Matematiska vetenskaper
Chalmers tekniska högskola

Tentamen: 2018-08-31 kl $08^{30} - 12^{30}$
Matematisk statistik för K, TMA074

Lärare och Jour: David Bolin, telefon 772 53 75.

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. En maskin har tre komponenter, A, B och C. Komponent A går sönder med sannolikhet 0.02. Om komponent A redan har gått sönder kommer också komponent B gå sönder med sannolikhet 0.2. Sannolikheten att komponent B går sönder givet att komponent A är hel är 0.026. Komponent C går sönder sannolikhet 0.01, men om komponent A redan har gått sönder går den sönder med sannolikhet 0.1. Om både komponent A och B redan har gått sönder kommer också komponent C gå sönder med sannolikhet 0.5.

   (a) Vid en undersökning av maskinen finner man att komponent B är trasig, vad är sannolikheten att komponent A också är det? (2p)

   (b) Går komponent A och komponent B sönder oberoende av varandra? (1p)

   (c) Vad är sannolikheten att alla tre komponenterna går sönder i maskinen? (2p)

2. En tvådimensionell slumpvariabel (X, Y) har täthetsfunktion

$$f(x,y) = \begin{cases} \frac{12}{19}(1 + x^2 + xy) & \text{om } 0 \leq x \leq 1 \text{ och } 0 \leq y \leq 1, \\ 0 & \text{för övrigt.} \end{cases}$$

   (a) Beräkna den marginella täthetsfunktionen $f(x)$ för X. (2p)

   (a) Beräkna $\mathsf{P}(X > Y)$. (3p)

3. Kristin vill testa Lisebergs nya berg- och dalbana. En gång i minuten kommer en vagn som maximalt kan ta 18 personer, men alla sittplatser blir inte alltid tagna. Antag att antalet fyllda platser per vagn varierar enligt följande fördelning.

Antal fyllda platser	13	14	15	16	17	18
sannolikhet	0.01	0.04	0.05	0.1	0.3	0.5

   (a) Beräkna väntevärde och varians för antalet fyllda sittplatser i en vagn. (2p)

   (b) Antag att det står 1035 personer framför Kristin i kön, vad är sannolikheten att hon kommer behöva vänta mer än en timme på att få åka? (3p)

4. I en kurs införs ett nytt kemiexperiment som studenterna måste klara av. Av erfarenhet vet man att försöket kan misslyckas med en viss sannolikhet, oberoende av hur skickliga studenterna är.

   (a) Under första året experimentet ingår i kursen utför 62 studenter det oberoende av varandra. Vad är fördelningen för det totala antalet lyckade försök? (1p)

   (b) Av de 62 studenterna klarar 55 studenter försöket. Förklara hur momentmetoden fungerar i allmänhet, och använd metoden för att skatta den okända parametern i fördelningen från uppgift (a). (2p)

   (c) Nästa gång kursen ges utför 50 studenter försöket. Eftersom försöket är detsamma kan vi anta att parametern från (b) inte har förändrats till detta tillfälle. Använd resultatet från (b) för att skatta sannolikheten att åtminstone 45 studenter klarar experimentet. (2p)

5. Man vill bygga ett nytt vindkraftverk och gör därför under ett års tid mätningar en gång i veckan av den maximala dagliga vindhastigheten. Från dessa 52 mätningar, som kan antas vara oberoende, beräknas $\bar{x} = 6.58$, $\sum_{i=1}^{52} x_i^2 = 2848.5$ och $s^2 = 11.7$. Mätningarna utförs för att skatta hur stor del av tiden kraftverket kommer kunna producera el, eftersom den typ av vindkraftverk som ska byggas endast producerar el om blåser mellan 3m/s och 25m/s.

   Som ofta för mätningar av vindhastighet visar det sig att datan väl kan modelleras med en Rayleighfördelning. Denna fördelning har en parameter $a > 0$ och täthetsfunktion

$$f(x) = \begin{cases} \frac{x}{a} e^{-\frac{x^2}{2a}} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

   Om X är Rayleighfördelad gäller att $\mathsf{E}(X) = \sqrt{a\pi/2}, \mathsf{V}(X) = (2 - \pi/2)a$ och $\mathsf{V}(X^2) = 4a^2$.

   (a) Härled maximum likelihood-skattningen a^*_{ML} av parametern a. (3p)

   (b) Använd resultatet från (a) för att skatta sannolikheten att vindkraftverket kommer generera el en slumpvis vald dag. Om du inte lyckades lösa (a), antag att $a = 25$. (2p)

   (c) Är a^*_{ML} en väntevärdesriktig skattning av a? Om du inte lyckades lösa uppgift (a), beskriv hur man undersöker detta i allmänhet. (3p)

   (d) Beräkna ett konfidensintervall för a med approximativ konfidensgrad 95%. (2p)

6. William Sealy Gosset började 1899 arbeta för bryggeriet Guinness och jobbade bland annat med att undersöka vilken sorts korn som gav bäst skörd. I sina försök utvecklade Gosset den statistiska teorin som krävs för analys av små stickprov, och publicerade 1908 uppsatsen “The Probable error of a Mean” där han introducerade t-fördelningen. Eftersom Guinness inte ville avslöja att de använde avancerade statistiska metoder publicerade Gosset under namnet Student, och det är därför fördelningen kallas för Students t-fördelning.

   I tabellen nedan finns data från Gossets uppsats. Den visar resultat från ett försök med 11 kornsorter där det undersöktes ifall torkning av kornet innan det såddes gav en bättre skörd.

År	1899							1900			
Kornsort	1	2	3	4	5	6	7	8	9	10	11
Inte torkad	1903	1935	1910	2496	2108	1961	2060	1444	1612	1316	1511
Torkad	2009	1915	2011	2463	2180	1925	2122	1482	1542	1443	1535

   Datan som visas i tabellen är vikt (i enheten pund) korn per tunnland.

   (a) Förklara vilket sorts t-test som bör användas för att undersöka om torkning ger en ökad skörd, samt vilka antaganden som då krävs. Skriv också ner nollhypotes och mothypotes för testet. (2p)

   (b) Utför testet från (a) för att underöka om torkning ger en ökad skörd. Använd signifikansnivå 90%. (2p)

   (c) Som ses i tabellen gjordes de sju första försöken under år 1899, och de fyra sista under år 1900. En möjlighet är att torkning fungerade bättre under en av de två säsongerna (kanske vädret var lite annorlunda ett av åren så att torkning passade bättre då). Vi vill nu undersöka om detta var fallet. Börja med att testa om ett poolat t-test kan användas för denna undersökning. (3p)

   (d) Beroende på resultatet från (c), utför antingen ett poolat eller ett icke-poolat t-test för att undersöka om året hade en signifikant påverkan på den förvändade skillnaden som torkning gav. (3p)

Lycka till!