Moment metoden

1. $\lim E[X] = 16 \cdot cx1 = \frac{\theta}{2}$

2: Sätt $E[X] = \bar{X}$

3: $\frac{\hat{\theta}}{2} = \bar{X} \Rightarrow \hat{\theta} = \bar{X} 2$

4: $E[\bar{X} 2] = 2 E[X] \Rightarrow 2 \cdot \frac{\hat{\theta}}{2} = \hat{\theta}$

Theory of Statistics - Complete Notes

Hampus Beijer

May 16, 2025

Introduktion

Dessa anteckningar täcker hela kursen Theory of Statistics I, inklusive alla föreläsningar från 1 till 14. Materialet innehåller grundläggande begrepp i sannolikhetsteori, statistisk inferens, Bayesiansk analys och hypotesprövning.

1 Föreläsning 1: Grundläggande sannolikhetsbegrepp

1.1 Mängdlära

- **Utfallsrum (S)**: Mängden av alla möjliga utfall
- **Händelse**: En delmängd av utfallsrummet
- **Operationer**:
  - Union $A \cup B$: Utfall i A eller B eller båda
  - Snitt $A \cap B$: Utfall i både A och B
  - Komplement A^c: Utfall inte i A
- **De Morgans lagar**:

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

1.2 Sannolikhetstolkningar

- **Frekventistisk**: Långsiktig relativ frekvens

$$P(A) = \lim_{n \to \infty} \frac{\text{Antal gånger A inträffar}}{n}$$

- **Bayesiansk**: Grad av tro
- **Axiom**:

SV X är diskret, $-3, -2, -1, 0, 1, 2, 3$, med lika $p(x) = \frac{1}{7}$

Vad är p.f av $Y = X^2 - X$

x	y
-3	12
-2	6
-1	2
0	0
1	0
2	2
3	6

då är ys pf

y	
$y=12$	$1/7$
$y=6$	$2/7$
$y=0$	$2/7$
$y=2$	$2/7$

Om detta är korrekt ska dessa summa till 1.

$\frac{1}{7} + \frac{2}{7} \cdot 3 = \frac{7}{7} = 1$

- $P[r(X) = y]$ är sannolikheten att transformationen $r(X)$ resulterar i värdet y
- $r : r(x) = y$ innebär alla x-värden som uppfyller $r(x) = y$
- f(x) är sannolikheten att $X = x$
- Summationen $\sum f(x)$ är summan av alla f(x) för dessa x värden är $g(y)$

Diskret fall, exempel: Antag att X är ett tärningskast med $X \in 1, 2, 3, 4, 5, 6$ med likformig fördelning, $f(x) = 1/6$. Låt $Y = r(X) = X$. Då finns det två olika möjligtheter (Y=0 eller Y=1). $g(y)$ kan du beräknas som Y=1 eller Y=1 enligt:

- $g(0) = \sum_{x:xjämt} = f(2) + f(4) + f(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$
- $g(1) = \sum_{x:xjämt} = f(1) + f(3) + f(5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$

Regeln

$P(|X| \le x) = p$
$= P(-x \le X \le x)$

$P(|X| \ge x) =$
$P(X > x) + P(X \le -x)$

$P(X^2 \ge y)$
$= P(|X| \ge \sqrt{y})$

Glöm aldrig att beräkna stödet för det ny G!

3.3 Funktion av kontinuerligt SV

När man arbetar med en kontinuerlig SV är stegen annrolunda för att finna fördelningen, så hur hittar man då fördelningen för Y=r(X) då X är känd?

1: Förutsättningar Funktionen $Y = r(x)$ måste vara en 1:1 funktion (inverterbar)

2: CDF för Y CDFn för Y betecknas $G(y)$ och skrivs som:

$$G(y) = P(Y \le y) = P[r(X) \le y] = P[X \le s(y)] = \int_a^{s(y)} f(x)dx = F[s(y)] - F[a] \tag{3}$$

där F(X) är CDFn för X.

3: PDF för Y Derivera G(y) med avseende på y då får man täthetsfunktionen (PDF) för Y:

$$f_Y(y) = f_X(s(y)) \left| \frac{ds(y)}{dy} \right|, \alpha < y < \beta \tag{4}$$

där $s(y) = r^{-1}(y)$. Denna metod är förutsätter att inversen exiterar. Metoden kallas för variablentransformation och använder kedjeregeln för att besvara sannolikhet:

4 Föreläsning 4: Bivariata och multivariata fördelningar

4.1 Bivariat fördelning

En bivariat fördelning är en samling av alla sannolikheter för händelser där både X och Y är iblandade, dvs $Pr[(X, Y) \in C]$ där C är en delmängd av möjliga utfall (t.ex X=0 och Y=¿ 1). Riktiga exempler på detta kan vara: Antal läkarbesök X och akutbesök Y, där C kan vara scenarier som "inga läkarbesök men minst ett akutbesök".

4.2 Joint Probablity Function: Diskreta variabler och gemensam sannolikhetsfunktion

För diskreta variabler defineras den gemensamma sannolikhetsfunktionen:

$$f(x,y) = P(X = x, Y = y) \tag{5}$$

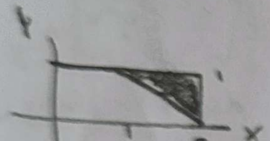
Sannolikheten för händelse C beräknas då genom att summera $f(x,y)$ över alla möjliga utfall i C:

$$Pr[(X,Y) \in C] = \int\int_C f(x,y) = \sum_{(x,y)\in C} f(x,y) \tag{6}$$

Då finns följande normeringsvillkor som säger att summan av alla möjliga sannolikheter måste summeras till 1:

$$\sum_{All x,y} f(x,y) = \sum_x \sum_y f(x,y) = 1 \tag{7}$$

Sammanfattningsvis kan man säga att den bivariata fördelningen beskriver hur X och Y beter sig tillsammans.



