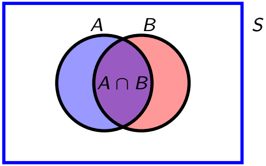
Plan for dagens undervisning

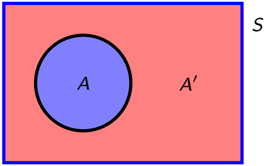
* Regne på eksempler/situasjoner som inkluderer
  - komplementærsetningen
  - betinget sannsynlighet
  - uavhengige hendelser
  - Bayes regel og total sannsynlighet
  - diskret stokastisk variabel, punktsannsynlighet og kumlativ fordeling
  - kontinuerlig stokastisk variabel, sannsynlighetstetthet og kumulativ fordeling
  - simultanfordeling
  - uavhengige stokastiske variabler

Additive regneregler for sannsynlighet

⋆ Sannsynlighet for union av to hendelser: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



⋆ Komplementærsetningen: $P(A') = 1 - P(A)$



Betinget sannsynlighet og uavhengige hendelser

Definisjon (Betinget sannsynlighet)

La A og B være hendelser i et utfallsrom S og anta $P(B) > 0$. Den betingede sannsynligheten for A gitt B er da

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

- $\star$ Multiplikasjonssetningen: $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$

Definisjon (Uavhengige hendelser)

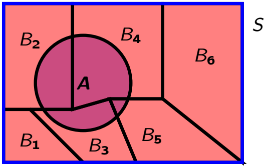
To hendelser A og B er uavhengige hvis

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

Hvis ikke sies hendelsene A og B å være avhengige.

Total sannsynlighet og Bayes regel

- Setningen om total sannsynlighet
  - la A være en hendelse i et utfallsrom S
  - la $B_1, B_2, \ldots, B_n$ være partisjon av S (dvs $B_i \cap B_j = \emptyset$, $i \neq j$ og $B_1 \cup B_2 \cup \ldots \cup B_n = S$)



$$P(A) = \sum_{i=1}^{n} P(A \cap B_i)$$
$$= \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

- Bayes regel

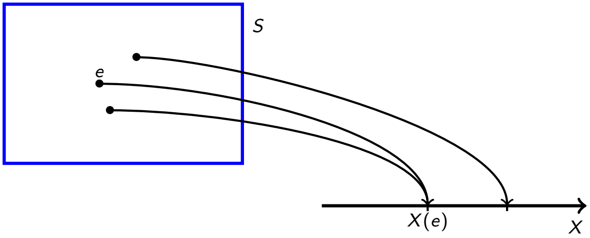
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

Stokastiske variabler

Definisjon (Stokastiske variabler)

En stokastisk variabel (SV) er en reell funksjon definert på et utfallsrom.

- La X være en stokastisk variabel definert på et utfallsrom S
  - $X(e) \in \mathbb{R}, e \in S$
  - $X : S \to \mathbb{R}$



- Notasjon: Bruker store bokstaver sent i alfabetet til å betegne SV, typisk X, Y, Z, U, V, eventuelt $X_1, X_2, \ldots$
- Man skiller mellom tre situasjoner:
  - $X(e)$ kan ta endelig mange verdier, f.eks. $0, 1, 2, 3, 4, 5$
  - $X(e)$ kan ta tellbart uendelig mange verdier, f.eks. $1, 2, \ldots$
  - $X(e)$ kan ta ikke-tellbart uendelig mange verdier, f.eks. $[-2, 3]$
- De to første situasjonene gir en diskret SV, og i den siste en kontinuerlig SV

For diskret SV: Punktsannsynlighet

- ⋆ For et tall x la
$$A_x = \{e \in S | X(e) = x\}$$
- ⋆ Vi lar da
$$P(X = x) = P(A_x) = P(\{e \in S | X(e) = x\})$$
- ⋆ Tilsvarende definerer vi
$$P(X \leq 3) = P(\{e \in S | X(e) \leq 3\})$$
$$P(a < X \leq b) = P(\{e \in S | a < X(e) \leq b\})$$

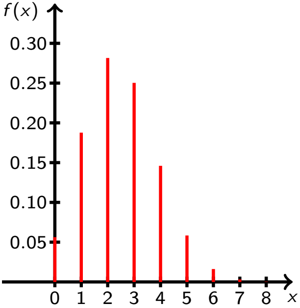
Definisjon (Punktsannsynlighet)

En funksjon $f(x)$ kalles punktsannsynlighet for en SV X dersom, for alle x,
$$f(x) = P(X = x)$$

