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 27 Maj 2013, Förmiddag 8.30-12.30.
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 (max 4 totalt).
Betygsgränser: För betyget ”3” fordras minst 15 poäng, för betyget ”4” minst 20 poäng och för betyget ”5” minst 25 poäng .

---

1. Du kastar tre vanliga sex-sidiga tärningar en gång.
   a. Vad är sannolikheten att poängtalen på alla tre tärningarna är lika? (1p)
   b. Låt **Y** vara definierad som summan av de tre poängtalen. Vad är då utfallsrummet för **Y**? (1p)
   c. Definiera **Y** som i c-uppgiften ovan. Vad är sannolikhetsfunktionen för **Y** i punkten **y=17**? (1p)
   d. Vad är den betingade sannolikheten att produkten av de tre poängtalen är lika med **3** givet att summan av poängtalen är **5**? (2p)
2. En likformigt fördelad stokastisk variabel på intervallet **[a,b]** har väntevärdet **0** och standardavvikelsen **1**. Bestäm med hjälp av denna information värdena på intervallgränserna **a** och **b**. (3p)
3. **Z** är livslängden på en apparat som består av två enheter **A** och **B** (d.v.s **Z=** tiden tills apparaten blir trasig första gången). Båda dessa enheter **A** respektive **B** har exponentialfördelade livslängder **X** respektive **Y** som kan antas oberoende av varandra och har väntevärden som är **2 år** respektive **3.5 år**.
   a. Antag att apparaten är trasig precis när någon av enheterna **A** och/eller **B** är trasig. Bestäm fördelningsfunktionen och sannolikhetstätheten för **Z**. (2p)
   b. Nu antar vi istället att apparaten är trasig först när båda enheterna **A** och **B** är trasiga. Bestäm under dessa nya förutsättningar fördelningsfunktionen och sannolikhetstätheten för **Z**. (2p)
4. På en cirkelskiva med enhetsradie och centrum i origo i ett ortogonalt koordinatsystem väljs en stokastisk punkt **(X,Y)** med likformig fördelning.
   a. Vad blir den tvådimensionella frekvensfunktionen för **(X,Y)**? (1p)
   b. Visa att kovariansen mellan **X** och **Y** är **0**. (2p)
   c. Visa att **X** och **Y** <u>inte</u> är oberoende stokastiska variabler. (2p)
5. I en Poissonprocess med okänd intensitet **c** pulser/tidsenhet observeras antalet ”pulser” **X** i ett tidsintervall av längden **20**. Då är $\hat{\mathbf{c}}$ **=X/20** både maximum likelihood-skattning och en momentskattning av intensitetsparametern **c**.
   a. Är $\hat{\mathbf{c}}$ väntevärdesriktig (unbiassed)? Motivera svaret! (1p)
   b. Vilket standardfel (standard error) har $\hat{\mathbf{c}}$? (1p)
   c. Ange en lämplig observerad skattning av sannolikheten att du i ett visst tidsintervall, som är disjunkt från det observerade intervallet och som har längden **3** tidsenheter, inte observerar någon ”puls” alls om du observerat att **x=12** ? (2p)

Fortsättning på baksidan!

6. Om tre händelser **A**, **B** och **C** är oberoende av varandra och sannolikheterna är **P(A)=0.2**, **P(B)=0.5** och **P(C)=0,6**. Vad är då:
   a. Sannolikheten att ingen av dem inträffar? (1p)
   b. Väntevärdet av **X**, där **X** är antalet av de tre händelserna som inträffar? (1p)
   c. Variansen för **X** (definierad som i b)? (1p)
7. Du skall observera **10** oberoende stokastiska variabler $X_1, \ldots, X_{10}$ med samma sannolikhetstäthet som beror på en okänd positiv parameter   och har formen

$$\mathbf{p(x;\ \ ) = 0.5\ \ \ exp(-\ \ |x|)\ ,\ x\ \ \ R\ .}$$

   a. Härled Maximum Likelihood-skattningen av (3p)
   b. Ange den observerade punktskattningen baserad på observation av stickprovet: 1.2 , 1.8 , -3.3 , 5.2 , 4.9 , -2.0 , 0.6 , 0.9 , -8.5 , 1.5 . (1p)
8. En forskare du känner har två gånger genomfört en datainsamling av **n=11** respektive **m=15** observationer som alla kan anses vara observationer av normalfördelade stokastiska variabler med samma okända varianser men med olika okända väntevärden i båda stickproven. Han har sedan räknat ut ett t-intervall för väntevärdet μ_1 i det första stickprovet som blev:

