Föreläsning 1

Mängden S av alla möjliga utfall u kallas försökets utfallsrum

En delmängd A till S kallas händelse

Ex:

$S = \{1,2,3,4,5,6\}$

$A = \{1,3,5\} = \{u \in S \mid u \text{ udda}\} = \{\text{udda utfall}\}$

$B = \{1,2\} = \{u \in S : u \leq 2\}$

om $u \in A \to$ "A inträffar"

$A \cup B$ inträffar $\Leftrightarrow$ A eller B inträffar

$A \cap B$ inträffar $\Leftrightarrow$ A och B inträffar

A^c inträffar $\Leftrightarrow$ A inte inträffar

Sannolikhetsmått

$\mathcal{P}(S)$ mängden av alla delmängder till S

Om $\mathbb{P} : \mathcal{P}(S) \to \mathbb{R}$ uppfyller

(i) $0 \leq \mathbb{P}(A) \leq 1 \quad \forall A$

(ii) $\mathbb{P}(S) = 1$

(iii) $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$ för alla följder av parvis disjunkta mängder $A_1, A_2, \ldots$ ($A_i \cap A_j = \emptyset$ då $i \neq j$)

då kallas $\mathbb{P}$ för sannolikhetsmått.

*Tag $A_1 = S \quad A_2 = A_3 = \ldots = \emptyset$

$\mathbb{P}(S) = \mathbb{P}(S) + \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset) \ldots \Rightarrow \mathbb{P}(\emptyset) = 0$

Proposition 3.1 $\forall$ händelser A & B

$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(A)$

Föreläsning 1

Mängden S av alla möjliga utfall u kallas försökets utfallsrum

En delmängd A till S kallas händelse

Ex:

$$S = \{1,2,3,4,5,6\}$$
$$A = \{1,3,5\} = \{u \in S \mid u \text{ udda}\} = \{\text{udda utfall}\}$$
$$B = \{1,2\} = \{u \in S : u \leq 2\}$$

om $u \in A$ → "A inträffar"

$A \cup B$ inträffar $\iff$ A eller B inträffar

$A \cap B$ inträffar $\iff$ A och B inträffar

A^c inträffar $\iff$ A inte inträffar

Sannolikhetsmått

$P(S)$ mängden av alla delmängder till S

Om $\mathbb{P} : P(S) \to \mathbb{R}$ uppfyller

(i) $0 \leq \mathbb{P}(A) \leq 1 \quad \forall A$

(ii) $\mathbb{P}(S) = 1$

(iii) $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$ för alla följder av parvis disjunkta mängder $A_1, A_2, \ldots$ ($A_i \cap A_j = \emptyset$ då $i \neq j$)

då kallas $\mathbb{P}$ för sannolikhetsmått.

*Tag $A_1 = S \quad A_2 = A_3 = \ldots = \emptyset$

$\mathbb{P}(S) = \mathbb{P}(S) + \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset) \ldots \implies \mathbb{P}(\emptyset) = 0$

Proposition 3.1 $\forall$ händelser A & B

$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$A \subseteq B \implies \mathbb{P}(A) \leq \mathbb{P}(A)$

Bevis:

A & A^c disjunkta & $A \cup A^c = S$

$\overset{(ii)}{\underset{(iii)}{\Longrightarrow}} 1 = \mathbb{P}(S) = \mathbb{P}(A) + \mathbb{P}(A^c) \Rightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

b) Enligt (iii)

$$\mathbb{P}(B) = \mathbb{P}((B \setminus A) \cup (A \cap B)) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$$

$$\Rightarrow \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

c) Enligt (iii) & b)

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

d) Om $A \subseteq B$ gäller $A \cap B = A$

$$\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A) \geq \mathbb{P}(A) \quad \text{enligt (i)}$$

Ex

Risk	dag
50%	lör
40%	sön
30%	båda

Chansen till uppehåll hela helgen?

Vad är S?

$S = \{(r,r), (r,u), (u,r), (u,u)\}$

Låt $A = \{r \text{ på } L\}$

$B = \{r \text{ på } S\}$

Vi söker $\mathbb{P}((A \cup B)^c) = 1 - \mathbb{P}(A \cup B) =$

$= 1 - (\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)) = 1 - (0,5 + 0,4 - 0,3) = 0,4$

Svar: 40%

För 3 händelser

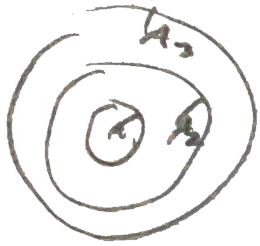
$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$$

Inklusion - Exklusionsformel prop 1.5

Kontinuitet hos sannolikheter

Antag $A_1 \subseteq A_2 \subseteq \cdot$ och skriv $A = \bigcup_{n=1}^{\infty} A_n$

Då gäller enligt (iii) att (om $A_0 = \emptyset$)



$$\begin{aligned} \mathbb{P}(A) &= \sum_{n=1}^{\infty} \mathbb{P}(A_n \setminus A_{n-1}) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(A_n) - \mathbb{P}(A_{n-1}) \\ &= \lim_{N \to \infty} \sum_{n=1}^{N} \mathbb{P}(A_n) - \mathbb{P}(A_{n-1}) \\ &= \lim_{N \to \infty} \mathbb{P}(A_N) \end{aligned}$$

Kortfattat:

Om $A_1 \subseteq A_2 \subseteq \dots$ och $A = \bigcup_{n=1}^{\infty} A_n$ gäller att

$\mathbb{P}(A_n) \to \mathbb{P}(A)$ då $n \to \infty$

Om istället

$A_1 \supseteq A_2 \supseteq \dots$ och $A = \bigcap_{n=1}^{\infty} A_n$ gäller enligt ovanstående

$\mathbb{P}(A_n^c) \to \mathbb{P}(A^c)$ och därmed

$\mathbb{P}(A_n) \to \mathbb{P}(A)$ då $n \to \infty$

Ex: välj ett tal på måfå i $[0,1]$

Detta betyder att $S = [0,1]$ och $\mathbb{P}(A) = \ell(A)$ = längden av A
(svårt att bevisa)

Låt $A_n = [0, \frac{1}{2} + \frac{1}{n}]$ $A = \bigcap_n A_n = [0, \frac{1}{2}]$

och $\mathbb{P}(A_n) \to \mathbb{P}(A)$

$(\ell(A_n) = \frac{1}{2} + \frac{1}{n}$ och $\ell(A) = \frac{1}{2})$

Föreläsning 2

$S = \{u_1, \dots, u_n\}$ $\quad p_i = \frac{1}{n}$

$$\mathbb{P}(A) = \frac{\text{Antal utfall i } A}{\text{totala antalet utfall}} = \frac{|A|}{n} = \frac{|A|}{|S|}$$

__Ex__ Slå tärning 3 ggr
$\mathbb{P}$(exakt en 6:a)?

$S = \{(x,y,z) : x,y,z \in \{1,2,3,4,5,6\}\}$

$|S| = 6^3 = 216$ $\quad A = \{(x,y,z) \in S : \text{exakt en } x,y,z \text{ är } 6\}$

$|A| = 1\cdot 5\cdot 5 + 5\cdot 1\cdot 5 + 5\cdot 5\cdot 1 = 75$

$\Rightarrow \quad \mathbb{P}(A) = \frac{75}{216} = \frac{25}{72}$

__Ex__ I en hatt ligger 3 kort r-r r-s s-s

Ett kort dras på måfå och läggs med på måfå vald sida upp. Vad är sannolikheten att den andra sidan har samma färg

kalla korten a b c

$S = \{a1, a2, b1, b2, c1, c2\}$

öker $\mathbb{P}(A)$ där $A = \{a1, a2, c1, c2\}$

$\mathbb{P}(A) = \frac{4}{6} = \frac{2}{3}$

Behöver kunna räkna antal element i en mängd
__kombinatorik__

Multiplikationsprincipen – Axiom

$n_1 n_2 \dots n_r$

__Ex__ Födelsedagsproblemet

En skolklass har r elever. Vad är sannolikheten olika födelse

$$S = \{(x_1, \dots, x_r) : \forall i : x_i \in \{1, \dots, 365\}\}$$

$$A = \{(x_1, \dots, x_r) \in S : \forall i \neq j : x_i \neq x_j\}$$

$$\mathbb{P}(A) = \frac{365 \cdot 364 \cdot \dots \cdot (365 - r + 1)}{365^r}$$

Val utan återläggning med ordning

Val med åter med ordning

$\binom{n}{r}$ - Välj r av n Utan återlägg utan ordning "n över r"

$$(n)_r = \binom{n}{r} r! \Rightarrow \binom{n}{r} = \frac{(n)_r}{r!} = \frac{n!}{r!(n-r)!}$$

Antal sätt att välja r element av n m. åter u. ordning är antal ickenegativa heltalslösningar till

$$x_1 + \dots + x_n = r$$

Visest ett $\binom{n-1+r}{r} = \binom{n-1+r}{n-1}$

Betingad Sannolikhet och oberoende

Ex två tärningar slås:

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$$

Ex Vad är sannolikheten att regn 1:a 2:a augusti? Kunskap om det regnar 1:a påverkar om regnar 2:a

$$\mathbb{P}(A_1 \cap A_2) \neq \mathbb{P}(A_1)\mathbb{P}(A_2)$$

$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot$ sannolikhet regn 2:a givet regn 1:a

Def: För två händelser A & B ges sannolikheten för B givet A av

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

Def Två händelser A & B sägs vara <u>oberoende</u> om $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

Prop: om A & B oberoende så är A & B^c oberoende

Bevis:
$$\begin{aligned}\mathbb{P}(A \cap B^c) &= \mathbb{P}(A) - \mathbb{P}(A \cap B) = \\ &= \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) \\ &= \mathbb{P}(A)(1 - \mathbb{P}(B)) = \\ &= \mathbb{P}(A)\mathbb{P}(B^c)\end{aligned}$$

Prop: Fixera A och låt för varje $B \in P(S)$
$$Q(B) = \mathbb{P}(B|A)$$

Då är Q ett sannolikhetsmått

Bevis:

(i) $0 \leq \mathbb{P}(A \cap B) \leq \mathbb{P}(A) \Rightarrow 0 \leq \underbrace{\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}}_{Q(B)} \leq 1$

(ii) $Q(S) = \mathbb{P}(S|A) = \frac{\mathbb{P}(A \cap S)}{\mathbb{P}(A)} = 1$

(iii) Låt $B_1, \ldots$ vara disjunkta då gäller

$$Q(\cup_n B_n) = \mathbb{P}(\cup_n B_n | A) = \frac{\mathbb{P}(\cup_n (A \cap B_n))}{\mathbb{P}(A)} =$$
$$= \frac{\sum_n \mathbb{P}(A \cap B_n)}{\mathbb{P}(A)} = \sum_n Q(B_n)$$

Betingad sannolikhet gäller alla räkneregler

Låt $X: S \to \mathbb{R}$ vara en sv. X diskret om V_X är ändlig eller uppräkneligt.

$$V_X = \{x_1, x_2, \dots\}$$

$$V_X = \{X(u) : u \in S\}$$

$$p(x_k) = \mathbb{P}(X = x_k) = \mathbb{P}(\{u \in S : X(u) = x_k\})$$

då blir $p : V_X \to [0,1]$ och p kallas frekvensfunktion (pmf)

Notera att varje funktion $p : \{x_1, \dots\} \to [0,1]$: $\sum_{k=1}^{\infty} p(x_k) = 1$ är pmf till en sv, nämligen den sv X för vilken $\mathbb{P}(X = x_k) = p(x_k)$

Def Funktionen $F : \mathbb{R} \to \mathbb{R}$ given av

$$F(x) = \mathbb{P}(X \leq x)$$

kallas X:s fördelningsfunktion

Om X diskret:

$$F(x) = \sum_{x_k \in V_X : x_k \leq x} p(x_k)$$

och

$$p(x_k) = F(x_k) - F(x_k -)$$

Ex Singla slant 4ggr X antal klave

$$V_X = \{0,1,2,3,4\}$$

$$p(k) = \binom{4}{k}\left(\frac{1}{2}\right)^4 = \binom{4}{k}\frac{1}{16}$$

k	$p(k)$	$F(k)$
0	1/16	1/16
1	1/4	5/16
2	3/8	11/16
3	1/4	15/16
4	1/16	1

Varje fördelnings funktion F är växande

$F(\infty) = 1 \qquad F(-\infty) = 0$

$\lim_{n\to\infty} F(x+\frac{1}{n}) = \lim_{n\to\infty} \mathbb{P}(X \le x + \frac{1}{n}) = \mathbb{P}(X \le x) = F(x)$

ty $\{\omega : X(\omega) \le x\} = \cap_n \{\omega : X(\omega) \le x + 1/n\}$

Andra fakta

- $\mathbb{P}(a < X \le b) = F(b) - F(a)$
- $\mathbb{P}(X > x) = 1 - F(x)$

Def en s.v. X kallas kontinuerlig om $\exists f : \mathbb{R} \to \mathbb{R}$ s.a $\forall x \in \mathbb{R}$:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

Funktionen f kallas för X:s täthet (pdf)

Definitionen ekvivalent med att säga F är deriverbar med f som derivata

Def 2.5 i boken lite fel, räcker inte att F är kontinuerlig

om X kont:

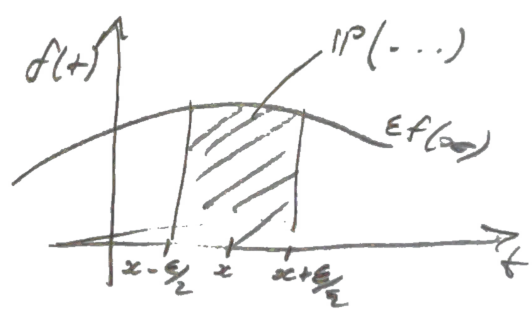
$$\mathbb{P}(X \in B) = \int_B f(t)\,dt \qquad \forall B \subseteq \mathbb{R}$$

OBS $\forall x$ gäller $\mathbb{P}(X = x) = \int_x^x f(t)\,dt = 0$

$$\mathbb{P}(a \le X \le b) = \mathbb{P}(a < X < b) = F(b) - F(a)$$

Tolkning av f:

$$\mathbb{P}(x - \varepsilon/2 < X < x + \varepsilon/2) \approx \varepsilon f(x)$$



Likformig fördelning

En kont. s.v. sägs vara likformigt fördelad på $[a,b]$ om

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases}$$

dvs om $$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & x \in [a,b] \\ 1, & x > b \end{cases}$$

kort skrivsätt: $X \sim \text{likf}[a,b]$

Obs: $\mathbb{P}(X \in [c,d]) = \frac{1}{b-a} \int_c^d dx = \frac{d-c}{b-a}$

Proposition: Om $X \sim \text{likf}[0,1]$ och $Y = a + (b-a)X$ gäller att $Y \sim \text{likf}[a,b]$

Bevis: Eftersom $F_X(x) = x \ \forall x \in [0,1]$ gäller $\forall y \in [a,b]$ att

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(a + (b-a)X \leq y) =$$
$$= \mathbb{P}\left(X \leq \frac{y-a}{b-a}\right) = \frac{y-a}{b-a}$$

Fördelning för funktioner av SV

Låt X vara SV & $g: \mathbb{R} \to \mathbb{R}$ en fkn

: Låt $Y = g(X)$ Vad är fördelningen för Y?

Fallet då X diskret:

$$V_x = \{x_1, x_2, \dots\}$$

$$V_Y = \{g(x_j) : x_j \in V_x\}$$

För varje $y_k \in V_Y$ låt $B_k = \{x_j \in V_x : g(x_j) = y_k\}$

Då gäller

$$p_Y(y_k) = \mathbb{P}(Y = y_k) = \mathbb{P}(X \in B_k) = \sum_{x_j \in B_k} p_x(x_j)$$

Fallet då X kont: Y kan bli diskret, kont eller annat beroende på g

Ex: Låt $X \sim$ Likf $[0,1]$

$$g(x) = \begin{cases} 0, & x < 2/3 \\ 1, & x \geq 2/3 \end{cases}$$

Med $Y = g(x) \Rightarrow V_Y = \{0,1\}$ och $p_Y(1) = 1/3$

Man jobbar ofta fall till fall

Ex Låt $X \sim$ likf $[0,2]$, dvs $f(x) = 1/2$ och $F(x) = x/2$

Låt $Y = X(2-X)$

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X(2-X) \leq y)$$

$$x(2-x) = y \Leftrightarrow x = 1 \pm \sqrt{1-y}$$

Alltså $F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X \leq 1 - \sqrt{1-y}) + \mathbb{P}(X > 1 + \sqrt{1-y}) =$

$$= \frac{1-\sqrt{1-y}}{2} + 1 - \frac{1+\sqrt{1-y}}{2} = 1 - \sqrt{1-y}$$

Ett generellt resultat:

Antag g strängt växande, Då finns g^{-1} och är växande. Alltså

$$F_Y(y) = \mathbb{P}(g(X) \le y) = \mathbb{P}(X \le g^{-1}(y)) = F_X(g^{-1}(y))$$

$$f_Y(y) = f_X(g^{-1}(y))\,(g^{-1})'(y)$$

Ex Låt $X \sim$ likf $[0,1]$ $Y = X^3 \Rightarrow g(x) = x^3$

så $g^{-1}(y) = y^{1/3}$ och $(g^{-1})'(y) = \frac{1}{3}y^{-2/3}$

Eftersom $f(x) = 1$ får vi

$$f_Y(y) = \frac{1}{3}y^{-2/3},\ y \in [0,1]$$

Väntevärde

Låt $X_1, X_2, X_3, \dots$ vara sv definierade på samma utfallsrum.