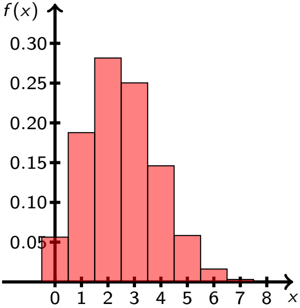
- ⋆ Merk at vi vil ha
  - $f(x) \geq 0$ for alle x
  - $\sum_x f(x) = 1$

Stolpediagram og sannsynlighetshistogram

Stolpediagram

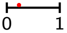


Sannsynlighetshistogram



For kontinuerlig SV: Sannsynlighetstetthet

- ⋆ Merk: punktsannsynligheter kan ikke brukes til å angi sannsynligheter for kontinuerlige SV
- ⋆ Eksempel:
  - Anta at vi slipper ei (infinitesimal) lita kule ned på intervallet $[0, 1]$ på en slik måte at alle posisjoner på intervallet er «like sannsynlige».
  - La X være posisjonen til kula.
  - Hva er da $P(X = 0.58)$?
  - Hvis vi tenker uniform sannsynlighetsmodell får vi



$$P(X = 0.58) = \frac{g}{m} = \frac{1}{\infty} = 0$$

Definisjon (Sannsynlighetstetthet)

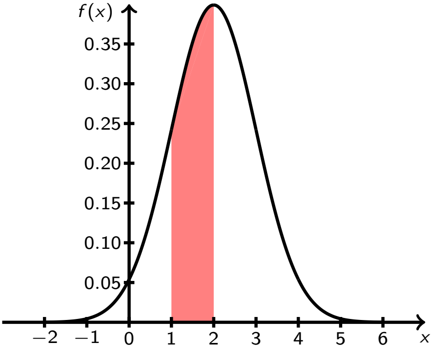
Sannsynlighetstettheten $f(x)$ for en kontinuerlig SV X er gitt ved

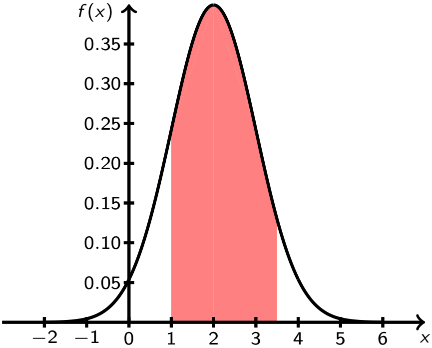
$$P(a < X \leq b) = \int_a^b f(x)dx$$

for alle $a < b$.

- ⋆ Merk at vi vil ha
  - $f(x) \geq 0$ for alle $x \in \mathbb{R}$
  - $\int_{-\infty}^{\infty} f(x)dx = 1$

Sannsynlighetstetthet





$$P(1 < X \leq 2) = \int_1^2 f(x)dx \qquad P(1 < X \leq 3.5) = \int_1^{3.5} f(x)dx$$

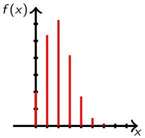
- Merk: For kontinuerlige SV får vi

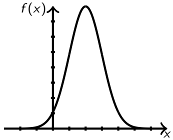
$$P(X \leq x) = P(X < x) + P(X = x) = P(X < x)$$
$$P(a < X \leq b) = P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b)$$

Kumulativ fordeling

- ⋆ Husk: punktsannsynlighet og sannsynlighetstetthet

$f(x) = P(X = x)$ $\qquad$ $P(a < X \leq b) = \int_a^b f(x)dx$





Definisjon (Kumulativ fordeling)

Kumulativ fordeling for en stokastisk variabel X er

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

- ⋆ Merk: $F(x)$ er gitt fra $f(x)$
  - diskret SV: Hvis mulige verdier for X er $0, 1, 2, \ldots$

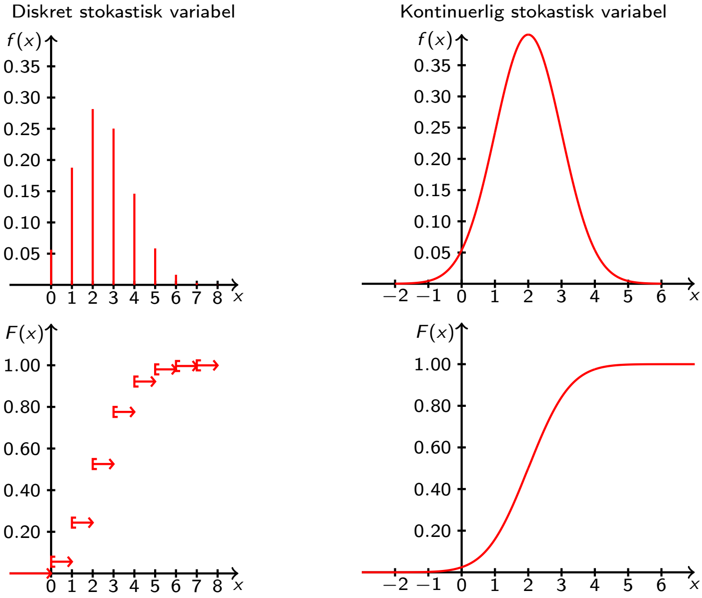
$$F(x) = \sum_{t=0}^{x} f(t) \qquad \Leftrightarrow \qquad f(x) = F(x) - F(x-1)$$