$$\mu_1\ \ \ \mathbf{[12.3\ ,\ 15.8]\ \ \ (95\%)}\ .$$

Han hade också räknat ut ett likadant t-intervall för väntevärdet μ_2 i det andra stickprovet och resultatet blev:

$$\mu_2\ \ \ \mathbf{[6.4\ ,\ 11.1\ ]\ \ \ (95\%)}\ .$$

a. Nu skulle han vilja se hur ett observerat konfidensintervall för differensen $=\mu_2 - \mu_1$ mellan de båda väntevärdena med konfidensgraden **99%** baserat på alla **26** observationerna ser ut? Men han har tappat bort de båda stickprovens observationsvärden och inte heller sparat deras aritmetiska medelvärden och stickprovs-standardavvikelser. Han frågar därför dig (som har rykte om dig att kunna räkna bra) om du kan hjälpa honom på något sätt? Kan du det? (3p)
b. Eftersom du var så duktig på att hjälpa honom i a-uppgiften så undrar han också om du kan hjälpa honom att rekonstruera hur slutsatsen i testet skulle blivit om du använt ett tvåsidigt t-test, med signifikansnivån **1%,** för att testa nollhypotesen att de båda väntevärdena är lika, dvs **=0**. Med hjälp av enkel titt på resultatet i a svarar du snabbt…..? Ja vad blir svaret och vad var det för regel du använde (1p)

Lycka Till!

LÖSNINGSFÖRSLAG TMA 321 MATEMATISK STATISTIK 2013-05-27
AV OLLE NERMAN, MV, CHALMERS TEKNISKA HÖGSKOLA

1) a. ANTAL MÖJLIGA FALL $m = 6^3 = 216$, ANTAL GYNNSAMMA FALL $g = 6$ $\Rightarrow$ P(ALLA TÄRN. VISAR LIKA) $= \frac{6}{216} = \frac{1}{36}$

b. $\{$MÖJLIGA UTFALL PÅ $Y\} = \{3, 4, 5, \ldots, 18\}$

c. $\{Y=17\} = \{(5,6,6), (6,5,6), (6,6,5)\}$ HAR TRE ELEMENT $\Rightarrow$
$p(y) = \frac{3}{216}$ FÖR $y = 17$.

d. $\{Y=5\} = \{(1,2,2), (2,1,2), (2,2,1), (3,1,1), (1,3,1), (1,1,3)\}$
$\{$UTFALL MED PRODUKTEN $= 3\} = \{(3,1,1), (1,3,1), (1,1,3)\}$ $\Rightarrow$

$$P(\text{PRODUKTEN} = 3 \mid Y=5) = \frac{3/216}{6/216} = \frac{3}{6} = \frac{1}{2}$$

2) $E[X] = \frac{a+b}{2}$ OCH $VAR[X] = \frac{(b-a)^2}{12}$, DÄR X BETECKNAR DEN STOK. VARIABELN, FRÅN VILLKOREN

$$\frac{a+b}{2} = 0 \qquad \sqrt{\frac{(b-a)^2}{12}} = 1$$

FÖLJER FÖRST $a = -b$ OCH SEDAN $\sqrt{\frac{(2b)^2}{12}} = 1$, D.V.S. ATT $b^2 = \pm\sqrt{3}$. EFTERSOM $a < b$ FÖLJER ATT $a = -\sqrt{3}$ OCH $b = \sqrt{3}$.

