Tentamentsskrivning i **Matematisk Statistik TMA321**

Tid: den 3 juni, 2019
Hjälpmedel: Typgodkänd miniräknare, egenhändigt skriven formelsamling om två A4 fram och bak (dvs 4 sidor), samt utdelade tabeller.

Tentamen består av 8 frågor om sammanlagt 50 poäng. Preliminära betygsgränser är satta till:
betyg "3": 20 till 29 poäng
betyg "4": 30 till 39 poäng
betyg "5": 40 eller fler poäng.

OBS! Alla lösningar skall vara väl redovisade och motiverade. Talen är ej ordnade efter svårighetsgrad.

1. Låt X, Y, Z vara oberoende $Exp(\lambda)$-fördelade slumpvariabler. Låt K_1 vara en rektangel med sidlängderna X och Y, och låt K_2 vara en kvadrat med sidlängd Z.

   (a) Beräkna sannolikheten att omkretsen av K_2 är större än omkretsen av K_1. (2p)

   (b) Beräkna sannolikheten att omkretsen av K_1 överstiger λ. Beräkna även sannolikheten att omkretsen av K_2 överstiger λ. (2p)

   (c) Låt K_3 vara en rektangel med sidlängder X och $1/Y$. Beräkna täthetsfunktionen för arean av K_3. (2p)

2. I en population har 0.1% av alla genen g1 och 0.5% har genen g2. Båda dessa gener ökar risken för en specifik cancerform som vi här kallar JT's sjukdom. Sannolikheten att en person

   (i) Utan g1,g2 utvecklar JT's sjukdom är 1%.

   (ii) Med g1 men utan g2 utvecklar JT's sjukdom är 9%.

   (iii) Med g2 men utan g1 utvecklar JT's sjukdom är 4%.

   (iv) Med g1 och g2 utvecklar JT's sjukdom är 19%.

   Vi antar att individer i populationen bär på generna oberoende av varandra. Vi väljer ut en individ slumpmässigt (dvs likformigt i populationen).

   (a) Vad är sannolikheten att denna individ kommer utveckla JT's sjukdom? (2p)

   (b) Givet att individen utvecklade JT's sjukdom, vad är sannolikheten att hen bär på gen g1? (2p)

   (c) Hur stor andel av alla som utvecklar JT's sjukdom bär på både genen g1 och genen g2? (2p)

3. Låt $Z \sim N(0,1)$, dvs Z har en standardiserad normalfördelning.

   (a) Beräkna den momentgenererande funktionen för Z^2. (2p)

   (b) Låt $Z_1, \ldots, Z_n$ vara i.i.d. (dvs oberoende och likafördelade) $N(0,1)$. Beräkna den momentgenererande funktionen för $Z_1^2 + \cdots + Z_n^2$. (2p)

   (c) Låt N vara likformigt fördelad på mängden $0, 1, \ldots, n-1$ och låt

   $$Y := Z_1^2 + \ldots + Z_N^2.$$

   Bestäm den momentgenererande funktionen för Y. (2p)

4. Låt X vara en slumpvariabel med sannolikhetsfunktion

   $$\mathbb{P}(X = k) = (1 - e^{-\alpha})e^{-\alpha(k-1)} \text{ för } k = 1, 2, \ldots,$$

   där $\alpha > 0$.

   (a) Hitta momentskattaren (MME:n) för α. (3p)

   (b) Hitta maximum likelihoodskattaren (MLE:n) för α. (3p)

   (c) Antag att vi har fått följande dataserie av utfall från slumpvariabeln X :

   6 13 16 9 21 17 10

   Vad blir det numeriska värdet på skattarna ovan? (1p)

5. Lisa-Bert mäter halten av dioxiner i nio markprover. Lisa-Bert antar att data kommer från en normalfördelning med okända μ och σ^2. Hennes dataserie blev som följer:

prov nr:	1	2	3	4	5	6	7	8	9
halt:	11.1	11.5	10.1	8.5	9.3	8.9	12.4	9.4	10.7

   (a) Hitta ett 95% konfidensintervall för μ. (3p)

   (b) Hitta ett 99% konfidensintervall för σ^2. (3p)

6. Låt $X_1, X_2, \ldots, X_n$ vara i.i.d. $U[0, \theta]$ och låt $\hat{\theta} = \max(X_1, \ldots, X_n)$ vara en skattare för θ.

   (a) Bestäm täthetsfunktionen för slumpvariabeln $\hat{\theta}$. (2p)

   (b) Beräkna väntevärdet av $\hat{\theta}$. Är skattaren väntevärdesriktig? (2p)

   (c) Är $\hat{\theta}$ en konsistent skattare? (3p)

7. I Solna vill Gubb-Jan undersöka hur råttor absorberar substansen Q från födan de intar. Gubb-Jan gör därför en serie av kontrollerade experiment där råttorna får olika doser av substansen Q, och sedan tas biopsier på råttorna för att se hur mycket av ämnet som har tagits upp. Gubb-Jan fick följande mätserie:

Dos i maten:	30	40	50	60	70	80	90	100
Upptagen mängd:	4.87	6.09	6.38	5.89	6.99	6.37	8.48	6.85

Data kunde sammanfattas med $S_{xx} = 4200$, $S_{yy} = 7.51$ och $S_{xy} = 134.4$.

Gubb-Jan ansätter en linjär regressionsmodell $y = \beta_0 + \beta_1 x$ där x är dosen i maten och y upptagen mängd. Hjälp Gubb-Jan med följande uppgifter:

(a) Skatta β_0, β_1. (2p)

(b) Skapa ett 95% konfidensintervall för β_1 och testa huruvida $\beta_1 = 0$ på 95%-nivån. (2p)

(c) Ange förklaringsgraden och beräkna residualerna. Verkar Gubb-Jans modell rimlig? (2p)

8. Arne är ansvarig för inköp av brädor till ett större bygge. Han vet att brädorna generellt sätt håller hög kvalité, men då bygget är stort misstänker Arne att en del brädor ändå är defekta.

   Antag att sannolikheten att en bräda är defekt är 0.0001, och antag att bygget köper in 2 000 000 brädor totalt.

   Låt X vara antalet brädor (av de 2 000 000 som köptes in) som är defekta.

   (a) Vilken fördelning har X? Motivera noggrannt. (2p)

   (b) En del av bygget använder 20 000 av brädorna. Vad är sannolikheten att tre eller fler av dessa är defekta brädor? (2p)

   (c) Vad är sannolikheten att det totala antalet defekta brädor av de 2 000 000 som köps in överstiger 220? (2p)