TDAB01 SANNOLIKHETSLÄRA OCH STATISTIK

TENTAMEN 2022-01-05

Lärare

Jose M. Peña. Nås via Teams **kl 16-17** (eller på `jose.m.pena@liu.se` om Teams inte fungerar).

Betyg

För full poäng i varje delfråga krävs tydliga och väl motiverade svar.
Maximalt antal poäng: 20 poäng
Betyg 5 = 18-20 poäng
Betyg 4 = 14-17 poäng
Betyg 3 = 10-13 poäng
Betyg U = 0-9 poäng

Tillåtna Hjälpmedel

Miniräknare med tomt minne. Tabell- och formelsamling (ingår i tentamen). Slides på kurswebbsidan (som pdf eller utskrivna, med eller utan egna anteckningar). Kursboken.

Instruktioner

- Ladda ner tentan från LISAM från kl 14.00. Lämna in dina svar i LISAM senaste kl 18.15. Om du har fått förlängd tid, v.v. och mejla dina svar till `jose.m.pena@liu.se`.
- Om du har frågor, v.v. och använd Teams för att kontakta mig **kl 16-17** (eller `jose.m.pena@liu.se` om Teams inte fungerar).
- Tentan är individuell, dvs hjälp från eller kommunikation med andra är inte tillåten.
- Tentan är anonym, dvs skriv ej ditt namn någonstans.
- Du rekommenderas lösa uppgifterna med papper och penna, fota lösningarna och ladda upp dem. Det är också OK att använda Word eller LaTeX.

Uppgifter

(1) (2 p) Låt X och Y vara kontinuerliga variabler med simultan täthetsfunktion

$$f(x,y) \propto \exp\Big[-\frac{x^2}{2}\Big]\exp\Big[-\frac{y^2}{2}\Big].$$

Är X or Y beroende eller oberoende?

Lösning: Obs. $f(x,y)$ är produkten av två standard normal täthetsfunktioner (efter man lägger till normalisering konstanter). Då, $f(x) = \int f(x,y)dx$ är standard normal, och lika för $f(y)$. Då, $f(x,y) = f(x)f(y)$. Då, X och Y är oberoende.

(2) (2 p) Låt X och Y vara binära variabler. Ge en simultanfördelning för (X,Y) där X and Y är beroende och, dessutom, de marginalfördelningarna för X och Y är likformiga fördelningar, dvs sannolikheten är konstant över utfallsrummet.

Lösning: p(X=0,Y=0)=p(X=1,Y=1)=0.1, p(X=0,Y=1)=p(X=1,Y=0)=0.4.

(3) (4 p = 2 + 2) En person som är frisk en dag är frisk dagen efter med sannolikhet 0.9. En person som är sjuk en dag är sjuk dagen efter med sannolikhet 0.8. Det finns dock ett undantag till den här regeln: När en person blir sjuk, hen är sjuk minst två dagar.

Ange transitionsmatrisen som tillhör Markovkedjan. Om man är frisk idag, vad är sannolikheten att man är frisk om tre dagar?

Lösning: Det finns tre tillstånd: Frisk, sjuk första dagen, sjuk fortsättning. Då,

$$P = \begin{pmatrix} 0.9 & 0.1 & 0 \\ 0 & 0 & 1 \\ 0.2 & 0 & 0.8 \end{pmatrix}.$$

Svaret på frågan är $p_{1,1}^{(3)}$, dvs (1,1)-elementen av P^3.

(4) (2 p) Slumpvariabeln X är normal fördelad med $E(X) = -3$ och $var(X) = 4$. Beräkna (a) $p(X = -3)$, (b) $p(X \leq 2.39)$, (c) $p(-2.39 < X < 2.39)$, och (d) värdet a så att $p(X > a) = 0.33$.

Lösning:

(a) $p(X = -3) = 0$.

(b) $p(X \leq 2.39) = p(\frac{X+3}{2} \leq \frac{2.39+3}{2}) = p(Z \leq 2.695) = \Phi(2.695) = 0.9965$ med $Z \sim \mathcal{N}(0,1)$.

(c) $p(-2.39 < X < 2.39) = p(\frac{-2.39+3}{2} < \frac{X+3}{2} < \frac{2.39+3}{2}) = p(0.305 < Z < 2.695) = \Phi(2.695) - \Phi(0.305) = 0.9965 - 0.6179$.

(d) $0.33 = p(X > a) = 1 - p(X \leq a) \Rightarrow p(X \leq a) = 0.67 = p(\frac{X+3}{2} \leq \frac{a+3}{2}) = p(Z \leq \frac{a+3}{2}) = \Phi(\frac{a+3}{2}) \Rightarrow \frac{a+3}{2} = 0.44 \Rightarrow a = -2.12$.

(5) (3 p) En fabrik producerar en viss vara. Varans kvalitetsindikatorn följer en normal fördelning $N(\mu = 3, \sigma = 3)$. Varan godkänns om kvalitetsindikatorn tillhör intervallet väntevärdet plus/minus två standardavvikelser, dvs om kvaliteten inte avviker från väntevärdet med mer än två standardavvikelser.

Beräkna sannolikheten att det finns fler än 990 godkända varor i ett stickprov med 1000 varor. Anta att varorna är oberoende.

Lösning: Varan är godkänd om $p(\mu - 2\sigma \leq X \leq \mu + 2\sigma)$ där X är kvalitetsindikatorn, dvs om $p(-2 \leq Z \leq 2)$ med $Z \sim \mathcal{N}(0,1)$. Låt $p = \Phi(2) - \Phi(-2)$. Då, $Y \sim Bin(1000, p)$ där Y är antal varor godkända i stickprovet. Pga centrala gränsvärdessatsen, $Y \sim N(1000p, \sqrt{1000p(1-p)})$. Då, svaret är $p(Y > 990.5)$, som man kan räkna genom standarisering av en normal fördelad variabel.

(6) (5 p) I kursen, har ni lärt er maximum likelihood metoden för punktskattning. Det finns andra metoder dock för punktskattning, tex momentmetoden. Läs avsnitt 9.1.1 om momentmetoden i kursboken. Använd momentmetoden för att härleda två estimatorer för väntevärdet av en Poisson fördelad variabel.

Lösning: Om X är Poisson fördelad, då $E[X] = var(X)$. Momentmetoden säger att $E[X]$ kan skattas som $\bar{X}$ och $var(X)$ som S^2. Då, $\overline{X}$ och S^2 är båda estimator av $E[X]$ enligt momentmetoden.

(7) (2 p) Som en fortsättning av den föregående uppgiften, bygg ett 95 % tvåsidigt konfidensintervall för väntevärdet av Poisson fördelningen. Anta att vi har beräknat medelvärdet av ett stickprov av storlek 500 från fördelningen och fått att medelvärdet är 2.

Lösning: Låt X vara Poisson fördelad. Låt $\theta = E[X]$ och $\hat{\theta} = \bar{X}$. Då, $\bar{X} \pm z_{0.025} \cdot \frac{\sigma}{\sqrt{n}}$ är ett 95%-igt konfidensintervall för θ. Dessutom, $\sigma^2 = var(X) = E[X]$ och $E[X]$ skattas som $\bar{X}$, som är lika med 2. Nu kan vi bygga intervallet.

Lycka till!