3. a. $\lambda_X = \frac{1}{2}$ OCH $\lambda_Y = \frac{2}{7}$. Z = MINIMUM AV TVÅ OBEROENDE EXPONENTIALFÖRDELADE VARIABLER ÄR LIVSLÄNGDEN I a-DELEN. DENNA VARIABEL Z ÄR OCKSÅ EXPONENTIALFÖRDELAD MED PARAMETER $\lambda_X + \lambda_Y$ TY:

$$F_Z(z) = P(Z \le z) = 1 - P(Z > z) = 1 - P(X > z) \cdot P(Y > z)$$
$$= 1 - e^{-\left(\frac{1}{2} + \frac{2}{7}\right)z}$$

$f_Z(z) = \left(\frac{1}{2} + \frac{2}{7}\right) e^{-\left(\frac{1}{2} + \frac{2}{7}\right)z}$ ÄR SANNOLIKHETSTÄTHETEN FÖR Z

b. NU ÄR LIVSLÄNGDEN Z' IS TÄLLET MAXIMUM AV X OCH Y

$$F_{Z'}(z) = P(\{X \le z\} \cap \{Y \le z\}) = F_X(z) \cdot F_Y(z)$$
$$= \left(1 - e^{-\frac{1}{2}z}\right)\left(1 - e^{-\frac{2}{7}z}\right)$$

$$\frac{d}{dz} F_{Z'}(z) = f_{Z'}(z) = \frac{1}{2} e^{-\frac{1}{2}z} + \frac{2}{7} e^{-\frac{2}{7}z} - \left(\frac{1}{2} + \frac{2}{7}\right) e^{-\left(\frac{1}{2} + \frac{2}{7}\right)z}$$

ÄR SANNOLIKHETSTÄTHETEN FÖR Z'

FORTS: LÖSNINGSFÖRSLAG TMA 321 MATEMATISK STATISTIK 2018-05-27
AV OLLE NERMAN, MV, CHALMERS TEKNISKA HÖGSKOLA

4. a. $f(x,y) = c$ OCH $\iint_{x^2+y^2\le 1} f(x,y)\,dx\,dy = 1 \Rightarrow c = \frac{1}{\pi}$ (TY π ÄR AREAN AV $\{x^2+y^2 \le 1\}$.

b. $KOV[X,Y] = \iint_{x^2+y^2\le 1} xy \frac{1}{\pi}\,dx\,dy \underset{\uparrow}{=} 0$ ($E[X]=E[Y]=0$ P.G.A SYMMETRIN)

T.EX. POLÄRA KOORDINATER

c. TÄTHETERNA FÖR X OCH Y ÄR $f_X(x) = \frac{2\sqrt{1-x^2}}{\pi}$, $-1\le x\le 1$ RESP. $f_Y(y) = \frac{2\sqrt{1-y^2}}{\pi}$ $-1 \le y \le 1$, MEN DET BETYDER T.EX. ATT

$f_Y(y) f_X(x) > 0 = f(x,y)$ NÄR BÅDE $|x|$ OCH $|y|$ $> \frac{1}{\sqrt{2}}$ OCH < 1.

SLUTSATSEN ÄR ATT X OCH Y ÄR BEROENDE

---

5 a. $E[X/20] = \frac{E[X]}{20} = \frac{20\cdot c}{20} = c$. ALLTSÅ ÄR $\hat{c} = \frac{X}{20}$ VÄNTEVÄRDESRIKTIG

b. $\sqrt{VAR[\hat{c}]} = \frac{\sqrt{VAR[X]}}{20} = \frac{\sqrt{20c}}{20} = \frac{c}{\sqrt{20}} \left(\approx \frac{\hat{c}}{\sqrt{20}}\right)$

TEORETISK S.E. — SKATTAD S.E.

c. $P(\text{INTERVALL AV LÄNGD 3 TOMT}) = \frac{(3c)^0 e^{-3c}}{0!} = e^{-3c}$

$\hat{c} = \frac{12}{20} \Rightarrow e^{-3c}$ SKATTAS AV $e^{-3\hat{c}} = e^{-9/5}$

---

6 a. $P(A^c \cap B^c \cap C^c) = P(A^c)\cdot P(B^c)\cdot P(C^c) = (1-0{,}2)\cdot(1-0{,}5)\cdot(1-0{,}6) = 0{,}16$

b. $X = X_A + X_B + X_C$ DÄR $X_A = \begin{cases} 0 & \text{OM } A^c \text{ INTR.} \\ 1 & \text{OM } A \text{ INTR.} \end{cases}$ O.S.V. $\Longrightarrow$

