Matematiska vetenskaper **Tentamen: 2016-01-04 kl** $14^{00} - 18^{00}$
Chalmers tekniska högskola **Matematisk statistik för K (TMA073 och TMA074)**

Lärare och Jour: David Bolin, telefon 070-2324231.

Tillåtna hjälpmedel: Formelsamling och valfri miniräknare med tömda minnen.

Korrekt, väl motiverad lösning ger poängen som är indikerad i parentes vid vardera uppgift. Totalt kan man få 40 poäng och betygsgränserna för betyg 3, 4 och 5 är 16, 24 och 32 poäng.

1. (a) I rollspel är det vanligt med tolvsidiga tärningar (så kallade T12or). En T12a har 1 till 12 på sidorna och alla sidor har samma sannolikhet att komma upp. Beräkna väntevärde och varians för utfallet av ett kast med en T12a. (2p)

   (b) Antag att du ska ta tåget till Stockholm och att det är 30% sannolikhet för att det kommer snöa vid avgångstiden. Antag också att det är 70% sannolikhet för att tåget blir försenat om det snöar, samt att det är 10% sannolikhet för att tåget blir försenat om det inte snöar. Antag att du kommer till stationen och ser att tåget är försenat, vad är sannolikheten för att det snöar? (3p)

2. En maskin tillverkar gem och klipper därför upp bitar av metalltråd vars längder blir oberoende normalfördelade med väntevärde 10cm och standardavvikelse 0.5cm. Om en trådbit avviker mer än 1cm från 10cm sorteras den bort.

   (a) Hur stor är sannolikheten för att en bit sorteras bort? (2p)

   (b) För varje gem som maskinen tillverkar klipper den upp trådbitar tills att den lyckas göra en med acceptabel längd. Vad är sannolikheten att den måste klippa till fler än två bitar innan den lyckas? (3p)

3. Miljöförvaltningen ska uppdatera sina instrument som används för att mäta halter av luftförorening och mäter därför halten PM2.5 (enhet $\mu g/m^3$) i ett prov med två olika mätutrustningar. Mätningarna ger följande resultat:

Mätutrustning 1:	9.9017	9.2693	8.9876	8.2544	8.7757
Mätutrustning 2:	9.7930	9.9830	9.5546	10.0427	9.5683

   Alla mätningar görs på samma prov, vars halt PM2.5 är 10 $\mu g/m^3$, och man kan anta att båda mätutrustningarna ger oberoende normalfördelade mätningar med samma varians.

   (a) Undersök om den förväntade halten skiljer sig signifikant mellan mätutrustningarna. (3p)

   (b) Undersök om den första mätutrustningen har ett systematiskt fel i sina mätningar. (2p)

4. I en bostadsrättsförening undersökte man energiförbrukningen hos de olika typerna av lägenheter i föreningen. Man valde ut sex lägenheter och fick följande resultat:

Antal rum :	1	1	2	2	3	3
Bostadsyta m^2 :	37	45	54	54	67	72
Energiförbrukning MWh :	9.8	10.3	10.1	11.2	12.0	12.3

   En rimlig modell är att energiförbrukningen (y_i) beror linjärt av bostadsytan (x_i), alltså att $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, där $\varepsilon_i \sim \mathsf{N}(0, \sigma^2)$ är oberoende.

   (a) Skatta parametrarna i modellen. (3p)

   (b) Ange ett 95% konfidensintervall för energiökningen per kvadratmeter. (1p)

   (c) Ange ett 95% konfidensintervall för grundförbrukningen β_0. (1p)

5. En restaurang anordnar under december månad julbord. Från tidigare år vet de att 90% av besökarna brukar äta julskinkan som serveras. Efter den första veckan tycker personalen att det verkar gå åt mindre skinka än vanligt och räknar därför under en dag hur många av besökarna som äter av skinkan. Resultatet är att av totalt 25 besökare så äter 20 av skinkan.

   (a) Antalet besökare som äter skinkan kan antas vara binomialfördelat där parametern p beskriver sannolikheten att en besökare äter av skinkan. Använd direktmetoden för testa om $p < 0.9$. (2p)

   (b) Enligt centrala gränsvärdessatsen (CGS) kan vi normalapproximera binomialfördelningen om n är tillräckligt stort ($np(1-p) > 10$). Vad säger CGS och varför kan vi använda den på binomialfördelningen? (2p)

   (c) Testa om $p < 0.9$ genom att använda ett test baserat på normalapproximation. (2p)

   Personalen på restaurangen misstänker att den eventuella nedgången beror på att besökarna har blivit mer miljömedvetna och de funderar därför på att byta till en kravmärkt julskinka. Dagen efter det första testet serverar de därför istället en kravmärkt skinka och räknar igen hur många av besökarna som äter av skinkan. Resultatet blir att 27 av totalt 30 besökare äter den kravmärkta skinkan.

   (d) Låt $\tilde{p}$ beteckna sannolikheten för att en besökare äter den kravmärkta skinkan och låt som tidigare p vara sannolikheten för att en besökare äter den vanliga skinkan. Vi kan skatta $p_d = \tilde{p} - p$ med $p_d^* = \tilde{p}^* - p^*$, där p^* samt $\tilde{p}^*$ är de vanliga skattarna för p och $\tilde{p}$. Vad är $\mathsf{E}(p_d^*)$ samt $\mathsf{V}(p_d^*)$? (2p)

   (e) Finns det någon anledning att tro att fler besökare väljer att äta av skinkan om den är kravmärkt? Använd ett 95% konfidensintervall för p_d för att undersöka påståendet, det går bra att normalapproximera här. (2p)

6. Antag att X är en kontinuerlig stokastisk variabel med täthetsfunktion

$$f(x) = c \cdot \begin{cases} 1 - \frac{x}{r} & 0 \leq x \leq r \\ 0 & \text{annars} \end{cases}$$

   där $r > 0$ är en parameter i fördelningen och c är en normeringskonstant.

   (a) Beräkna normeringskonstanten c och beräkna sedan $\mathsf{E}(X)$ och $\mathsf{V}(X)$. (4p)

   (b) Vad är fördelningsfunktionen för X? (2p)

   (c) Använd momentmetoden för att ta fram en skattare $\hat{r}$ av parametern r baserat på ett stickrov $x_1, \ldots, x_n$ från fördelningen. (2p)

   (d) Är momentskattaren $\hat{r}$ väntevärdesriktig? (2p)

Lycka till!