$f(x,y) = \begin{cases} \frac{3}{2}y^2, & 0 \le x \le 2 \\ & 0 \le y \le 1 \end{cases}$

Beräkna $P(X+Y>2)$

Stödet visar oss att olikheten enbart gäller om $y=1 \Rightarrow x+1>2$ pss då måste x vara $>2 \Rightarrow S[1;2]$ vilket blir nya stödet och $y > 2-x$

$\int_1^2\int_{x-2}^1 \frac{3}{2}y^2 dx dy$

$\int_1^2 \frac{3}{2}\left[\frac{y^3}{3}\right]_{x-2}^1 =$

$\int_1^2 \frac{3}{2}\left[\frac{1^3}{3} - \frac{(x-2)^3}{3}\right]$

$= \int_1^2 \left[\frac{1}{2} - \frac{(x-2)^3}{2}\right] dx$

$= \left[\frac{x}{2} - \left(\frac{(x-2)^4}{2\cdot 4}\cdot(-1)\right)\right]_1^2$

$= \left[\frac{2}{2} - \frac{0}{8} - \left(\frac{1}{2} - (-1)\cdot\frac{(1-2)^4}{8}\right)\right]$

$= 1 - 0 - \frac{1}{2} + \left(-\frac{1}{8}\right)$

$= \frac{4}{8} - \frac{1}{8} = \frac{3}{8}$

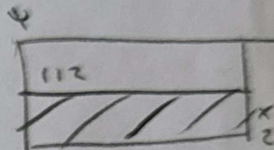
4.3 Continuous joint distributions: Kontinuerlig bivariat fördelning

Den kontinuerliga bivariata fördelningen (joing pdf) deineras enligt:

$$Pr[(X,Y) \in C] = \int\int_C f(x,y) \tag{8}$$

Där $f(x,y) \geq 0$ är den gemensamma täthetsfunktionen för X,Y.

- För en skild variabler (unvariat) motsvaras sannolikheten av area under täthetskurvan.
- För två varaibler **bivariat** motsvarar sannolikheten av volymen under täthetsytan över regionen C, även kallas bottenplattan.
- Den totala volymen under täthetsytan 'r alltid 1: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$



Beräkna $P(Y \le \frac{1}{2})$

nytt stöd

$0 \le y \le 1/2$

$0 \le x \le 2$

$\int_0^{1/2}\int_0^2 \frac{3}{2}y^2 dx dy$

$= \int_0^{1/2} \frac{3}{2}\left[y^2 x\right]_0^2 =$

$\int_0^{1/2} \frac{3}{2}\left[y^2 2\right] = \int_0^{1/2} 3y^2 dy$

$= 3\left[y^3/3\right]_0^{1/2} = 3\left(\frac{(1/2)^3}{3}\right)$

$= 1/8$

4.4 Joint CDF

4.4.1 1: Gemensam fördelningsfunktion, joint cdf

4.4.2 2: Sannolikheter för rektangulära områden

$\boxed{P(X \le 1)} \Rightarrow \int_0^1 \frac{3}{2}y^2 dy = \frac{3}{2}\left[\frac{y^3}{3}\right]_0^1$

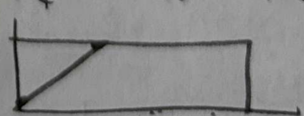
$= \left[\frac{y^3}{2}\right]_0^1 = \frac{1}{2}$

4.5 Gemensam fördelning

- Gemensam CDF:

$$F_{X,Y}(x,y) = P(X \le x, Y \le y)$$

$P(X = 3Y) = P(Y = \frac{X}{3})$



Alla lösningar ligger på linjen => det ligger oändligt många lösningar

- **Gemensam pdf (kontinuerlig):**

$$P((X,Y) \in C) = \iint_C f_{X,Y}(x,y)dxdy$$

4.6 Marginalfördelningar

Skriv kort om vad det är

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

4.7 Betingade fördelningar

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

4.8 Oberoende

X och Y är oberoende om:

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

$f(x,y) = \{c(x+y^2), \quad 0 \le x \le 1, \ 0 \le y \le 1$

$f(y) = \int_0^1 c(x+y^2)dx = c\left[\frac{x^2}{2} + y^2x\right]_0^1$

$= c\left[\frac{1}{2} + y^2\right] = f(y)$ då blir den betingade:

$f(x|y) = \frac{c(x+y^2)}{c(\frac{1}{2}+y^2)} = \frac{(x+y^2)}{(\frac{1}{2}+y^2)}$

4.9 Bivariat normal

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}\right)$$

Beräkna $f(x < \frac{1}{2} | Y = \frac{1}{2})$

Betingad fördelning:

$$X|Y = y \sim N\left(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y), \sigma_X^2(1-\rho^2)\right)$$

$= \int_0^{1/2} f(x | 1/2) = \frac{(x + \frac{1}{4})}{\frac{1}{2} + \frac{1}{4}}$

$= \int_0^{1/2} \frac{x+\frac{1}{4}}{\frac{3}{4}} = \int_0^{1/2} \frac{4}{3}\left[x + \frac{1}{4}\right]$

$= \frac{4}{3}\left[\frac{x^2}{2} + \frac{x}{4}\right]_0^{1/2} = \frac{4}{3}\left[\frac{0.5^2}{2} + \frac{0.5}{4}\right]$

