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.
Måndagen den 26 Maj 2014, 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 poäng totalt).
Betygsgränser: För betyget ”3” fordras minst 15 poäng, för betyget ”4” minst 20 poäng och för betyget ”5” minst 25 poäng .

---

1. En stokastisk variabel **X** är normalfördelad med väntevärdesparameter **μ** och standardavvikelseparamaeter ***σ***. Bestäm båda parametrarna **μ** och ***σ*** (med god approximation) från vetskapen att sannolikheten **P(X>5)=0.5** och sannolikheten P**(X>10)=0.01** . (3p)

2. En tvådimensionell stokastisk variabel **(X,Y)** har en sannolikhetstäthet **f(x,y)** som är proportionell mot **g** definierad av att:
   g(x,y)=xy, för **x** ∈ **[0,1] och y** ∈ **[0,2]**
   (g(x,y)= **0** för övrigt.)

   Bestäm:
   a. Konstanten **c** som gör att **f(x,y)=cg(x,y)** sannolikhetstäthet. (1p)
   b. Korrelationen mellan **X** och **Y**. (2p)
   c. Kovariansen mellan **Z=XY** och **X.** (2p)

3. a. Beräkna momentgenererande funktionen **M(t)** för en stokastisk variabel **X** som har utfallsrummet **Ω**x = **{-1,0,2}** och sannolikhetsfunktionen **p(x)=1/3** för alla **x** ∈ **Ω**x. (2p)
   b. Använd resultatet i a-delen för att ange momentgenererande funktionen för aritmetiska medelvärdet av två oberoende stokastiska variabler med samma fördelning som **X** . (2p)

4. Betrakta en Poissonprocess N(t)med intensiteten **3** pulser per tidsenhet och definiera händelserna **A=**att det inte kommer någon puls i intervallet **[0,5]** respektive **B=** att det inte kommer någon puls i intervallet **[0,7]**:
   a. Uttryck **A** och **B** i termer av processen **N(t).** (1p)
   b. Vad är betingade sannolikheten för händelsen **A** givet **B**, **P(A|B)** ? (1p)
   c. Vad är betingade sannolikheten för B givet **A**, **P(B|A)**? (2p)

5. Du misstänker att en viss ”Croupier” (han som snurrar roulette-hjulet och sätter fart på kulan) är skicklig på att kasta så att sannolikheten för **0** är <u>lite mindre än den borde</u> (den skall vara **1/37** om han inte fuskar). Därför tänker du med stort tålamod hålla reda på hur många nollor som resulterar från **2000** Roulette-omgångar (hela semestern går nog tyvärr åt). Syftet är att försöka bevisa hans fusk med statistisk hypotesprövning.

VÄND!

a. Vad är den naturliga nollhypotesen för dig i denna situation och vilken fördelning har **X** = antalet nollor bland de **2000** omgångarna under antagande av denna för dig naturliga nollhypotesen? (1p)
b. Hur kan du approximera fördelningen för **X** under nollhypotesen? (1p)
c. Vilka utfall på **X**, stora, små (eller både stora och små), bör du välja som förkastelseområde vid ett formellt hypotestest? (Diskutera gärna) (1p)
d. Använd approximationsidéerna i b. (och resultatet i c) för att bestämma förkastelseområdet vid approximativa signifikansnivån $\boldsymbol{\alpha}$= **0.05** (2p)

6. I ett viss enkelt genetiskt korsningsförsök med **200** upprepningar, så säger teorin att varje enskild avkomma i korsningsförsöket skall ha **0**, **1** eller **2** kopior av en viss genetisk kod, med sannolikheter som är $\mathbf{(1-p)^2}$, $\mathbf{2p(1-p)}$ respektive $\mathbf{p^2}$. Här är **p** en okänd sannolikhet som vi vill bestämma. Dessutom gäller att alla försökets **200** ”avkommor” kan antas ha oberoende antal kopior av koden ifråga. Om du använder ett stickprov med antalet kopior $\mathbf{X_i}$ för alla försöken, i=1,2…,200. Vad blir då
   a. Momentskattningen av parametern **p**? (2p)
   b. Är skattninen i a-delen väntevärdesriktig? (motivera) (1p)
7. Du vet att visst slags mynt har vikter som följer en fördelning med väntevärdet 10.55 gram och standardavvikelse 0.2 gram. Du vill packa mynten i påsar om tusen i varje och får den strålande iden att först väga påsen och sedan slänga i mynt ett och ett, utan att räkna dem tills totalvikten överstiger påsens vikt + 10.5 kilo, då du direkt slutar.
   a. Vad är chansen att du får fler mynt än 1000 i påsen (approximativt)? (2p)
   b. Vad är chansen att du får fler mynt än 999 i påsen (approximativt)? (1p)
   c. Använd svaren i a och b delen för att approximera chansen att det blir exakt 1000 mynt i påsen. (1p)
8. I en linjär regressionsmodell av oberoende normalfördelade stokastiska variabler (**Y**) med avseende på inställningsvariabeln (**x**) antas väntevärdena på **Y**-variablerna följa det linjära sambandet **a+bx** för två okända reella parametrar **a** och **b**. Varianserna för Y-variablerna <u>är kända och alla antas vara</u> **=0.25** . Du observerar de tre **y**-variabler till **6.10**, **1.70**, respektive **3.85** vid **x**-inställningarna **10.0** , **4.0**, respektive **7.0** .

   a. Beräkna en observerad punktskattning av regressionslinjens värde vid inställningen **x=5.5** . (1p)
   b. Beräkna ett symmetriskt konfidensintervall för riktningskoeffecienten **b** med konfidensgraden **95 %**. (2p)
   c. Testa på signifikansnivån **5 %** nollhypotesen att **b=0** mot den tvåsidiga alternativa hypotesen att $\mathbf{b\neq 0}$. Vad blir din slutsats? (motivera). (1p)

Lycka Till!