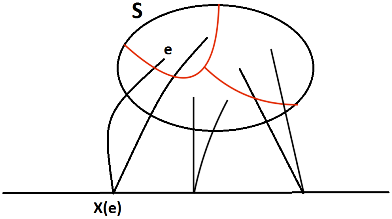
Repetisjon 5

Stokastisk Variabel (SV)

En **stokastisk variabel** X er en reel funksjon på utfallsrommet S

$$X : S \to \mathcal{R}$$

$$X(e) = x$$



- Stor bokstav X for den stokastiske variabelen
- Liten bokstav x for en bestemt realisering $X(e) = x$

Stokastisk Variabel

To situasjoner:

- **Diskret** hvis mengden av mulige realisering er endelig eller tellbar
- **Kontinuerlig** hvis mengden av mulig realisering er ikke tellbart, typisk $\mathcal{R}$, eller (a, ∞) eller (a, b)

Diskret Stokastisk variabel - I

Hvordan beskriver vi en diskret SV X

- Sannsynlighet fordeling (punkt sannsynlighet)

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

  - Merk: $P(X = x) = P(e \in \mathcal{S} | X(e) = x)$
- Egenskaper til f
  - $0 \leq f(x) \leq 1$ for alle mulige verdier av x
  - $\sum_x f(x) = 1$
- Kumulative sannsynlighet fordeling

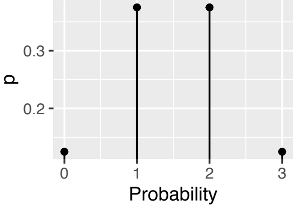
$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$$

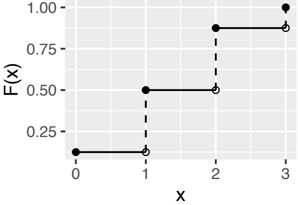
  - Egenskaper til F
  - $0 \leq F(x) \leq 1$
  - $F(x)$ er en voksende trappe-funsjon

Diskret Stokastisk variabel - II

Kast 3 mynter: $X = \{\text{Antall Kroner}\}$

x	p
0	0.125
1	0.375
2	0.375
3	0.125





- $f(x)$ er definert bare på de verdiene som X kan ta mens $F(x)$ er definert over hele $\mathcal{R}$
- Differanse mellom "$<$" of "$\leq$" er viktig
    - I vårt eksampel er
    - $P(X < 0) = 0$
    - $P(X \leq 0) = 1/8$

Kontinuerlig stokastisk variabel - I

- Sannsynlighet fordeling (sannsynlighetstetthet)

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

- Egenskaper til f
  - $f(x) \geq 0$ for alle $x \in \mathcal{R}$
  - $\int_{-\infty}^{+\infty} f(x)dx = 1$
- kumulative sannsynlighet fordeling

Kontinuerlig stokastisk variabel - I

- Sannsynlighet fordeling (sannsynlighetstetthet)

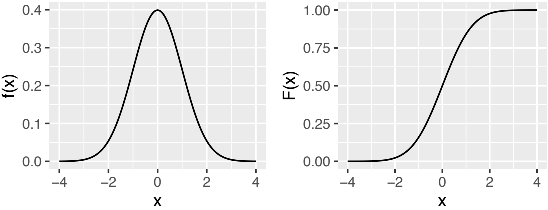
$$P(a < X < b) = \int_a^b f(x)dx$$

- Egenskaper til f
  - $f(x) \geq 0$ for alle $x \in \mathcal{R}$
  - $\int_{-\infty}^{+\infty} f(x)dx = 1$
- kumulative sannsynlighet fordeling

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

- Egenskaper til F
  - $0 \leq F(x) \leq 1$
  - $F(x)$ er en stigende, kontinuerlig funksjon

Kontinuerlig Stokastisk variabel - II



- $f(x) > 1$ er muling. Dette er ikke en sannsynlighet
- Differanse mellom "$<$" of "$\leq$" er ikke viktig

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

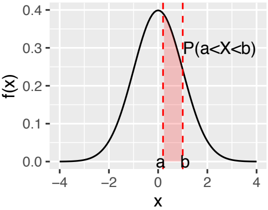
- For en kontinuerlig SV er

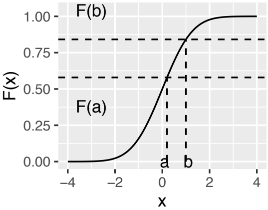
$$f(x) = \frac{dF(x)}{dx}$$

der den deriverte eksisterer.

