- Stokastisk variabel: **Onsdag**
- Diskret sannsynsfordeling: **Onsdag**
- Kontinuerleg sannsynsfordeling:**Onsdag**
- Kummulativ sannsynsfordeling:**Onsdag**
- Diskret simultanfordeling **I dag**
- Kontinuerleg simultanfordeling **I dag**
- Marginal sannsynsfordeling **I dag**
- Betinga sannsynsfordeling **I dag**
- Statistik uavhengig **I dag**

Stokastisk variabel

Ein *stokastisk variabel er ein funksjon som assosierer eit reelt tal med kvart element i utfallsrommet.* $X : S \to \Re$

Notasjon:

- X, Y: Store bokstavar på stokastiske variable
- x, y: Små bokstavar på utfall (datapunkt)

Diskret stokastisk variabel:

Dersom den stokastiske variabelen gjev telbart antall utfall.

Kontinuerleg stokastisk variabel

Dersom den kontinuerlege variabelen har utfall på kontinuerleg skala.

Definisjon

Paret $(x, f(x))$ blir kalla sannsynsfordelinga til den diskret stok. var. X dersom

- $0 \leq f(x)$
- $\sum_{\forall x} f(x) = 1$ (summen over alle mogelege x)
- $f(x) = P(X = x)$

Kontinuerleg sannsynsfordeling

Definisjon

Funksjonen $f(x)$ definert for alle reelle tal $x \in \Re$ blir kalla sannsynsfordelinga til den kontinuerlege stok. var. X dersom

- $0 \leq f(x)$
- $\int_{-\infty}^{\infty} f(x)dx = 1$
- $P(a \leq X \leq b) = \int_a^b f(x)dx$

Kummulativ sannsynsfordeling

Definisjon

Den *kummulative fordelinlgsfunksjonen* $F(x)$ *for ein stok. var.* X *er:*

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

- *Diskret:* $F(x) = \sum_{t \leq x} f(t)$
- *Kontinuerleg* $F(x) = \int_{-\infty}^{x} f(t)dt$

Våre data

- Høgde
- Høgde mor
- Høgde far
- Kjønn (mann = 1, kvinne = 2)
- Hårfarge som 5-åring (lys = 1, brun = 2,svart = 3, raud= 4)
- Augefarge (blå= 1, brun= 2, grøn= 3)
- Kva er ditt forhold til programmering/Matlab?
  - A Programmering kan eg ikkje (Har ikkje hatt ITGK eller liknande).
  - B Kan programere, men kjenner ikkje Matlab.
  - C Har hatt ITGK med Matlab, men liker det ikkje.
  - D OK, det går nok bra med Matlab-øvingar.
  - E Kjempe bra!

Simultan sannsynsfordeling for diskret stokastisk variabel

Funksjonen $f(x, y)$ er ei simultan sannsynsfordeling for dei diskret stokastiske variablane X og Y dersom

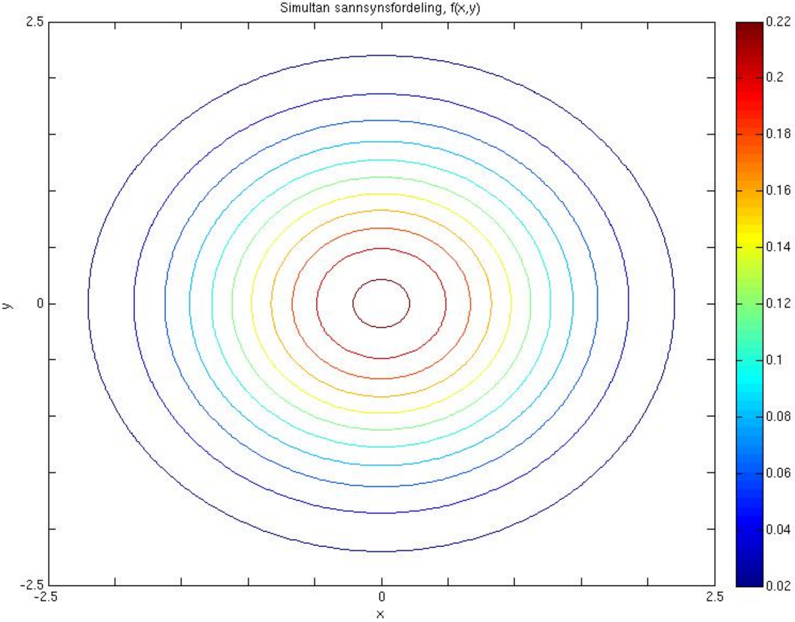
1. $f(x, y) \geq 0$ for alle (x, y)
2. $\sum_{\forall x} \sum_{\forall y} f(x, y) = 1$
3. $P(X = x, Y = y) = f(x, y)$

