Bayes Teorem. Teorem 2.14

La $B_1, B_2, \dots, B_n$ vere ein partisjon av S.

$\left(\bigcup_{i=1}^{n} B_i = S,\ B_i \cap B_j = \phi,\ i \neq j.\ P(B_i) > 0,\ i = 1, 2, \dots, n \right.$

La A vere ei hending i S slik at $P(A) > 0$

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

Eks. Test på sjukdom som 2% av populasjonen lir av (Korona)

La $A \sim$ tilfeldig testa person har positiv respons

$B \sim$ tilfeldig ~~testa~~ valgt person har sjukdommen.

Under B: $P(B) = 0.02$, $P(A|B) = 0.99$, $P(A'|B) = 0.01$

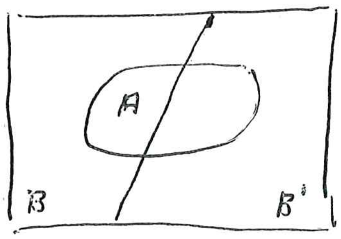
Under B' $P(B') = 0.98$. $P(A|B') = 0.01$, $P(A'|B') = 0.99$

$P(A|B)$ blir kalla sensitivitet.

$P(A'|B')$ blir kalla ~~sensitivitet~~, spesifisitet

Vil finne $P(B|A)$.

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



$$= \frac{0.99 \cdot 0.02}{0.99 \cdot 0.02 + 0.01 \cdot 0.98} = 0.67$$

$P(B) = 0.01 \Rightarrow P(B|A) = \frac{1}{2}$.

$\Rightarrow A = (A \cap B) \cup (A \cap B')$

Kap 3. Stokastiske (tilfeldige) variable og sannsyns-fordelingar.

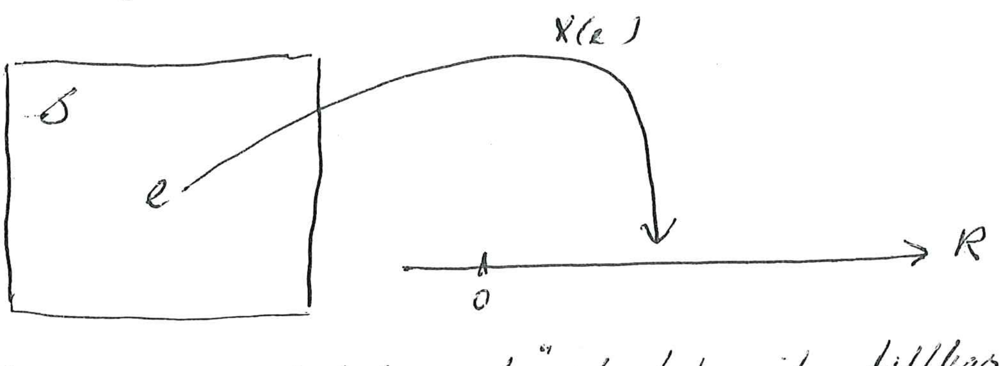
3.1. Stokastiske variable.

Eks. Eit oljeselskap skal vurdere om utvinning er lønnsomt. Seismiske undersøkingar indikerer at brønnen kan vere av 4 typar med følgjande sannsyn

Utfallsrom	$P(e_i)$	Gevinst $X(e_i)$
e_1	0.2	-20
e_2	0.1	5
e_3	0.2	5
e_4	0.5	10

Definisjon 3.1

Ein stokastisk variabel er ein funksjon frå utfallsrommet og inn i R som knyter tal til (hendingar eller utfall i utfallsrommet).



Dersom variabelen høgst tek eit tellbart tal av verdiar, seier vi at variabelen er diskret.

Dersom ein stokastisk variabel kan ta verdiar på ein kontinuerleg skala (alle verdiar i R eller ei delmengde av R) seiast variabelen å vere kontinuerleg.

Diskrete sannsynsfordelingar

Funksjonen $f(x) = P(X=x) \stackrel{a}{=} P\{e \mid e \in S \text{ og } X(e)=x\}$ blir kalla punktsannsynet til den diskret variabelen X.

Eks. Oljeselskap. $P(X=5) = P(e \mid e \in S \text{ og } X(e)=5)$

$= P(e_2) + P(e_3) = 0.3.$

Eit punktsannsyn må oppfylle

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

Mengda av ordna par $(x, f(x))$ blir kalla ei sannsynsfordeling for X.

Eks. Oljeselskap.

x	-20	5	10
$P(X=x)$	0.2	0.3	0.5

Eksempel 3.4

Hos ein bilselgar er 50% av dei selde bilane elbilar. La X vere talet på selde elbilar for dei 4 første selde bilane

Finn $P(X=k)$, $k = 0, 1, 2, 3, 4$

La $B_0 \sim$ sel ikkje elbil, $B_1 \sim$ sel elbil

$S = \{(B_0 B_0 B_0 B_0), (B_1, B_0, B_0, B_0), \ldots (B_1, B_1, B_1, B_1)\}$

i alt 16 enkeltutfall alle like sannsynlege.

$$P(X = k) = \frac{\binom{4}{k}}{16}, \quad k = 0, 1, \ldots, 4.$$ Vi får

k	0	1	2	3	4
$P(X=k)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

$$P(X \le 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = \underline{\frac{15}{16}}$$

