Dagens tekst

Kap 3: Stokastiske variable og sannsynsfordelingar

- Stokastisk variabel:
- Diskret sannsynsfordeling:
- Kontinuerleg sannsynsfordeling:
- Kummulativ sannsynsfordeling:
- Diskret simultanfordeling
- Kontinuerleg simultanfordeling
- Marginal sannsynsfordeling
- Betinga sannsynsfordeling
- Statistik uavhengig

Kap 4: Matematisk forventning

- Forventningsverdi
- Varians / standard avvik
- Kovarians / korrelasjon
- Viktige reknereglar

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)

Eksempel augefarge

$X : \{\text{blå}, \text{brun}, \text{grøn}\} \to \{1, 2, 3\}$

Diskret sannsynsfordeling

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

Eksempel augefarge

	blå	brun	grøn
x	1	2	3
$f(x)$	$\frac{81}{130}$	$\frac{21}{130}$	$\frac{28}{130}$

Kontinuerleg sannsynsfordeling

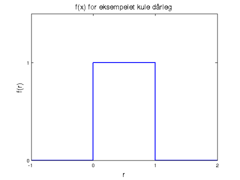
Definisjon

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

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

Eksempel, kule dårleg

Like sannsynleg å treffe alle avstandar mellom 0 og 1.



Simultan sannsynsfordeling, diskret

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

Eksempel, hår- og augefarge

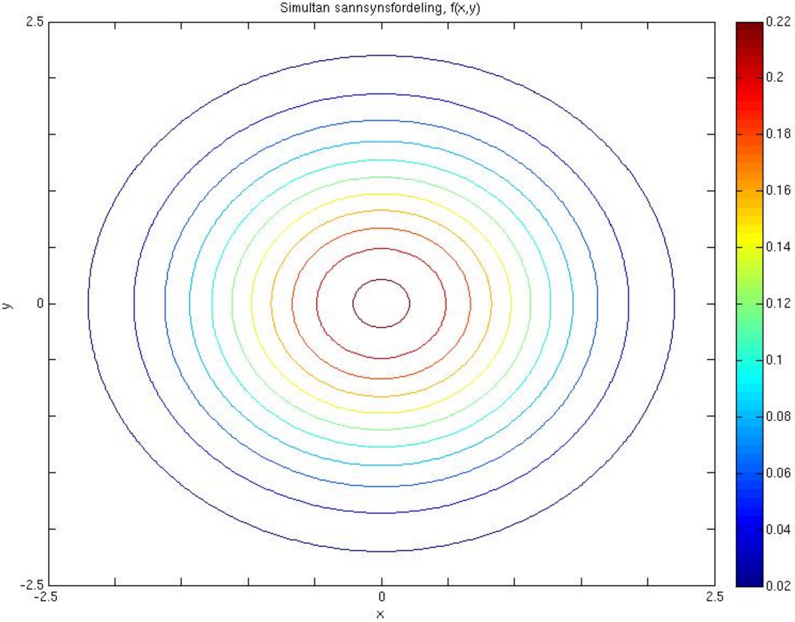
$f(x,y)$	Hår(Y):	lyst	brunt	svart	raudt
Auge (X):	x/y	1	2	3	4
Blå	1	$60/130$	$17/130$	$1/130$	$3/130$
Brun	2	$4/130$	$9/130$	$8/130$	0
Grøn	3	$14/130$	$13/130$	0	$1/130$

