Tentamen i MSG100 Sannolikhetsteori 1, Göteborgs Universitet. Deltentamen 2, 7.5 hp.
Tid: Torsdagen den 28 Augusti 2014, kl. 8.30-12.30.
Examinator och jour: Olle Nerman, telefon 772 3565.
Hjälpmedel: Miniräknare, egen formelsamling (4 A4-sidor på 2 blad) och till skrivningen medhörande tabeller.
Betygsgränser: För betyget G fordras 12 poäng, för betyget VG 20 poäng.

---

1. Den empiriska fördelningsfunktionen för ett stickprov av **n=200** kontinuerligt fördelade stokastiska variabler observerades i punkten **x=3,3** till värdet **0,57**.
   a. Vad betyder utsagan ovan (d.v.s.vad har man gjort för att beräkna siffran **0,57**)? (1p)
   b. Beräkna ett symmetriskt observerat konfidensintervall med approximativ konfidensgrad **95%** för fördelningsfunktionens värde i punkten **x=3,3**. (2p)

2. Lite teori:
   a. Redogör för skillnaden mellan en teoretisk och en observerad punktskattning. (1p)
   b. Vad är förutsättningarna för att en punktskattnings teoretiska standardavvikelse kan anses vara ett direkt mått på skattningens precision, och vad brukar punktskattningen då kallas? (1p)
   c. Vad menas med att en följd av punktskattningar av en parameter baserade på större och större stickprov är konsistent? (1p)

3. Brownsk Rörelse.
   a. Redogör för vad som menas med en standard Brownsk rörelse (utan drift). (1p)
   b.Vilken fördelning har en standard Brownsk rörelse **{B(t)}** utan drift evaluerad i tidpunkten **t=8** ? (1p)
   c. Vilken fördelningsfunktion har maximum av **B(t)**, definierad som i b-delen, över tids-intervallet **[0,8]**? (Motivera svaret) (2p)
   d. Vilken fördelningsfunktion har minimum av **(B(t)**, definierad som i b-delen, över tids-intervallet **[0,8]**? (Motivera svaret) (1p)

4. I en Poissonprocess med okänd intensitet **c** pulser/tidsenhet observeras antalet ”pulser” **X** i ett tidsintervall av längden **40**. Då är $\hat{\mathbf{c}}$ **=X/40** både en maximum likelihood-skattning och en momentskattning av intensitetsparametern **c**.
   a. Är $\hat{\mathbf{c}}$ väntevärdesriktig (unbiassed)? (Motivera svaret!) (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 **5** tidsenheter, inte observerar någon ”puls” alls, om du observerat att **x=52** . (2p)
5. Du skall observera **10** oberoende stokastiska variabler $\mathbf{X_1,\dots,X_{10}}$ med samma sannolikhetstäthet som beror på en okänd positiv parameter $\boldsymbol{\theta}$ och har formen

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

   a. Härled Maximum Likelihood-skattningen av $\boldsymbol{\theta}$. (2p)

   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)

6. Hur kan man förvandla en sekvens av oberoende likformigt fördelade slumptal simulerade på intervallet **[0,1]** till

   a. en sekvens av oberoende standard normalfördelade stokastiska variabler (1p)
   b. en sekvens av oberoende exponentialfördelade slumptal med intensitetsparameter=**4**? (2p)
   c. en Poissonprocess med intensitetsparameter **7**? (1p)

7. Stickprovsvariansen är som bekant en väntevärdesriktig punktskattning av den teoretiska variansen. Detta medför att längden i kvadrat på konfidensintervall beräknade med t-fördelningsmetoden enkelt kan jämföras med (den förutbestämda) längden i kvadrat på motsvarande intervall när man känner den teoretiska variansen. Hur ser kvoten ut mellan väntevärdet av längden i kvadrat på ett enstickprovsbaserat symmetriskt t-intervall vid konfidensgrad **95%** vid stickprovsstorlek **20** och (den deterministiska) längdkvadraten för motsvarande normalfördelningsbaserade intervall när teoretiska variansen är känd? (3p)

8. I en linjär regression med svarsvariabler **y** och inställningsvariabler **x** blev medelvärdet av **y**-variablerna **13.4** och medelvärdet av **x**-variablerna var **0.37**. Skattningen av riktningskoeffecieten $\boldsymbol{\beta}$ blev **0.42**.

   a. Beräkna en observerad punktskattning av interceptet $\boldsymbol{\alpha}$**.** (1p)
   b. Beräkna en punktskattning av väntevärdet vid **x**-inställningen **7.5** . (1p)

9. 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 eventuellt med olika okända väntevärden i respektive stickprov. Han har sedan räknat ut ett t-intervall för väntevärdet $\boldsymbol{\mu_1}$ i det första stickprovet som blev:

$$\boldsymbol{\mu_1 \in [12.3\ ,\ 15.8]\quad (95\%)}\ .$$

Han hade också räknat ut ett likadant t-intervall för väntevärdet $\boldsymbol{\mu_2}$ i det andra stickprovet och resultatet blev:

$$\boldsymbol{\mu_2 \in [6.4\ ,\ 11.1\ ]\quad (95\%)}\ .$$

   a. Nu skulle han vilja se hur ett observerat konfidensintervall för differensen $\Delta=\boldsymbol{\mu_2 - \mu_1}$ 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 (= vad blir det för intervall)? (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 $\Delta=\mathbf{0}$. 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!