Tentamen i TMA321 Matematisk Statistik, Chalmers Tekniska Högskola.

Hjälpmedel: Valfri räknare, egenhändigt handskriven formelsamling (4 A4-sidor på 2 blad) och till skrivningen medhörande tabeller.

Fredagen den 24 Augusti 2012, Eftermiddag 14.00-18.00.

Examinator och Jour: Olle Nerman, tel. 7723565, rum 3056, MV, Chalmers.

Tentamen består av 8 frågor med sammanlagt 32 poäng+ eventuella bonuspoäng från dina två inlämningsuppgifter.

Betygsgränser: För betyget "3" fordras minst 15 poäng, för betyget "4" minst 20 poäng och för betyget "5" minst 25 poäng .

---

1. Du kastar två vanliga sex-sidiga tärningar en gång.
   a. Vad är sannolikheten att poängtalen på de båda tärningarna är lika? (1p)
   b. Vad är den betingade sannolikheten att de båda poängtalen är lika givet att summan av poängtalen är **10**? (1p)
   c. Låt **Y** vara definierad som absolutbeloppet av skillnaden av de två poängtalen. Vad är då utfallsrummet för **Y**? (1p)
   d. Bestäm väntevärdet för **Y** definierad i c-uppgiften. (1p)
2. Antag att är Bernouilli-fördelad med parameter **p=0.3** .
   a. Bestäm momentgenererande funktionen för . (2p)
   b. Om , , …, är oberoende stokastiska variabler med samma fördelning som , vilken momentgenererande funktion får då = + +…+ ? (2p)
   c. Beräkna variansen för definierad som i b-delen med hjälp av momentgenererande funktionen för . (2p)

3. Äpplen som slumpmässigt väljs från ett stort importerat parti väger i genomsnitt (väntevärdet) **150** gram, och vikten varierar enligt en fördelning med en standardavvikelse som är **30** gram.
a. Du antar (approximerar med) att successivt valda äpplen har en vikt som är oberoende av varandra och följer en fördelning med dessa parametrar. Du väljer nu **33** äpplen. Vad är (approximativt) sannolikheten att dessa tillsammans väger minst **5** Kg? (2p)
b. Hur många äpplen behöver du minst ta för att sannolikheten för att de tillsammans skall väga minst **5** Kg, skall vara större än **0.95?** (approximera på lämpligt sätt) (2p)

4. I ett normalfördelningsstickprov med **10** observationer har någon beräknat stickprovsmedelvärdet och stickprovsstandardavvikelsen till **16.35** och **0.47**.
   a. Du ombeds att förvandla informationen till ett konfidensintervall för det bakomliggande väntevärdet **μ** för de enskilda observationerna med konfidensgraden **99%**. (2p)

b. Du ombeds testa hypotesen **H_0: väntevärdet=15** med signifikansnivån **5%** mot Hypotesen **H_1**: väntevärdet **> 15.** Vad blir din slutsats? (2p)

5. Du observerar en Poissonprocess under **10** timmar till **x(10)=58** händelsetidpunkter i processen under observationstiden.
   a. Bestäm en observerad Maxumum Likelihood - punktskattning av intensiteten **c** (uttryckt i pulser per timme) i Poissonprocessen baserad på **x(10)**. (2p)
   b. Vad är teoretiska standardfelet i skattningen ovan? (1p)
   c. Vad blir en lämplig observerad punktskattning av det teoretiska standardfelet i b-uppgiften ut? (1p)