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

Kontinuerlig Stokastisk variabel - II





- $f(x) > 1$ er muling. Dette er ikke en sannsynlighet
- Differanse mellom "$<$" of "$\leq$" er ikke viktig

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

- For en kontinuerlig SV er

$$f(x) = \frac{dF(x)}{dx}$$

der den deriverte eksisterer.

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

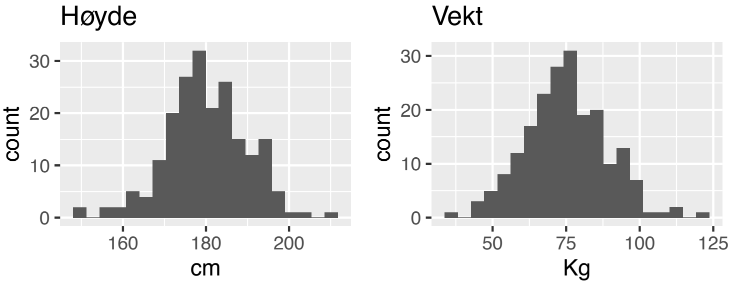
Exempel

- Vi har registrert vekt og høyde av 202 atleter

$$X = \{\text{Vekt}\}$$

$$Y = \{\text{Høyde}\}$$

- Vi kan se på fordeling av de to variablene hver for seg



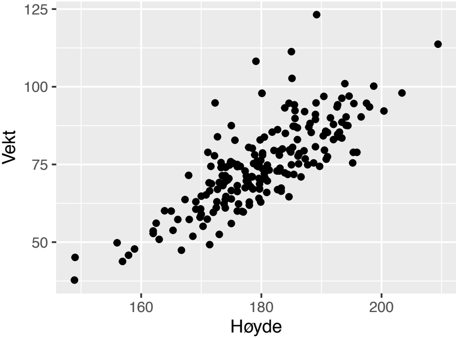
Exempel

Hva hvis vi ønsker å si noe om forholdet mellom X og Y? - Er de høyeste også tyngeste?

Exempel

Hva hvis vi ønsker å si noe om forholdet mellom X og Y?

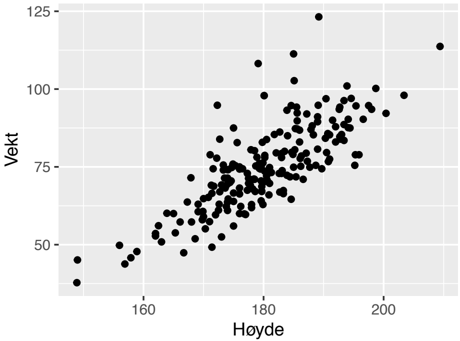
- Er de høyeste også tyngeste?



Exempel

Hva hvis vi ønsker å si noe om forholdet mellom X og Y?

- Er de høyeste også tyngeste?



- Vi trenger en måte å beskrive X of Y samtidig: vi trenger en simultan fordeling!

Planen for idag

- Simultan fordeling
- Marginal fordeling
- Betinget fordeling
- Uavhengige stokastisk variabler

Example for $f(x, y)$

		X			**h(y)**
		0	1	2	
Y	0	0.07	0.2	0.07	
	1	0.27	0.27	0	
	2	0.12	0	0	
g(x)					

Example for $f(x, y)$

		X			**h(y)**
		0	1	2	
Y	0	0.07	0.20	0.07	**0.34**
	1	0.27	0.27	0.00	**0.54**
	2	0.12	0.00	0.00	**0.12**
g(x)		0.46	0.47	0.07	**1.00**