$= \frac{1}{3}$

5 Föreläsning 5: Väntevärde och varians

3 SV från $X \sim U_n$, X_1, X_2, X_3 där $E[X] = 5$, Beräkna $E[2X_1 - 3X_2 + X_3 - 4]$

$= 2E[X_1] - 3E[X_2] + E[X_3] - 4$

$= 2 \cdot 5 - 3 \cdot 5 + 5 - 4$

$= -4$

5.1 Väntevärde

$$E[X] = \begin{cases} \sum_x xf(x) & \text{diskret} \\ \int_{-\infty}^{\infty} xf(x)dx & \text{kontinuerlig} \end{cases}$$

5.2 Varians

$$\text{Var}(X) = E[(X-\mu)^2] = E[X^2] - \mu^2$$

$= \int_{-\infty}^{\infty} x^2 f(x) - \mu^2$

5.3 Kovarians och korrelation

$$\text{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

$$\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

Betingat väntevärde/Varians

X and Y have a joint dist, pdf $= f(x,y) = \{x+y, \quad 0 \le x \le 1, \ 0 \le y \le 1$

Beräkna $E[Y|X] = \int y f(y|x) dy$

1: $f(x) = \int_0^1 (x+y)dy = [xy + y^2/2]_0^1 = [x + \frac{1}{2}]$

2: $f(y|x) = \frac{f(x,y)}{f(y)} = \frac{x+y}{x+1/2} = \frac{2x+2y}{2x+1}$

3: $\int_0^1 y\left(\frac{2x+2y}{2x+1}\right) = \int \frac{2xy+2y^2}{2x+1}dy = \frac{1}{2x+1}\int_0^1 (2xy+2y^2)dy = \frac{1}{2x+1}\left[\frac{2xy^2}{2} + \frac{2y^3}{3}\right]_0^1 = \frac{1}{2x+1}\left[\frac{2x}{2} + \frac{2}{3}\right] = \frac{1}{2x+1}\left[\frac{6x+4}{6}\right] = \frac{1}{2x+1}\left[\frac{3x+2}{3}\right]$ [Omskriv för enkelhetens skull]

$= \frac{3x+2}{3(2x+1)} = E[Y|X]$

$\boxed{\text{Var}[Y|X] = E[Y^2|X] + E^2[Y|X]}$

$E[Y^2|X] = \int_0^1 y^2\left(\frac{2x+2y}{2x+1}\right)dy = \int_0^1 \frac{2xy^2+2y^3}{2x+1}$

$= \frac{1}{2x+1}\int_0^1 (2xy^2+2y^3)dy = \frac{1}{2x+1}\left[\frac{2xy^3}{3} + \frac{2y^4}{4}\right]_0^1 = \frac{1}{2x+1}\left[\frac{2x}{3} + \frac{2}{4}\right]$

$= \frac{1}{2x+1}\left[\frac{8x+6}{12}\right] = \frac{1}{2x+1}\left[\frac{4x+3}{6}\right] = \frac{4x+3}{6(2x+1)}$

$= E[Y^2|X] \longrightarrow$

$X \sim uni(a,b)$ är en uni SV. Bestäm dess mgf: $E[e^{tX}] = \int_a^b \frac{e^{tx}}{b-a}dx = \left[\frac{e^{tx}}{t(b-a)}\right]_a^b = \frac{e^{tb}}{t(b-a)} - \frac{e^{tb}}{t(b-a)} = \frac{e^{tb}-e^{ta}}{t(b-a)}$

Betingad varians

$Var[Y|X] = \frac{4X+3}{6(2X+1)} - \left(\frac{3X+2}{3(2X+1)}\right)^2$

$= \frac{4x+3}{6(2x+1)} - \frac{(3x+2)^2}{9(2x+1)^2} = \underline{\text{Variansen}}$

X är en SV med mgf: $\psi(t) = e^{t^2+3t}$

Beräkna $E[X]$ och $Var[X]$ för X!

$\psi'(0) = e^{t^2+3t}(2t+3) = e^0(2\cdot 0+3) = 3$

Detta är väntevärdet aka första momentet

$\psi''(0) =$ /produktregeln/

$= e^{t^2+3t}(2t+3)^2 + e^{t^2+3t}(2)$

$= e^0(2\cdot 0+3)^2 + e^0\cdot 2$

$= 1\cdot 9 + 1\cdot 2 = 11 = E[X^2]$ andra momentet

$Var[X] = E[X^2] - E[X]^2 = 11 - 3^2 = 2$

$\hookrightarrow E[X] = 3$

$Var[X] = 2$

Median

X är SV $f(x) = e^{-x}, x>0$

Finn medianen

$P(X \le m) = \int_0^m e^{-x}dx = \frac{1}{2}$

$= \left[-1e^{-x}\right]_0^m = -e^{-m} - (-e^0)$

$= 1 - e^{-m} = \frac{1}{2} \Rightarrow$

$e^{-m} = 1 - \frac{1}{2} \Rightarrow e^{-m} = \frac{1}{2}$

$-m = \ln(1/2)$

$= -m = \ln 1 - \ln 2$

$= -m = -\ln 2$

$m = \ln 2$

$\boxed{m = 0.69}$

5.4 Momentgenererande funktion

$\psi(t)$ $M_X(t) = E[e^{tX}]$

Moment:

$$E[X^n] = M_X^{(n)}(0)$$

6 Föreläsning 6: Kontinuerliga fördelningar

6.1 Gammafördelning

$$X \sim \text{Gamma}(\alpha, \beta)$$

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$$

$$E[X] = \alpha/\beta, \quad \text{Var}(X) = \alpha/\beta^2$$

Derivata

Kedjeregeln = Inre · yttre

Produktregeln $(fg)' = f'g + f\cdot g'$

Kovarians

om X och Y är negativt korrel är $Var[X+Y]$ större än $Var[X-Y]$

Sats:

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

$Var[X-Y] = Var[X] + Var[Y] - Cov(X,Y)$

ex: $Var[X] = 4$, $Var[Y] = 9$, $Cov = -1$

DVS: $Var[X-Y] > Var[X+Y]$

6.2 Betafördelning

gå igenom Poisson, och Skriv uppgift 5.4.8

$$X \sim \text{Beta}(\alpha, \beta)$$

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

$$E[X] = \frac{\alpha}{\alpha+\beta}$$

7 Föreläsning 7: Gränsvärdessatser

7.1 Markovs olikhet

För $X \geq 0$:

$$P(X \geq t) \leq \frac{E[X]}{t}$$

7.2 Chebyshevs olikhet

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

alternativ formel

mäter hur stor sannolikhet att SV X avviker från sitt medelvärde med minst ett visst värde t

7.3 De stora talens lag

För oberoende likafördelade X_i med $E[X_i] = \mu$:

$$\bar{X}_n \xrightarrow{P} \mu$$

$$P(|X - E[X]| \geq t) \leq \frac{Var[X]}{t^2}$$

$X \sim gamma(\alpha, \beta)$ visa att $cX \sim gamma(\alpha, \beta/c)$

$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}\cdot x^{\alpha-1}\cdot e^{-x\beta}$

$Y = cX$ och $X = Y/c$

$f(s(y))\left|\frac{ds(y)}{dy}\right|$

$= \frac{\beta^\alpha}{\Gamma(\alpha)}\cdot(y/c)^{\alpha-1}\cdot e^{-\frac{\beta y}{c}}\cdot\frac{1}{c} = \frac{\beta^\alpha}{\Gamma(\alpha)}\cdot\frac{y^{\alpha-1}}{c^{\alpha-1}}\cdot e^{-\frac{\beta y}{c}}\cdot\frac{1}{c}$

$= \frac{\beta^\alpha}{\Gamma(\alpha)}\cdot\frac{y^{\alpha-1}}{c^{\alpha-1}\cdot c}\cdot e^{-\frac{\beta y}{c}} = \frac{1}{\Gamma(\alpha)}\cdot\frac{\beta^\alpha}{1}\cdot\frac{1}{c^\alpha}\cdot y^{\alpha-1}\cdot e^{-\frac{\beta y}{c}}$

$= \frac{1}{\Gamma(\alpha)}\cdot\frac{\beta^\alpha}{c^\alpha}\cdot y^{\alpha-1}\cdot e^{-\frac{\beta y}{c}} = \frac{1}{\Gamma(\alpha)}\cdot(\beta/c)^\alpha\cdot y^{\alpha-1}\cdot e^{-\frac{\beta y}{c}}$

$= \frac{(\beta/c)^\alpha}{\Gamma(\alpha)}\cdot y^{\alpha-1}e^{\frac{\beta y}{c}} \Rightarrow cX \sim gamma(\alpha, \beta/c)$

Hur stort måste ett stickprov vara från en given fördelning för att $P(x)$ ska vara minst 0.99 att medelvärdet kommer att inom 2 standardavvikelser från medelvärdet av fördelningen?

$n=?$	Vi söker: $P(\|\bar{X}_n-\mu\|\leq 2\sigma)\geq 0.99$

Chebychevs olikhet ger!

$P(\|X-\mu\|\geq t)\leq Var[X]/t^2$

sätt $t=2\sigma$

$P(\|X-\mu\|\geq 2\sigma)\leq \frac{\sigma^2}{4\sigma^2}$ | sätt nu $\bar{X}_n$ |

$P(\|\bar{X}_n-\mu\|\geq 2\sigma^2)\leq \frac{1}{4n}$

$P(\|\bar{X}_n-\mu\|\leq 2\sigma^2)\geq 1-\frac{1}{4n}$

$1-\frac{1}{4n}=0.99 \Rightarrow n=25 \Rightarrow n\geq 25$

Chebychevs Olikhet

X är en S.V med
- $E[X]=10$
- $P(X\leq 7)=0.2$
- $P(X\geq 13)=0.3$

Bevisa: $Var[X]\geq \frac{9}{2}$

$P(\|X-\mu\|\geq t)\leq Var[X]/t^2$

$P(\|X-10\|\geq t)$ $\|t=3\|$

$\Rightarrow P(X\geq 13)+P(X\leq 7)\leq \frac{Var[X]}{3^2}$

$= 0.3+0.2\leq \frac{Var[X]}{9}$

$= 0.5\cdot 9\leq Var[X] \Rightarrow Var[X]\geq \frac{9}{2}$

Markovs Olikhet

X är en S.V för vilket $P(X\geq 0)=1$ och $P(X\geq 10)=\frac{1}{5}$

Bevisa att $E[X]\geq 2$

$P(X\geq t)\leq \frac{E[X]}{t}$

$P(X\geq 10)\leq \frac{E[X]}{10}$

$\frac{1}{5}\leq E[X]/10$

$2\leq E[X] \Rightarrow E[X]\geq 2$

Väntevärde/varians

Betingat väntevärde

Och

Betingad varians

Saknas för diskreta

Och kontinuerliga fall

$X_1,...X_n$ formar ett stickprov $\bar{X}$ som är $exp(\beta)$ fördelat. Finn $\bar{X}_n$ fördelning.

exp mgf $=\left(\frac{\beta}{\beta-t}\right)$	$\varphi(t)=E[e^{t\bar{X}}]$

$=E[e^{t/n\cdot X_1}\cdots e^{t/n X_n}]$

$\Rightarrow \left(\frac{\beta}{\beta-t/n}\right)^n = \left(\frac{\beta n}{\beta n-t}\right)^n$

Man ser alltså att vi har en gamma fördelning

DVS: $\bar{X}_n \sim gamma(n,\beta n)$

Skriv om Momentmetoden och hur den används för olika fördelningar

7.4 Centrala gränsvärdessatsen

För oberoende likafördelade X_i med $E[X_i] = \mu$, $\text{Var}(X_i) = \sigma^2$:

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$$

8 Föreläsning 8: Simuleringsmetoder

Oviktigt

8.1 Monte Carlo-integration

$$E[g(X)] \approx \frac{1}{n}\sum_{i=1}^{n} g(X_i)$$

8.2 Invers transformationsmetod

För $U \sim U(0,1)$, sätt $X = F^{-1}(U)$.

