LINKÖPINGS UNIVERSITET
Matematiska institutionen
Matematisk statistik

Utbildningskod: TAMS24
Modul: TEN1
Datum: 2019-08-23
Institution: MAI

Tentamen i statistik

TAMS24/TEN1 2019-08-23

Tillåtna hjälpmedel:

- en miniräknare (ingen dator);
- Formel- och tabellsamling i matematisk statistik (från MAI);
- Formel- och tabellsamling i matematisk statistik, TAMS65;
- TAMS24: Notations and Formulas (av Xiangfeng Yang).

Betygsgränser (tillräckliga): 8-11 poäng ger betyg 3; 11.5-14.5 ger betyg 4; 15-18 ger betyg 5. Dina lösningar måste vara fullständiga, välmotiverade, noggrant nedskrivna och avslutade med ett tydligt svar. Var försiktig med vad som är stokastiskt eller ej. Antaganden du gör måste vara explicita. Approximationer är tillåtna om rimliga och tydligt motiverade. Uppgifterna är i slumpmässig ordning.

Lösningar dyker upp på hemsidan någon timme efter skrivningens sluttid.

En gång för länge länge sedan, i ett land långt långt borta, så fanns en Flamingofarm med namnet Exodus. Farmen sköttes av herrarna Rick och Gary. När Exodus startades så köptes en större mängd flamingor från en kvinna boendes i träskmarkerna i Louisiana. Flamingorna flyttades till ett lämpligt habitat i i södra Florida. Målet med farmen var att exportera flamingor både till världens djurparker och till privata intressenter. Rick och Gary producerade stolt reklam där man visade upp sin kompetens i flamingoavling. Efter en inte allt för lång tid så fick de ett samtal från sin första intresserade klient. Men händelserna som följde var inte vad de förväntade sig...

1. Ett Danskt fikabrödsföretag — med ett hemligt recept på munkar — vill köpa en större mängd flamingor varje månad av någon outgrundlig anledning. De är beredda att betala en viss summa per kilogram flamingo, så tyngre flamingor renderar mer vinst. Gary tar ut ett stickprov av flamingor och mäter varje individs vikt:

$$2.69 \quad 2.90 \quad 3.23 \quad 3.52 \quad 2.65 \quad 3.71 \quad 3.46 \quad 3.05$$

Antag att stickprovet kommer från en normalfördelning med varians 0.0625 och okänt väntevärde μ. De olika mätningarna antas oberoende.

   (a) Testa hypotesen $H_0 : \mu = 3.0$ mot $H_1 : \mu \neq 3.0$ på signifikansnivån 5%. (2p)

   (b) Vad är testets styrka vid $\mu = 3.1$? (1p)

   (c) Vilken är den högsta konfidensgrad som vi kan välja för att med just detta stickprov förkasta H_0? Är det ett rimligt förfarande att använda denna kalkyl när man väljer signifikansnivån? (2p)

2. När man använde resultatet från föregående övning för att avgöra vad man ska fakturera det danska företaget så blev resultatet inte riktigt som önskat. För att undvika ett ”faster disaster” så bestämde sig Gary och Rick för att inte anta att variansen är känd. Använd samma mätningar som i förra övningen och svara på följande frågor.

   (a) Testa antagandet att variansen är 0.0625 mot mothypotesen att variansen i själva verket är större. Använd signifikansnivån 0.05. (2p)

   (b) Antag att variansen är okänd. Finn ett konfidensintervall för μ med konfidensgraden 99%. (1p)

3. Det visade sig att det fanns lite finstilt text där det stod att alla fåglar skulle slaktas innan de skickades. Uppenbarligen förtvivlade så satte Rick och Gary igång en process för att så humant som möjligt avliva fåglarna medelst FLAMINGO DECAPITATOR 2000™ (en 'shovel headed killing machine'). Bladen är väldigt vassa, men behöver vässas efter en viss tid gått för att behålla effektiviteten ('the strike of the beast').

   Företaget *metal command* som säljer bladen påstår att tiden det tar innan det är nödvändigt med vässning (under antagande om en viss användning) är exponentialfördelad med väntevärde 1.0 dagar. Exodusfarmen bestämmer sig för att testa detta genom att använda 50 identiska maskiner parallellt (och på samma sätt) under 2.5 dagar. De noterar vid varje 6-timmars-intervall hur många maskiner som man behövt ta ur bruk för vässning (dessa maskiner hålls sedan ur bruk för att inte påverka resultatet).

Tid (timmar)	< 6	< 12	< 18	< 24	< 30	< 36	< 42	< 48	< 54	< 60
Frekvens:	11	20	26	32	36	39	42	44	46	47

   Använd ett lämpligt test med signifikansnivå 10% för att se om vi kan förkasta hypotesen att tiden är Exp($\mu = 1.0$)-fördelad. (2p)

4. Företaget *pleasures of the flesh* kontaktade Exodus och erbjöd sig att sälja tillväxthormon som är speciellt designat för fåglar av samma typ som flamingor. Företaget påstår att tillväxten hos fåglarna är linjärt beroende på mängden hormon som ges. Ett experiment för att undersöka rimligheten i påståendet genomfördes, men när Gary och Rick studerade residualplottar från en linjär regression så ser de tecken på något kvadratiskt.