Def: Man säger att $X_1, X_2, \dots$ är oberoende om det $\forall B_1, B_2, \dots \subseteq \mathbb{R}$ gäller att händelserna $\{X_1 \in B_1\}, \{X_2 \in B_2\}, \dots$ är oberoende

Låt N vara stort och $X_1, \dots, X_n$ vara oberoende och alla ha samma förd som X. Antag $V_X = \{x_1, x_2, \dots\}$

Låt $N_k = |\{j : X_j = x_k\}|$ vi förväntar oss att ha

$N_k \approx Np(x_k)$ vilket medför

$$\text{medelvärdet av } X_j = \sum_{k=1}^{\infty} x_k \frac{N_k}{N} \approx \sum_{k=1}^{\infty} x_k p(x_k)$$

(Detta är tyngdpunkten av tallinjen om massorna $p(x_k)$ läggs ut på pos x_k)

Def **Väntevärdet** av en diskret sv X ges av

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} x_k p(x_k)$$

För kont. X gäller

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx$$

(formellt krävs $\sum_{k=1}^{\infty} |x_k| p(x_k) < \infty$ resp. $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$)

Ex Satsa en marker på rött på roulette.
Låt X vara vinsten. Då gäller $\mathbb{P}(X=1) = 18/37$
och $\mathbb{P}(X=-1) = 19/37$
Alltså:

$$\mathbb{E}[X] = (-1)\frac{19}{37} + 1 \cdot \frac{18}{37} = -\frac{1}{37}.$$

Ex Vv av diskret likformig: $V_X = \{1, 2, \dots, n\}$ och
$\mathbb{P}(X=k) = 1/n \quad \forall k = 1, \dots, n$

$$\mathbb{E} = \sum_{k=1}^{n} k \frac{1}{n} = \frac{n+1}{2}$$

Exvis är Vv av tärningsslag $\frac{7}{2}$

Ex $X \sim \text{likf}(a, b)$

$$\mathbb{E}[X] = \int_a^b x \frac{1}{b-a} dx = \frac{b+a}{2}$$

Ex: $X \sim \text{likf}(0,1)$ låt $Y = X^2$
$F_Y(y) = \mathbb{P}(X \leq \sqrt{y}) = \sqrt{y}$ för $y \in (0,1)$
Alltså $f_Y(y) = 1/(2\sqrt{y})$

$$\mathbb{E}[Y] = \int_0^1 y \frac{1}{2\sqrt{y}} dy = \frac{1}{3}\left[y^{3/2}\right]_0^1 = \frac{1}{3}$$

Notera alltså
$\mathbb{E}[X^2] = 1/3 \quad \mathbb{E}[X] = 1/2$
$\mathbb{E}[X]^2 = \frac{1}{4}$

Proposition 2.9: Antas X är kont och $V_X = [0,\infty)$ Då gäller

$$\mathbb{E}[X] = \int_0^\infty (1-F(x))dx = \int_0^\infty \mathbb{P}(X>x)dx$$

Bevis:

$$\int_0^\infty \mathbb{P}(X>x)dx = \int_0^\infty \int_x^\infty f(t)\,dt\,dx =$$

$$= \int_0^\infty \left(\int_0^t dx\right) f(t)\,dt = \int_0^\infty t\cdot f(t)\,dt = \mathbb{E}[X] \quad \square$$

Diskret variant: Antas $V_X = \{0,1,2,\ldots\}$

$$\mathbb{E}[X] = \sum_{n=1}^\infty \mathbb{P}(X\geq n) = \sum_{n=0}^\infty \mathbb{P}(X>n)$$

Bevis:

$$\sum_{n=1}^\infty \mathbb{P}(X\geq n) = \sum_{n=1}^\infty \sum_{k=n}^\infty p(k) = \sum_{k=1}^\infty \sum_{n=1}^k p(k) = \sum_{k=1}^\infty k\,p(k) \quad \square$$

Ex: ovan $F_Y(y) = \sqrt{y}$, $y\in(0,1)$

$$\mathbb{E}[Y] = \int_0^1 (1-\sqrt{y})\,dy = \frac{1}{3} \quad \text{då } y\in(0,1)$$

Ex: Låt X antal tärningskast till 6:a

$$\mathbb{P}(X>n) = \left(\frac{5}{6}\right)^n$$

Alltså

$$\mathbb{E}[X] = \sum_{n=0}^\infty \left(\frac{5}{6}\right)^n = \frac{1}{1-\frac{5}{6}} = 6$$

Väntevärden av fkn av sv

Sats: Låt $g : \mathbb{R} \to \mathbb{R}$ o X en sv.
Om X diskret gäller
$$\mathbb{E}[g(X)] = \sum_{k=1}^{\infty} g(x_k) p(x_k)$$
Om X kont. gäller
$$\mathbb{E}[g(x)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

Bevis: Diskret: Skriv $Y = g(X)$ $V_Y = \{y_1, y_2, \dots$
$$\mathbb{E}[Y] = \sum_j y_j \cdot \mathbb{P}(Y = y_j) = \sum_j y_j \sum_{k : g(x_k) = y_j} p(x_k) =$$
$$= \sum_j \sum_{k : g(x_k) = y_j} g(x_k) p(x_k) = \sum_{k=1}^{\infty} g(x_k) p(x_k)$$
Specialfall för kont. i boken.

Ex Låt a, b vara konstanter X en sv. Då är
$$\mathbb{E}[aX + b] = \int_{-\infty}^{\infty} (ax + b) f(x) dx = a \int_{-\infty}^{\infty} x f(x) dx + b = a\mathbb{E}[X] + b$$
Ex Låt $X \sim \text{likf}(0,1)$
$$g(x) = \begin{cases} 0 & x < 2/3 \\ 1 & x \geq 2/3 \end{cases}$$
$$\mathbb{E}[g(x)] = \int_{-\infty}^{\infty} g(x) f(x) dx = \int_{2/3}^{1} dx = \frac{1}{3}$$

Oändliga väntevärden

I allmänhet kräver vi $\int_{-\infty}^{\infty} |x| f(x)\,dx < \infty$ för att def $\mathbb{E}[x]$

Detta då

$$\int_{-\infty}^{\infty} x f(x)\,dx = \int_{0}^{\infty} x f(x)\,dx - \int_{-\infty}^{0} (-x) f(x)\,dx = \infty - \infty = ?$$

Men om $X \geq 0$ försvinner andra termen

Säger då $\mathbb{E}[X] = \infty$

Analogt i diskret fall.

<u>Ex:</u> Låt X vara antalet kast t.o.m klave

Sammanlagd förlust före första klaven blir $2^X - 1$

Det gäller $\mathbb{P}(X=k) = \left(\frac{1}{2}\right)^k \quad k = 1, 2, \ldots$ så

$$\mathbb{E}[2^X - 1] = \sum_{k=1}^{\infty} (2^k - 1)\left(\frac{1}{2}\right)^k = -1 + \sum_{k=0}^{\infty} 1 = \infty$$

Varians

Låt X vara sv med $\mathbb{E}[X] = \mu < \infty$

<u>Definition:</u> $\mathrm{Var}(X) = \mathbb{E}[(X-\mu)^2]$

Man kan ha $\mathrm{Var}(X) = \infty$ Man skriver $\mathrm{Std}(X) = \sqrt{\mathrm{Var}(X)}$

Vanlig beteckning $\sigma^2 = \mathrm{Var}(X)$

Steiners formel $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

<u>Bevis:</u> kont:

$$\mathrm{Var}(X) = \int (x-\mu)^2 f(x)\,dx = \int x^2 f(x)\,dx - 2\mu \int x f(x)\,dx + \mu^2 \int f(x)\,dx = \mathbb{E}[X^2] - \mu^2$$

$\square$

Ex $X \sim \text{Unif}[a,b] \implies \mathbb{E}[X] = \frac{a+b}{2}$

$$\mathbb{E}[x^2] = \int_a^b x^2 \frac{1}{b-a} dx = \frac{b^3-a^3}{3(b-a)}$$

Alltså
$$Var(X) = \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

Prop: $Var(aX+b) = a^2 Var(X)$

Bevis
$$Var(aX+b) = \mathbb{E}[(aX+b-\mathbb{E}[aX+b])^2] =$$
$$= \mathbb{E}[(aX+b-(a\mu+b))^2] =$$
$$= a^2 \mathbb{E}[(X-\mu)^2] = a^2 Var(X)$$

Teoretiskt viktiga olikheter:

Markovs olikhet: Antag $X \geq 0$. Då gäller $\forall a > 0$ att
$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

__Bevis:__ Låt
$$g(x) = \begin{cases} 0 & x < a \\ a & x \geq a \end{cases}$$

då är $g(X) \leq X$ så $\mathbb{E}[g(X)] \leq \mathbb{E}[X]$. Eftersom $g(X)$ är a med sannolikhet $\mathbb{P}(X \geq a)$ och 0 annars
$$\mathbb{E}[g(X)] = a\,\mathbb{P}(X \geq a)$$ $\blacksquare$

Chebyshevs olikhet: Skriv $\mu = \mathbb{E}[X]$ och $\sigma^2 = Var(X)$

Det gäller $\forall \varepsilon > 0$ att
$$\mathbb{P}(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

__Bevis:__ Enligt Markovs olikhet gäller att
$$\mathbb{P}(|X-\mu| \geq \varepsilon) = \mathbb{P}((X-\mu)^2 \geq \varepsilon^2) \leq \frac{\mathbb{E}[(X-\mu)^2]}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$

Ex Längden X av på måfå vald svensk man är en sv med vv 180cm och std 5cm. Sätt en gräns på andelen svenska män över 210 cm.

Vi har $\mu = 180$, $\sigma^2 = 5^2 = 25$ så

$$\mathbb{P}(X \geq 210) \leq \mathbb{P}(|X-\mu| \geq 30) \leq \frac{25}{30^2} = \frac{1}{36}$$

<u>Indikatorfördelning</u>: Låt A vara en händelse. Låt

$$X(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

Då kallas X indikatorn för A.

Skriv $p = \mathbb{P}(X=1)$. Då gäller $\mathbb{E}[X] = p$.

Eftersom $\mathbb{E}[X^2] = p$, får vi $Var(X) = p - p^2 = p(1-p)$

Man skriver ofta $X = I_A$

<u>Binomialfördelning</u>: Låt $A_1, A_2, \ldots$ vara oberoende händelser med sannolikhet p. Summan

$$X = \sum_{k=1}^{n} I_{A_k} \quad \text{sägs vara binomialfördelad}$$

med parametrar n & p. Skrivsätt $X \sim Bin(n,p)$

Man kan välja ut k kast på $\binom{n}{k}$ sätt

Sannol. att exakt dessa ger krona är $p^k(1-p)^{n-k}$

Alltså

$$\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

OBS: att $k\binom{n}{k} = n\binom{n-1}{k-1}$

<u>Därför</u>

$$\mathbb{E}[X] = \sum_{k=0}^{n} k\binom{n}{k} p^k (1-p)^{n-k} = np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} =$$

$$= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = np$$

där sista likheten följer av att termerna är frekvensfunktn. för $Bin(n-1,p)$ och summerar då till 1

<u>Ex</u>: Lag A & B möts i bäst av sju matcher. Antag sannol för A vinner 0,6 och att matcherna oberoende. Vad är sannorl. att A vinner serien.

Låt X vara antal matcher A vinner.
Då är $X \sim Bin(7, 0{,}6)$

$$\mathbb{P}(X \geq 4) = \sum_{k=4}^{7} \binom{7}{k} 0{,}6^k \cdot 0{,}4^{7-k} \approx 0{,}71 \qquad \square$$

Två observationer

- $X_1 \sim Bin(n_1, p)$, $X_2 \sim Bin(n_2, p)$ o. X_1, X_2 oberoende $\Rightarrow X_1 + X_2 \sim Bin(n_1 + n_2, p)$

- $X \sim Bin(n, p) \Rightarrow n - X \sim (n, 1-p)$

<u>Geometrisk fördelning</u> En slant ger klave med sannorl. p. Den singlas till första klave. Låt X vara antalet kast som krävdes. Vi får

$$\mathbb{P}(X = k) = p(1-p)^{k-1}, \quad k = 1, 2, \ldots$$

Man säger X är geometriskt fördelad med par. p
$X \sim Geo(p)$

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} \mathbb{P}(X > k) = \sum_{k=0}^{\infty} (1-p)^k = \frac{1}{p}$$

$$Var(X) = \frac{1-p}{p^2}$$

<u>Ex</u> X = antal tärningskast till första 6:an
$X \sim Geo(\frac{1}{6})$

Ex: Spela 100 rader på lotto varje vecka.
Låt X vara antal veckor tills sju rätt
Vi har $X \sim Geo(p)$ där p är sannol att få sju rätt en given vecka
Det finns $\binom{35}{7}$ rader

$$p = \frac{100}{\binom{35}{7}} \Rightarrow E[X] = \frac{\binom{35}{7}}{100} \approx 67245$$

67245 veckor $\approx$ 1293 år

Poissonfördelning

$$\mathbb{P}(X=k) = e^{-\lambda}\frac{\lambda^k}{k!} \qquad k = 0,1,2\ldots$$

Sägs X vara Poissonsfördelad med par. λ
$X \sim Poi(\lambda)$ (obs $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$ enligt taylor)
typ: incidenter per tidsperiod, oberoende incidenter
- Antal vulkanutbrott i världen per år
- fordon som passerar på enslig skogsväg
- Antal sönderfall på viss tid
- Ben brutet i VG i januari

$$\mathbb{E}[X] = e^{-\lambda}\sum_{k=0}^{\infty} k\frac{\lambda^k}{k!} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = \lambda$$

$$\mathbb{E}[X(X-1)] = e^{-\lambda}\sum_{k=2}^{\infty} k(k-1)\frac{\lambda^k}{k!} = \lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

Alltså blir $\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X] = \lambda^2 + \lambda$
Alltså $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

Ex på enslig skogsväg passerar genomsnitt 5.3 fordon per dag.

Sannolikhet att högst två passerar en given dag?

Rimligt $X \sim Poi(5.3)$

$$\mathbb{P}(X \leq 2) = e^{-5.3}\left(\frac{5.3^0}{0!} + \frac{5.3^1}{1!} + \frac{5.3^2}{2!}\right) \approx 0,102.$$

Ex Sverige har 10 milj invånare.

Sker c.a. 200.000 benbrott per år.

Sannolikheten att person lever 100 år utan benbrott

Sannol att få exakt 2 benbrott.

Verkar rimligt att $X \sim Pois(2)$ ty $\left(\frac{200\,000 \cdot 100}{10\,000\,000} = 2\right)$

$$\mathbb{P}(X = 0) = e^{-2}\left(\frac{2^0}{0!}\right) = e^{-2} \approx \frac{1}{8}$$

$$\mathbb{P}(X = 2) = e^{-2}\left(1 + \frac{2^1}{1!} + \frac{2^2}{2!}\right) = 2e^{-2} \approx \frac{1}{4}$$

Ex När n är stort och p litet gäller

$Bin(n,p) \approx poi(np)$

Som ex. Låt c konstant, låt $X \sim Bin(n, c/n)$ och låt $n \to \infty$. För ett k som inte ändrar sig gäller då

$$\mathbb{P}(X = k) = \binom{n}{k}\left(\frac{c}{n}\right)^k\left(1 - \frac{c}{n}\right)^{n-k}$$

$$\frac{\binom{n}{k}}{n^k} \xrightarrow{n\to\infty} \frac{1}{k!}$$

$$\left(1 - \frac{c}{n}\right)^{n-k} \xrightarrow{n\to\infty} e^{-c}$$

$$\Rightarrow \mathbb{P}(X = k) \longrightarrow \frac{c^k}{k!}e^{-c}$$

dvs när n stort är X $Poi(c)$ fördelad

Exponential fördelning.