9 Föreläsning 9-10: Bayesiansk inferens

9.1 Bayes sats

$$p(\theta \mid x) = \frac{f(x \mid \theta)p(\theta)}{f(x)} \propto \mathcal{L}(\theta; x)p(\theta)$$

viktigt

9.2 Konjugerade priorer

- Binomial - Beta prior
- Normal - Normal prior (känd varians)
- Poisson - Gamma prior

9.3 Prediktiv posterior

$$p(x_{ny} \mid x) = \int f(x_{ny} \mid \theta)p(\theta \mid x)d\theta$$

10 Lecture 10: Bayesian Analysis

10.1 Non-informative Priors

- Bernoulli model: $\theta \sim \text{Beta}(1,1)$ (uniform prior)
- Normal model: $p(\theta) = 1$ (improper prior)

$$\int p(\theta)d\theta = \infty \quad \text{(must verify posterior is proper)}$$

1. $P(A) \geq 0$ för alla händelser A
2. $P(S) = 1$
3. För disjunkta händelser $A_1, A_2, \ldots$:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

1.3 Betingad sannolikhet

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{om } P(B) > 0$$

1.4 Bayes sats

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$$

För en partition $B_1, \ldots, B_k$:

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{\sum_{j=1}^{k} P(A \mid B_j)P(B_j)}$$

1.5 Kombinatorik

- **Permutationer:** Ordade arrangemang

$$P(n, k) = \frac{n!}{(n-k)!}$$

- **Kombinationer:** Oordade urval

$$C(n, k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- **Multinomial:**

$$\binom{n}{n_1, \ldots, n_k} = \frac{n!}{n_1! \cdots n_k!}$$

2 Föreläsning 2: Slumpvariabler,

Föreläsningen täcker:

- Slumpvariabler (SV)
- Diskreta fördelningar (PF)
- Kontinuerliga fördelningar (PDF)
- Kumulativ fördelningsfunktion (CDF)

2.1 Stokastiska variabler (SV)

- **Definition:** En funktion $X : S \to \mathbb{R}$ som avbildar utfall på reella tal
- **Exempel:** Tärningskast

$$X(s) = \begin{cases} 0 & \text{om } s \in \{1,2,3\} \\ 1 & \text{om } s \in \{4,5,6\} \end{cases}$$

2.2 Fördelingstyper

En SV är diskret om den endast kan ta ett ändligt eller räknebart antal olika värden $x_1, x_2, \ldots$. En slumpvariabel är kontinuerlig om den kan ta varje värde inom ett givet intervall. För kontinuerliga och diskreta sannolikhetsfördelningar finns följande defitioner:

- **Diskret (Ändligt/uppräkneligt stöd.):** En diskret variabel har en sannolikhetsförunktion som på engelska är **probablitiy function, p.f**

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

$\int_{-\infty}^{\infty} f(x)dx = 1$

- **Kontinuerlig:** Sannolikhetsfunktionen för en kontinuerlig variabel kallas för en **Probablity Density Function, PDF** eller på svenska **Täthetsfunktion**, och är som ett histogram med små bin widths. Täthetsfunktion defineras:

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

Täthetsfunktionen för en kontinuerlig variabeö har följand egenskaper:

- $P(a \leq X \leq b) = \int_a^b f(x)dx$
- $f(x) \geq 0$ för alla x
- $\int_a^b f(x)dx = 1$

2.3 Viktiga fördelningar

- **Likformig diskret** ($X \sim U(a, b)$):

$$f(x) = \begin{cases} \frac{1}{b-a+1} & a, .., b \\ 0 & \text{annars} \end{cases}$$

- **Bernoulli** ($X \sim \text{Bern}(p)$):

$$P(X = 1) = p, \quad P(X = 0) = 1 - p$$

- **Binomial** ($X \sim \text{Bin}(n, p)$):

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- **Likformig kontinuerlig** ($X \sim U(a, b)$):

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{annars} \end{cases}$$

- **Normal** ($X \sim N(\mu, \sigma^2)$):

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

2.4 Kumulativ fördelningsfunktion (CDF)

Kumulativa fördelningsfunktionen eller **Cumulative Distribution Function (CDF)** för en kontinuerligt och diskret slumpvariabel $F(x)$ defineras:

$$F(x) = P(X \leq x), -\infty \leq x \leq \infty$$

En CDF har följande egenskaper

- Icke-avtagande: Om $x_1 \leq x_2$ då är $F(x_1) \leq F(x_2)$
- $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$
- Alltid kontinuerlig från höger $F(x) = F(x^+) = P(X \leq x)$
- Interval sannolikheter: $P(x_1 < X \leq x_2) = F(x_2) - F(x_1)$
- Skillnaden mellan strikta och svaga olikheter är viktiga $P(X < x) = F(x^-)$ och $P(X \leq x) = F(x) = F(x^-) + P(X = x)$

Kontinuerliga variabler har en **relation mellan pdf och cdf:**

$$F(x) = \int_{-\infty}^{x} f(t)dt \text{ och omvänt gäller} \frac{dF(x)}{d} = f(x)$$

3 Föreläsning 3: Funktioner av stokastiska variabler

3.1 Kvantiler

1: Definition av p-kvantil: p-kvantilen $F^{-1}(p)$ är det minsta x så att $F(x) \geq p$. Denna definiton fungerar för alla olika typer av CDF:er även om de är diskontinuerliga eller 1:1 över hela stödet för X.

