Tentamen i MSG100 Sannolikhetsteori 1, Göteborgs Universitet. Deltentamen 2, 7.5 hp. Tid: Onsdagen 2012-12-19 kl. 8.30-12.30. Examinator: Olle Nerman.

Jour: Olle Nerman, telefon: 031-7723565.

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. Vid observation av **25** oberoende stokastiska variabler $X_1, X_2, \ldots, X_{25}$ från en och samma exponentialfördelning, blev medelvärdet **2,8** . Skatta på lämpligt sätt:
   a. Intensitetsparametern **λ i** exponentialfördelningen . (1p)
   b. Väntevärdet i exponentialfördelningen. (1p)
   c. Teoretiska medianen i exponentialfördelningen. (1p)
   d. Sannolikheten för ett värde över **6** i exponentialfördelningen. (1p)
   e. Standardfelet i skattningen i b-uppgiften (diskutera förutsättningarna här). (1p)

2. Ett uppåt begränsat konfidensintervall med konfidensgrad **99%** för en riktningskoefficient **β** i en vanlig linjär regressionsmodell (där residualerna är normalfördelade och oberoende med väntevärden=**0** och standardavvikelser= **σ**) observerades till **(-∞ ; 0,58)**.
   Kan du från detta avgöra om nollhypotesen $H_0: \beta=0$ skall förkastas i ett ensidigt test som har signifikansnivån **1%** och mothypotesen $H_1: \beta<0$? (Motivera) (3p)

3. Den empiriska fördelningsfunktionen för ett stickprov av **n=200** kontinuerligt fördelade stokastiska variabler observerades i punkten **x=5,3** till värdet **0,67**.
   a. Vad betyder utsagan ovan (d.v.s. vad har man gjort för att räkna ut siffran **0,67**)? (1p)
   b. Beräkna ett symmetriskt observerat konfidensintervall med approximativ konfidensgrad **95%** för fördelningsfunktionens värde i punkten **x=5,3**. (2p)

4. Bensinåtgången i miljöbilar tenderar att bygga på en glädjekalkyl i den meningen att verkliga åtgången av bensin i vardagskörning oftast är högre än den som testbilarna hade när miljöbilsklassningen erhölls. En mindre dieseldriven bil av ett visst märke/modellkombination **A** och en liknande dieseldriven bil av märke/modellkombination **B** har båda uppgivit en dieselförbrukning som är på gränsen för miljöbilsklassningen. Du som jobbar i redaktionen för en viss biltidning vill hjälpa läsarna att välja mellan modellerna och skaffar **10** slumpvalda nya bilar av vardera slaget och låter **10** olika förare köra varsin bil samma sträcka två gånger, en gång var med en (unik) bil av vardera modellen och med oberoende slumpmässig tidsordning mellan **A** och **B** bilarna för de olika förarna. Skillnaden (med tecken) i drivmedelsförbrukningen per mil mellan **A**-bilen och **B**-bilen i varje förares två körningar registrerades och genomsnittet av de **10** differenserna blev uppmätt till **-3.6 ml** (=milliliter)och stickprovsstandardavvikelsen baserad på de enskilda skillnaderna, blev observerad till **s=2,1 ml**. Du antar att observationerna av enskilda differenser kommer från ett oberoende normalfördelningsstickprov.
   a. Betrakta först nollhypotesen att det inte är någon skillnad på de två bilmodellernas genomsnittsförbrukning och den tvåsidiga mothypotesen (alternativa hypotesen) att det är en skillnad. Genomför ett test med signifikansnivån **5%**. Vilken slutsats drar du? (2p)

b. Du misstänker istället, redan innan försöket, att **A**-bilarna drar mer bränsle än **B**-bilarna, och är ute efter att bevisa detta, men med en tuffare signifikansnivå på **1%**. Du använder därför istället ett ensidigt test. Hur ser då din mothypotes och ditt förkastelseområde ut? och vilken slutsats drar du av observationerna? (2p)

5. Betrakta en Poissonprocess **{X(t), t>0}** med intensiteten **3** (pulser/tidenhet).
   a. Vilken fördelning får **X(3)**= antalet pulser i intervallet **(0,3]** om du betingar med att antalet pulser i intervallet **(0,2]** är **4,** d.v.s. med att **X(2) = 4**? (2p)
   b. Med samma definitioner som i a-delen:
   vilken fördelning får **X(2)** betingat med att **X(3)=3**? (2p)

6. Du har ett oberoende stickprov $X_1, X_2, \ldots, X_n$ på en stokastisk variabel med standardavvikelse=σ (och godtyckligt okänt väntevärde μ).

   a. Vilket väntevärde har differensen $X_i - X_j$, där **i** och **j** är två olika tal, sådana att $1 \le i<j \le n$? (1p)
   b. Visa att summan av kvadraterna av alla parvisa differenser $(X_i - X_j)$, sådana att $1 \le i<j \le n$, har väntevärdet $n(n-1)\sigma^2$. (2p)
   c. Om du definierar **Y** som summan av kvadraterna av alla parvisa differenser mellan olika X_i och X_j sådana att $1 \le i<j \le n$ dividerat med $n(n-1)$, så blir alltså skattningen **Y** enligt utsagan i b-delen en väntevärdesriktig punktskattning av variansen σ^2. Visa att denna skattning sammanfaller med den vanliga stickprovsvariansen, d.v.s att $Y= S^2$! (c *är ganska svår!*) (2p)

7. Om man observerar en Brownsk rörelse **{X (t)}** för alla **t** i ett intervall, t. ex. för alla **t** i intervallet **[0,1]**, så kan man uppskatta variansparametern σ^2 med godtyckligt stor noggrannhet. Detta kan man förstå genom att använda att tillskotten i processen över successiva disjunkta intervall av fix längd **1/n** är ett stickprov för varje fixt heltal **n**. Vi betraktar specialfallet med intervallet **[0,1]** och definierar ett stickprov $Y_1, Y_2, \ldots, Y_n$ genom succesiva differenser:

   $$Y_1=X(1/n)-0,\ Y_2= X(2/n)-X(1/n), \ldots,\ Y_n=X(1)-X((n-1)/n)$$

   a. Vilken gemensam fördelning har de enskilda Y_i-variablerna? (1p)
   b. Vilken egenskap hos Brownsk rörelse gör att de är oberoende av varandra? (1p)
   c. Skatta (räcker med teoretiska skattningen) variansparametern σ^2 i den Brownska rörelsen med hjälp av Maximum Likelihood-metoden baserad på $Y_1, Y_2, \ldots, Y_n$ för fixt **n**. (1p)
   d. Visa att skattningen av variansparametern σ^2 från c-delen är väntevärdesriktig. (1p)
   e. Beräkna teoretiska standardfelet (medelfelet i bokens terminologi) av skattningen från c-delen (som funktion av **n**). (1p)
   f. Visa att skattningen är konsistent i den meningen att den konvergerar i sannolikhet mot den riktiga variansparametern σ^2 när intervallängderna **1/n** konvergerar mot **0** (d.v.s. när $n \to \infty$). (1p)

Lycka till!
Olle