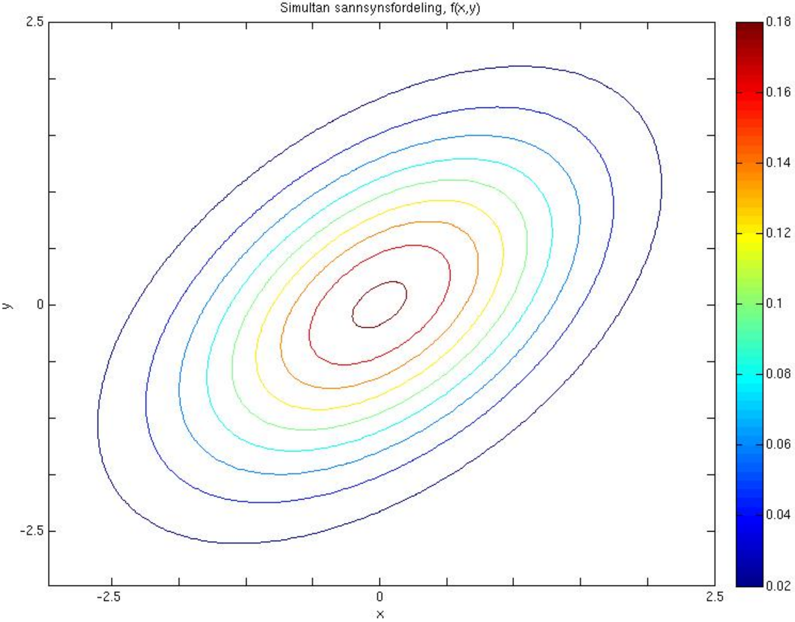
Simultan sannsynsfordeling, kontinuerleg

Simultan sannsynsfordeling for kontinuerleg stokastisk variabel

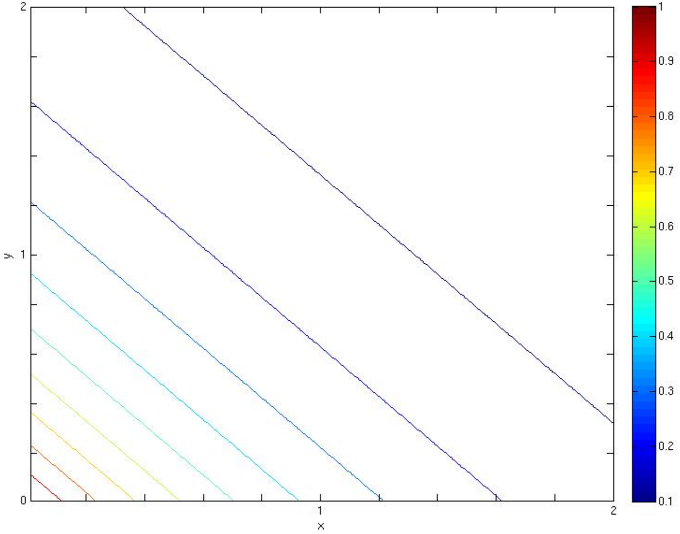
Funksjonen $f(x, y)$ er ei simultan sannsynsfordeling for dei kontinuerlege stokastiske variablane X og Y dersom

1. $f(x, y) \geq 0$ for alle (x, y)
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) = 1$
3. $P((X, Y) \in A) = \int \int_A f(x, y) dxdy$


Simultan sannsynsfordeling, f(x,y)
2.5
0
-2.5
-2.5
0
2.5
x
y
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02


Simultan sannsynsfordeling, f(x,y)
2.5
0
-2.5
y
-2.5
0
2.5
x
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02

Simultanfordeling elektrisk komponent eksempel



Marginalfordeling

Definisjon

Dersom $f(x, y)$ er simltanfordelinga til (X, Y) er *marginalfordelingane til X og Y hhv:*

for diskrete stok.var:

- $g(x) = \sum_{\forall y} f(x, y)$, *og*
- $h(y) = \sum_{\forall x} f(x, y)$.

for kontinuerlege stok.var:

- $g(x) = \int_{-\infty}^{\infty} f(x, y)dy$, *og*
- $h(y) = \int_{-\infty}^{\infty} f(x, y)dx$.