Om kontinuerlig:

$x = F^{-1}(p)$ och $p = F(x)$

2: Specialfall för kontinuerlig CDF Om CDFn, $F(x)$ är kontinuerlig och 1:1 (dvs invers finns) så är p-kvantilen helt enkelt inversen av CDFn $x = F^{-1}(p =)$. Här kan man direkt lösa p = F(x):

$$x = F^{-1}(p) och p = F(x) \tag{1}$$

p är sannolikheten mellan 0 och 1 vilket representerar den kumulativa fördelningen upp till värdet x. Det är andelen av fördelningen som ligger till vänster om x. Exempelvis, p=0.975 innebär att 97.5% av data $\leq x$. +x är kvantilvärden eller gränsvärdet i fördelningen som motsvarar sannolikheten p. Exempel med median:

Vi har en SV $X \sim exp(\lambda)$ som har täthetsfunktionen $f(x) = \lambda e^{-\lambda x}, x > 0$. Konsekvent är då CDF $F(x) = 1 - e^{-\lambda x}$. **Vi söker medianen m = 0.5.** Teorin säger då att vi kan finna denna genom $F(m = 0.5) = 1 - e^{-\lambda x} = 0.5$. Lös sedan ut **m** och då fås $+e^{-\lambda m} = -0.5$ så $e^{-\lambda m} = 0.5$ sen $-\lambda m = ln(1/2)$ som kan skrivas som $-\lambda m = -ln(2)$ och $m = ln(2)/\lambda$. **3:Kända kvantiler**

- Median: $F^{-1}(0.5)$
- Första kvantil: $F^{-1}(0.25)$
- Tredje kvantil: $F^{-1}(0.75)$

3.2 Funktioner av en diskret SV

För att förstå funktioner av slumpvariabler. Föreställ dig följande scenario.

1:Situationen Vi har fördelningen för X t.ex sannolikhetsfunktionen för f(x) är känd. Fördelningen för Y=r(X) där r(.) ären given funktion (linjär, invertering, log, logit osv). Målet är då att finna funktionen av SV Y som är en funktion av SV X, Y = r(X).

2: Transformeringsexempler

- $Y = a + bX$ (Linjär transformation)
- $Y = 1/X$ (Invertering)
- $Y = ln(X)(Logaritm)$
- $Y = log(X/1 - X)(Logit - transform)$

3. Metod för att hitta sannolikhetsfunktion (PF) för Y=r(X) när X är en diskret SV:

$$g(y) = P(Y = y) = P[r(X) = y] = \sum_{x:r(x)=y} = f(x) \tag{2}$$

där,

- $g(y)$ är sannolikhetsfunktionen för Y. Dvs, den anger sannolikheten att Y antar värdet y

Föreläsning 9

Binomial

Anta att andelen defekta produkter θ är kända, 0.1 eller 0.2 och prior p.f av θ är $P(0.1) = 0.7$, $P(0.2) = 0.3$
Anta att 8 produkter väljs ut där 2 är defekta. Finn θ posterior.

Likelihood: $P(x|\theta) = p^x (1-p)^{n-x}$

$x = 2$
$n = 8 \Rightarrow P(0.1|2) = 0.1^2 \cdot (1-0.1)^{8-2} \cdot 0.7 = 0.00372$

$P(0.2|2) = 0.2^2 \cdot (1-0.2)^{8-2} \cdot 0.3 = 0.00315$

Uniform

Anta en enstaka observation, x, ska väljas från uniform fördelning på intervallet $[\theta - \frac{1}{2}; \theta + \frac{1}{2}]$ där θ är okänt. Prior fördelning θ är uniform på intervallet $[10; 20]$. Om den valda observationen x är 12 vad är då posterior?

$X \sim uni(\theta - \frac{1}{2}; \theta + \frac{1}{2})$, $n = 1$

$\theta \sim uni(10, 20)$

$$f(x|\theta) = \frac{1}{\theta + \frac{1}{2} - \theta + \frac{1}{2}}$$

$$P(\theta) = \frac{1}{20-10} = 0.1$$

$$P(\theta|12) \propto f(12|\theta) \cdot P(\theta) = 1 \cdot 0.1 = 0.1$$

$x - \frac{1}{2} < \theta < x + \frac{1}{2} \iff$

$\theta - \frac{1}{2} < x < \theta + \frac{1}{2}$

$11.5 < \theta < 12.5$

$10 < \theta < 20$

Normalfördelningen

Anta att ~~ett~~ OSU på 100 obs tas från $N(\theta, \sigma = 2)$ där θ är prior fördelning som är $N(\mu_0, v_0^2)$. Bevisa att hur stor std för priorn blir, posterior std kommer alltid $\leq \frac{1}{5}$