En kont. sv. som har täthet

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0$$

Sägs vara exponentialfördelad med par λ

Kort: $X \sim \exp(\lambda)$

Uppstår av livslängd av saker som inte åldras

Vi ser att $F(x) = 1 - e^{-\lambda x}$, dvs $\mathbb{P}(X > x) = e^{-\lambda x}$

$$\mathbb{E}[X] = \int_0^\infty x \lambda e^{-\lambda x} dx = \int_0^\infty e^{-\lambda x} dx = \frac{1}{\lambda}$$

Det gäller också att

$$\mathbb{E}[X^2] = \int_0^\infty x^2 \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

så

$$Var(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Ex LED-lampa har livslängd 10 år. hur stor sannol. att den håller minst 10 år? Rimligt (?) anta lampans inte åldras. därmed livslängd $X \sim \exp(\lambda)$

$\mathbb{E}[X] = 10$ dvs. $\lambda = 0{,}1$

$$\mathbb{P}(X > 10) = e^{-0{,}1 \cdot 10} = e^{-1} \approx 0{,}37$$

Allmänt: $X \sim \exp(\lambda) \Rightarrow \mathbb{P}(X > \mathbb{E}[X]) = e^{-1}$

Exponentialfördelningen uppfyller glömskeegenskapen:

$$\mathbb{P}(X > t + x \mid X > t) = \mathbb{P}(X > x)$$

ty $\mathbb{P}(X > t + x \mid X > t) = \dfrac{\mathbb{P}(X > t + x, X > t)}{\mathbb{P}(X > t)} = \dfrac{\mathbb{P}(X > t + x)}{\mathbb{P}(X > t)} =$

$$= \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = \mathbb{P}(X > x)$$

Vice versa: antag att

$$\mathbb{P}(X>t+x \mid X>t) = \mathbb{P}(X>x) \qquad \forall x,t$$

Skriv $G(x) = \mathbb{P}(X>x)$ precis som ovan

$$\mathbb{P}(X>t+x \mid X>t) = \frac{\mathbb{P}(X>t+x)}{\mathbb{P}(X>t)}$$

antags vara lika med $\mathbb{P}(X>x)$ vilket ger

$G(x+t) = G(t)G(x) \quad \forall x,t$

subtrahera $G(x)$ från båda sidor, dela med t och låt $t \to 0$ och få

$$G'(x) = G'(0)G(x) \quad \forall x$$

Denna har unika lösn.

$$G(x) = e^{-\lambda x} \quad \text{där } \lambda = -G'(0) > 0 \quad \text{ty } G \text{ avtagande}$$

Alltså är X exponentialfördelad med par. λ

X har glömskeegenskapen omm $X \sim \exp(\lambda)$

<u>Ex</u> exp & geo-fördelning

Låt $X \sim \exp(\lambda)$ $Y = \lceil X \rceil$ Då gäller

$$\mathbb{P}(Y=k) = \mathbb{P}(X \in (k-1,k]) = \mathbb{P}(X>k-1) - \mathbb{P}(X>k) =$$
$$= e^{-\lambda(k-1)} - e^{-\lambda k} = e^{-\lambda(k-1)}(1-e^{-\lambda}) =$$
$$= p(1-p)^{k-1} \quad \text{med} \quad p = 1-e^{-\lambda}$$

<u>Kort:</u> $X \sim \exp(\lambda) \Rightarrow \lceil X \rceil \sim \mathrm{Geo}(1-e^{-\lambda})$ $\square$

prop: Om $X_1 \sim \exp(\lambda_1)$ & X_1, X_2 oberoende
$X_2 \sim \exp(\lambda_2)$
så gäller
$\min(X_1, X_2) \sim \exp(\lambda_1 + \lambda_2)$

Bevis:

$$\mathbb{P}(\min(X_1, X_2) > x) = \mathbb{P}(X_1 > x, X_2 > x) =$$

$$= \mathbb{P}(X_1 > x)\,\mathbb{P}(X_2 > x) = e^{-\lambda_1 x} e^{-\lambda_2 x} = e^{-(\lambda_1 + \lambda_2)x} \qquad \square$$

Låt $T_1, \dots, T_n$ vara oberoende & $\exp(\lambda)$-fördelade

$X = \sum_{k=1}^{n} T_k$ sägs vara gammafördelad med par n & λ

Kort $X \sim \Gamma(n, \lambda)$

prop: $\mathbb{P}(X > x) = e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}$

Bevis: återkommer

Poissonprocessen Låt $0 < \tau_1 < \tau_2 < \dots$
vara följd av tidpunkter då en viss typ av händelser inträffar.

Skriv $T_1 = \tau_1$ $T_2 = \tau_2 - \tau_1$ $T_3 = \tau_3 - \tau_2$ ...
för tidsmellanrum. Om $T_1, T_2, \dots$ är oberoende och $\exp(\lambda)$-fördelade kallas $\{\tau_1, \dots\}$ för en Poissonprocess med intensitet λ.

OBS: För en Poi-process med int λ gäller $\tau_n \sim \Gamma(n, \lambda)$

~~Låt~~

Låt $X(t) = \max\{n : \tau_n \leq t\}$ = antal impulser som kommit vid tid t.

Enligt prop ovan

$$\mathbb{P}(X(t)=n) = \mathbb{P}(\tau_n \leq t, \tau_{n+1} > t) =$$
$$= \mathbb{P}(\tau_{n+1} > t) - \mathbb{P}(\tau_n > t) =$$
$$= e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

Alltså:

$X(t) \sim Poi(\lambda t)$ Tack vare glömskeegenskapen gäller

$X(s) - X(t) \sim Poi(\lambda(s-t)) \quad \forall t < s$

Ex Lågtrafikerad väg, snitt fem bilar/timme

Sannolikhet att minst två passerar på en kvart?

Rimligt anta Poissonprocess $\Rightarrow \lambda = 5$ bilar/timme

Vi är intress $X(1/4) \sim Poi(5/4)$

$$\mathbb{P}(X(1/4) \geq 2) = 1 - e^{-5/4}\left(\frac{(5/4)^0}{0!} + \frac{(5/4)^1}{1!}\right) \approx 0,36 \qquad \square$$

Betrakta två oberoende Poissonprocesser, P_1 & P_2 med intensiteter λ_1 & λ_2

- Tiden T_{P_1} till nästa impuls i P_1 är $\exp(\lambda_1)$ fördelad
- — " — T_{P_2} — " — P_2 — " — $\exp(\lambda_2)$ — " —
- Dessa tider är oberoende

Tiden till nästa impuls i någon a. dem är $\min(T_1, T_2) \sim \exp(\lambda_1 + \lambda_2)$

Alltså: Sammanvägda processen är Poissonprocess med intensitet $\lambda_1 + \lambda_2$

Konsekvens: $X_1(t) + X_2(t) \sim Poi(\lambda_1 + \lambda_2)$

M a.

$Y_1 \sim Poi(\lambda_1)$ $Y_2 \sim Poi(\lambda_2)$

Y_1, Y_2 oberoende $\Rightarrow$ $Y_1 + Y_2 \sim Poi(\lambda_1 + \lambda_2)$

Direktbevis:

$$\mathbb{P}(Y_1 + Y_2 = k) = \sum_{j=0}^{k} \mathbb{P}(Y_1 = j)\,\mathbb{P}(Y_2 = k-j) =$$

$$= \sum_{j=0}^{k} e^{-\lambda_1} \frac{\lambda_1^j}{j!} e^{-\lambda_2} \frac{\lambda_2^{k-j}}{(k-j)!} =$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{\sum_{j=0}^{k} \binom{k}{j} \lambda_1^j \lambda_2^{k-j}}{k!} = e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^k}{k!} \quad \square$$

där sista likheten är binomialsatsen

Ex lågtrafikväg. 5 bil/timme 2 mc/timme

Sannol att exakt 3 passerar på en halvtimme?

Rimligt 2 Poiprocesser oberoende med int 5 & 2

$\Rightarrow$ Poi-process med int 7

$X(t) \Rightarrow X(\frac{1}{2}) \sim Poi(\frac{7}{2})$ Alltså

$$\mathbb{P}(X(\tfrac{1}{2}) = 3) = e^{-7/2} \frac{(7/2)^3}{3!} \approx 0{,}22$$

Om det passerar exakt n fordon på tid t vad är betingade sannol att k är mc

Skriv $X_b(t)$ $X_m(t)$

$$\mathbb{P}(X_m(t) = k \mid X(t) = n) = \frac{\mathbb{P}(X_m(t) = k, X(t) = n)}{\mathbb{P}(X(t) = n)} =$$

$$= \frac{\mathbb{P}(X_m(t) = k)\,\mathbb{P}(X_b(t) = n-k)}{\mathbb{P}(X(t) = n)} = \frac{e^{-2t} \frac{(2t)^k}{k!} e^{-5t} \frac{(5t)^{n-k}}{(n-k)!}}{e^{-7t} \frac{(7t)^n}{n!}} =$$

$$= \frac{\frac{2^k}{k!} \cdot \frac{5^{n-k}}{(n-k)!}}{\frac{7^n}{n!}} = \binom{n}{k} \left(\frac{2}{7}\right)^k \left(\frac{5}{7}\right)^{n-k}$$

Normalfördelning: Uppstår approximativt av summan av många små slumpmässiga bidrag.

<u>Def</u>: En kont sv X sägs vara normalfördelad med parametrar μ, σ^2 om dess täthet ges av

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

Kort: $X \sim N(\mu, \sigma^2)$

För $X \sim N(0,1)$ skrivs tätheten $\varphi(x)$ & fördelningsfktn $\Phi(x)$

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

Prop: Låt $Z \sim N(0,1)$ & $X = \mu + \sigma Z$ Då är $X \sim N(\mu, \sigma^2)$

Omvänt $X \sim N(\mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$

Bevis $F_X(x) = \mathbb{P}(\mu + \sigma Z \leq x) = \mathbb{P}\left(Z \leq \frac{x-\mu}{\sigma}\right) =$

$$= \int_{-\infty}^{\frac{x-\mu}{\sigma}} \varphi(t)\,dt$$

sätt $t = \frac{s-\mu}{\sigma} \Rightarrow dt = \frac{1}{\sigma} ds$ och

$t = \frac{x-\mu}{\sigma} \Longleftrightarrow s = x$

$$\rightarrow \int_{-\infty}^{x} \underbrace{\frac{1}{\sigma}\varphi\left(\frac{s-\mu}{\sigma}\right)}_{\text{tätheten för } N(\mu, \sigma^2)} ds \qquad \square$$

Låt $Z \sim N(0,1)$ Vi har $\mathbb{E}[Z] = \int_{-\infty}^{\infty} x\varphi(x)\,dx$
integralen är 0 ty φ är jämn
så $\mathbb{E}[Z] = 0$

Vidare är $\mathbb{E}[Z^2] = \int_{-\infty}^{\infty} x^2 \varphi(x)\,dx$ Fnk. $x\varphi(x)$ har
primitiv $-\varphi(x)$
Partiell integration ger

$$\mathbb{E}[Z^2] = \underbrace{\left[-x\varphi(x)\right]_{-\infty}^{\infty}}_{0} + \underbrace{\int_{-\infty}^{\infty} \varphi(x)\,dx}_{1} = 1$$

Alltså $\mathrm{Var}(Z) = 1$
För $X \sim N(\mu, \sigma^2)$ skriv $X = \mu + \sigma Z$
få

$$\mathbb{E}[X] = \mu \qquad \mathrm{Var}(X) = \sigma^2$$

OBS att

$$\mathbb{P}(X \le x) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

tack vare symmetri gäller

$\Phi(-x) = 1 - \Phi(x)$ Så man behöver bara $\Phi(x)$ för $x > 0$

Ex Längden X på måfå man är approx
normalfördelad med vv 180 cm standardav. 5 cm

tex $\mathbb{P}(X \le 176) \approx \Phi\left(\frac{176-180}{5}\right) = \Phi(-0{,}8) = 1 - \Phi(0{,}8) \approx 0{,}2$

$\mathbb{P}(X \ge 210) \approx 1 - \Phi\left(\frac{210-180}{5}\right) = 1 - \Phi(6) \approx 9{,}9 \cdot 10^{-10}$

Normalfördelningar är i tillämpningar alltid approximativ
och fungerar dåligt för extrema sannolikheter.

Ex Marmeladbolaget burkar med 400g standardavv. 4g hur mycket ska bolaget sikta på att fylla så vikten blir minst 400g med 0,95 sannolikhet?

Vad man siktar på är rimligt att ha som vv

Skriv μ $X \sim N(\mu, 4^2)$

Välj μ så $P(X > 400) = 0{,}95$

$$P(X>400) = 1 - \Phi\left(\frac{400-\mu}{4}\right) = \Phi\left(\frac{\mu-400}{4}\right) = 0{,}95$$

Nu är $\Phi^{-1}(0{,}95) \simeq 1{,}64$

$$\Rightarrow \frac{\mu-400}{4} \approx 1{,}64 \Rightarrow \mu \approx 406{,}6g$$

Lite standardvärden på Φ^{-1}

x	Φ^{-1}
0,9	1,28
0,95	1,64
0,975	1,96
0,99	2,33
0,995	2,58

Flerdim. fördelningar

Om X, Y två sv def på samma utfallsrum kallas (X, Y) för 2Dimensionell sv.

Den bivariata fördelningsfunktionen för (X,Y) ges av

$$F(x,y) = \mathbb{P}(X \leq x, Y \leq y), \quad (x,y) \in \mathbb{R}^2$$

Vi har

$$F(x,\infty) = \mathbb{P}(X \leq x, Y < \infty) = \mathbb{P}(X \leq x) = F_X(x)$$

$$F(\infty, y) = F_Y(y)$$

Diskreta tvådim sv.

X, Y diskret $\rightarrow$ (X, Y) diskret

OBS

$$V_X \times V_Y = \{(x_j, y_k) : x_j \in V_X, y_k \in V_Y\} \supseteq V_{X,Y}$$

Vi antar ändå $V_{X,Y} = V_X \times V_Y$

den bivariata frekvensfunk (sannolikhetsfunktion, pdf) (X,Y)

$$p(x_j, y_k) = \mathbb{P}(X = x_j, Y = y_k), \quad (x_j, y_k) \in V_X \times V_Y$$

Direkt obs:

$$p_X(x_j) = \mathbb{P}(X = x_j, Y \in V_Y) = \sum_{y_k \in V_Y} p(x_j, y_k)$$

och $p_Y(y_k) = \sum_{x_j \in V_X} p(x_j, y_k)$

Ex Antag (X,Y) antar alla par av heltal (x,y) : $1 \leq x \leq y \leq 4$ med samma sannolikhet

Det finns $4+3+2+1 = 10$ värden

$$p(x,y) = \frac{1}{10}, \quad 1 \leq x \leq y \leq 4$$

$$p_X(x) = \sum_{y=x}^{4} \frac{1}{10} = \frac{5-x}{10} \quad x = 1,2,3,4$$

$$p_Y(y) = \sum_{x=1}^{y} \frac{1}{10} = \frac{y}{10}, \quad y = 1,2,3,4$$

$\square$

Ex slå två tärningar

X första

Y summa

$p(x,y)$ punkt (x,y) möjligt om

$x \in \{1, \dots, 6\}$

$y \in \{1+x, \dots, 6+x\}$

$$p(x,y) = \frac{1}{36} \qquad \begin{array}{l} 1 \le x \le 6 \\ 1+x \le y \le x+6 \end{array}$$

Ex Låt $X \sim Poi(\lambda)$ givet $X = x$ låt $Y \sim Bin(x,p)$

vi får $p(x,y) = e^{-\lambda} \frac{\lambda^x}{x!} \binom{x}{y} p^y (1-p)^{x-y}$, $0 \le y \le x < \infty$

Marginalfördel för Y blir

$$p_Y(y) = \sum_{x=y}^{\infty} p(x,y) = e^{-\lambda} \sum_{x=y}^{\infty} \frac{\lambda^x}{x!} \binom{x}{y} p^y (1-p)^{x-y} =$$

$$= \frac{e^{-\lambda} \lambda^y p^y}{y!} \sum_{x=y}^{\infty} \frac{(\lambda(1-p))^{x-y}}{(x-y)!} = \frac{e^{-\lambda} (\lambda p)^y}{y!} e^{\lambda(1-p)} =$$

$$= e^{-\lambda p} \frac{(\lambda p)^y}{k!}$$ M.a.o $Y \sim Poi(\lambda p)$

Konsekvens: om man i en Poi process med int λ räknar med impulser m.a. sannol p får man en Poi-process med int λp

(s.k. uttunnad poi-process)

Sats X, Y diskreta oberoende

$\iff p(x_j, y_k) = p_X(x_j) p_Y(y_k) \quad \forall x_j \in V_X, y_k \in V_Y$

Bevis
Om X, Y ober. gäller

$$p(x_j, y_k) = \mathbb{P}(X = x_j, Y = y_k) =$$

$$= \mathbb{P}(X = x_j) \mathbb{P}(Y = y_k) =$$

$$= p_X(x_j) p_Y(y_k)$$

å andra sidan om p kan faktoriseras
får vi $\mathbb{P}(X \in A, Y \in B) = \sum_{x_j \in A} \sum_{y_k \in B} p(x_j, y_k) =$

$$= \sum_{x_j \in A} p_X(x_j) \sum_{y_k \in B} p_Y(y_k) = \mathbb{P}(X \in A) \mathbb{P}(Y \in B)$$

Kont tvådim s.v.: paret (X, Y) kont. om
$\exists$ täthetsfunk. $f: \mathbb{R}^2 \to \mathbb{R}^2$:

$$\mathbb{P}((X, Y) \in B) = \iint_B f(x, y)\, dx\, dy \quad \forall B \subseteq \mathbb{R}^2$$

Några obs?