6. Redogör för hypotesprövningsbegreppen:
   a. **Nollhypotes**. (1p)
   b. **Mothypotes**. (1p)
   c. **p-värde**. (1p)
7. Låt **Y** vara maximum av **10** oberoende likformigt fördelade stokastiska variabler på intervallet **[0,5]**.
   a. Bestäm fördelningsfunktionen och sannolikhetstätheten för **Y**. (3p)
   b. Bestäm väntevärdet för **Y**. (2p)
8. I en standard linjär regressionsmodell med inställningsvariabler **x** och svarsvariabler **Y** beror variansen av skattade linjens läge i en fix punkt, säg punkten **x_0**, på aritmetiska medelvärdet och "stickprovs"-variansen hos **x**-variablerna.
   a. Hur ser beroendet av dessa ut? (1p)
   b. Om du vill ha hög precision i skattningen av linjens läge i punkten **x_0=10**, hur skall du då välja dina **x**-inställningar i försöket. Diskutera. (2p)
   c. Vad kan det finnas för praktiska risker med att sprida isär x-inställningarna väldigt mycket? (1p)

LÖSNINGAR TILL TMA 321 MATEMATISK STATISTIK F/KF
2012-08-24 . OLLE NERMAN

---

1. a. ANTAL MÖJLIGA = 36 , ANTAL GYNNSAMMA = 6 KLASSISK S. DEF $\Rightarrow$ $P =$

b. A = POÄNGTALEN LIKA , B = SUMMAN ÄR 10 , SÖKT $P(A|B) = P(A \cap B)/P(B) =$
$= (1/36)/(3/36) = \frac{1}{3}$

c. UTFALLSRUMMET FÖR $Y = \{$MÖJLIGA UTFALL$\} = \{0,1,2,3,4,5\}$

d. $E[Y] = \sum_{y=0}^{5} y\, p_Y(y) = 1 \cdot \frac{10}{36} + 2\frac{8}{36} + \ldots + 5\frac{2}{36} = \frac{70}{36} = \frac{35}{18}$

---

2. a. $M_X(t) = E[e^{tX}] = (1-p) + pe^t$

b. OBER. $\Rightarrow$ $M_Y(t) = M_{X_1}(t) M_{X_2}(t) \ldots M_{X_{10}}(t) = ((1-p) + pe^t)^{10}$

c. $M_Y'(t) = 10pe^t((1-p)+pe^t)^9$ OCH $M_Y''(t) = 10 \cdot 9\, p^2 ((1-p)+pe^t)^8 e^t + 10p(1-$

$\Longrightarrow E[Y] = M_Y'(0) = 10p$ OCH $E[Y^2] = M_Y''(0) = 10 \cdot 9 \cdot p^2$

$\therefore$ $VAR[Y] = E[Y^2] - (E[Y])^2 = 90p^2 + 10p - 100p^2 = 10p(1-p)$

---

3. a. CENTRALA GRÄNSVÄRDESSATSEN $\Rightarrow$ VIKTSUMMAN $X \approx N(33 \cdot 150; 33 \cdot 30$
(I GRAM). SÖKT SANNOLIKHET ÄR

$$P(X > 5000) = 1 - P(X \le 5000) \approx$$
$$1 - \Phi\left(\frac{5000 - 4950}{30 \cdot \sqrt{33}}\right) = 1 - \Phi(0{,}29) \approx 0{,}386$$

b. Y_n SUMMAN AV VIKTERNA OM VI TAR n ÄPPLEN $\Rightarrow$

$P(Y_n > 5000) \ge 0{,}95$ FÖR n STÖRRE ÄN LÖSNINGEN TILL

$$1 - \Phi\left(\frac{5000 - n \cdot 150}{\sqrt{n}\, 30}\right) = 0{,}05$$

$\Longleftrightarrow$

$$\Phi\left(\frac{n \cdot 150 - 5000}{\sqrt{n}\, 30}\right) = 0{,}95$$

$\Longleftrightarrow$ $$\frac{n \cdot 150 - 5000}{\sqrt{n}\, 30} = 1{,}65$$

LÖSNING GER ETT n STRAX ÖVER 35
$\therefore$ 36 ÄPPLEN BEHÖVS!