$x_1, \dots, x_{100}$ från OSU $\sim N(\theta, \sigma = 2)$

Prior $\theta \sim N(\mu_0, v_0^2)$

Posterior $\theta | x \sim N(\mu_1, v_1^2)$

Sats 7.3.3 ger formeln

$$v_1^2 = \frac{\sigma^2 \cdot v_0^2}{\sigma^2 + n v_0^2}$$

$$= \frac{4 \cdot v_0^2}{4 + 100 v_0^2} = /1 : v_0^2/$$

$$= \frac{\frac{4}{v_0^2} \cdot v_0^2}{\frac{4}{v_0^2} + 100 \frac{v_0^2}{v_0^2}} = \frac{4}{\frac{4}{v_0^2} + 100}$$

$$\rightarrow \frac{4}{100} \text{ om } v_0^2 \rightarrow +\infty$$

dvs:

$$\frac{4}{\frac{4}{v_0^2} + 100} \leq \frac{4}{100} \xrightarrow{} (std)^2 = \sqrt{\frac{4}{100}} = \frac{2}{10} = \frac{1}{5}$$

om $v_0^2 > 0$ så $v_1 \leq \frac{1}{5}$

Anta att ett OSU från en normal-fördelning där $N(\theta, \sigma = 2)$ med prior fördelning $\theta \sim N(\mu_0, \sigma = 1)$. Vad är det minsta antalet observationer som stickprovet måste innehålla för att kunna minska std för posterior för θ med 0.1?

Formel:

$$v_1^2 = \frac{\sigma^2 \cdot v_0^2}{\sigma^2 + n v_0^2} = \frac{4 \cdot 1}{4 + n \cdot 1}$$

$$\frac{4}{4+n} \leq 0.1^2 \Rightarrow 4 \leq 0.1(4+n)$$

$$\Rightarrow \frac{4 - 0.04}{0.01} \leq n \Rightarrow n \geq 3.96$$

Gamma fördelningen

Anta att $x_1, \dots, x_n$ formar ett stickprov från en fördelning där täthetsfunktionen $f(x|\theta)$ är:

$$f(x|\theta) = \theta x^{\theta-1} \quad 0<x<1$$

anta även att θ är okänd ($\theta>0$) och prior fördelningen av θ är gamma fördelat med α, β ($\alpha>0$ och $\beta>0$). Bestäm väntevärdet och varians för posterior fördelningen.

Posterior ~ gamma(?, ?) med $E[\theta] = \frac{\alpha}{\beta}$ och $Var[\theta] = \frac{\alpha}{\beta^2}$

Prior blir : $\cancel{\frac{\beta^\alpha}{\Gamma(\alpha)}} \cdot \theta^{\alpha-1} \cdot e^{-\beta\theta}$

denna kan strykas!

$p(\theta|x) \propto f(x|\theta) \cdot p(\theta)$

$p(\theta|x) \propto \theta x^{\theta-1} \cdot \theta^{\alpha-1} \cdot e^{-\beta\theta}$

$\propto \theta^{\alpha} \cdot x^{\theta-1} \cdot e^{-\beta\theta}$

$\propto \theta^{\alpha} \cdot e^{(\theta-1)\ln x} \cdot e^{-\beta\theta} \propto e^{-\theta(\beta-\ln x)} \cdot \theta^{\alpha+1-1}$

DVS $\Rightarrow \alpha = \boxed{\alpha+1}$ och $\beta = \boxed{\beta-\ln x}$

Så väntevärdet och varians blir:

$$E[\theta|x] = \frac{\alpha+1}{\beta-\ln x} \quad \text{och} \quad Var[\theta|x] = \frac{\alpha+1}{(\beta-\ln x)^2}$$

Om man vill ta hänsyn till stickprovet måste man skriva om likelihoodfunktionen så den kan inkludera alla x ist för bara ett sådant. Det ger följande resultat!

$$E[\theta|x] = \frac{\alpha+1}{\beta-\sum\ln x_i} \quad \text{och} \quad Var[\theta|x] = \frac{\alpha+1}{(\beta-\sum\ln x_i)^2}$$

A physicist makes 25 independent measurements of the specific gravity of a certain body. He knows that the limitations of his equipment are such that the std of each measurement is σ units

a) By using chebyshev's inequality find the lower bound for the probability that the average of his measurements will differ from the actual specific gravity of the body by less than $\frac{\sigma}{4}$ units

b) By using the CLT find an approximate value of the probability in (a)

$$P\left(|\bar{X}_n - \mu| < \frac{\sigma\sqrt{n}}{4\sqrt{n}}\right) =$$

$$P\left(|Z| < \frac{\sqrt{n}}{4}\right) \approx$$

$$\Phi(1.25) - \Phi(-1.25)$$

$$= 2 \cdot \Phi(1.25) - 1$$

$$= 2 \cdot 0.89944 - 1 = 0.7888$$

10.2 Normal Data Analysis

- Known variance, uniform prior:

$$\theta|x \sim N\left(\bar{x}, \frac{\sigma^2}{n}\right)$$

- Known variance, normal prior $\theta \sim N(\mu_0, \tau_0^2)$:

$$\frac{1}{\tau_n^2} = \frac{n}{\sigma^2} + \frac{1}{\tau_0^2}$$
$$\mu_n = w\bar{x} + (1-w)\mu_0$$
$$\text{where } w = \frac{n/\sigma^2}{n/\sigma^2 + 1/\tau_0^2}$$

10.3 Poisson Model

$$\text{Likelihood: } p(y|\theta) \propto \theta^{\sum y_i} \exp(-\theta n)$$
$$\text{Conjugate prior: } \theta \sim \text{Gamma}(\alpha, \beta)$$
$$\text{Posterior: } \theta|y \sim \text{Gamma}\left(\alpha + \sum y_i, \beta + n\right)$$

11 Lecture 11: Estimation Methods

11.1 Bayes Estimators

$$\text{Quadratic loss: } L(a,\theta) = (\theta - a)^2 \Rightarrow \text{Posterior mean}$$
$$\text{Linear loss: } L(a,\theta) = |\theta - a| \Rightarrow \text{Posterior median}$$

