

Tentamen i Sannolikhetslära och statistik (TDAB01), 6 hp

Tid:	08:00-12:00
Tillåtna hjälpmedel:	Miniräknare med tomt minne. Tabell- och formelsamling (delas ut tillsammans med tentamen)
Examinator:	Mattias Villani, tel. 070 – 0895205
Jourlärare:	Sarah Alsaadi, tel. 013 – 284647
Betyg:	Maximalt antal poäng: 20 poäng. Varje delfråga ger maximalt 5 poäng. Betyg 5 = 17-20 poäng Betyg 4 = 12.5-16.5 poäng Betyg 3 = 9-12 poäng

För full poäng krävs tydliga och väl motiverade svar.

- Givet en slumpvariabel X som bara kan anta värdena 3, 4, 6, 7 och 9 med sannolikheterna $P(X = 3) = 1/3$, $P(X = 4) = 1/4$, $P(X = 6) = 1/6$ och $P(X = 7) = 1/7$.
 - Beräkna $P(X = 9)$.
Lösning: $P(X = 9) = 3/28$.
 - Beräkna $P(4 \leq X < 7)$.
Lösning: $5/12$.
 - Beräkna $E(X)$ och $Var(X)$.
Lösning: $E(X) = 139/28 \approx 4.96$ och $Var(X) = E(X^2) - E(X)^2 = 3163/784 \approx 4.04$.
- Ett försäkringsbolag säljer försäkringar till 200 kunder under ett givet år. Anta att kunderna drabbas av olyckor oberoende av varandra och låt X_i vara en slumpvariabel som betecknar ersättningsbeloppet till kund i . Anta även att $E(X_i) = 100$ och $Var(X_i) = \sigma^2$.
 - Om $\sigma = 100$, vad är sannolikheten att bolaget under ett givet år måste betala ut mer än 21000 kronor?
Lösning: Låt $Y = \sum_{i=1}^{200} X_i$. Vi utnyttjar CGS och får att $P(Y > 21000) = P(\frac{Y - 200 \cdot 100}{\sigma \sqrt{200}} > \frac{21000 - 200 \cdot 100}{\sigma \sqrt{200}}) \approx P(Z > \frac{100}{\sigma \sqrt{2}})$ där $Z \sim N(0, 1)$. Insättning av $\sigma = 100$ ger att $P(Y > 21000) \approx 0.22$.
 - Bestäm σ så att sannolikheten att bolagets sammanlagda utbetalning under det givna året ska överstiga 21000 är 1%.
Lösning: Vi har att $P(Z > \frac{100}{\sigma \sqrt{2}}) = 1 - \Phi(\frac{100}{\sigma \sqrt{2}})$ och att $P(Z > \frac{100}{\sigma \sqrt{2}}) = 0.01$. Tillsammans ger ekvationerna att $\sigma = \frac{100}{2.33 \cdot \sqrt{2}}$.
- Låt $N(t)$ vara en Poissonprocess med intensitet $\lambda = 5$.

- (a) Vilken fördelning har $N(2)$?

Lösning : $N(2) \sim Po(5 * 2)$.

- (b) Beräkna $P(N(2) - N(1) = 2)$

Lösning : $P(N(2) - N(1) = 2) = e^{-5} * 5^2 / 2!$.

- (c) Beräkna $P(N(2) = 3 | N(1) = 1)$.

Lösning : $P(N(2) = 3 | N(1) = 1) = P(N(2) - N(1) = 2) = e^{-5} * 5^2 / 2!$.

- (d) Låt T vara tiden mellan den andra och tredje händelsen i $N(t)$. Vilken fördelning har T ?

Lösning : $T \sim exp(5)$ (väntevärde $1/5$).

4. Låt $X_1, \dots, X_n | \lambda$ vara ett oberoende stickprov från en fördelning med täthet

$$f(x) = \lambda^2 x e^{-\lambda x},$$

för $x > 0$ och $f(x) = 0$ annars. Parametern λ är strikt positiv.

- (a) Härled maximum likelihood estimatorn av λ .

Lösning (2 p): Likelihood funktionen ges av

$$L(\lambda) = \prod_{i=1}^n \lambda^2 x_i e^{-\lambda x_i} = \lambda^{2n} \left(\prod_{i=1}^n x_i \right) e^{-\lambda \sum_{i=1}^n x_i}$$

och log-likelihooden

$$\log L(\lambda) = 2n \log \lambda + \log \left(\prod_{i=1}^n x_i \right) - \lambda \sum_{i=1}^n x_i$$

med första derivata

$$\frac{d \log L(\lambda)}{d\lambda} = \frac{2n}{\lambda} - \sum_{i=1}^n x_i = \frac{2n}{\lambda} - n\bar{x}.$$

Sätt första derivatan lika med noll och lös för λ för att få maximumet

$$\hat{\lambda}_{ML} = \frac{2}{\bar{x}}.$$

För att verifiera att detta faktiskt är ett maximum kontrollerar vi att andraderivatan är negativ

$$\frac{d^2 \log L(\lambda)}{d\lambda^2} = -\frac{2n}{\lambda^2} < 0$$

för alla $\lambda > 0$.

- (b) Skatta λ med momentmetoden.

Lösning (1 p): Väntevärdet för X kan beräknas genom integration, $E(X) = \int_0^\infty x f(x) dx$, eller genom att upptäcka att fördelningen $f(x) = \lambda^2 x e^{-\lambda x}$ är en $\Gamma(2, \lambda)$ fördelning. Tabell- och formelsamlingen ger då direkt

$$E(X) = \frac{\alpha}{\lambda} = \frac{2}{\lambda}.$$

Om vi löser $E(X) = \bar{x}$ med avseende på λ får vi momentskattningen

$$\hat{\lambda}_{MOM} = \frac{2}{\bar{x}},$$

vilket råkar vara exakt samma estimator som ML-estimatorn.

- (c) Härled posteriorfördelningen för λ givet en konjugerad apriorifördelning.
Lösning (2 p): Vi använder Bayes sats och likelihooden från Uppgift 4a)

$$\begin{aligned} p(\lambda|x_1, \dots, x_n) &\propto p(x_1, \dots, x_n|\lambda)p(\lambda) \\ &\propto \lambda^{2n} e^{-\lambda n\bar{x}} \cdot p(\lambda). \end{aligned}$$

Vi ser på formen av likelihoodfunktionen (“ λ upphöjt i något multiplicerat med $\exp()$ av λ ggr en konstant”) att $\Gamma(a, b)$ -fördelningen är konjugerad:

$$p(\lambda|x_1, \dots, x_n) \propto \lambda^{2n} e^{-\lambda n\bar{x}} \cdot \lambda^{a-1} e^{-b\lambda} = \lambda^{2n+a-1} e^{-\lambda(n\bar{x}+b)},$$

vilket är proportionellt mot $\Gamma(2n + a, n\bar{x} + b)$ -fördelningen. En Gamma-fördelad prior ger alltså en Gamma-fördelad posterior vilket visar att priorn faktiskt är konjugerad (till den likelihood vi har).

LYCKA TILL!

MATTIAS