---

4. a. t-INTERVALL (SYMMETRISKT). FRIHETSGRADTAL = 9
GER

$$\mu = \bar{x} \pm 3{,}25 \cdot \frac{s_x}{\sqrt{10}} = 16{,}35 \pm \frac{3{,}25 \cdot 6{,}47}{\sqrt{10}} \quad (99\%)$$

TY $F_{t(9)}(3{,}25) = 0{,}995$ UR TABELL.

b. FÖRKASTA OM $T = \sqrt{n}\,\frac{\bar{X} - \mu_0}{S} \ge c$, DÄR $F_{t(9)}(c) = 0{,}95$
$\mu_0 = 15$, $n = 10$. OBSERVERAT $t \approx 9{,}1$ OCH UR TABELL $c \approx 1{,}83$
$\Rightarrow$ VI FÖRKASTAR $H_0: \mu = \mu_0 = 15$ MOT $H_1: \mu > 15$

(LÖSN TMA [illegible])

FÖRB

5. a. $P_{X(10)}(k) = \frac{(10c)^k e^{-10c}}{k!} \qquad k = 0, 1, 2, \ldots$

$L(c) = \quad - \| -$

$\frac{d \ln L(c)}{dc} = 0 \iff \frac{d}{dc}\left(k \cdot \ln(10c) - 10c - \ln k!\right) = 0$

$\iff \frac{k}{c} - 10 = 0 \Rightarrow c = \frac{k}{10}$;

ML-ESTIMATOR (TEOR. PUNKTSKATTNING) $\hat{c} = \frac{X(10)}{10}$

OBSERVERAD P.S. $\hat{c} = 5{,}8$

b. STANDARDFELET ÄR $\sqrt{VAR[\hat{c}]} = \sqrt{VAR\left[\frac{X(10)}{10}\right]}$

$= \sqrt{\frac{VAR(X(10))}{100}} = \sqrt{\frac{c \cdot 10}{100}} = \sqrt{\frac{c}{10}}$

c. UPPSKATTAT STANDARDFEL ÄR $\sqrt{\frac{\hat{c}}{10}} = \sqrt{0{,}58}$

6. SE KURSBOKEN

7. a. $F_Y(t) = F_X(t)^{10} = \begin{cases} 0 & \text{om } t \le 0 \\ (t/5)^{10} & \text{om } 0 < t < 5 \\ 1 & \text{om } t \ge 5 \end{cases}$

(HÄR BETECKNAR X EN LIKF. FÖRD PÅ INT. $[0,5]$.)

$f_Y(t) = F_Y'(t) = \begin{cases} 10 \cdot \left(\frac{t}{5}\right)^9 & 0 < t < 5 \\ 0 & \text{F.Ö.} \end{cases}$

b. $E[Y] = \int_0^5 t \cdot 10 \left(\frac{t}{5}\right)^9 dt = \ldots = \frac{10 \cdot 5}{11} = 4\frac{6}{11}$

8 a. MED b_0 INTERCEPTET OCH b_1 LUTNINGEN SÅ SKATTAR $b_0 + b_1 x_0$ MED $\hat{b}_0 + \hat{b}_1 x_0$, DÄR $\hat{b}_0$ OCH $\hat{b}_1$ ÄR MINSTAKVADRATSKATTNINGARNA

$VAR[\hat{b}_0 + \hat{b}_1 x_0] = \sigma^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1} (x_i - \bar{x})^2}\right]$

b. OM MAN SÄTTER $\bar{x} = x_0$ SÅ MINIMERAS UTTRYCKET! (OM DETTA INTE HELT MÖJLIGT AV PRAKTISKA SKÄL SÅ SKALL MAN HÅLLA $\bar{x}$ NÄRA x_0 OCH SAMTIDIGT SPRIDA UT x_i-VÄRDENA)

c. MAN FÅR INGEN KOLL PÅ MODELLANTAGANDET ATT MAN HAR EN LINJÄR REGRESSIONSFUNKTION! (M.M)