Matematisk statistik för D, I, Π och Fysiker

Föreläsning 4

Johan Lindström

18 september 2018

Repetition
- Standardfördelningar
- Funktioner av en stokastiska variabel
- Normalfördelningen
- Grafisk presentation

Tvådimensionel stokastisk variabel (X, Y)
- Marginal fördelningar

Betingade fördelningar
- Total sannolikhet & Bayes sats

Oberoende stokastiska variabler

Exempel

Repetition
- Standardfördelningar
- Funktioner av en stokastiska variabel
- Normalfördelningen
- Grafisk presentation

Tvådimensionel stokastisk variabel (X, Y)
- Marginal fördelningar

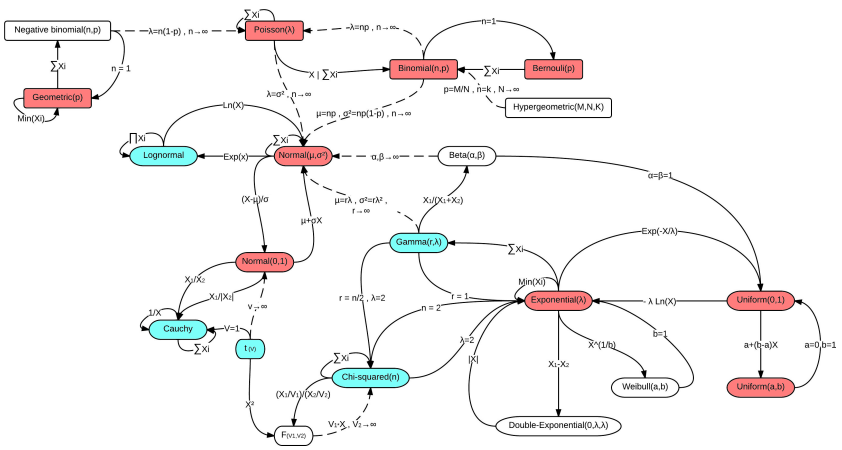
Betingade fördelningar
- Total sannolikhet & Bayes sats

Oberoende stokastiska variabler

Exempel

Standardfördelningar (Kap. 3.4, 3.6)

		E(X)	V(X)
För första gången	$X \in \text{ffg}(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Geometrisk	$X \in \text{Ge}(p)$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
Binomial	$X \in \text{Bin}(n,p)$	np	$np(1-p)$
Poisson	$X \in \text{Po}(\mu)$	μ	μ
Rektangel	$X \in \text{R}(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential	$X \in \text{Exp}(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normal	$X \in \text{N}(\mu,\sigma)$	μ	σ^2



en.wikipedia.org/wiki/Relationships_among_probability_distributions

FMSF05/MASC01 — Probability Theory

Funktioner av en stokastiska variabel (Kap. 3.10)

Givet en s.v. X. Vilken fördelning får $Y = g(X)$?

Kontinuerligt Y : Uttryck $F_Y(y)$ som funktion av $F_X(y)$.

$$F_Y(y) = \text{P}(Y \leq y) = \text{P}(g(X) \leq y) \stackrel{?}{=} F_X(\ldots)$$

$$f_Y(y) = \frac{d}{dy} F_Y(y)$$

Diskret Y : Räkna ut sannolikhetsfunktionen

$$p_Y(k) = \sum_{j:g(j)=k} p_X(j)$$

"lägga ihop $p_X(j)$ för alla j sådana att $g(j) = k$"

Normalfördelning (Kap. 3.6c, 6)

$$\Phi(x) = F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

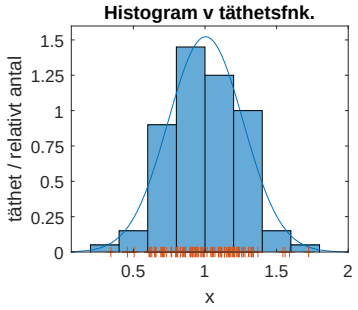
där $\Phi(x)$ räknas ut numeriskt eller fås från tabell.

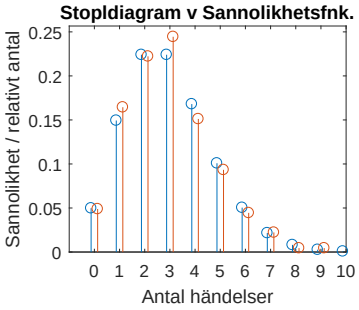
Standardiserad Normalfördelning (Kap. 6.3)

Om $X \in N(\mu, \sigma)$, med $E(X) = \mu$ och $V(X) = \sigma^2$, så är

$$\frac{X - \mu}{\sigma} \in N(0, 1)$$

Grafisk presentation av data (Kap. 10.2)





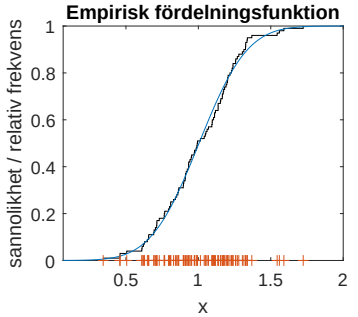

```
histogram(X, 'Normalization', 'pdf')
```

