MVE090 Matematisk statistik Z2

Skriftlig tentamen fredag den 28 augusti 2015 kl 8.30 – 12.30

Lärare och jour: Patrik Albin, telefon 070 6945709.

Hjälpmedel: Beta eller häftet *Tommy Norberg: Formler och tabeller till matematisk statistik på universitet och tekniska högskolor* eller fyra handskrivna A4-sidor (xerox-kopior, datautskrifter etc. är ej tillåtna) – endast ett av dessa tre hjälpmedel är alltså tillåtet och eleven väljer själv vilket alternativ den vill använda (innan tentan börjar).

Betygsgränser: 12, 18 resp. 24 poäng för betyg 3, 4 resp. 5.

Motiveringar: alla svar och lösningar skall motiveras såvida inget att anges.

1. Bestäm sannolikheten att vid kast av fem stycken vanliga tärningar (Yatzy) resultatet blir så kallad kåk ("full house" på engelska), dvs. två stycken tärningar visar ett visst antal prickar (lika många på båda tärningarna) och de tre andra tärningarna visar ett visst annat antal prickar (lika många på alla tre tärningarna) . **(5 poäng)**

2. Hur många gånger N måste man kasta ett välbalanserat mynt (dvs. sannolikhet för klave är 1/2) för att sannolikheten skall vara minst 90% att man fått minst två stycken klave under de N kasten? **(5 poäng)**

3. Låt $\lambda > 0$ vara ett givet tal. Bestäm $E[Y|X=x]$ för en tvådimensionell kontinuerlig stokastisk variabel (X,Y) med täthetsfunktion

$$f_{X,Y}(x,y) = \frac{\lambda^3(x+y)}{2}\,\mathrm{e}^{-\lambda(x+y)} \quad \text{för } x,y \geq 0 \quad \text{och} \quad f_{X,Y}(x,y) = 0 \quad \text{för övrigt.}$$

(5 poäng)

4. Utanför en fotbollsarena önskar en teknolog finna ett konfidensintervall inom vilket längden av fotbollssupportar ligger med 95% säkerhet genom att registrera längden x_i för ett antal testsupportrar $i = 1, \ldots, n$. Beskriv hur detta kan gå till. **(5 poäng)**

5. Utanför en fotbollsarena önskar en teknolog finna en linjär modell $y = a + b\,x$ som på bästa sätt beskriver sambandet mellan längd x och vikt y för fotbollssupportar genom att registrera längd x_i och vikt y_i för ett antal testsupportrar $i = 1, \ldots, n$. Enligt Beta ges bästa linjen av $a = \bar{y} - b\,\bar{x}$ och $b = [\sum_{i=1}^n x_i\,(y_i - \bar{y})]/[\sum_{i=1}^n x_i\,(x_i - \bar{x})] = [\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})]/[\sum_{i=1}^n (x_i - \bar{x})^2]$. Hjälp teknologen härleda dessa formler! **(5 poäng)**

6. Låt $X_1, \ldots, X_n$ vara ett stickprov på en kontinuerlig stokastisk variabel X med täthetsfunktion $f(x) = \lambda^{-1}\,\mathrm{e}^{-(x-\mu)/\lambda}$ för $x \geq \mu$ och $f(x) = 0$ för $x < \mu$. Bestäm skattningen enligt maximum likelihood metoden av parametrarna $\lambda > 0$ och $\mu \in \mathbb{R}$. **(5 poäng)**

Lycka till!

MVE090 Matematisk statistik Z2

Lösningar till tentamen fredag den 28 augusti 2015

1. $6 \cdot 5 \cdot \binom{5}{2} \cdot (\frac{1}{6})^5$.

2. Sannolikheten för högst en klave på N kast är $\binom{N}{0} \cdot (\frac{1}{2})^N + \binom{N}{1} \cdot (\frac{1}{2})^N = (N+1)\,(\frac{1}{2})^N$ som är $7/64 > 10\%$ för $N = 6$ och $8/128 < 10\%$ för $N = 7$, så svaret är $N = 7$.

3. Eftersom

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \ldots = \frac{\lambda\,(1+\lambda x)}{2}\,\mathrm{e}^{-\lambda x} \quad \text{för } \ x \geq 0,$$

följer att

$$E[Y|X = x\,] = \int_{-\infty}^{\infty} y\, f_{Y|x}\,dy = \int_{-\infty}^{\infty} y\,\frac{f_{X,Y}(x,y)}{f_X(x)}\,dy = \ldots = (2+\lambda\,x)/(\lambda+\lambda^2\,x).$$

4. Det finns flera sätt – ett är att passa en normalfördelning till data enligt exempel 7.2.4 i boken av Milton och Arnold och sedan låta intevallet vara området mellan 2,5% och 97,5% percentilerna för den passade fördelningen – ett annat är att låta intevallet vara området mellan 2,5% och 97,5% percentilerna för den empiriska fördelningen ("relative cumulative frequency ogive") för data.

5. Se avsnitt 11.1 i boken av Milton och Arnold.

6. Man ser lätt att för varje givet λ-värde likelihood funktionen $\prod_{i=1}^{n} f(x_i)$ maximeras för $\mu = \min(x_1, \ldots, x_n)$. För varje givet μ-värde visar vidare en enkel derivering att likelihood funktionen maximeras för $\lambda = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu) = \bar{x} - \mu$.