  - kontinuerlig SV:

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

- ⋆ Vi kan ikke alltid finne en formel for $F(x)$

Visualisering av $f(x)$ og $F(x)$



Simultanfordeling

Definisjon (Simultanfordeling)

Hvis vi har to diskrete SV X og Y er deres simultanfordeling (simultan punktsannsynlighet) $f(x,y)$ gitt ved

$$f(x,y) = P(X = x, Y = y).$$

Hvis vi har to kontinuerlige SV X og Y er deres simultanfordeling (simultan sannsynlighetstetthet) $f(x,y)$ gitt ved

$$P((X,Y) \in A) = \iint_A f(x,y) dx\, dy,$$

der $A \subseteq \mathbb{R}^2$.

- ⋆ Merk:
  - noen ganger skriver vi $f_{XY}(x,y)$ i stedet for $f(x,y)$ for å understreke at det er simultanfordelingen for de SV X og Y
  - når X og Y er diskrete SV må $f(x,y)$ oppfylle

$$f(x,y) \geq 0$$
$$\sum_x \sum_y f(x,y) = 1$$

  - når X og Y er kontinuerlige SV må $f(x,y)$ oppfylle

$$f(x,y) \geq 0$$
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx\, dy = 1$$

Marginalfordeling

- La $f_{XY}(x,y)$ være simultanfordelingen til de to stokastiske variablene X og Y
  - hva er da (marginal)fordelingen til X, $f_X(x)$?

- Hvis X og Y er diskrete SV

$$f_X(x) = P(X=x) = P\left(\bigcup_y (X=x, Y=y)\right) = \sum_y P(X=x, Y=y) = \sum_y f_{XY}(x,y)$$
$$f_Y(y) = P(Y=y) = P\left(\bigcup_x (X=x, Y=y)\right) = \sum_x P(X=x, Y=y) = \sum_x f_{XY}(x,y)$$

- Hvis X og Y er kontinuerlige SV

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

Betinget fordeling

Betinget sannsynlighet: $P(A|B) = \frac{P(A\cap B)}{P(B)}$
Multiplikasjonssetningen: $P(A \cap B) = P(A|B)P(B)$
Bayes regel: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

Definisjon (Betinget fordeling)

La X og Y være SV med simultanfordeling $f_{XY}(x, y)$. Den betingede fordelingen for Y gitt $X = x$ er da

$$f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)} \quad \textit{hvis} \quad f_X(x) > 0,$$

og den betingede fordelingen for X gitt $Y = y$ er

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)} \quad \textit{hvis} \quad f_Y(y) > 0.$$

- ⋆ Multiplikasjonssetningen for fordelinger

$$f_{XY}(x, y) = f_{Y|X}(y|x)f_X(x)$$
$$f_{XY}(x, y) = f_{X|Y}(x|y)f_Y(y)$$

- ⋆ Bayes regel for fordelinger

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}$$

- ⋆ Setningen om total sannsynlighet

$$f_X(x) = \sum_y f_{X|Y}(x|y)f_Y(y) \quad \text{(diskret SV)}$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(y) \quad \text{(kontinuerlig SV)}$$

Uavhengige stokastiske variabler

Hendelsene A og B er uavhengige hvis
$P(A|B) = P(A) \Leftrightarrow P(A \cap B) = P(A)P(B)$

Definisjon (Uavhengige stokastiske variabler)

To stokastiske variabler X og Y er uavhengige hvis og bare hvis

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y) \quad \textit{for alle} \quad (x, y).$$

- ⋆ Hvis X og Y er uavhengige og $f_X(x) > 0$ får vi også

$$f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{f_X(x) \cdot f_Y(y)}{f_X(x)} = f_Y(y)$$

- ⋆ Tilsvarende hvis X og Y er uavhengige og $f_Y(y) > 0$ får vi

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{f_X(x) \cdot f_Y(y)}{f_Y(y)} = f_X(x)$$