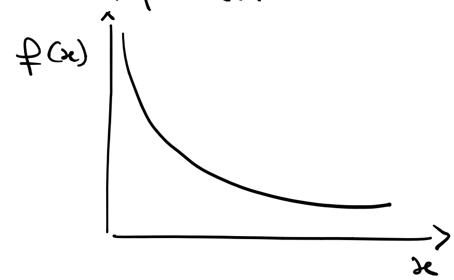
Eks 6/
Eks 7

Viktige kontinuerlig forventning

Eksponensiel fordeling

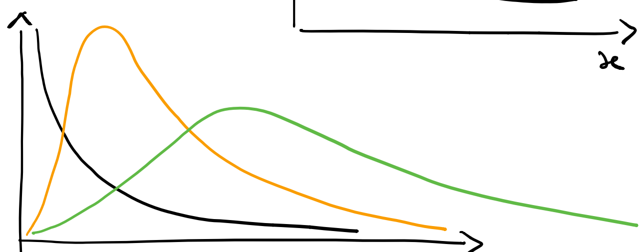
X : tid mellom to hendelser i en Poisson prosess

$$\boxed{f(x) = \lambda e^{-\lambda x}} \qquad x > 0$$



Gamma fordeling

To parametre α β



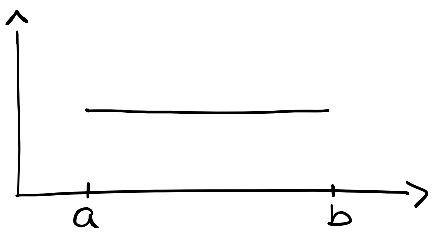
$\alpha = 1$, $\beta = 1/\lambda$ $\rightarrow$ Eksponensiell (λ)

$\alpha = \nu/2$, $\beta = 2$ $\longrightarrow$ χ^2_ν

Uniform

$f(x)$ konstant i

en intervall



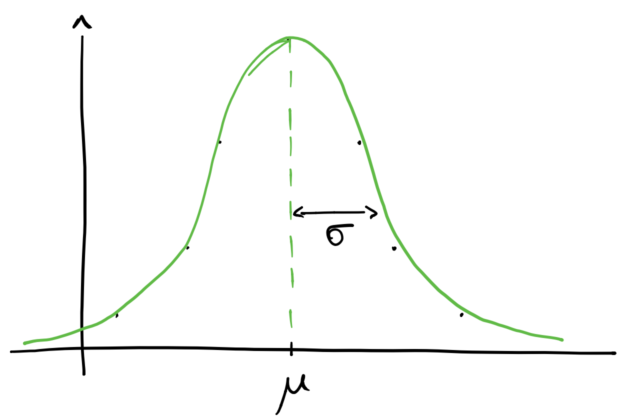
Normal fordeling

To param

μ - forventningsverdi

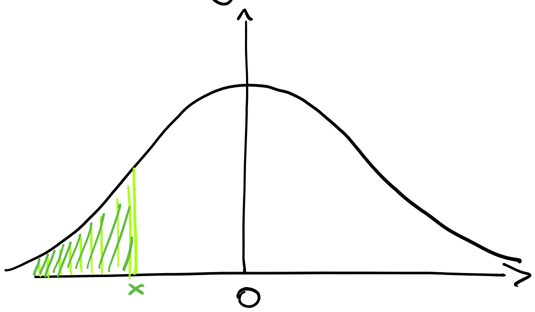
σ^2 - varians

$X \sim N(\mu, \sigma^2)$



- linear kombinasjon av uavhengige normal sv er normal fordelt

- $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$ standard normal fordeling



Dette finner vi i tabeller

August 2015 opp 1 Eks 8

Forhold mellom noen fordelinger

→ Hypergeometrisk $\simeq$ binomial når N er stor i forhold til n

→ Binomisk $\simeq$ poisson når n stor og p liten

→ Binomisk $\simeq$ normal når n er stor

Transformasjoner av SV

3 tilfeller

1) $Y = u(x)$ $u(\cdot)$ er en-entydig

2) $Y = \max(x_1, x_2, \dots x_n)$

$Y = \min(x_1, x_2, \dots x_n)$

$Y = X_{(i)}$ ordensvariabel

3) $Y = a + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$

1)

Start fra $F_Y(y)$

Teorem $\quad g_Y(y) = f(w(y))\,|w'(y)| \quad w(y) = u^{-1}(y)$

2) Start fra $F_Y(y)$

$V = \max(x_1, \dots x_n) \qquad F_V(v) = [F_X(v)]^n$

$U = \min(x_1, \dots x_n)$

3 Moment genererende funksjon

$$\boxed{M_X(t) = E(e^{tX}) = \int e^{tx} f(x)\,dx}$$

$\boxed{\text{Regneregler}}$

→ $M_{X+a}(t) = e^{at} M_X(t)$

→ $M_{aX}(t) = M_X(at)$

→ $X_1, X_2, \ldots X_n$ uav.

$Y = X_1 + X_2 + \ldots + X_n$

$M_Y(t) = M_{X_1}(t) M_{X_2}(t) \ldots M_{X_n}(t)$

$\boxed{\text{Egenskaper}}$

→ $M_X^{(r)}(0) = \mu_r$

→ En til en forhold mellom $M_X(t)$ og $f_X(x)$

Eksam August 2000