$E[X] = P(A)+P(B)+P(C) = 0{,}2+0{,}5+0{,}6 = 1{,}3$

c. MED BET. SOM I b-DELEN, X_A, X_B OCH X_C OBEROENDE $\Rightarrow$

$VAR[X] = VAR[X_A] + VAR[X_B] + VAR[X_C] = 0{,}2(1-0{,}2) + 0{,}5(1-0{,}5) + 0{,}6(1-0{,}6)$

$= 0{,}16 + 0{,}25 + 0{,}24 = 0{,}65$

---

7. a. $L(\theta) = \prod_{i=1}^{n} \left(0{,}5\,\theta e^{-\theta|x_i|}\right) = 0{,}5^n \cdot \theta^n \cdot e^{-\theta \sum_{i=1}^{n}|x_i|}$

$\ln L(\theta) = n \ln 0{,}5 + n \ln\theta - \theta \sum_{i=1}^{n}|x_i|$

$\frac{d \ln L(\theta)}{d\theta} = \frac{n}{\theta} - \sum_{i=1}^{n}|x_i| = 0 \Rightarrow \theta = \frac{n}{\left(\sum_{i=1}^{n}|x_i|\right)}$ $(n=10)$ $\hat{\theta} = \frac{n}{\sum_{i=1}^{n}|X_i|}$

b OBSERVERA $\sum_{i=1}^{10}|x_i| = 1{,}2+1{,}8+3{,}3+5{,}2+4{,}9+2{,}0+0{,}6+0{,}9+8{,}5+1{,}5 = 29{,}9$

$\Rightarrow$ OBSERVERAT $\hat{\theta} = \frac{10}{29{,}9}$

FORTS. LÖSNINGSFÖRSLAG TMA 321 MATEMATISK STATISTIK 2013-05-27
AV OLLE NERMAN, MV, CHALMERS TEKNISKA HÖGSKOLA

8. a. KALLA OBSERVATIONERNA I FÖRSTA STICKPROVET FÖR $X_1, \ldots, X_{11}$ OCH DE I ANDRA FÖR $Y_1, \ldots, Y_{15}$. UR DE TVÅ t-INTERVALLEN

$$\mu_1 = \bar{x} \pm 2{,}228 \frac{s_x}{\sqrt{11}} \quad (95\%)$$

OCH

$$\mu_2 = \bar{y} \pm 2{,}145 \frac{s_y}{\sqrt{15}} \quad (95\%)$$

(BYGGER PÅ t-FÖRDELNINGAR MED $11-1=10$ RESP $15-1=14$ FRIHETSGRADER) KAN VI RÄKNA UT ATT

$$\bar{x} = \frac{12{,}3+15{,}8}{2} = 14{,}05 \quad s_x = \frac{\sqrt{11}}{2{,}228}\left(\frac{15{,}8-12{,}3}{2}\right) = 2{,}6051$$

$$\bar{y} = \frac{6{,}4+11{,}1}{2} = 8{,}75 \quad s_y = \frac{\sqrt{15}}{2{,}145}\left(\frac{11{,}1-6{,}4}{2}\right) = 4{,}243$$

DETTA GER ATT

$$s_p = \sqrt{s_p^2} = \sqrt{\frac{(11-1)\, 2{,}6051^2 + (15-1)\cdot 4{,}243^2}{(11+15-2)}} = \sqrt{13{,}275} \approx 3{,}651$$

99% TVÅSTICKPROVS KONFIDENSINTERVALL FÖR Δ

$$\Delta = (\bar{y}-\bar{x}) \pm 2{,}797\, s_p \sqrt{\frac{1}{11}+\frac{1}{15}} \quad (99\%)$$

(BYGGER PÅ FRIHETSGRADSTAL I t-FÖRD. $= 24$.)

$\Longrightarrow$ OBSERVERAT KONFIDENSINTERVALL

$$\Delta = -5{,}3 \pm 4{,}053 \quad (99\%)$$

b. VI ANVÄNDER REGELN ATT FÖRKASTA $H_0: \Delta = 0$ OM OCH ENDAST OM $0 \notin$ KONFIDENSINTERVALLET ($\alpha = 0{,}01$ SVARAR MOT KONFIDENSGRAD 99%). EFTERSOM $0 \notin$ INTERVALLET I a-DELEN SÅ FÖRKASTAR VI H_0 (TILL FÖRMÅN FÖR ATT $\mu_X > \mu_Y$).