- $F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s, t)\, dt\, ds$
- $f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$
- $\mathbb{P}(X \in A) = \mathbb{P}((X, Y) \in A \times \mathbb{R}) = \int_A \int_{-\infty}^{\infty} f(x, y)\, dy\, dx$

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

- $f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$

Ex Låt (X,Y) ha täthet $f(x,y) = c(x+3y)$
$0 \le x \le 1$
$0 \le y \le 1$ vad är c & vad är f_X & f_Y

Det måste gälla att $\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\,dy\,dx = 1$

så $1 = c\int_0^1\int_0^1 (x+3y)\,dx\,dy =$

$$= c\left(\int_0^1 x\,dx + 3\int_0^1 y\,dy\right) = 2c \Rightarrow c = \frac{1}{2}$$

$$f_X(x) = \frac{1}{2}\int_0^1 (x+3y)\,dy = \frac{1}{2}\left(x+\frac{3}{2}\right) \quad 0 \le x \le 1$$

& $f_Y(y) = \frac{1}{2}\int_0^1 (x+3y)\,dx = \frac{1}{2}\left(3y+\frac{1}{2}\right), \; 0 \le y \le 1$

Ex: Låt $D = \{(x,y) : x^2+y^2 = 1\}$ $(X,Y) \sim \text{likf}(D)$
dvs

$$f(x,y) = \begin{cases} \frac{1}{\pi}, & (x,y) \in D \\ 0, & \text{annars} \end{cases}$$

Låt $Z = |(X,Y)| = \sqrt{X^2+Y^2}$ vad är förd för Z

$$F_Z(z) = \mathbb{P}((X,Y) \in B_z) = \frac{1}{\pi}\text{Area}(B_z)$$

där $B_z = \{(x,y) : x^2+y^2 \le z^2\} \Rightarrow \text{Area}(B_z) = \pi z^2$

så $F_Z(z) = z^2, \; 0 \le z \le 1$

och

$f_Z(z) = 2z, \; 0 \le z \le 1$

forts:

$$f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy = \frac{1}{\pi}\int_{-\sqrt{1+x^2}}^{\sqrt{1+x^2}} dy = \frac{2}{\pi}\sqrt{1-x^2}$$

$\boxed{}$

Om $f(x,y) = f_X(x) f_Y(y)$ $\forall$ (x,y) gäller

$$\mathbb{P}(X \in A, Y \in B) = \int_A \int_B f(x,y)\,dy\,dx = \int_A f_X(x)\,dx \int_B f_Y(y)\,dy =$$

$$= \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B) \quad \text{dvs } X, Y \text{ ober}$$

Omvänt: X, Y ober $\Rightarrow$ $f(x,y) = f_X(x) f_Y(y)$ för (nästan) alla (x,y)

Funktioner av två sv

Om (X, Y) & $g: \mathbb{R}^2 \to \mathbb{R}$ vad är förd. $g(X,Y)$

Avgör från fall till fall

Ex: $(X,Y) \sim \text{likf}([0,1]^2)$ Låt $A = XY$ för $0 < a < 1$

$$F_A(a) = \mathbb{P}(XY \leq a) = \int_0^1 \int_0^{\min(a/x,\,1)} dy\,dx =$$

$$= \left(a + a\int_a^1 \frac{1}{x}\,dx\right) = a - a\ln(a) = a(1 - \ln(a))$$

Derivering ger

$$f_A(a) = -\ln(a), \quad 0 < a < 1$$

$\square$

OSL i två dim

- Diskret $\mathbb{E}[g(X,Y)] = \sum_{(x_j, y_k) \in V_{X,Y}} g(x_j, y_k)\,p(x_j, y_k)$
- kont: $\mathbb{E}[g(X,Y)] = \int_{\mathbb{R}}\int_{\mathbb{R}} g(x,y) f(x,y)\,dx\,dy$

Bevis diskret Låt $V_g = g(V_{X,Y})$

$$\mathbb{E}[g(X,Y)] = \sum_{z \in V_g} z \, \mathbb{P}(g(X,Y) = z) =$$

$$= \sum_{z \in V_g} z \sum_{(x_j, y_k) : g(x_j, y_k) = z} p(x_j, y_k) =$$

$$= \sum_{z \in V_g} - - - -$$

$$= \sum_{(x_j, y_k) \in V_{X,Y}} g(x_j, y_k) p(x_j, y_k)$$

Följdsats: $\forall \; X, Y$ gäller $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Bevis kont. Enl. osv

$$\mathbb{E}[X+Y] = \int_{\mathbb{R}} \int_{\mathbb{R}} (x+y) f(x,y) \, dx \, dy =$$

$$= \int_{\mathbb{R}} x \int_{\mathbb{R}} f(x,y) \, dy \, dx + \int_{\mathbb{R}} y \int_{\mathbb{R}} f(x,y) \, dx \, dy =$$

$$= \int_{\mathbb{R}} x f_X(x) \, dx + \int_{\mathbb{R}} y f_Y(y) \, dy = \mathbb{E}[X] + \mathbb{E}[Y]$$

Ex Låt $I_1, \ldots, I_n$

$\mathbb{P}(I_k = 1) = 1 - \mathbb{P}(I = 0) = p \qquad X = \sum_{k=1}^{n} I_k$ Följdsats $\Rightarrow$

$$\Rightarrow \mathbb{E}[X] = \sum_{k=1}^{n} \mathbb{E}[I_k] = np$$

Ex visar detta att $X \sim Bin(n,p) \rightarrow \mathbb{E}[X] = np$

Betingade fördelningar

Antag X & Y diskreta

$$p_{Y|X}(y|x) = \mathbb{P}(Y=y \mid X=x) = \frac{p(x,y)}{p_X(x)}, \quad (x,y) \in V_{X,Y}$$

Ex

$x \backslash y$	0	1
0	1/10	1/4
1	1/5	9/20

$$p_{Y|X}(0|1) = \frac{p(1,0)}{p_X(1)} = \frac{1/5}{1/5 + 9/20} = \frac{4}{13} \qquad p_{X|Y}(1|1) = \frac{p(1,1)}{p_Y(1)} = \frac{9/20}{1/4 + 9/20} = \frac{9}{14}$$

Ex två tärning

X första tärning

Y summa

$p_{Y|X}(y|x) = 1/6, \quad x+1 \le y \le x+6$ på får vi en TSL

$$p_Y(9) = \sum_{x \in V_X} p_{Y|X}(y|x)\, p_X(x) = \frac{1}{6}\left(0 + 0 + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right) = \frac{1}{9}$$

i allmänhet $\quad p_Y(y) = \dfrac{6 - |y-7|}{36}$

Motsvarighet för kont. sv, Vill betinga på $\{X = x\}$ även om $P(X=x)=0$ Antag (X,Y) kont

$$\Rightarrow \mathbb{P}(Y \in B \mid X = x) \approx \mathbb{P}(Y \in B \mid X \in x \pm \Delta x) = \frac{P(Y \in B,\ X \in x \pm \Delta x)}{P(X \in x \pm \Delta x)} =$$

$$= \frac{\int_B \int_{x-\Delta x}^{x+\Delta x} f(x,y)\,dx\,dy}{\int_{x-\Delta x}^{x+\Delta x} f_X(x) \cdot 1\,dx} \approx \frac{2\Delta x \int_B f(x,y)\,dy}{2\Delta x\, f_X(x)} = \int_B \frac{f(x,y)}{f_X(x)}\,dy$$

Def täthetsfunktionen för y givet $X=x$ som

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

Då får vi per def

$$P(Y\in B \mid X=x) = \int_B f_{Y|X}(y|x)\,dy = \int_B \frac{f(x,y)}{f_X(x)}\,dy$$

Notera:

$$P(Y\in B) = \int_{\mathbb{R}}\int_B f(x,y)\,dy\,dx = \int_{\mathbb{R}} f_X(x)\int_B f_{Y|X}(y|x)\,dy\,dx =$$

$$= \int_{\mathbb{R}} P(Y\in B \mid X=x)\, f_X(x)\,dx$$

Detta är en kont. variant av totala sannolikhetslagen

Annan variant

$$f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx = \int_{\mathbb{R}} f_{Y|X}(y|x) f_X(x)\,dx$$

Lite intim parerbes Y disk. X kont.

Antag $\exists f : V_Y \times \mathbb{R} : \forall y \in V_X \ \forall B \subseteq \mathbb{R}$:

$$P(X\in B, Y=y_k) = \int_B f(x,y_k)\,dx$$

Def:

$$P(Y=y_k \mid X=x) = \frac{f(x,y_k)}{f_X(x)}$$

TSL:

$$\mathbb{P}(Y=y_k) = \int_{\mathbb{R}} \mathbb{P}(Y=y_k \mid X=x)\, f_X(x)\,dx$$

Vi def också

$$f_{x|y}(x|y) = \frac{f(x,y_k)}{p_Y(y_k)}$$

$$\Rightarrow f_X(x) = \sum_{y_k \in V_y} f_{x|y}(x|y_k)\, p_Y(y_k)$$ $\square$

Om (X,Y) kont X,Y obor

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

Ex $(X,Y) \sim \text{Likf}(D)$ $D = \{(x,y): x^2+y^2<1\}$

tidigare $f(x,y) = \frac{1}{\pi}$, $(x,y) \in D$

$$f_X(x) = \frac{2}{\pi}\sqrt{1-x^2}, \quad 0<x<1$$

Alltså

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{1}{2\sqrt{1-x^2}}, \quad -\sqrt{1-x^2} < y < \sqrt{1-x^2}$$

Menv. den bet. fördelningen för Y givet $X=x$ är $\text{likf}(-\sqrt{1-x^2}, \sqrt{1-x^2})$

Ex Välj $X \sim \text{likf}(0,1)$, givet $X=x$ $Y \sim \text{Likf}(0,x)$

Vi har $f_X(x)=1$ $0<x<1$ $f_{Y|X}(y|x) = \frac{1}{x}$ $0<y<x$

Vad är $\mathbb{P}(Y<X^2)$?

$$\mathbb{P}(Y<X^2) = \int_0^1 \mathbb{P}(Y<x^2 | X=x)\, f_X(x)\,dx =$$

$$= \int_0^1\int_0^{x^2} \frac{1}{x}\,dy\,dx = \int_0^1 x\,dx = \frac{1}{2}$$ $\square$

Om $T_1, \dots$ är $\exp(\lambda)$-fördelad & oberoende
är
$$X_n = \sum_{k=1}^{n} T_k \sim \Gamma(n,\lambda)$$ härdat att
$$\mathbb{P}(X_n > x) = $$
$$= e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}.$$

Sant för $n=1$, så antag för $n=m$ obs att

$$X_{m+1} = X_m + T_{m+1}$$
$$\mathbb{P}(X_{m+1} > x) = \int_0^\infty \mathbb{P}(X_{m+1} > x \mid T_{m+1} = t) f_T(t)\, dt =$$
$$= \int_0^\infty \mathbb{P}(X_m > x - t \mid T_{m+1} = t) f_T(t)\, dt =$$
$$= \int_0^x \lambda e^{-\lambda t} e^{-\lambda(x-t)} \sum_{k=0}^{m-1} \frac{(\lambda(x-t))^k}{k!} dt + e^{-\lambda x}$$
Men
$$\int_0^x \frac{\lambda^{k+1}(x-t)^k}{k!} dt = \frac{(\lambda x)^{k+1}}{(k+1)!}$$
Så hela uttrycket blir
$$e^{-(\lambda x)} \sum_{k=0}^{m} \frac{(\lambda x)^k}{k!}$$
$\blacksquare$

Prop: om X & Y ober gäller

a) $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

b) $Var(X+Y) = Var(X) + Var(Y)$

Bevis a) kont.
$$\mathbb{E}[XY] = \iint xy f(x,y)\, dx\, dy = \int x f_X(x)\, dx \int y f_Y(y)\, dy = \mathbb{E}[X]\mathbb{E}[Y]$$
för b):
$$Var(X+Y) = \mathbb{E}[(X+Y-\mu_x-\mu_y)^2] = \mathbb{E}[(X-\mu_x)^2] + \mathbb{E}[(Y-\mu_y)^2] +$$
$$+ 2\mathbb{E}[(X-\mu_x)(Y-\mu_y)] = Var(X) + Var(Y)$$
ty sista termen är noll då X, Y ober

Mer generellt $X_1, \ldots, X_n$ medför

$$\mathbb{E}\left[\prod X_k\right] = \prod \mathbb{E}[X_k], \quad Var\left(\sum X_k\right) = \sum Var(X_k)$$

<u>Ex</u> $X \sim Bin(n,p)$. Skriv $X = \sum_{k=1}^{n} I_k$ där I_k ober.

$$Var(X) = \sum_{k=1}^{n} Var(I_k) = np(1-p)$$

$\square$

Låt X_1, X_2 ober & lika fördelade med $\mathbb{E}[X_1] = \mu$ & $Var(X_1) = \sigma^2 < \infty$

Skriv $$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

$$Var(\overline{X}_n) = \frac{1}{n^2}\sum_{k=1}^{n} Var(X_k) = \frac{\sigma^2}{n}$$

<u>Stora talens lag</u>: För varje $\epsilon > 0$ gäller

$$\lim_{n\to\infty} \mathbb{P}(|\overline{X}_n - \mu| \geq \epsilon) = 0$$

<u>Bevis</u>: Enligt Chebyshevs olikhet gäller

$$\mathbb{P}(|\overline{X}_n - \mu| \geq \epsilon) \leq \frac{Var(\overline{X}_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$$

Betingade väntevärden

$\mathbb{E}[Y|X=x]$ är vv av den sv som har samma fördelning som Y givet $X=x$

Def:

- Y diskret: $\mathbb{E}[Y|X=x] = \sum_{y_k \in V_Y} y_k \, p_{Y|X}(y_k|x)$
- Y kont: $\mathbb{E}[Y|X=x] = \int_{-\infty}^{\infty} y \, f_{Y|X}(y|x) \, dy$

TSL:

- X diskret: $\mathbb{E}[Y] = \sum_{x_j \in V_X} \mathbb{E}[Y|X=x_j] \, p(x_j)$
- X kont: $\mathbb{E}[Y] = \int_{-\infty}^{\infty} \mathbb{E}[Y|X=x] \, f_X(x) \, dx$

Bevis kont:

$$\int_{-\infty}^{\infty} \mathbb{E}[Y|X=x] f_X(x) dx = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy \right) f_X(x) dx =$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} y f(x,y) \, dy \, dx \underset{\uparrow \text{OSL}}{=} \mathbb{E}[Y] \quad \square$$

Ex $X \sim \text{likf}(0,1)$ $Y \sim \text{likf}(0,X)$

förut: $f(x,y) = \frac{1}{x}$ $0 < y < x < 1$ Enl OSL.

$$\mathbb{E}[Y] = \int_0^1 \int_0^x y f(x,y) \, dy \, dx = \int_0^1 \int_0^x y \cdot \frac{1}{x} \, dy \, dx = \frac{1}{2} \int_0^1 x \, dx =$$

$$= \frac{1}{4}$$

Alt. $\mathbb{E}[Y] = \int_0^1 \mathbb{E}[Y|X=x] f_X(x) dx = \int_0^1 \frac{x}{2} dx = \frac{1}{4}$

Def $\mathbb{E}[Y|X]$ är den s.v som antar värdet $\mathbb{E}[Y|X=x]$ då $X=x$.

M.a.o. om funktionen g ges av $g(x)=\mathbb{E}[Y|X=x]$ så är $\mathbb{E}[Y|X]=g(X)$

kortform av TSL för v.v.

Antag (X,Y) kont. (eller diskret). Enl. TSL, CSL

$$\mathbb{E}[Y]=\int \mathbb{E}[Y|x]f_X(x)\,dx=\int g(x)f_X(x)\,dx=\mathbb{E}[g(x)]=\mathbb{E}[\mathbb{E}[Y|X]]$$

Alltså $\mathbb{E}[\mathbb{E}[Y|X]]=\mathbb{E}[Y]$

Ex Låt $X\sim \mathrm{Geo}(p)$ & sedan $Y\sim \mathrm{Bin}(X,r)$. Vad är $\mathbb{E}[Y]$?

$$\mathbb{E}[Y]=\mathbb{E}[\mathbb{E}[Y|X]]=\mathbb{E}[Xr]=\frac{r}{p}$$

Ex $X_1, X_2, \ldots$ oberoende & $\mathbb{E}[X_k]=\mu \quad \forall k$

Låt N oberoende av X_k:na & pos. heltalsvärd

$$\mathbb{E}\left[\sum_{k=1}^{N}X_k\right]=\mathbb{E}\left[\mathbb{E}\left[\sum_{k=1}^{N}X_k \,\Big|\, N\right]\right]=\mathbb{E}[N\mu]=\mu\,\mathbb{E}[N] \qquad \square$$

OBS att $\mathbb{P}(A)=\mathbb{E}[\mathbb{P}(A|X)]$

ty. h.l. $=\int \mathbb{P}(A|X=x)f_X(x)\,dx=P(A)$

Kovarians & korrelation: X, Y 2st s.v.

<u>Def</u>: $Cov(X,Y) = \mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$

Ett mått på hur X & Y samvarierar:

$Cov(X,Y) > 0$ → positivt beroende

$Cov(X,Y) < 0$ → Negativt beroende

Formel: $Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

Om X, Y ober. gäller $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

$\Rightarrow Cov(X,Y) = 0$ om X, Y ober.

<u>OBS</u>: X, Y ober. $\Longrightarrow Cov(X,Y) = 0$

inte nödvändigtvis omvänt.