11.2 MLE Examples

- Poisson: $\hat{\theta} = \bar{x}$
- Bernoulli: $\hat{\theta} = \bar{x}$
- Normal (known σ^2): $\hat{\mu} = \bar{x}$

11.3 Method of Moments

Solve system:

$$\mu_j(\theta) = m_j \quad \text{for } j = 1, ..., k$$

12 Lecture 12: Sampling Distributions

12.1 Chi-Square Distribution

- $X \sim \chi^2_m$ is *Gamma*$(m/2, 1/2)$
- Properties:

$$E(X) = m$$
$$Var(X) = 2m$$

- Sum of squares:

$$\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1} \quad \text{(unknown } \mu\text{)}$$

12.2 t-Distribution

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

13 Lecture 13: Confidence Intervals

13.1 Normal Mean CI

$$\bar{X} \pm t_{n-1,\alpha/2} \frac{S}{\sqrt{n}}$$

13.2 Normal Variance CI

$$\left(\frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}} \right)$$

13.3 Fisher Information

$$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} \ln f(X|\theta)\right]$$
$$I_n(\theta) = nI(\theta)$$

14 Lecture 14: Hypothesis Testing

14.1 Test Statistics

- Normal (known σ^2):

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$$

Data från antal kalorier i en korv sandw in. 20 olika märken:

186, 181, ..., 132.

Anta att dessa 20 observationer från ett observerat stickprov av 20 ober S.V från normalfördelningen med okänd μ och σ^2. Skapa ett 90% KI för medelvärdet av kalorier μ i en korv.

KI = $P(A < U < B)$

Vi finner en pivotal kvantitet i detta fall U från t-förd. då σ^2 är okänd

$U = t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t(n-1)$: DVS

$P(A < \frac{\bar{x} - \mu}{s/\sqrt{n}} < B) = \gamma$

där

$\boxed{A = \bar{x} - t \cdot s/\sqrt{n} \text{ och } B = \bar{x} + t \cdot \frac{s}{\sqrt{n}}}$

DVS

$P(\bar{x} - t \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t \cdot \frac{s}{\sqrt{n}}) = 0.9$

om $\bar{x} = 156.85$ och $s = 22.64$

$156.85 \pm 1.729 \cdot \frac{22.64}{\sqrt{20}}$

- t-test (unknown σ^2):

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

- F-test (variance comparison):

$$V = \frac{S_X^2/(m-1)}{S_Y^2/(n-1)} \sim F_{m-1,n-1}$$

14.2 Likelihood Ratio Test

$$\Lambda = \frac{f(x|\hat{\theta}_0)}{f(x|\hat{\theta})}, \quad -2\ln\Lambda \overset{a}{\sim} \chi_k^2$$

14.3 Error Types

Type I	Reject H_0 when true
Type II	Fail to reject H_0 when false

Fisher information

$X \sim Poi(\theta)$, θ är okänd. Finn $I(\theta)$ för s.v. X!

$$L(\theta) = \frac{e^{-\theta} \cdot \theta^x}{x!}$$

$$\ln L(\theta) = \ln\left(\frac{e^{-\theta} \cdot \theta^x}{x!}\right)$$

$$= \ln(e^{-\theta} \cdot \theta^x) - \ln(x!)$$

$$= \ln(e^{-\theta}) + \ln\theta^x - \ln(x!)$$

$$= -\theta + x\ln\theta - \ln(x!)$$

$$\boxed{= x\ln\theta - \ln(x!) - \theta = \ln L(\theta)}$$

$$\frac{d\ln L(\theta)}{d} = \frac{x}{\theta} - 1$$

$$\boxed{\frac{d\ln L(\theta)}{d^2} = -\frac{x}{\theta^2}}$$

Fisher information:

$$-E[d/d^2 \ln L(\theta)] = -E[-x/\theta^2]$$

$$= -\left(-\frac{E[X]}{\theta^2}\right) = \frac{\theta}{\theta^2} = \frac{1}{\theta} = I(\theta)$$

MLE skattningars asymptotiska egenskaper

$X_1, \dots, X_n$ formar ett slumpmässigt stickprov från en fördelning med pdf:

$f(x|\theta) = \theta x^{\theta-1}$, $0 < x < 1$, $\theta > 0$ och okänd

Bestäm den asymptotiska fördelningen av MLE $\hat{\theta}$

MLE asymptotiska egenskaper:

$$\boxed{\hat{\theta} \sim N\left(\theta, \frac{1}{I(\theta) \cdot n}\right)}$$

$L(\theta) = \theta^n \cdot \Sigma x_i^{\theta-1}$, $\ln L(\theta) = \ln(\theta^n \cdot \Sigma x_i^{\theta-1})$

$$\boxed{= n\ln\theta + (\theta-1) \cdot \Sigma \ln x_i}$$

$$d/d\theta \ln L(\theta) = \frac{n}{\theta} + \Sigma \ln x_i$$

MLE blir

$$\frac{n}{\theta} = \Sigma \ln x_i$$

$$\boxed{\hat{\theta} = \frac{n}{\Sigma \ln x_i}}$$

andra derivata blir

$$\frac{d\ln L(\theta)}{d^2\theta} - \frac{n}{\theta^2}$$

$$I(\theta) = -E\left[\frac{d\ln L(\theta)}{d^2}\right]$$

$$I(\theta) = -E\left[-\frac{n}{\theta^2}\right] = \frac{n}{\theta^2}$$

DVS MLE skattningars asymptotiska egenskaper blir:

$$\hat{\theta} \overset{asymp}{\sim} N\left(\theta, \frac{1}{\frac{n}{\theta^2}}\right)$$

DVS

$$\hat{\theta} \overset{asymp}{\sim} N\left(\theta, \frac{\theta^2}{n}\right)$$