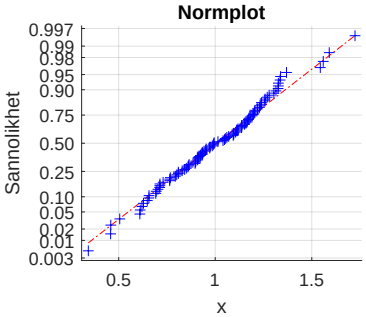
```
n = hist(Y, 0:10);
stem(0:10, n/sum(n))
```


Empirisk fördelningsfunktion

Om den s.v. X har fördelningsfunktion $F_X(x)$ då gäller att

$$Y = F_X(X) \in R(0, 1) \qquad F_Y(y) = y, \quad 0 < y < 1$$

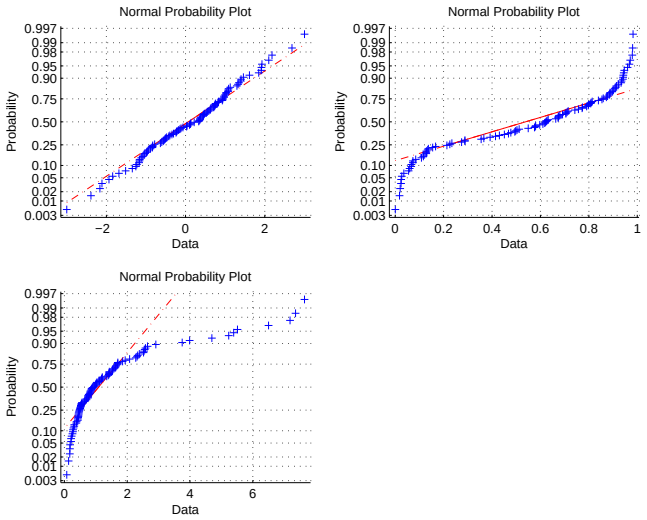





```
n=length(x);
stairs(sort(x), (1:n)/n)
```

```
normplot(x)
qqplot(x, ...inv((1:n)/n))
```

Exempel normplot



Repetition
- Standardfördelningar
- Funktioner av en stokastiska variabel
- Normalfördelningen
- Grafisk presentation

Tvådimensionel stokastisk variabel (X, Y)
- Marginal fördelningar

Betingade fördelningar
- Total sannolikhet & Bayes sats

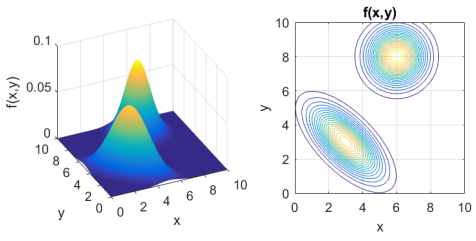
Oberoende stokastiska variabler

Exempel

Tvådim. stokastisk variabel (X, Y) (Kap. 4)

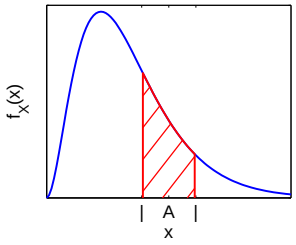
En 2D (bivariate) stokastisk variabel beskrivs av en
Simultan fördelnings-, sannolikhets- eller täthetsfunktion

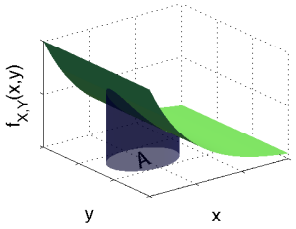
$$F_{X,Y}(x,y) = \mathsf{P}(X \leq x, Y \leq y) \qquad \text{(Def. 4.2)}$$
$$p_{X,Y}(j,k) = \mathsf{P}(X = j, Y = k) \qquad \text{(Def. 4.3)}$$
$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) \qquad \text{(Def. 4.4)}$$



Sannolikheten att hamna i A är integral av tätheten över A

$$P(X \in A) = \int_A f_X(x)\,dx \qquad P((X,Y) \in A) = \iint_A f_{X,Y}(x,y)\,dxdy$$

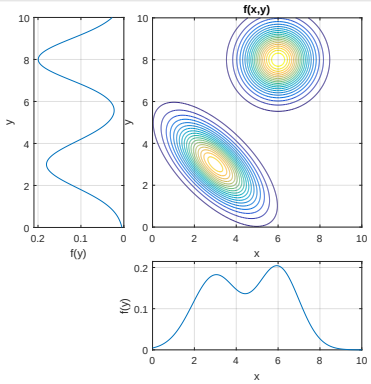




Marginella sannolikhets- & täthetsfunktioner (Kap. 4.3, 4.4)

Genom att summera/integrera över en av de s.v. fås

$$p_X(j) = \sum_k p_{X,Y}(j,k) \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$



Repetition
- Standardfördelningar
- Funktioner av en stokastiska variabel
- Normalfördelningen
- Grafisk presentation