Betinga fordeling

Definisjon

La X og Y vere to stokastiske variable (kont. eller diskrete) med simultan sannsynsfordelin $f(x,y)$ og marginal sannsynsfordeling hhv $g(x)$ og $h(y)$.
Den *betinga fordelinga til* X *gjeve at* $Y = y$ *er*

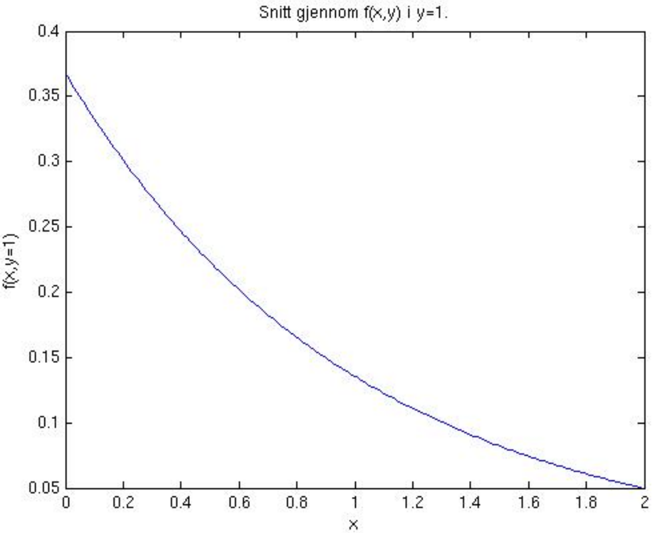
$$f(x|y) = f(x,y)/h(y)$$

($h(y) > 0$*)*
Tilsvarande er den betinga fordelinga til Y gjeve at $X = x$

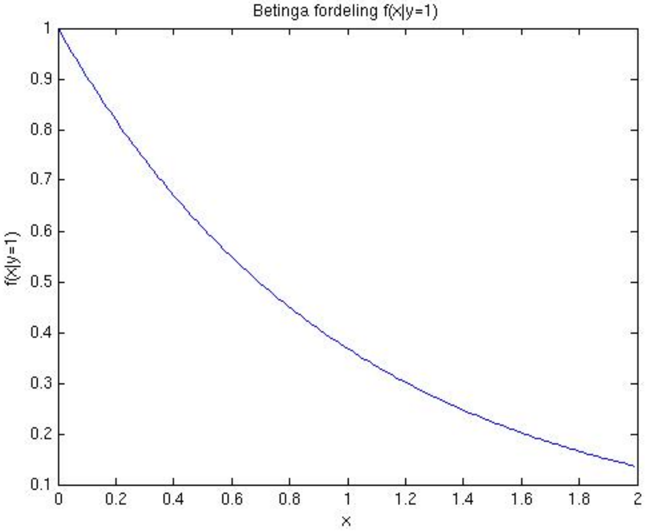
$$f(y|x) = f(x,y)/g(x)$$

$(g(x) > 0)$

Snitt gjennom simultanfordeling i $Y = 1$



To elektriske komponentar, betinga fordeling



Statistisk uavhengig

Definisjon

La X og Y vere to stokastiske variable med simultan sannsynsfordelin $f(x, y)$ og marginal sannsynsfordeling hhv $g(x)$ og $h(y)$.
X og Y er *statistisk uavhengige dersom, og berre dersom*

$$f(x, y) = g(x) \cdot h(y)$$

Generelt statistisk uavhengighet

Definisjon

$X_1, X_2, \ldots, X_n$ er *statistisk uavhengige dersom*

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) \ldots f_n(x_n)$$

Dagens tekst

Kap 3: Stokastiske variable og sannsynsfordelingar

- Stokastisk variabel: $X : S \to \Re$
- Diskret sannsynsfordeling: $f(x)$ slik at $P(X = x) = f(x)$
- Kontinuerleg sannsynsfordeling:
  $f(x)$ slik at $P(a < X < b) = \int_a^b f(x)$
- Kummulativ sannsynsfordeling: $F(x)$ slik at $P(X \leq b) = F(b)$
- Diskret simultanfordeling $P(X = x \cap Y = y) = f(x, y)$
- Kontinuerleg simultanfordeling
  $P((X, Y) \in A) = \int \int_A f(x, y) dxdy$
- Marginal sannsynsfordeling: For ein stok.var. aleine
- Betinga sannsynsfordeling $f(x|y) = f(x, y)/h(y)$
- Statistik uavhengig dersom $f(x_1, x_2) = f_1(x_1) f_2(x_2)$