Några obs:

$Cov(X, a) = 0$

$Cov(X, X) = Var(X)$

$Cov(X,Y) = Cov(Y,X)$

Formeln ger också att kovariansen är bilinjär:

$$Cov(a_1X_1 + a_2X_2, b_1Y_1 + b_2Y_2) = a_1b_1 Cov(X_1,Y_1) + a_1b_2 Cov(X_1,Y_2) + a_2b_1 Cov(X_2,Y_1) + a_2b_2 Cov(X_2,Y_2)$$

<u>Ex</u>:

$$Cov(X+a, Y+b) = Cov(X,Y) + Cov(X,b) + Cov(a,Y) + Cov(a,b) = Cov(X,Y)$$

Variansen för en summa:

$$Var(X+Y) = Cov(X+Y, X+Y) = Var(X) + 2Cov(X,Y) + Var(Y)$$

Ex $\Theta \sim \text{likf}(0, 2\pi)$: $X = \cos\theta$, $Y = \sin\theta$

$$\mathbb{E}[X] = \int_0^{2\pi} \cos\theta \frac{1}{2\pi} d\theta = 0, \quad \mathbb{E}[Y] = \int_0^{2\pi} \sin\theta \frac{1}{2\pi} d\theta = 0$$

$$\mathbb{E}[XY] = \int_0^{2\pi} \cos\theta \sin\theta \frac{1}{2\pi} d\theta = 0$$

Alltså $\text{Cov}(X, Y) = 0 - 0 \cdot 0 = 0$ men X, Y är ej oberoende

---

Ex: Låt X, Y oberoende $\text{likf}(0,1)$. Låt $Z = XY$ $W = X + Y$. Bestäm $\text{Cov}(Z, W)$

$$\mathbb{E}[Z] = \mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y] = \frac{1}{4}$$

$$\mathbb{E}[W] = \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 1$$

$$\mathbb{E}[ZW] = \mathbb{E}[XY(X+Y)] = \mathbb{E}[X^2]\,\mathbb{E}[Y] + \mathbb{E}[X]\,\mathbb{E}[Y^2] = \frac{1}{3}$$

Alltså

$$\text{Cov}(Z, W) = \frac{1}{3} - \frac{1}{4} \cdot 1 = \frac{1}{12} \quad \text{pos. beroende}$$

Hur starkt är det positiva beroendet?

Schwarz olikhet: $\forall$ s.v. X & Y gäller

$$\mathbb{E}[XY]^2 \leq \mathbb{E}[X^2]\,\mathbb{E}[Y^2]$$

Bevis $\forall$ t konstant gäller

$$0 \leq \mathbb{E}[(X - tY)^2] = \mathbb{E}[X^2] - 2t\,\mathbb{E}[XY] + t^2\,\mathbb{E}[Y^2]$$

H.L. blir minst då $t = \dfrac{\mathbb{E}[XY]}{\mathbb{E}[Y^2]}$ insättning i H.L. ger

$$0 \leq \mathbb{E}[X^2] - 2\frac{\mathbb{E}[X,Y]^2}{\mathbb{E}[Y^2]} + \frac{\mathbb{E}[XY]^2}{\mathbb{E}[Y^2]} = \mathbb{E}[X^2] - \frac{\mathbb{E}[XY]^2}{\mathbb{E}[Y^2]}$$

förläng med $\mathbb{E}[Y^2]$ och då är det klart

$\square$

Def korrelationskoefficienten för (X,Y) ges av
$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$
Vi har:
$$-1 \le \rho(X,Y) \le 1$$
enligt schwarz anv på $X-\mu_X$ & $Y-\mu_Y$. Vi har också

- $\rho(X,Y) = \rho(aX, bY)$ skalningsvarians
- $Y = aX+b$ för konstant $a>0 \iff \rho(X,Y)=1$
- $Y = aX+b$ — " — $a<0 \iff \rho(X,Y)=-1$

Def: om $\rho(X,Y)=0$ kallas X & Y okorrelerade

Ex X,Y ober. likf$(0,1)$, $Z = XY$, $W = X+Y$
$Cov(Z,W) = \frac{1}{12}$ vad är $\rho(Z,W)$?
$Var(W) = Var(X+Y) = Var(X) + Var(Y) = \frac{1}{6}$ ty
variansen av likf$(0,1) = 1/12$
$Var(Z) = Var(XY) = \mathbb{E}[X^2Y^2] - \mathbb{E}[XY]^2 = \mathbb{E}[X^2]\mathbb{E}[Y^2] - (\mathbb{E}[X]\mathbb{E}[Y])^2 =$
$= \left(\frac{1}{3}\right)^2 - \left(\frac{1}{4}\right)^2 = \frac{7}{144}$ ty $\mathbb{E}[X^2] = \frac{1}{3}$
Alltså $\rho(Z,W) = \frac{1/12}{\sqrt{\frac{1}{6}\cdot\frac{7}{144}}} = \sqrt{\frac{6}{7}}$

Ex: Bästa linjära prediktor: Välj a & b för att minimera
$\mathbb{E}[(Y-(aX+b))^2]$. Antag $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ & $Var(X) = Var(Y) = 1$
då är $\rho = Cov(X,Y) = \mathbb{E}[XY]$ &
$\mathbb{E}[(Y-(aX+b))^2] = \mathbb{E}[Y^2] - 2\mathbb{E}[Y(aX+b)] + \mathbb{E}[(aX+b)^2] =$
$= 1 - 2a\rho + a^2 + b^2$
detta min med $a = \rho$ $b=0$. Man ser i detta fall är bästa linjära prediktor av Y givet X lika med ρX. Allmänt ges den av
$$\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(X-\mu_X)$$
$\square$

Centrala Gränsvärdessatsen

$X_1, X_2, \ldots$ oberoende & likafördelade med $\mathbb{E}[X_1] = \mu$ & $Var(X_1) = \sigma^2 < \infty$. Låt

$$S_n = \sum_{k=1}^{n} X_k$$

Då gäller $\forall x \in \mathbb{R}$ att

$$\mathbb{P}\left(\frac{S_n - n\mu}{\sigma \sqrt{n}} \leq x \right) \longrightarrow \Phi(x)$$

då $n \to \infty$. M.a.o. S_n är approx $N(n\mu, n\sigma^2)$-fördelad för stora n. (Oavsett X_k:nas fördelning)

<u>Ex</u> genomsnitt 100 jordbävningar per år

Vad är sannol. att 110 inträffar ett år

Rimligt att anta jb Poi-process med int 100

jb ~ Poi(100) lite jobbigt att räkna på

Alt. tider $T_1, T_2, \ldots$ mellan jb, oberoende & exp(100) fördelade. Låt

$$S_n = \sum_{k=1}^{n} T_k$$

Vi söker $\mathbb{P}(S_{110} \leq 1)$. Vi har $\mathbb{E}[T_k] = \frac{1}{100} = 0{,}01$

$\sigma^2 = Var(T_k) = \frac{1}{100^2}$ Alltså enligt CGS

$$\mathbb{P}(S_{110} \leq 1) = \mathbb{P}\left(\frac{S_{110} - 110 \cdot 0{,}01}{0{,}01 \cdot \sqrt{100}} \leq \frac{1 - 110 \cdot 0{,}01}{0{,}01\sqrt{100}} \right) \approx \Phi\left(-\frac{0{,}1}{\sqrt{0{,}011}}\right) \approx$$

$\approx 0{,}17$

Ex $X \sim Bin(n,p)$ med n stor & p icke extremt ($\frac{1}{n} << p << 1-\frac{1}{n}$ vi tänker $n \to \infty$ & p fixt) Skriv

$$X = \sum_{k=1}^{n} I_k .$$

Vi har $\mu = \mathbb{E}[I_k] = p$ & $\sigma^2 = Var(I_k) = p(1-p)$. enl. CGS

$$\mathbb{P}\left(\frac{X-np}{\sqrt{np(1-p)}} \leq x\right) \to \Phi(x)$$

Exvis. slå tärning 700 ggr & låt X antal 6:or. Vad är $\mathbb{P}(X \geq 100)$ (X diskret $\Rightarrow \mathbb{P}(X > 99,5)$)

$$\mathbb{P}(X > 99,5) = 1 - \mathbb{P}\left(\frac{X - 700\cdot\frac{1}{6}}{\sqrt{700\cdot\frac{1}{6}\cdot\frac{5}{6}}} \leq \frac{99,5 - 700\cdot\frac{1}{6}}{\sqrt{700\cdot\frac{1}{6}\cdot\frac{5}{6}}}\right) \approx$$

$$\approx 1-\Phi(-1,74) = \Phi(1,74) \approx 0,959 \qquad \square$$

OBS: CGS fungerar även utan likafördelning under vissa enkla förutsättningar, även utan oberoende under tuffare krav. I båda fallen måste variansen justeras. Se upp med extrema sannolikheter

Statistik

Samtliga parametrar okända, skatta parametrar utifrån data. Skrivsätt iid = independent and identically distributed = oberoende & likafördelade

Def: Låt $X_1, X_2, \ldots, X_n$ vara iid och förd. som X. Då kallas $X_1, \ldots, X_n$ för ett stickprov på X (eller F_X)

Antag att F_X beror på en parameter θ. Exvis $X \sim Poi(\theta)$, $X \sim exp(\theta)$, $X \sim N(\theta, 1)$ eller $\mathbb{P}(X=1) = 1 - \mathbb{P}(X=0) = \theta$. Parametern kan vara flerdimensionell t.ex. $\theta = (\mu, \sigma^2)$ & $X \sim N(\mu, \sigma^2)$

En funktion av $X_1, \ldots, X_n$ som anv. till skatta θ kallas punktskattning (estimator) av θ. Standard beteckningar $\hat{\theta}$, $\hat{\theta}_n$, $\hat{\theta}(X_1, \ldots, X_n)$

OBS $\hat{\theta}$ är s.v. Efter data $X_1 = x_1, \dots$ observerats får man ett givet värde $\hat{\theta}(x_1, \dots, x_n)$ ett estimat av θ

Def: Om det, oavsett det korrekta värdet på θ, gäller att

$$\mathbb{E}[\hat{\theta}] = \theta$$

kallas $\hat{\theta}$ för en väntevärdesriktig skattning av θ. (unbiased) (vvr)

Def: Om

$$\lim_{n \to \infty} \mathbb{P}(|\hat{\theta}_n - \theta| > \varepsilon) = 0$$

kallas $\hat{\theta}$ för en konsistent skattning av θ

Prop: Om $\hat{\theta}_n$ är vvr & $Var(\hat{\theta}_n) \to 0$ så är $\hat{\theta}_n$ konsistent.

Bevis: enl. Chebyshevs olikhet:

$$\mathbb{P}(|\hat{\theta}_n - \theta| > \varepsilon) \le \frac{Var(\hat{\theta}_n)}{\varepsilon^2} \to 0$$

$\square$

Låt $X_1, \dots, X_n$ vara stickprov på en sv X med $\mathbb{E}[X] = \mu$ och $Var(X) = \sigma^2 < \infty$, Kom ihåg att

$$\bar{X} = \frac{1}{n} \sum_{k=1}^{n} X_k$$

Vi vet att $\mathbb{E}[\bar{X}] = \mu$ & $Var(\bar{X}) = \frac{\sigma^2}{n}$ så enl. prop är $\bar{X}$ en konsistent skattning av μ $\square$

Man ska se konsistens som mycket önskvärt krav och som i sig själv inte räcker för att säga att en skattning är bra.

Def: Om $\hat{\theta}$ & $\tilde{\theta}$ är två vvr skattningar av θ &

$$Var(\hat{\theta}) < Var(\tilde{\theta}) \quad \forall \theta$$

säger man att $\hat{\theta}$ är mer effektiv än $\tilde{\theta}$

Skattning av varians: Låt $X_1, \dots, X_n$ vara stickprov på en s.v. X med $\mathbb{E}[X] = \mu$ & $Var(X) = \sigma^2 < \infty$

man brukar skatta σ^2 med

$$s^2 = \frac{1}{n-1} \sum_{k=1}^{n} (X_k - \overline{X})^2$$

Prop: s^2 är vvr, om $\mathbb{E}[X^4] < \infty$ gäller också att s^2 är konsistent.

Delbevis: Lätt att se $s^2 = \frac{1}{n-1}\left(\sum_{k=1}^{n} X_k^2 - n\overline{X}^2\right)$ Det gäller att

$$\mathbb{E}[\overline{X}^2] = Var(\overline{X}) + \mathbb{E}[\overline{X}]^2 = \frac{\sigma^2}{n} + \mu^2$$

Vi har också att $\mathbb{E}[X_k^2] = Var(X_k) + \mathbb{E}[X_k]^2 = \sigma^2 + \mu^2$ Detta ger

$$\mathbb{E}[s^2] = \frac{1}{n-1}\left(n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right) = \sigma^2$$

Konfidensintervall Låt $X_1, \dots, X$ vara sp på en sv X vars förd. beror av okänd par. θ. Om T_1 & T_2 är två funktioner av $X_1, \dots, X_n$ &

$$\mathbb{P}(T_1 \leq \theta \leq T_2) = q$$ så kallas $[T_1, T_2]$ konfidensintervall för θ av konfidensgrad q

Man skriver $T_1 \leq \theta \leq T_2 \quad (q)$

OBS: Sannol. q gäller innan T_1, T_2 observerats. När vi obs. $T_1 = t_1$, $T_2 = t_2$ och skriver $t_1 \leq \theta \leq t_2$ är detta ett påstående som är antingen sant eller falskt, men som, innan detta observerats, hade sannol. q att få värden t_1 & t_2 som skulle göra detta sant.

Låt $X \sim \mathrm{Likf}(0, \theta)$, θ okänd & $X_1, \dots, X_n$ ett sp på X. Vi vet att $\mathbb{E}[\overline{X}] = \frac{\theta}{2}$ så $\hat{\theta} = 2\overline{X}$ är en vvr skattning på θ.

Alt. Skriv $M = \max\{X_1, \dots, X_n\}$, skulle vara dumt att skatta θ mindre än M.

Gör skattning baserat på M

$$F_M(x) = \mathbb{P}(M \le x) = \mathbb{P}(X \le x)^n = \left(\frac{x}{\theta}\right)^n$$

$$f_M(x) = \frac{1}{\theta^n} n x^{n-1}, \quad x \in (0, \theta)$$

$$\mathbb{E}[M] = \int_0^\theta x \frac{1}{\theta^n} n x^{n-1} dx = \frac{n}{\theta^n} \int_0^\theta x^n dx = \frac{n}{n+1} \theta$$

Alltså $\tilde{\theta} = \frac{n+1}{n} M$ är en annan vvr skattning av θ

Vilken är mest effektiv?

$$Var(\hat{\theta}) = \frac{4}{n} Var(X) = \frac{\theta^2}{3n}$$

$$\mathbb{E}[M^2] = \frac{n}{\theta^n} \int_0^\theta x^{n+1} dx = \frac{n}{n+2} \theta^2$$

$$Var(M) = \left(\frac{n}{n+2} - \frac{n^2}{(n+1)^2}\right) \theta^2$$

$$Var(\tilde{\theta}) = \left(\left(\frac{n+1}{n}\right)^2 \frac{n}{n+2} - 1\right) \theta^2 = \frac{\theta^2}{n(n+2)}$$

Alltså $\tilde{\theta}$ är mycket mer effektiv än $\hat{\theta}$

Konfidensintervall: Verkar bra att låta T_1, T_2 vara funktioner av M. Vi gör ett symmetriskt konf. interval med konfidensgrad 95% dvs tar

$\mathbb{P}(T_1 \ge \theta) = \mathbb{P}(T_2 \le \theta) = 0{,}025$ Vi har

$\mathbb{P}(M \le x) = \frac{x^n}{\theta^n} = 0{,}025$ då $x = 0{,}025^{1/n} \theta$

Alltså:

$$0{,}025 = \mathbb{P}(M \le 0{,}025^{1/n} \theta) = \mathbb{P}\left(\theta \ge \frac{M}{0{,}025^{1/n}}\right)$$

P.s.s.

$$\mathbb{P}\left(\theta \le \frac{M}{0{,}975^{1/n}}\right) = 0{,}025$$

Alltså $\frac{M}{0{,}975^{1/n}} \le \theta \le \frac{M}{0{,}025^{1/n}}$ (95%)

Konfidensintervall för μ i normalfördelning

Vi har $X_1, \ldots, X_n$ sp på $N(\mu, \sigma^2)$.

Börja med att anta σ^2 är känd

Basera på $\bar{X}$. det gäller att $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

dvs $$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Skriv $z_\alpha = \phi^{-1}(1-\alpha)$. Då gäller

$$\mathbb{P}\left(-z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) = 1-\alpha$$

Lös ut μ och få det symmetriska konfidensintervallet

$$\mu \in \bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Uppåt begr konf. interv. : Utnyttja att

$$\mathbb{P}\left(-z_\alpha \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right) = 1-\alpha$$

Lös ut μ och få

$$\mu \leq \bar{X} + z_\alpha \frac{\sigma}{\sqrt{n}} \qquad (1-\alpha)$$

Om σ^2 okänd ersätt σ^2 m. s^2. Det gäller att

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$$

Härma nu rakt av

$$\mu \in \bar{X} \pm F^{-1}_{t_{n-1}}\left(1-\frac{\alpha}{2}\right)\frac{s}{\sqrt{n}} \quad \text{med konfgrad } 1-\alpha$$

~~Ex~~ Elnätets mäts spänning en gång om dagen i en vecka.