Vikt (kg)	2.1344	2.3870	2.5861	2.8209	3.0492	3.2521	3.6765	4.0815
Hormon (mg/kg)	0.1250	0.2500	0.3750	0.5000	0.6250	0.7500	0.8750	1.0000

Betrakta följande två modeller:

$$\text{Modell 1:} \quad Y = \beta_0 + \beta_1 x + \epsilon$$

och

$$\text{Modell 2:} \quad Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon,$$

där $\epsilon \sim N(0, \sigma^2)$ och olika mätningar antas vara oberoende. Följande kalkyler har genomförts.

Modell 1:

i	$\widehat{\beta}_i$	$d(\widehat{\beta}_i)$
0	1.8036	0.0784
1	2.1242	0.1241

Variansanalys

	Frihetsgrader	Kvadratsumma
REGR	1	2.9610
RES	6	0.0607
TOT	7	3.0217

Modell 2:

i	$\widehat{\beta}_i$	$d(\widehat{\beta}_i)$
0	2.0456	0.0813
1	0.9627	0.3315
2	1.0324	0.2876

Variansanalys

	Frihetsgrader	Kvadratsumma
REGR	2	3.0047
RES	5	0.0170
TOT	7	3.0217

(a) Är termen i modell 2 som motsvarar x^2 meningsfull? Utför ett test på 5%-nivån. Hur tolkar du ditt resultat? (2p)

(b) Hitta ett 99% konfidensintervall för β_1 (använd modell 2). Skiljer det sig från motsvarande konfidensintervall för β_1 i modell 1? Hur tolkar vi detta resultat? (1p)

5. Exodus stöter på ett problem när en viss fiskart dyker upp i stort antal och konkurerar med fåglarna om maten. Den första planen för att lösa problemet baserade sig på en kemisk metod som kallas *Chemi-kill*, men på grund av oro för vad effekten blir på fåglarna så genomförde man inte detta. Som tur var så fick farmen besök av deras nära vän Susan som berättade att hon i en dröm hört någon berätta för henne om problemet och att lösningen var att introducera pirayor i habitatet.

   Gary kontaktar en Peruansk specialist Maria, som påstår att det finns två speciellt aggressiva typer av rödmagade pirayor. Rick och Gary importerar samma mängd av båda sorterna och utför följande experiment.

   Två likadana vattentankar fylls med 300 exemplar (i varje) av den problematiska fisksorten. Sedan följer en lektion i våld när samma mängd pirayor av de olika sorterna tillsätts tankarna (en sort i varje). Experimentet skulle kunna fortsätta tills det ingen kvar var, men man bestämmer sig för att avbryta efter 1 dygn och räkna resterande fiskar. I den första tanken var det 200 kvar och i den andra 180. Låt p_1 vara sannolikheten att en fisk blir uppäten i den första tanken och p_2 sannolikheten att en fisk blir uppäten i den andra.

   (a) Föreslå väntevärdesriktiga skattningar för p_1, p_2, and $p_2 - p_1$. Hitta sedan 95%-iga konfidensintervall p_1, p_2 and $p_2 - p_1$. Kan vi förkasta hypotesen om att de båda sorterna är lika aggressiva? (2p)

   (b) Efter ytterligare ett samtal till Maria så berättar hon att det är hennes erfarenhet att den andra sorten är lite mer aggressiv än den första. Vid signifikansnivån 5%, kan vi förkasta hypotesen om att båda sorterna är lika aggressiva om vi har mothypotesen att den andra sorten (motsvarande p_2) är mer aggressiv än den första? (1p)

6. På Exodusfarmen så planerar man ett test för att verifiera om en viss typ av situation föreligger. För att verkligen förstå testet så går Rick och Gary igenom beviset, men de fastnar vid följande del. Låt $\boldsymbol{p} = (p_1, p_2, \ldots, p_k)^T$ vara en sannolikhetsvektor, där $k \geq 2$ är ett heltal. Antag att $\boldsymbol{Y} = (Y_1\ Y_2\ \cdots\ Y_k)^T \sim N(\boldsymbol{0}, \boldsymbol{C})$, där

$$\boldsymbol{C} = \begin{pmatrix} 1-p_1 & -\sqrt{p_1p_2} & -\sqrt{p_1p_3} & \cdots & -\sqrt{p_1p_k} \\ -\sqrt{p_2p_1} & 1-p_2 & -\sqrt{p_2p_3} & \cdots & -\sqrt{p_2p_k} \\ -\sqrt{p_3p_1} & -\sqrt{p_3p_2} & 1-p_3 & \cdots & -\sqrt{p_3p_k} \\ \vdots & \vdots & & \ddots & \vdots \\ -\sqrt{p_kp_1} & -\sqrt{p_kp_2} & -\sqrt{p_kp_3} & \cdots & 1-p_k \end{pmatrix}.$$

   Det står i beviset att det nu följer att $\boldsymbol{Y}^T\boldsymbol{Y} \sim \chi^2(k-1)$. Bevisa detta! (2p)

Trevlig helg!