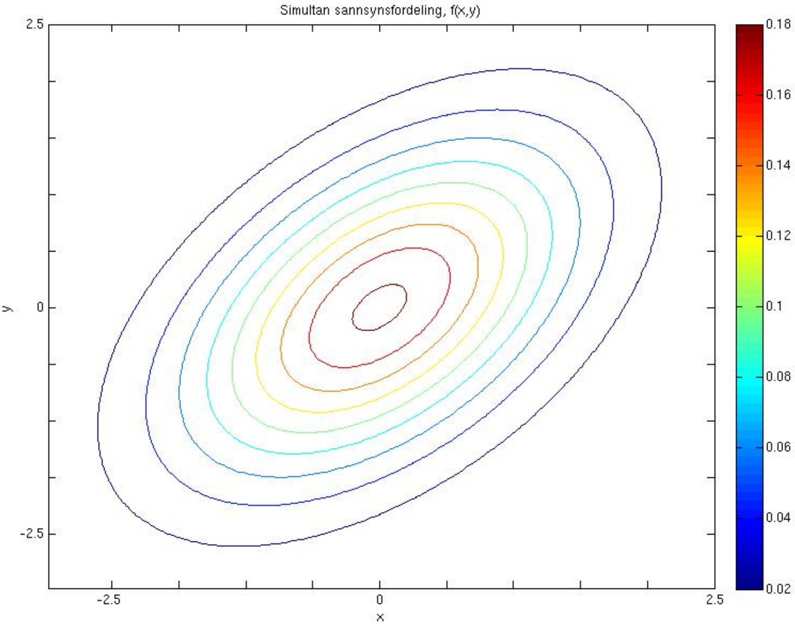
Simultan sannsynsfordeling, kontinuerleg

Simultan sannsynsfordeling for kontinuerleg stokastisk variabel

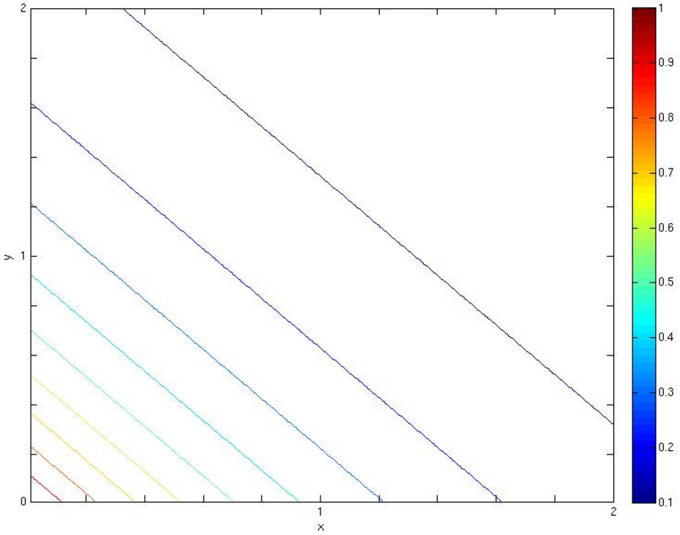
Funksjonen $f(x, y)$ er ei simultan sannsynsfordeling for dei kontinuerlege stokastiske variablane X og Y dersom

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:

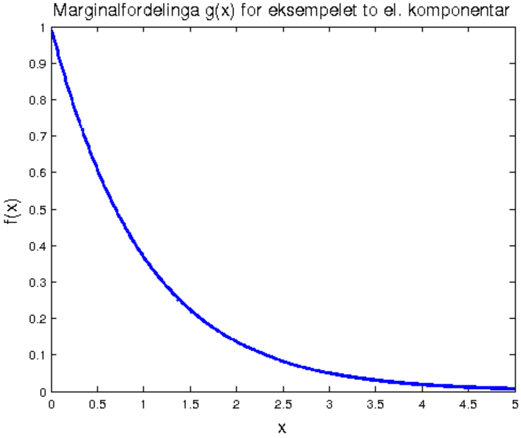
- $g(x) = \int_{-\infty}^{\infty} f(x, y) dy$, *og*
- $h(y) = \int_{-\infty}^{\infty} f(x, y) dx$.