<u>Def. 3.5</u>

Kumulativ fordelings-funksjon, $F(x)$, til ein diskret variabel X med sannsyn $f(x)$ er:

$$F(x) = P(X \le x) = \sum_{t \le x} f(t), \quad -\infty < x < \infty$$

<u>Eks. Bilselgar</u>

$F(0) = f(0) = \frac{1}{16}$

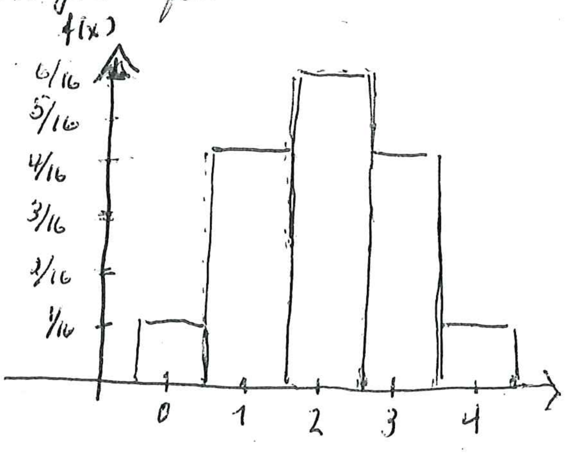
$F(1) = f(0) + f(1) = \frac{5}{16}$

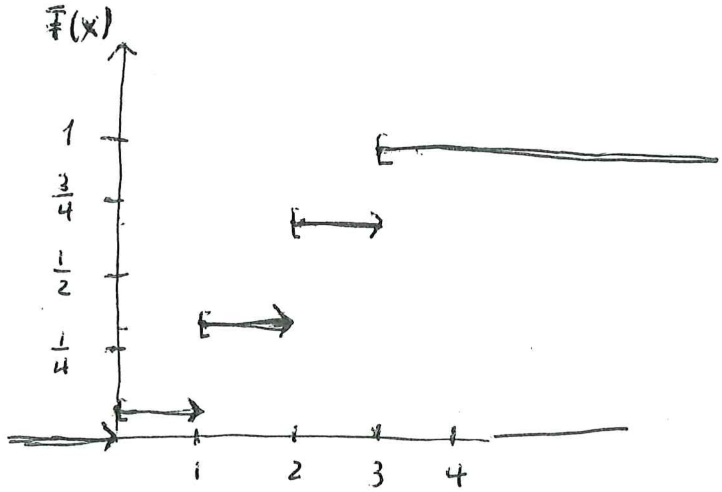
$F(2) = f(0) + f(1) + f(2) = \frac{11}{16}$

$F(3) = f(0) + f(1) + f(2) + f(3) = \frac{15}{16}$

$F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = 1$

Grafisk form



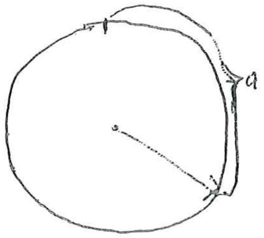


3.3 Kontinuerlege sannsynsfordelingar

Eks. X = høgda av tilfeldig valgt mann

Y = levetida til ei lyspære

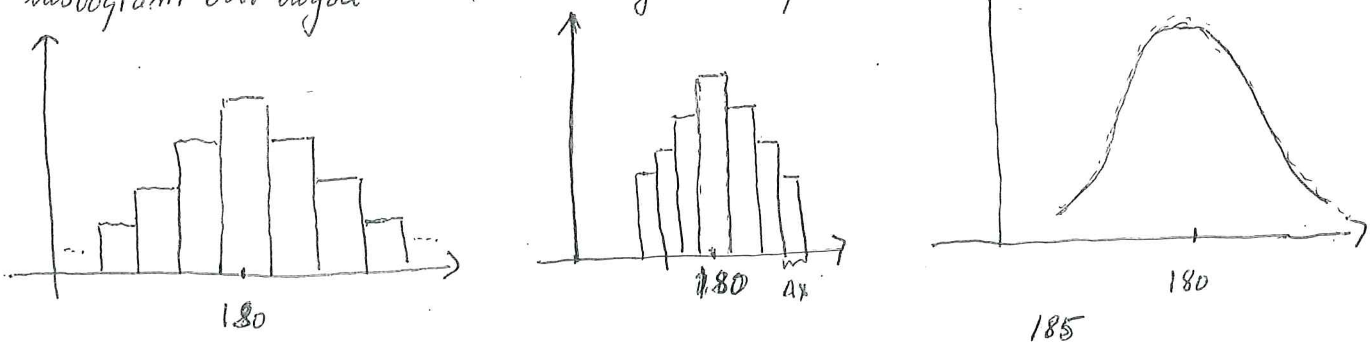
Z = avstanden, a, målt i radianar frå eit visst startpunkt til der ei pil stansar $0 \leq a \leq 2\pi$



For kontinuerleg fordelte variable er alle punktsannsyn 0.

Men naturleg at $P(0 \leq Z \leq a) = \frac{a}{2\pi}$

histogram over høgde → mange observasjonar — uendeleg mange



$$P(175 \leq X \leq 185) \approx \sum_{175 \leq x \leq 185} f(x)\Delta x \longrightarrow \int_{175}^{185} f(x)\,dx$$

Legg merke til at vi har:

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

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