Tvådimensionel stokastisk variabel (X, Y)
- Marginal fördelningar

Betingade fördelningar
- Total sannolikhet & Bayes sats

Oberoende stokastiska variabler

Exempel

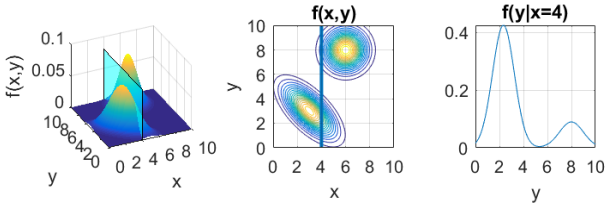
Betingade fördelningar för X givet Y (Kap. 4.8)

Precis som för sannolikheter, $P(A|B) = \frac{P(A \cap B)}{P(B)}$, fås

Betingade fördelningar (Def. 4.7, 4.8)

Diskret: $$p_{X|Y}(j|k) = \frac{p_{X,Y}(j,k)}{p_Y(k)}$$

Kontinuerlig: $$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$



Total sannolikhet & Bayes sats (Kap. 4.8)

Total sannolikhet

$$P(A) = \sum_i P(A|H_i)P(H_i)$$

Bayes sats

$$P(H_k|A) = \frac{P(A|H_k)P(H_k)}{\sum_i P(A|H_i)P(H_i)}$$

För sannolikhets- och täthetsfunktionerna har vi nu

$$p_X(j) = \sum_k p_{X|Y}(j|k)p_Y(k)$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(y)\mathrm{d}y$$

$$p_{Y|X}(l|j) = \frac{p_{X|Y}(j|l)p_Y(l)}{\sum_k p_{X|Y}(j|k)p_Y(k)}$$

$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y)f_Y(y)}{\int_{-\infty}^{\infty} f_{X|Y}(x|z)f_Y(z)\mathrm{d}z}$$

Repetition
Standardfördelningar
Funktioner av en stokastiska variabel
Normalfördelningen
Grafisk presentation

Tvådimensionel stokastisk variabel (X, Y)
Marginal fördelningar

Betingade fördelningar
Total sannolikhet & Bayes sats

Oberoende stokastiska variabler

Exempel

Oberoende stokastiska variabler (Kap. 4.5)

Oberoende

Händelserna A och B är oberoende

$$\Longleftrightarrow$$

$$P(A \cap B) = P(A)P(B)$$

X och Y är **oberoende** stokastiska variabler

$$\Longleftrightarrow$$

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) \quad \text{för alla } (x,y)$$

$\Longleftrightarrow$	$\Longleftrightarrow$
$p_{X,Y}(j,k) = p_X(j)p_Y(k)$	$f_{X,Y}(x,y) = f_X(x)f_Y(y)$
för alla (j,k)	**för alla** (x,y)

Repetition
- Standardfördelningar
- Funktioner av en stokastiska variabel
- Normalfördelningen
- Grafisk presentation

Tvådimensionel stokastisk variabel (X, Y)
- Marginal fördelningar

Betingade fördelningar
- Total sannolikhet & Bayes sats

Oberoende stokastiska variabler

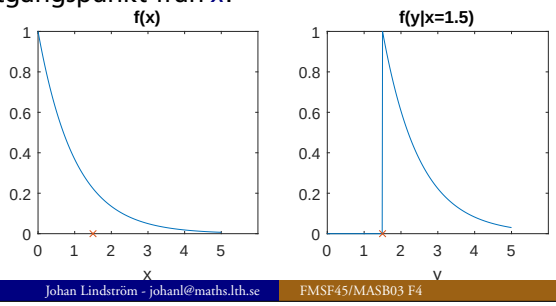
Exempel

Ex: Radioaktivt sönderfall / Kösystem / Lagerhantering

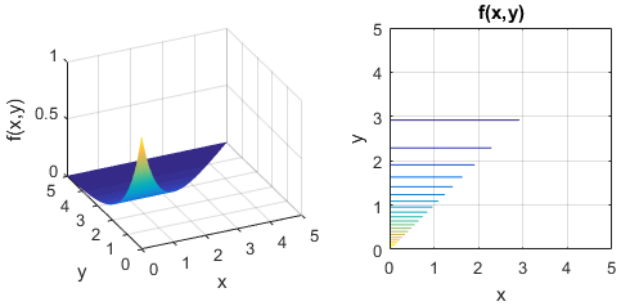
Antag att tiden mellan händelser beskrivs av en exponentialfördelning (med $\lambda = 1$).
Tiden till första händelsen är nu $X \in \text{Exp}(1)$ med

$$f_X(x) = e^{-x}, \quad 0 \leq x$$

Om vi vet att den första händelsen inträffar vid x är den **betingade tiden** till den andra händelsen exponentialfördelad med utgångspunkt från x.



Ex: Radioaktivt sönderfall / Kösystem / Lagerhantering

Bestäm

1. Den gemensamma tätheten $f_{X,Y}(x,y)$
2. Marginal tätheten $f_Y(y)$
3. Den betingade tätheten $f_{X|Y}(x|y=T)$
4. Avgör om X och Y är oberoende