

Om kurset

Generelt:

- Grunnleggende metodefag.
- Skal lære om grunnleggende ting i sannsynlighetsregning og statistikk.
- Viktig grunnlagsfag for andre kurs senere i studiet.
- Flere av temaene som tas opp i kurset vil starte som repetisjon av ting dere har hatt om i grunnkurset i sannsynlighetsregning og statistikk - vi vil imidlertid nå gå dypere inn i tingene og kreve mer forståelse/bedre beherskelse.
- Svært viktig å jobbe med stoffet på egen hånd - å regne oppgaver er et "must".

Innhold:

- Stokastiske variable, sannsynlighetsfordelinger. Repetisjon.
- Forventningsverdi og varians. Repetisjon, pluss mer om forventningsverdi/variens til funksjoner av stokastiske variable.
- Sannsynlighetsfordelinger. Repetisjon, pluss om flere fordelinger. Sentralgrenseteoremet.
- Poissonprosesser.
- Funksjoner av stokastiske variable.
- Ekstremverdivariable.
- Estimering, sannsynlighetsmaksimeringsprinsippet.
- Konfidensintervall og hypotesetesting, i andre situasjoner enn omhandlet i grunnkurset.

Sannsynlighet

En kort repetisjon av de grunnleggende tingene (se eventuelt kapittel 2 i boka for en grundigere repetisjon).

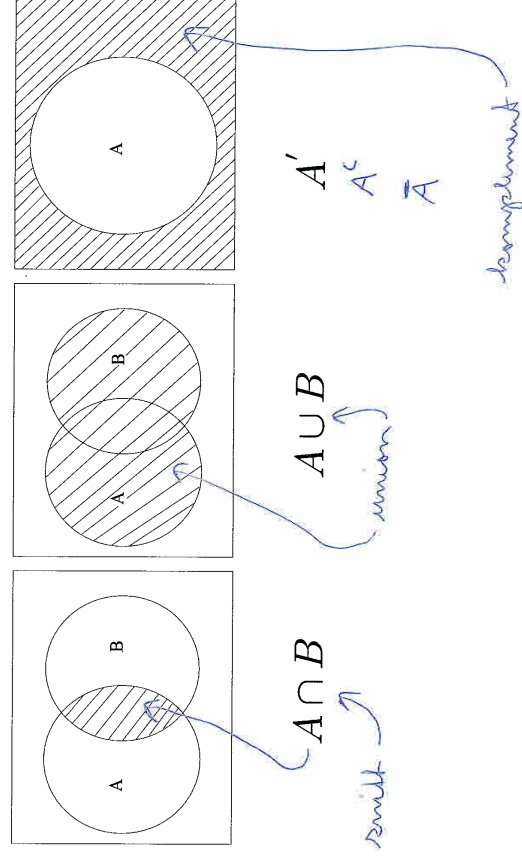
Noen grunnleggende definisjoner:

- Stokastisk forsøk: Et eksperiment der resultatet er underlagt tilfeldigheter.
- Utfallsrom, S : Mengden av mulige utfall av et stokastisk forsøk.
- Hendelse: En delmengde av utfallsrommet, $A \subset S$.
- Enkeltutfall, e : Et enkelt element i S .

Eksempler:

- Kast med terning: $S = \{1, 2, 3, 4, 5, 6\}$,
 $A = \{1, 3, 5\}$, $e = \{2\}$.
"Gulde"
- Levetid for lyspære: $S = \{t; t \geq 0\}$,
 $A = \{t; 0 \leq t < 5\}$. $\square$

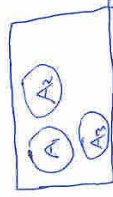
Snitt, union og komplement (illustrert ved Venn-diagram):



Sannsynlighet for en hendelse:

- $0 \leq P(A) \leq 1$. $\hookrightarrow P(A)$
- $P(S) = 1$.
- Dersom $A_1, A_2, \dots$ alle er gjensidig utelukkende (dvs $A_i \cap A_j = \emptyset$ for alle $i \neq j$) så er:

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$



Eksempel: Kaster en mynt to ganger.

Utfallsrom, $S = \{KK, KM, MK, MM\}$. La $A =$ "minst en mynt" $= \{KM, MK, MM\}$. Da er

$$P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

siden alle enkeltutfallene er like sannsynlige. □

Noen regneregler:

- $P(A') = 1 - P(A)$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Betinge sannsynlighet:

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

"gitt" "antall"

Dette er sannsynligheten for at A inntreffer når vi vet at B har inntruffet.