Resultat:

230.7, 226.9, 228.6, 232.2, 227.3, 227.0, 229.1

Rimligt att anta $N(\mu, \sigma^2)$ fördelade

Vi har $\bar{X} = 228.9$ och $s^2 = 2.014^2$ så

$s/\sqrt{n} = \frac{2.014}{\sqrt{7}} = 0.761$ För t_6-förd gäller t.ex.

$F_{t_6}^{-1}(0.975) = 2.46$, $F_{t_6}^{-1}(0.995) = 3.71$, $F_{t_6}^{-1}(0.9) = 1.44$

Alltså

$\mu = 228.9 \pm 2.46 \cdot 0.761 = 228.9 \pm 1.9$ (95%)

$\mu = 228.9 \pm 3.71 \cdot 0.761 = 228.9 \pm 2.8$ (99%)

$\mu \leq 228.9 + 1.44 \cdot 0.761 = 230.0$ (90%)

Om vi nu vet att $\sigma^2 = 2^2$ & vill ha 95% symmetriskt konfidensintervall utnyttjar vi att $z_{0.975} = 1.96$. och får

$\mu = 228.9 \pm 1.96 \cdot \frac{2}{\sqrt{7}} = 228.9 \pm 1.5$ (95%)

□

Det gäller att $F_{t_n}^{-1}(a) \to \Phi^{-1}(a)$ då $n \to \infty$

Tumregel: $n \geq 100 \to t_n \approx N(0,1)$

Konfidensintervall för σ^2 hoppar vi över. baserar på

$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

Prediktion (σ^2 känd) Vad kommer spänningen vara nästa dag? Vi har $X_1, \dots, X_n$ sp på $N(\mu, \sigma^2)$ & undrar vad vi kan säga om ny obs. Y

$$Y - \overline{X} \sim N\left(0, \sigma^2 + \frac{\sigma^2}{n}\right) \left(t_y \quad Var(Y - \overline{X}) = Var(Y) + Var(\overline{X}) = \sigma^2 + \frac{\sigma^2}{n}\right.$$

Alltså

$$P\left(-z_{\alpha/2} \le \frac{Y - \overline{X}}{\sigma\sqrt{1 + 1/n}} \le z_{\alpha/2}\right) = 1 - \alpha$$

Vi får ett prediktionsintervall

$$Y = \overline{X} \pm z_{\alpha/2}\, \sigma \sqrt{1 + \tfrac{1}{n}} \qquad (1-\alpha)$$

I vårt ex med $\sigma^2 = 2^2$ $n = 2$

$$Y = 2\cancel{28}.9 \pm 1.96 \cdot 2\sqrt{\tfrac{3}{2}} = 228.9 \pm 4.2 \quad (95\%)$$

Reflektioner:

- Normalfördelningsantagandet alltid lite fel, ffa i svansen
- Ju större n desto mer normal $\overline{X}$ & därmed klara mer extrema konfidensgrader
- Om beroende mellan obs: kan $\overline{X}$ fortfarande bli normal men konfidensgraden kan bli helt fel ty $Var(\overline{X})$ blir helt annan.
- Det finns test (goodness-of-fit) och normalfördelningsplot för att kolla normalfördelning

Konfidensintervall för p — Undersökning m. 10 000 deltagare säger 8.4 % stödjer A.

Ge konfidensintervall för hela populationen

Låt p vara andelen väljare som stödjer A

Låt $X_1, \ldots, X_n$ vara indikatorerna om resp. väljare stödjer A. $\sum_k X_k \sim Bin(n,p)$ Enl CGS är

$$\bar{X} \simeq N\left(p, \frac{p(1-p)}{n}\right) \quad \text{dvs} \quad \frac{\bar{X}-p}{\sqrt{\frac{p(1-p)}{n}}} \simeq N(0,1)$$

Så om $z_{\alpha/2} = \Phi^{-1}(1-\frac{\alpha}{2})$ gäller

$$\mathbb{P}\left(-z_{\alpha/2} \leq \frac{\bar{X}-p}{\sqrt{\frac{p(1-p)}{n}}} \leq z_{\alpha/2}\right) \simeq 1-\alpha$$

Man kan lösa ut p ur olikheten och bilda konfidensintervall för p. Lite bökigt. Man kan utnyttja att $\frac{\bar{X}}{p}$ är approx 1 med stor sannolikhet och få att

$$\frac{\bar{X}-p}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}} \simeq N(0,1)$$

Detta är De-Moivre-Laplace gränsvärdessats

Här är lätt att lösa ut p

$$p = \bar{X} \pm z_{\alpha/2}\sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \qquad (1-\alpha)$$

I vårt exempel $\bar{X} = 0.084$ För 95% dvs: $z_{0.025} = 1.96$

$$\Rightarrow p = 0.084 \pm 1.96\sqrt{\frac{0.084 \cdot 0.916}{10\,000}} = 0,084 \pm 0.0054$$

dvs $8.4\% \pm 0.54\%$

Reflektioner

- Inga problem i svansarna
- Mycket bra skattningar av p bland
  - dem som svarar
  - och talar sanning

Svarsbortfall, lögner & felvar är stora problem i opinionsundersökningar.

Maximum Likelyhood (ML): Punktskatta θ genom att maximera sannolikheten att få de data man fick.

Mer precist: $X_1, \dots, X_n$ sp på sv m. frekv.fkn. $p(x) = p_\theta(x)$ beror av θ $X_i = x_i$ finn θ som

$$\max \prod_{k=1}^{n} p_\theta(x_k) =: L(\theta; x_1, \dots, x_n)$$

Detta max Skrivet $\hat{\theta}$ kallas ML-skattning av θ

Ex $X \sim Bin(n, \theta)$. Finn ML-s av θ

$$L(\theta; x) = p_\theta(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

Om $x = 0$ är L avtagande så max fås då $\theta = 0$

Om $x = n$ — " — växande — — $\theta = 1$

Annars Lös

$$L'(\theta; x) \propto x\theta^{x-1}(1-\theta)^{n-x} - (n-x)\theta^x(1-\theta)^{n-1-x} = 0$$

$\Rightarrow x(1-\theta) - (n-x)\theta = 0 \Rightarrow$ Lösn $\theta = \frac{x}{n}$

$\Rightarrow \hat{\theta} = \frac{x}{n}$ $\square$

OBS Som underlättar $L(\theta; X)$ max precis då $l(\theta; X) =: \ln L(\theta; X)$ maximeras.

Ex: Låt $X \sim Geo(\theta)$ ML-s θ. $L(\theta; X) = \theta(1-\theta)^{x-1}$ så

$$l(\theta; x) = \ln(\theta) + (x-1)\ln(1-\theta)$$

$$\Rightarrow l'(\theta; x) = \frac{1}{\theta} - \frac{x-1}{1-\theta} = 0 \Rightarrow \theta = \frac{1}{x}$$

ML-s av θ är alltså $\hat{\theta} = \frac{1}{x}$

Ex $X_1, \dots, X_n$ sp på $Poi(\theta)$

$$L(\theta) = \prod_k e^{-\theta}\theta^{x_k}/x_k! \quad \text{Alltså}$$

$$\ell(\theta) = -n\theta + \left(\sum_k x_k\right)\ln(\theta) - \sum_k \ln(x_k!)$$

$$\Rightarrow \ell'(\theta) = -n + \frac{\sum_k x_k}{\theta} = 0$$

$$\Rightarrow \hat{\theta} = \frac{1}{n}\sum_k x_k = \bar{x}$$

$\square$

För kont. förd. maximerar istället tätheten av sp som fkn. av θ

Ex $X_1, \dots, X_n$ sp på $N(\mu, \sigma^2)$ ML-S μ σ^2

$$L(\mu, \sigma) = \prod_k \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x_k-\mu)^2}{\sigma^2}}$$

$$\Rightarrow \ell(\mu, \sigma) = -\frac{n}{2}\ln(2\pi) - n\ln\sigma - \frac{1}{2}\frac{\sum_k (x_k-\mu)^2}{\sigma^2}$$

Alltså $\frac{\partial \ell}{\partial \mu} = \frac{\sum_k (x_k - \mu)}{\sigma^2}$ som är noll då $\mu = \bar{x}$

Vidare är

$$\frac{\partial \ell(\hat{\mu}, \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\sum_k (x_k - \hat{\mu})^2}{\sigma^3}$$

Detta är noll då $\sigma^2 = \frac{1}{n}\sum_k (x_k - \hat{\mu})^2$ dvs

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{k=1}^n (x_k - \bar{x})^2$$

$\square$

Hypotesprövning / tester:

Slantsingel

Vill testa nollhypotesen

$H_0 : p = \frac{1}{2}$

mot alternativhypotesen

$H_a : p \neq \frac{1}{2}$

Fixera litet tal $\alpha > 0$. Beteckna data x & en mängd B_0: $\mathbb{P}_{H_0}(x \notin B_0) = \alpha$
och $x \notin B_0$ talar till förmån för H_a över H_0.
Om det sedan visar sig att $x \in B_0$ förkastar vi H_0 till förmån av H_a på signifikansnivån α.. Om $x \in B_0$ accepterar vi H_0

I ex n slantsinglingar. X antal klave
Om X avviker mycket från $\frac{n}{2}$ talar detta till förmån för H_A över H_0. Enl. CGS gäller under H_0

$$\frac{X - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \approx N(0,1) \quad \text{med } z_{\alpha/2} = \Phi^{-1}(1 - \frac{\alpha}{2}) \text{ gäller alltså}$$

$$\mathbb{P}_{H_0}\left(\frac{X - \frac{n}{2}}{\sqrt{n/4}} \notin \pm z_{\alpha/2}\right) \approx \alpha$$

Dvs $\mathbb{P}_{H_0}\left(X \notin \frac{n}{2} \pm z_{\alpha/2}\sqrt{\frac{n}{4}}\right) \approx \alpha$

Vi förkastar H_0 till förmån för H_A om X utanför intervallet

Ex med $n = 100$ & $\alpha = 0{,}05 \Rightarrow z_{0{,}025}\sqrt{\frac{100}{4}} \approx 10$

Och vi förkastar om X avviker med minst 10 från 50

Med $n = 10\,000$ & $\alpha = 0{,}01$ $\quad z_{0{,}005}\sqrt{\frac{n}{4}} \approx 129$

vi förkastar om X avviker med minst 129 från 5000

Ex $X_1, \ldots, X_n$ sp på $N(\mu, \sigma^2)$

Testa $H_0 : \mu = \mu_0$ mot $H_A : \mu \neq \mu_0$

Om H_0 sann bör $\overline{X}$ ligga nära μ_0 $\Rightarrow$ stora avvikelser talar för H_A

Under H_0: $$\frac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

Alltså:

$$\mathbb{P}_{H_0}\left(\overline{X} \in \mu_0 \pm F^{-1}_{t_{n-1}}\left(1 - \frac{\alpha}{2}\right) \frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

och vi förkastar H_0 till förmån av H_A på sign. α om $\overline{X} \notin \mu_0 \pm F^{-1}_{t_{n-1}}\left(1 - \frac{\alpha}{2}\right)\frac{S}{\sqrt{n}}$

Konkret fall: Burkar innehålla 400g

Välj 10 på måfå Burkars vikt $N(\mu, \sigma^2)$-förd.

Vill testa $H_0 : \mu = 400$ mot $H_A \neq 400$

5% sign.

$\Rightarrow$ $\overline{X} \in 400 \pm F^{-1}_{t_9}(0{,}975)\frac{S}{\sqrt{10}}$

Slår upp att $F^{-1}_{t_9}(0{,}975) \approx 2.26$ ger

$|\overline{X} - 400| \geq 0{,}71\, S$ kanske mer intressant

av $H_0 : \mu \geq 400$ mot $H_A : \mu < 400$

Generös mot H_0: testa $H_0 : \mu = \mu_0$ mot $H_A : \mu < \mu_0$

$\mathbb{P}_{H_0}\left(\overline{X} \leq \mu_0 - F^{-1}_{t_{n-1}}(1-\alpha)\frac{S}{\sqrt{n}}\right) = \alpha$

Förkasta H_0 på sign. α om $\overline{X} < \mu_0 - F^{-1}_{t_{n-1}}(1-\alpha)\frac{S}{\sqrt{n}}$

För burkarna $H_0 : \mu = 400$ mot $H_A : \mu < 400$ på sign. $\alpha = 5\%$ a

$\overline{X} < 400 - F^{-1}_{t_9}(0{,}95)\frac{S}{\sqrt{10}} \approx 400 - 0{,}58S$

Korrespondens mellan test & konf. intervall

Antag θ endim. Skapa konf. intervall

$$T_1 \le \theta \le T_2 \text{ med grad } 1-\alpha$$

$\Rightarrow \quad \mathbb{P}(T_1 \le \theta \le T_2) = 1-\alpha$

Vill testa $H_0: \theta = \theta_0$ mot alt. hypotes H_A. Förkasta H_0 om $\theta \notin [T_1, T_2]$. Detta ger att test på sig. nivå α ty $\mathbb{P}_{H_0}(\theta_0 \notin [T_1, T_2]) = \alpha$

Konfidensintervallet väljs: $\theta_0 \notin [T_1, T_2]$ talar till förmån för H_A framför H_0

Ex $X_1, \dots, X_n$ sp på $N(\mu, \sigma^2)$. Skapa symmetriskt konfidensintervall för μ:

$$\mu \in \overline{X} \pm F^{-1}_{t_{n-1}}\left(1-\frac{\alpha}{2}\right)\frac{s}{\sqrt{n}} \text{ med konfgrad } 1-\alpha$$

Enl. ovan förkasta $H_0: \mu = \mu_0$ till förmån av $H_A: \mu \neq \mu_0$ om

$$\mu_0 \notin \overline{X} \pm F^{-1}_{t_{n-1}}\left(1-\frac{\alpha}{2}\right)\frac{s}{\sqrt{n}}$$

$\square$

Att jämföra två stickprov

Testa om medicin sänker blodsocker. Dubbelblind studie

$X_1, \dots, X_n$ sp tagit medicin

$Y_1, \dots, Y_m$ sp tagit placebo

Antag $X_k \sim N(\mu_x, \sigma_x^2)$

$Y_k \sim N(\mu_y, \sigma_y^2)$

Vill testa $H_0: \mu_x = \mu_y$ mot $H_A: \mu_x \neq \mu_y$ (eller $H_A: \mu_x < \mu_y$) / göra konfidensintervall för $\mu_x - \mu_y$. Basera på $\overline{X} - \overline{Y}$

Fallet σ_x^2 & σ_y^2 kända. OBS

$$\overline{X} - \overline{Y} \sim N\left(\mu_x - \mu_y, \frac{1}{n}\sigma_x^2 + \frac{1}{m}\sigma_y^2\right)$$

Ger konf. intervall

$$\mu_x - \mu_y \in \overline{X} - \overline{Y} \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{1}{n}\sigma_x^2 + \frac{1}{m}\sigma_y^2}$$

Fallet σ_x^2 & σ_y^2 okända. Antag σ_x^2, σ_y^2 lika

Skatta σ^2 med den poolade stickprovsvariansen.

$$s_p^2 = \frac{\sum_{k=1}^{n}(X_k-\bar{X})^2 + \sum_{k=1}^{m}(Y_k-\bar{X})^2}{n+m-2} = \frac{(n-1)s_x^2+(m-1)s_y^2}{n+m-2}$$

Det visar sig

$$\frac{\bar{X}-\bar{Y}-(\mu_x-\mu_y)}{s_p\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim t_{n+m-2}$$

Ger symmetriskt konf. intervall

$\mu_x-\mu_y \in \bar{X}-\bar{Y} \pm F^{-1}_{t_{n+m-2}}(1-\frac{\alpha}{2})\,s_p\sqrt{\frac{1}{n}+\frac{1}{m}}$ med

konfgrad $1-\alpha$. Nedåt begr. konf. inter

$$\mu_x-\mu_y \geq \bar{X}-\bar{Y}-F^{-1}_{t_{n+m-2}}(1-\alpha)\,s_p\sqrt{\frac{1}{n}+\frac{1}{m}}$$

__Ex__ Böjhållfasthet vid olika temp.

$700°C$: 147 140 121 138 120 137

$800°C$: 193 227 201 212 207

Låt μ_x & μ_y vara förväntad hållf. vid 700 resp 800 grader

99% symmetriskt konf. intervall för $\mu_x-\mu_y$

Antar normalfördelad med samma varians

$\bar{X} = 132.8$, $s_x^2 = 10.83^2$, $\bar{Y} = 208.0$ $s_y^2 = 12.77^2$

$$s_p^2 = \frac{5s_x^2+4s_y^2}{9} = 11.73^2$$

Konf.intervallet blir

$$\mu_y-\mu_x \in \bar{Y}-\bar{X} \pm F^{-1}_{t_9}(0.995)\,s_p\sqrt{\tfrac{1}{6}+\tfrac{1}{5}} = 75.2 \pm 3.25\cdot 11.73\sqrt{\tfrac{1}{6}+\tfrac{1}{5}}$$

$= 75.2 \pm 23.1$

Betyder att $H_0: \mu_x = \mu_y$ kan förkastas på 1% sig. nivå

Tjuvkikande I alla datamängder finns slumpmässigt uppkomna mönster. Kan ge hypoteser.