Diskret marginalfordeling

Eksempel, hår- og augefarge

$f(x,y)$	Hår(Y):	lyst	brunt	svart	raudt	$g(x)$
Auge (X):	x/y	1	2	3	4	
Blå	1	$60/130$	$17/130$	$1/130$	$3/130$	$81/130$
Brun	2	$4/130$	$9/130$	$8/130$	0	$21/130$
Grøn	3	$14/130$	$13/130$	0	$1/130$	$28/130$
$h(x)$		$78/130$	$39/130$	$9/130$	$4/130$	

$g(x)$: Marginalfordeling til X, augefarge
$h(x)$: Marginalfordeling til Y, hårfarge


Marginalfordelinga g(x) for eksempelet to el. komponentar
f(x)
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

$g(x) = \exp(-x)$

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*

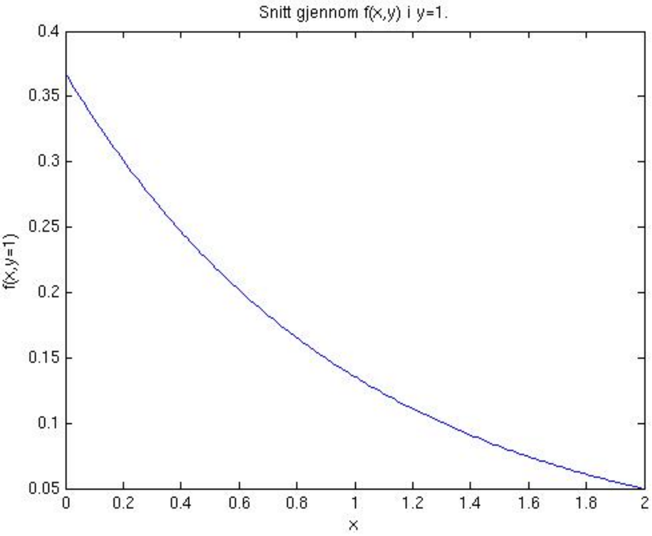
$$f(x|y) = f(x,y)/h(y)$$

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

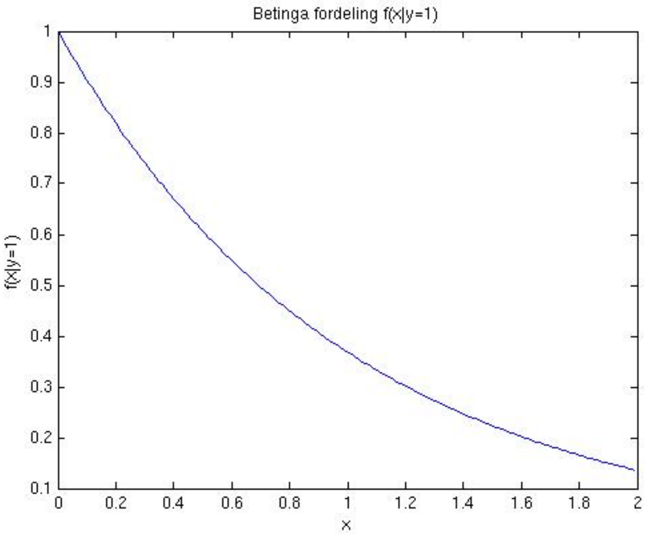
$$f(y|x) = f(x,y)/g(x)$$

$(g(x) > 0)$

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*

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

- 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 $f(x_1, x_2) = f_1(x_1) f_2(x_2)$

Viktige omgrep

- Forventningsverdi
- Varians / standard avvik
- Kovarians / korrelasjon

Viktige rekneregLar

Forventningsverdi

Definisjon 4.1

La X vere ein stok.var. med sannsynsfordeling $f(x)$. Forventningsverdien til X er då

$$\mu = E(X) = \sum_x xf(x)$$

dersom X er diskret, og

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

dersom X er kontinuerleg.

Dagens tekst

Kap 4: Matematisk forventning

Viktige omgrep

- Forventningsverdien til X = gj.snittet til ∞ data trekt frå $f(x)$
- Varians / standard avvik
- Kovarians / korrelasjon

Viktige reknereglar