Eksempel:

Terningkast, la $A = \{1, 3, 5\} =$ "oddetall" og $B = \{4, 5, 6\} =$ "større enn 3". Da er $A \cap B = \{5\}$ og vi får at:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3}$$

Dvs vet vi at terningkastet gav et resultat "større enn 3" (og det er all informasjon vi har) er sannsynligheten for at resultatet ble et oddetall lik $1/3$. □

Stokastiske variable og sannsynlighetsfordelinger

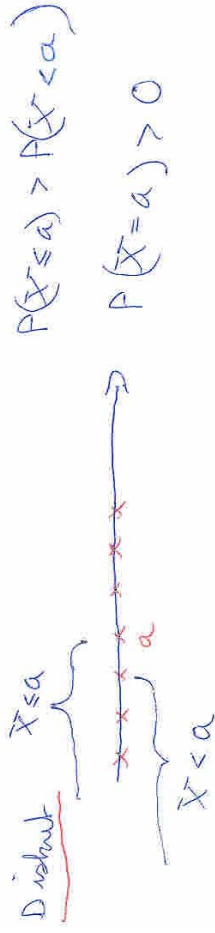
Dekkes av kapittel 3 i læreboka. *Tilfeldig variabel*

Definisjon: Stokastisk variabel er en variabel som får sin numeriske verdi bestemt av utfallet av et stokastisk forsøk. $\square$

Eksempler: Oljeinnhold i boreprøve, funksjonstid til elektronisk komponent, temperatur i morgen kl 12, antall feil i ny programkode, antall ulykker på en veistrekning, resultatet av et terningkast, osv (de fleste variable kan sees på som stokastiske). *mulige verdier: x x x x x x x* $\rightarrow \square$

Stokastiske variable er enten diskrete (f.eks. antall feil, antall ulykker, terningkast) eller kontinuerlige (f.eks. oljeinnhold, funksjonstid, temperatur). En diskret variabel kan kun anta et tellbart antall verdier, mens en kontinuerlig variabel kan anta enhver verdi i et intervall.

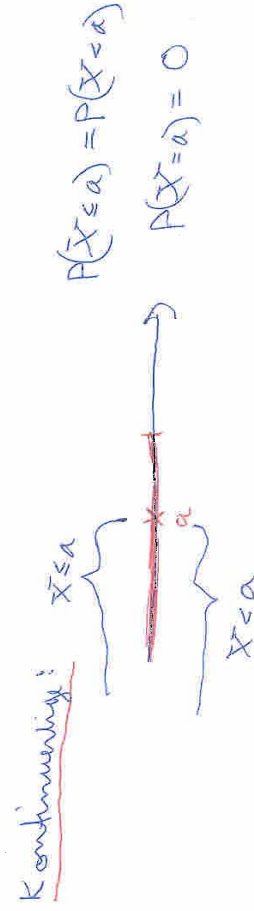
mulige verdier: x x x x x x x $\rightarrow$

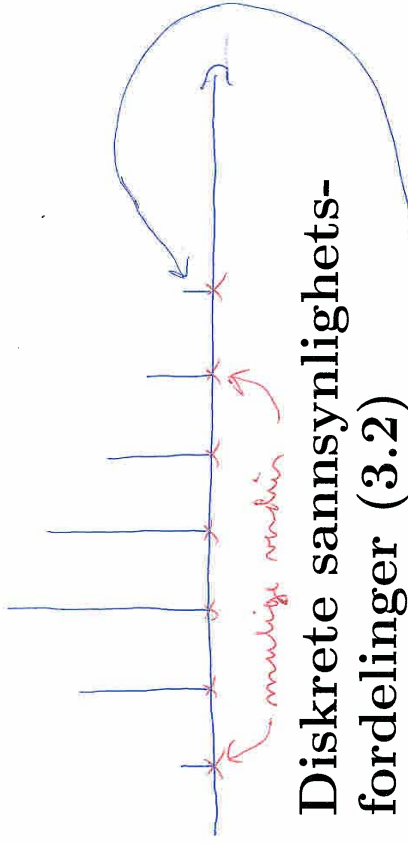


Notasjon: Stokastiske variable betegnes med store bokstaver, X, Y, Z , osv ("før forsøket er utført"), mens numerisk verdi til den stokastiske variabelen ("etter forsøket er utført") betegnes med små bokstaver, x, y, z , osv. $\square$

Diskrete og kontinuerlige stok. var. må behandles noe forskjellig. For diskrete variable gir det mening å snakke om $P(X = a)$, mens for kontinuerlige variable vil alltid $P(X = a) = 0$. For kontinuerlige variable ser vi bare på sannsynligheter for å få verdier i et intervall, f.eks. $P(a \leq X \leq b)$.