Bestäm alltid tester innan tittar på data.

Multipeltestning

Antag testar 100 nollhypoteser på 5% signifikansnivå.

I genomsnitt 5 gånger felaktigt förkasta nollhypotesen.

Kan bli många fel i ett statistikerliv.

Testa inte om du inte har goda skäl att tro att nollhypot. är falsk

---

Transformering av sv.

har sv X med förd F_1. Finn fkn g : $g(x)$ har förd F_2

Prop: om F_1, F_2 är inverbara & $X \sim F_1$ gäller

$$Y = F_2^{-1}(F_1(X)) \sim F_2$$

Bevis:

$$\mathbb{P}(Y \leq y) = \mathbb{P}(F_1(X) \leq F_2(y)) = \mathbb{P}(X \leq F_1^{-1}(F_2(y))) = F_2(y)$$

Specialfall: $U \sim \text{likf}(0,1)$ gäller att $F^{-1}(U) \sim F$ $\square$

Ex Använd $U \sim \text{likf}(0,1)$ till att simulera $\exp(\lambda)$-förd

$F(x) = 1 - e^{-\lambda x}$ Alltså

$$F^{-1}(u) = -\frac{\ln(1-u)}{\lambda}$$

Alltså är $-\frac{\ln(1-U)}{\lambda} \sim \exp(\lambda)$ $\square$

Minns $X_1, \ldots, X_n$ ober $\exp(\lambda)$-förd. medför $\sum_{k=1}^{n} X_k \sim \Gamma(n, \lambda)$ vi hade

$$F(x) = 1 - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}$$

Derivering ger

$$f(x) = \frac{1}{(n-1)!} \lambda^n x^{n-1} e^{-\lambda x}$$

Mer allmänt: $\alpha, \lambda > 0$ att $X \sim \Gamma(\alpha, \lambda)$ om

$$f(x) = \frac{1}{c} \lambda^{\alpha} x^{\alpha - 1} e^{-\lambda x}$$

<u>χ^2-fördelning.</u> Om $Z_1, \ldots, Z_n$ är ober $\sim N(0,1)$-förd säger man att

$$X = \sum_{k=1}^{n} Z_k^2$$

är χ^2-fördelad med n frihetsgrader

vi får $\mathbb{E}[X] = n$, $Var(X) = 2n$

Intressant faktum:

$$\mathbb{P}(Z_1^2 + Z_2^2 \leq x) = \mathbb{P}((Z_1, Z_2) \in B(0, \sqrt{x})) =$$

$$= \frac{1}{2\pi} \iint_{B(0,\sqrt{x})} e^{-\frac{1}{2}(t_1^2 + t_2^2)} dt_1 dt_2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\sqrt{x}} r e^{-\frac{1}{2} r^2} dr d\theta =$$

$$= \left[-e^{-\frac{1}{2} r^2} \right]_0^{\sqrt{x}} = 1 - e^{-\frac{1}{2} x}.$$ Alltså $Z_1^2 + Z_2^2 \sim \exp(1/2)$

dvs $\chi_2^2 = \exp(1/2)$. Det följer att om n jämt gäller

$$\sum_{k=1}^{n} Z_k^2 \sim \Gamma(n/2, 1/2)$$

Detta gäller även då n är udda

dvs $\chi_n^2 = \Gamma(n/2, 1/2)$

Låt $\mathbb{U} = (U_1, \ldots, U_n)$ ober ≐ $N(0,1)$

Låt M $n \times n$ ortogonal matris.

Låt $\mathbb{X} = M\mathbb{U}$. Då är $X_1, \ldots, X_n$ ober ≐ $N(0,1)$.

Låt

$$M = [c_1 \ldots c_n] \quad \text{där} \quad c_1 = [1/\sqrt{n} \ \ldots \ 1/\sqrt{n}]^T = (1/\sqrt{n})\mathbb{1}$$

på är

$$\mathbb{X} = \sum_{k=1}^{n} U_k c_k$$

Vi får

$$\overline{X} = \frac{1}{n}\mathbb{1}^T\mathbb{X} = \frac{1}{\sqrt{n}} U_1 \quad \text{ty } c_2, \ldots, c_n \text{ ortogonala mot } c_1$$

Alltså

$$\sum_{k=1}^{n}(X_k - \overline{X})^2 = (\mathbb{X} - \overline{X}\mathbb{1})^T(\mathbb{X} - \overline{X}\mathbb{1}) =$$

$$= \left(\mathbb{X} - \frac{1}{\sqrt{n}}U_1\mathbb{1}\right)^T\left(\mathbb{X} - \frac{1}{\sqrt{n}}U_1\mathbb{1}\right) =$$

$$= \left(\sum_{k=2}^{n} U_k c_k\right)^T\left(\sum_{k=2}^{n} U_k c_k\right) = \sum_{k=2}^{n} U_k^2$$

Det följer att

$$\sum_{k=1}^{n}(X_k - \overline{X})^2 \sim \chi^2_{n-1} \text{ ≐ är oberoende av } \overline{X}$$

dvs $(n-1)s^2 \sim \chi^2_{n-1}$ ≐ ober av $\overline{X}$

Med X_k i ∼ generellt $N(\mu, \sigma^2)$, standardisera ≐ få

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

Konfidensintervall för σ^2.

Låt $X_1, \ldots, X_n$ vara sp på $N(\mu, \sigma^2)$

Det gäller att

$$F^{-1}_{\chi^2_{n-1}}(\alpha/2) \leq \frac{(n-1)s^2}{\sigma^2} \leq F^{-1}_{\chi^2_{n-1}}(1-\alpha/2) \text{ med sannol. } 1-\alpha$$

Vi får

$$\frac{(n-1)s^2}{F^{-1}_{\chi^2_{n-1}}(1-\alpha/2)} \leq \sigma^2 \leq \frac{(n-1)s^2}{F^{-1}_{\chi^2_{n-1}}(\alpha/2)}$$

med konfidensgrad $1-\alpha$

Ett uppåt begränsat konf. intervall ges av

$$\sigma^2 \leq \frac{(n-1)s^2}{F^{-1}_{\chi^2_{n-1}}(\alpha)}$$

Felintensitet Låt T positiv s.v. Se som en livstid

Skriv $G(t) = \mathbb{P}(T > t)$. Givet överlevnad till tid t.

Vad är risken för haveri alldeles strax?

$$\mathbb{P}(T \in (t, t+\Delta t) \mid T > t) = \frac{\mathbb{P}(T \in (t, t+\Delta t))}{\mathbb{P}(T > t)} \approx \frac{f(t)\,\Delta t}{G(t)}$$

Def: Felintensiteten för T ges av

$$r(t) = \frac{f(t)}{G(t)}$$

Ofta lättare att specifiera $r(t)$ än $f(t), G(t)$

Obs att

$$r(t) = -\frac{d}{dt} \ln(G(t))$$

Alltså

$$\ln(G(t)) = -\int_0^t r(s)\,ds$$

dvs $G(t) = e^{-\int_0^t r(s)\,ds}$

Ex: • $T \sim \exp(\lambda) \Leftrightarrow r(t) = \lambda$

• $T \sim \text{likf}(0,1) \Rightarrow ?$

$$r(t) = \frac{f(t)}{G(t)} = \frac{1}{G(t)} = \frac{1}{1-t}, \quad t \in (0,1)$$

kraftigt åldrande

Ex: Låt
$$G(t) = \frac{1}{(1+t)^2} \quad t \geq 0$$

En s.k. tungsvansad fördelning

$$r(t) = \frac{f(t)}{G(t)} = \frac{2/(1+t)^3}{1/(1+t)^2} = \frac{2}{1+t}$$ Föryngring

Ex $r(t) = t, \quad t \geq 0$ Då blir

$$G(t) = e^{-\int_0^t r(s)dt} = e^{-\frac{1}{2}t^2}$$

(Ganska likt normalfördel. betingat på att vara > 0)

Summa av oberoende sv Om X, Y oberoende

Vad är förd. för $Z = X+Y$, Antag (X,Y) kont

$$F_Z(z) = \mathbb{P}(Z \leq z) = \mathbb{P}(X+Y \leq z) =$$
$$= \int_{-\infty}^{\infty} \mathbb{P}(x+Y \leq z \mid X=x) f_X(x)dx =$$
$$= \int_{-\infty}^{\infty} F_Y(z-x) f_X(x)dx$$

Under enkla villkor kan man derivera under integral.

$$f_Z(z) = \frac{d}{dz}\int_{-\infty}^{\infty} F_Y(z-x) f_X(x)dx = \int_{-\infty}^{\infty} f_Y(z-x) f_X(x)dx$$ om

$X, Y > 0 \Rightarrow$

$$f_Z(z) = \int_0^z f_Y(z-x) f(x)dx$$

om (X,Y) diskret

$$p_Z(z) = \sum_{x \in V_x} p_Y(z-x) p_X(x)$$

Redan utnyttjat detta genom visa

$$X_1 + X_2 \sim Poi(\lambda_1 + \lambda_2)$$

Ex Ev! CGS måste det gälla om

$X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$ & X_1, X_2 ober

medför att

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Kolla fallet $\mu_1 = \mu_2 = 0$ & $\sigma_1 = \sigma_2 = 1$

$$f_{X_1+X_2}(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(t-\frac{1}{2}z)^2 - \frac{1}{4}z^2} dt =$$

$$= \frac{1}{\sqrt{4\pi}} e^{-\frac{1}{4}z^2} \int \frac{1}{\sqrt{\pi}} e^{-(t-\frac{1}{2}z)^2} dt =$$

$$= \frac{1}{\sqrt{4\pi}} e^{-\frac{1}{4}z^2}$$

Under integralen står täthet för $N(z/2, 1/2)$ $\square$

Ex X, Y ober. & ~~exp~~ $\exp(\lambda)$-fördelade

$$f_{X+Y}(z) = \int_0^z \lambda e^{-\lambda(z-t)} \lambda e^{-\lambda t} dt = \lambda^2 e^{-\lambda z} \int_0^z dt = \lambda^2 z e^{-\lambda z}$$

Mao: $X + Y \sim \Gamma(2, \lambda)$

Momentgenererande funktion

För en sv X ges M_X av

$$M_X(t) = \mathbb{E}[e^{tX}], \quad t \in \mathbb{R}$$

Om X kont. & positiv med täthet f får man

$$M_X(t) = \int_0^\infty e^{tx} f(x)\,dx = \mathcal{L}f(-t)$$

Därför kan vi anta att om man vet alla M_X för alla t i en omgivning av noll så vet man även f.

Detta samma & gäller $\forall$ sv X.

Prop: Om X, Y ober gäller

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$

Bevis

$$M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX}]\,\mathbb{E}[e^{tY}]$$

Om vi får derivera innanför vv. (och ofta får man) får vi $M_X'(t) = \frac{d}{dt}\mathbb{E}[e^{tX}] = \mathbb{E}[Xe^{tX}]$

Detta ger $M_X'(0) = \mathbb{E}[X]$

Också

$$M_X''(t) = \frac{d^2}{dt^2}\mathbb{E}[e^{tX}] = \mathbb{E}[X^2e^{tX}]$$

Alltså $M_X''(0) = \mathbb{E}[X^2]$

Generellt

$$M_X^{(k)}(0) = \mathbb{E}[X^k], \quad k = 0, 1, \dots$$

<u>Ex</u> $X \sim Poi(\lambda)$

$$M_X(t) = \mathbb{E}[e^{tX}] = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} =$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = e^{\lambda(e^t - 1)}$$

om X_1, X_2 ober får vi

$$M_{X_1+X_2} = M_{X_1} M_{X_2} = e^{(\lambda_1+\lambda_2)(e^t-1)}$$

Viser att $X_1 + X_2 \sim Poi(\lambda_1 + \lambda_2)$

<u>Ex:</u> $X \sim N(\mu, \sigma^2)$

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

kvadratkomplettera exponent och få h.l.

$$e^{\mu t + \sigma^2 t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}} dx$$

funktionen i integral är täthet för $N(\mu+\sigma^2 t, \sigma^2)$

Så $M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$

Om $X_1 + X_2$ ober

$$M_{X_1+X_2} = e^{(\mu_1+\mu_2)t + (\sigma_1^2+\sigma_2^2)t^2/2}$$

dvs $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

En generalisering av inverterbarheten

Antag X & $X_1, X_2, \ldots$ är kont

Om $M_{X_n}(t) \to M_X(t)$ $\forall t$ så gäller

$$\mathbb{P}(X_n \leq x) \to \mathbb{P}(X \leq x) \quad \forall x$$

Beviskiss av CGS:

CGS: om $X_1, X_2, \ldots$ iid & fördelade som en SvX med $\mathbb{E}[X] = \mu$ $Var(X) = \sigma^2 < \infty$ gäller

$$\mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) \to \phi(x)$$

Räcker att visa för $\mu = 0$ & $\sigma^2 = 1$

Taylorutv. M_X kring 0 till ord 2:

$$M_X(s) \approx M_X(0) + sM_X'(0) + \frac{1}{2}s^2 M_X''(0) =$$

$$= 1 + s\mu + \frac{1}{2}s^2(\mu^2 + \sigma^2) = 1 + \frac{1}{2}s^2$$

Vill visa att

$$M_{S/\sqrt{n}}(t) \to M_{N(0,1)}(t) = e^{t^2/2}$$

Men $M_{S_n/\sqrt{n}}(t) = \mathbb{E}[e^{tS_n/\sqrt{n}}] = M_{S_n}\left(\frac{t}{\sqrt{n}}\right) = M_X\left(\frac{t}{\sqrt{n}}\right)^n$

$$\approx \left(1 + \frac{t^2}{2n}\right)^n \to e^{\frac{t^2}{2}} \qquad \square$$

p-värde Test av H_0 mot H_A

Testets p-värde, p, är det minsta tal α sådant att H_0 kan förkastas för H_A på sig.nivå α

Mer formellt: Låt x beteckna data. För varje α finn lämplig A_α : $\mathbb{P}_{H_0}(x \in A_\alpha) = \alpha$ Testet förkastar H_0 på signivå α om $x \notin A_\alpha$. Antag A_α:ina valda :

$\alpha_1 \leq \alpha_2 \Rightarrow A_{\alpha_1} \supseteq A_{\alpha_2}$

Då är

$$p = \min\{\alpha : x \notin A_\alpha\}.$$

Om θ endim parameter, gör konf.intervall.

$$T_1(x,\alpha) \leq \theta \leq T_2(x,\alpha) \qquad (1-\alpha)$$

Man förkastar $H_0: \theta = \theta_0$ på sign. α om $\theta_0 \notin [T_1(x,\alpha), T_2(x,\alpha)]$. Då är

$$p = \min\{\alpha : \theta_0 \notin [T_1(x,\alpha), T_2(x,\alpha)]\}$$

Typiskt är intervallgränserna kontinuerliga i α och då blir p lika med det α som ger $\theta_0 = T_1(x, \alpha)$ eller $\theta_0 = T_2(x, \alpha)$

<u>Ex</u> $X_1, \ldots, X_n$ sp på $N(\mu, \sigma^2)$. Testa $\mu = 0$ mot $\mu \neq 0$ antag $n = 10$, $\bar{x} = 1{,}7$, $s^2 = 1{,}5$. Vad är testets p-värde? konf. intervall

$$\mu \in \bar{x} \pm F^{-1}_{t_9}(1 - \alpha/2)\frac{s}{\sqrt{n}}$$

Punkten 0 hamnar på gränsen om $\bar{x} = F^{-1}_{t_9}(1 - \frac{\alpha}{2})\frac{s}{\sqrt{n}}$ dvs

$F^{-1}_{t_9}(1 - \frac{\alpha}{2}) = \frac{\bar{x}\sqrt{n}}{s} = \frac{1.7\sqrt{10}}{\sqrt{1.5}} = 4.39$

Alltså $1 - \frac{\alpha}{2} = F_{t_9}(4.39) = 1 - 0{,}0009$ så $\alpha = 0{,}0018$

Vi har alltså $p \approx 0.0018$

Linjär regression

- Ström funktion av spänning
- Sällan perfekta samband

Modell $Y_k = a + bx_k + \varepsilon_k$, $k = 1, \ldots, n$

där $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ är ober. $\in N(0, \sigma^2)$, a, b, σ okända

Talen x_k fixa tal

ML-skattning av a & b:

$Y_k \sim N(a + bx_k, \sigma^2)$

$$f_{Y_k}(y_k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_k - a - bx_k)^2}$$

Därför blir

$$L(a, b, \sigma; y_1, \ldots, y_n) = f_{Y_1, \ldots, Y_n}(y_1, \ldots, y_n) =$$
$$= \frac{1}{\sigma^n (2\pi)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(y_k - a - bx_k)^2}$$

Att maximera map. a & b innebär att minimera

$$\sum_{k=1}^{n}(Y_k - a - bx_k)^2 =: \mathcal{L}(a, b)$$

Alltså minsta kvadratmetoden

Ska lösa
$$\frac{\partial \mathcal{L}}{\partial a} = -2\sum_{k=1}^{n}(y_k - a - bx_k) = 0$$
$$\frac{\partial \mathcal{L}}{\partial b} = -2\sum_{k=1}^{n}x_k(y_k - a - bx_k) = 0$$

Detta ger $\hat{b} = \dfrac{\sum_{k=1}^{n}(y_k - \bar{y})(x_k - \bar{x})}{\sum_{k=1}^{n}(x_k - \bar{x})^2}$, $\hat{a} = \bar{y} - \hat{b}\bar{x}$

$S_{xx} = \sum_{k=1}^{n}(x_k - \bar{x})^2$ $\quad S_{xy} = \sum_{k=1}^{n}(y_k - \bar{y})(x_k - \bar{x})$

$S_{yy} = \sum_{k=1}^{n}(y_k - \bar{y})^2$ $\quad$ då blir $\hat{b} = \dfrac{S_{xy}}{S_{xx}}$

kan beräkna $\hat{b} \sim N\left(b, \dfrac{\sigma^2}{S_{xx}}\right)$

$$\hat{a} \sim N\left(a, \frac{\sigma^2 \sum_{k=1}^{n} x_k^2}{n S_{xx}}\right)$$

Om σ^2 känd, gör konf.intervall/test för a resp b med dessa

T.ex. $\frac{\hat{b}-b}{\sigma/\sqrt{S_{xx}}} \sim N(0,1)$

$$\hat{b} \in \hat{b} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{S_{xx}}} \quad (1-\alpha)$$

Om σ^2 okänd, ersätt med s^2

$$\frac{\hat{b}-b}{s/\sqrt{S_{xx}}} \sim t_{n-2}$$

och $b \in \hat{b} \pm F^{-1}_{t_{n-2}}(1-\frac{\alpha}{2}) \frac{s}{\sqrt{S_{xx}}} \quad (1-\alpha)$

Analogt för a

Vad är s^2?

$$s^2 = \frac{1}{n-2} \sum_{k=1}^{n} (y_k - \hat{a} - \hat{b}x_k)^2$$

Den är vvr för σ^2

Formel: $s^2 = \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right)$

Mer formler:

$$S_{xy} = \sum x_k y_k - \frac{1}{n}\left(\sum x_k\right)\left(\sum y_k\right)$$
$$S_{xx} = \sum x_k^2 - \frac{1}{n}\left(\sum x_k\right)^2$$

analogt y.