*skan x
likn x*





Diskrete sannsynlighetsfordelinger (3.2)

Definisjon: $f(x)$ er punktsannsynligheten til den diskrete stok. var. X dersom for alle mulige utfall x :

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$
3. $f(x) = P(X = x)$

□

Eksempel: Myntkast, kaster en mynt 3 ganger. La $X =$ "antall kron".

Utfallsrom	x	Utfallsrom	x
KKK	3	KMM	1
KKM	2	MKM	1
KMK	2	MMK	1
MKK	2	MMM	0

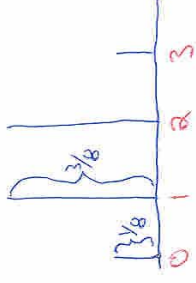
Vi ser at:

$$f(0) = P(X = 0) = 1/8$$

$$f(1) = P(X = 1) = 3/8$$

$$f(2) = P(X = 2) = 3/8$$

$$f(3) = P(X = 3) = 1/8$$



Dette er punktsannsynligheten til X . □

Definisjon: Den kumulative fordelingen til en diskret stok. var. X er:

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

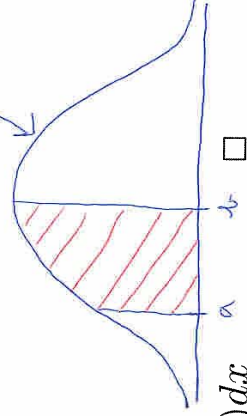
! eksempel:

$$\begin{aligned} F(2) &= P(X=0) + P(X=1) + P(X=2) \\ &= 1/8 + 3/8 + 3/8 = 7/8 \end{aligned}$$

Kontinuerlige sannsynlighetsfordelinger (3.3)

Definisjon: $f(x)$ er sannsynlighetstettheten til den kontinuerlige stok. var. X dersom:

1. $f(x) \geq 0$ for alle $x \in \mathcal{R}$
2. $\int_{-\infty}^{\infty} f(x) dx = 1$
3. $P(a < X < b) = \int_a^b f(x) dx$ $\square$

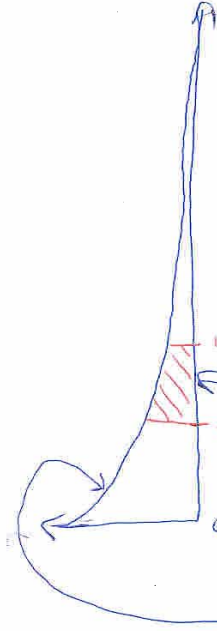


Merk:

Diskret: $P(a \leq X \leq b) = \sum_{a \leq x \leq b} f(x)$

Kontinuerlig: $P(a \leq X \leq b) = \int_a^b f(x) dx$

Merk også at for diskrete stok. var. er $P(X \leq a) \neq P(X < a)$, mens for kontinuerlige stok. var. er $P(X \leq a) = P(X < a) + P(X = a) = P(X < a)$.



Eksempel: La Y = "funksjonstid til elektronisk komponent". Anta at det er kjent at Y har sannsynlighetstetthet:

$$f(y) = \begin{cases} \frac{1}{5} e^{-y/5} & , y \geq 0 \\ 0 & , \text{ellers} \end{cases}$$

(Sjekk selv at $\int_{-\infty}^{\infty} f(y) dy = 1$). Vi kan nå regne ut sannsynligheter, f.eks:

$$\begin{aligned} P(Y \leq 5) &= \int_{-\infty}^5 f(y) dy = \int_0^5 \frac{1}{5} e^{-y/5} dy = \left[-e^{-y/5} \right]_0^5 \\ &= -e^{-1} + e^0 = \underline{\underline{0.63}} \\ P(4 < Y < 5) &= \int_4^5 f(y) dy = \int_4^5 \frac{1}{5} e^{-y/5} dy = \left[-e^{-y/5} \right]_4^5 \\ &= -e^{-1} + e^{-4/5} = \underline{\underline{0.08}} \quad \square \end{aligned}$$

Definisjon: Den kumulative fordelingen til en kontinuerlig stok. var. X er:

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

Funksjonstideksemplet:

$$F(y) = \int_{-\infty}^y f(t)dt = \int_0^y \frac{1}{5}e^{-t/5}dt = 1 - e^{-y/5}$$

$$F(5) = 1 - e^{-5/5} = \underline{\underline{0.63}} \quad \square$$

||

$$P(X \leq 5)$$

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