Ex Data ger

$S_{xy} = 91.323$ $\qquad \hat{b} = \frac{S_{xy}}{S_{xx}} = 0.548$

$S_{xx} = 166.628$ $\qquad \hat{a} = \bar{y} - \hat{b}\bar{x} = 75.74$

$S_{yy} =$

$$\Rightarrow y = 75.74 + 0.548x$$

också $s^2 = \frac{1}{4}\sum_1^6 (y_k - \hat{a} - \hat{b}x_k)^2 = 14.74$

Symmetriskt 95% konfintervall för b

$$b \in \hat{b} \pm F^{-1}_{t_4}(0.975)\frac{s}{\sqrt{S_{xx}}} = 0.548 \pm 2.776 \frac{\sqrt{14.74}}{\sqrt{166.6}} = 0.55 \pm 0.81$$

Om vi gör test av $H_0: b = 0$ mot $H_1: b \neq 0$ kan vi inte förkasta H_0 på 5% signnivå

Ex $n = 33$ $\sum xy = 41365$ $\sum x_k = 1104$

$\sum x^2 = 41086$ $\sum y_k = 1124$

$\sum y^2 = 43117$

Skatta $y = a + bx$ & ge 95% konfintervall för b

$S_{xy} = \sum xy - \frac{1}{n}\sum x \sum y = 3572$

$S_{xx} = 4152$

$\hat{b} = \frac{S_{xy}}{S_{xx}} = 0.86$ $\hat{a} = \bar{y} - \hat{b}\bar{x} = 5.29$

$y = 5.29 + 0.86x$

konfintervall $F^{-1}_{t_{31}}(0.975) = 2.04$ behöver s

$S_{yy} = 4833$ $s^2 = \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right) = 56.74$

$b \in 0.86 \pm 0.24$ (95%) $\square$

Prediktion i linjär regression

$Y = a + bx + \varepsilon$, Rimlig punktgissning $\hat{a} + \hat{b}x$

$D = Y - (\hat{a} + \hat{b}x)$ $\mathbb{E}(D) = 0$

$Var(D) = ?$ $\hat{a} = \bar{Y} - \hat{b}\bar{x}$

$D = Y - \bar{Y} - \hat{b}(x - \bar{x})$ obs Y & $\bar{Y}$ är oberoende

också Y & $\hat{b}$

Det gäller att $Cov(\hat{b}, \bar{Y}) = 0$ ty

$Cov(\bar{Y}, S_{xy}) = Cov(\bar{Y}, \sum xy - \frac{1}{n}\sum x_k \sum y_k) = \sum x_k Cov(\bar{Y}, Y_k) - \bar{x}\sum Cov(\bar{Y}, Y_k)$

$= \frac{1}{n}\sigma^2 \sum x_k - \bar{x}\sigma^2 = 0$

Vi får

$$\mathrm{Var}(D) = \sigma^2 + \frac{1}{n}\sigma^2 (x-\bar{x})^2 \mathrm{Var}(\hat{b}) = \left(1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{S_{xx}}\right)\sigma^2$$

Alltså

$$\frac{D}{\sigma\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{S_{xx}}}} \sim N(0,1)$$

kan anse att

$$\frac{D}{s\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{S_{xx}}}} \sim t_{n-2}$$

Ger prediktionsintervall

$$Y \in \hat{a}+\hat{b}x \pm F^{-1}_{t_{n-2}}(1-\alpha/2)\, s\sqrt{1+\frac{1}{n}+\frac{(x-\bar{x})^2}{S_{xx}}}$$ Med prediktionsgrad $1-\alpha$

Ex F16| Lin reg på (1,1) (2,2) (3,5)

Styrka. Styrkan är sannolikheten att förkasta nollhypotesen som funktion av vad sanningen verkligen är. konkret:

Testa $H_0: \theta=\theta_0$ mot $H_A: \theta\neq\theta_0$

$g(\theta_1) = \mathbb{P}_{\theta_1}(H_0 \text{ förkastas})$. Det gäller att $g(\theta_0)=\alpha$

Ex slantsingling. Test av $H_0: p=1/2$ mot $H_A: p\neq 1/2$. $n=100$

Låt X vara antal klave. Vi kom fram till att förkasta på 5% sign. nivå om $|X-50|\geq 10$. Vad är $g(0{,}7)$? om $p=0{,}7$ gäller

$$\frac{X-np}{\sqrt{np(1-p)}} = \frac{X-70}{\sqrt{21}} \approx N(0,1)$$

Alltså $\mathbb{P}_{0,7}(|X-50|\geq 10) = \mathbb{P}_{0,7}(X\geq 59{,}5) + \mathbb{P}_{0,7}(X\leq 40{,}5)$

Andra termen försumbar och

$$\mathbb{P}_{0,7}(X\geq 59{,}5) = \mathbb{P}_{0,7}\left(\frac{X-70}{\sqrt{21}} \geq \frac{59.5-70}{\sqrt{21}}\right) = 1-\Phi(-2.291) \approx 0.989$$

Alltså $g(0{,}7) \approx 0{,}989$ Med samma räkning ger $g(0{,}6) \approx 0{,}541$

Stickprov: $X_1, \ldots, X_n$ på $N(\mu, \sigma^2)$ med σ^2 känd.

Testa $H_0: \mu = 0$ mot $H_1: \mu > 0$ på 1% signivå.

Hur stort behöver n vara för $g(1) = 0.9$?

Testet förkastar H_0 om

$$\frac{\overline{X}}{\sigma/\sqrt{n}} \geq \Phi^{-1}(0.99) \approx 2.33$$

Vi ska alltså beräkna

$$\mathbb{P}_1\left(\frac{\overline{X}}{\sigma/\sqrt{n}} \geq 2.33\right).$$

Om $\mu = 1$ är $\frac{(\overline{X}-1)\sqrt{n}}{\sigma} \sim N(0,1)$ och

$$\mathbb{P}_1\left(\frac{\overline{X}\sqrt{n}}{\sigma} \geq 2.33\right) = \mathbb{P}_1\left(\frac{(\overline{X}-1)\sqrt{n}}{\sigma} \geq 2.33 - \frac{\sqrt{n}}{\sigma}\right) = \Phi\left(\frac{\sqrt{n}}{\sigma} - 2.33\right)$$

Eftersom $\Phi^{-1}(0.9) \approx 1.28$ är $HL \geq 0.9$ då

$$\frac{\sqrt{n}}{\sigma} - 2.33 \geq 1.28 \Rightarrow n \geq (3.61\sigma)^2$$

<u>Bayesiansk statistik.</u> Istället för att se okända parametrar som fixa tal, se dem som sv.

<u>Ex</u> Mynt slegs 10 ggr. antingen klave m.s. $2/3$ eller m.s. $1/3$

Om vi får klave 6 ggr vad ska vi tro om sannol. att myntet ger klave? Data $X \sim Bin(n, \theta)$. Vi fick $x = 6$, är $\theta = 1/3$ eller $= ?$

Vi kan tänka oss θ som sv med $\mathbb{P}(\theta = 1/3) = \mathbb{P}(\theta = 2/3) = \frac{1}{2}$

$$\Rightarrow \mathbb{P}(\theta = 2/3 \mid X = 6) = \frac{\mathbb{P}(X=6 \mid \theta = 2/3)\mathbb{P}(\theta = 2/3)}{\mathbb{P}(X=6 \mid \theta = 2/3)\mathbb{P}(\theta = 2/3) + \mathbb{P}(X=6 \mid \theta = 1/3)\mathbb{P}(\theta = 1/3)} =$$

$$= \ldots = \frac{4}{5}$$

Den fördelning vi tror innan experimentet $(\mathbb{P}(\theta = 2/3) = \frac{1}{2})$ kallas <u>prior</u>. och den betingade fördelningen givet data $(\mathbb{P}(\theta = 2/3 \mid X = 6) = \frac{4}{5})$ kallas posterior

Bayes formel för tätheter

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{\infty} f_{Y|X}(y|t)f_X(t)dt} = cf(x,y) =$$

$$= cf_{Y|X}(y|x)f_X(x)$$

De sista likheterna då x inte har med nämnaren att göra

I Bayesiansk statistik har vi alltid en parameter Θ vars prior $f_\Theta(t)$ är känd, liksom den betingade för data givet Θ: $f_{X|\Theta}(x|t)$

Vi vill veta posterior $f_{\Theta|X}(t|x)$. Att ge objektiv prior är oftast omöjligt så därför är Bayesiansk statistik ibland utesluten.

Däremot vanligt i AI.

Ex Om man vet $X \sim exp(\Theta)$ $0<\Theta<1$ och obs $X=x$.

Vad tror du om Θ? Kan tänka sig prior som likformig $f_\Theta(t)=1$ $0<t<1$

$\Rightarrow f_{\Theta|X}(t|x) = cf_{X|\Theta}(x|t)f_\Theta(t) = cte^{-tx}$, $0<t<1$

OBS: ska ses som fkn. av t. Känns igen som täthet för $\Gamma(2,x)$, men bara för $0<t<1$ så posterior är trunkerad gammafördel.

Ex Dataström se föreläsning TM3

Ex Slant singlas 3 ggr. krona alla 3. Vad ska vi tro om sannolikheten Θ?

ML-Skattning blir 0 men verkar orimligt. Bayesianskt antag att prior likf. dvs $f_\Theta(t)=1$, $0<t<1$

$f_{\Theta|X}(t|0) = cf_{X|\Theta}(0|t)f_\Theta(t) = c(1-t)^3$ (ser snabbt att $c=4$)

Man kan t.ex skatta Θ med väntevärdet i posterior

$\mathbb{E}[\Theta|X=0] = 4\int_0^1 t(1-t)^3 dt = \frac{1}{5}$

Man kan känna igen $f_{\Theta|X}(t|0)$ som en β-fördelning

Def: En sv Y sägs vara β-fördelad med par $a>0, b>0$ om
$$f_Y(y) = c\, y^{a-1}(1-y)^{b-1}, \quad 0<y<1 \qquad \text{kort } Y \sim \beta(a,b)$$
Obs att $\beta(1,1) = \text{likf}(0,1)$

Ex ovan är specialfall av följande obs. Låt $\Theta \sim \beta(a,b)$ och sedan $X_1, \dots, X_n$ ober. och
$$\mathbb{P}(X_k = x) = \begin{cases} \Theta, & x=1 \\ 1-\Theta, & x=2 \end{cases}$$

Låt n_1 vara antal k: $X_k = 1$ & $n_2 = n - n_1$. Då gäller att
$$f_{\Theta|X_1,\dots,X_n}(t|x_1,\dots,x_n) = f_{X_1,\dots,X_n}(x_1,\dots,x_n|t)\, f_\Theta(t) =$$
$$= c\, t^{n_1}(1-t)^{n_2} t^{a-1}(1-t)^{b-1} = c\, t^{a+n_1-1}(1-t)^{b+n_2-1}$$
Alltså: $\Theta | X_1 = x_1, \dots, X_n = x_n$ är $\sim \beta(a+n_1, b+n_2)$

Om istället $X \sim \text{Bin}(n, \Theta)$ dyker det bara upp en faktor $\binom{n}{n_1}$ ovan.
Den beror ej på $t \Rightarrow \Theta | X_1 = n_1 \sim \beta(a+n_1, b+n_2)$ även nu

Generalisering: Vektorn $Y = (Y_1, \dots, Y_{k-1})$ sägs vara Dirichletfördelad med parametrar $b_1>0, b_2>0, \dots, b_k>0$ om det för alla $y = (y_1, \dots, y_{k-1})$ sådana att $y_i \geq 0 \ \forall i$ och $Y_k = 1 - \sum_{i=1}^{k-1} y_i \geq 0$ gäller att
$$f(y) = c\, y_1^{b_1-1} y_2^{b_2-2} \cdots y_k^{b_k-1}$$
Betafördelningen är specialfallet $n=2$. Vi skriver
$Y \sim \text{Dir}(b_1, \dots, b_k)$

Antag att $\Theta = (\Theta_1, \dots, \Theta_{k-1}) \sim \text{Dir}(b_1, \dots, b_k)$ och sedan
$X = (X_1, \dots, X_n)$ oberoende: $\mathbb{P}(X_k = j) = \Theta_j, \quad j = 1, \dots, k$
$$f_{\Theta|X}(t_1, \dots, t_{k-1}|x) = c\, t_1^{n_1} \cdots t_k^{n_k} \cdot t_1^{b_1-1} \cdots t_k^{b_k-1} = c\, t_1^{n_1+b_1-1} \cdots t_k^{n_k+b_k-1}$$
Dvs $\Theta | X = x \sim \text{Dir}(b_1+n_1, \dots, b_k+n_k)$. Detta är exempel på konjugerade prior. Om data X har en viss fördelning $f_{X|\Theta}$ och det visar sig att posterior $f_{\Theta|X}$ är av samma typ som prior f_Θ säger man att f_Θ är en konjugerande prior till $f_{X|\Theta}$

Allför.: β-förd är konj. prior till binomial förd.

och mer generellt Dir-förd är konj. prior till multinomiala normal-förd.

Ex Låt $\Theta \sim N(0,1)$ & $X \sim N(\Theta,1)$

$f_{\Theta|X}(t|x) = c e^{-\frac{1}{2}(x-t)^2} \cdot e^{-\frac{1}{2}t^2} = c e^{-(t-x/2)^2}$

Dvs $\Theta|X \sim N(x/2, 1/2) \Rightarrow$ Normalförd. är konjugerad prior till Normalfördelningen med känd varians.

Problem med nämnaren

Latent Dirichlet allocation: Låt $\Theta_i \sim Dir(a_1, \ldots, a_k)$, $i = 1, \ldots, D$

Låt $\psi_i \sim Dir(b_1, \ldots, b_n)$ $i = 1, \ldots, k$ Allt ober.

$\forall i, j$, $i = 1, \ldots, D$, $j = 1, \ldots, L$ Låt $Z_{ij} \sim \Theta_i$ och sedan $W_{ij} \sim \psi_{Z_{ij}}$

Vad är $f_{Z,\psi,\Theta|W}(\psi,\Theta,z|w)$ Täljaren i bayes formel hyfsat enkel men nämnaren är

$$f_W(w) = \sum_Z \int \cdots \int f_{W|\psi,\Theta,Z}(w|\psi,\Theta,Z) f(\psi,\Theta,Z)\, d\psi\, d\Theta$$

Detta är k^{DL} termer som alla är integraler

Gibbs sampling: Både $f_{Z|\psi,\Theta,W}$, $f_{\psi|\Theta,Z,W}$ & $f_{\Theta|\psi,Z,W}$ är lätta (De två sista är Dirichlet och givet allt annat är Z_{ij} oberoende med lätt beräknade fördelningar)

Gör så här:

1. Starta med vilka värden som helst på ψ, Θ, Z
2. Uppdatera Z enligt $f_{Z|\psi,\Theta,W}$
3. Uppdatera ψ enligt $f_{\psi|\Theta,Z,W}$
4. uppdatera Θ enligt $f_{\Theta|\psi,Z,W}$
5. Upprepa 2–4

Detta konvergerar emot den rätta fördelningen.

Gör många uppdateringar