LINKÖPINGS UNIVERSITET
Matematiska institutionen
Matematisk statistik

Kurskod: TAMS24
Provkod: TEN1
Datum: 2019-01-09
Institution: MAI

Tentamen i statistik

TAMS24/TEN1 2019-01-09

Tillåtna hjälpmedel:

- en miniräknare (ingen computer);
- Formel- och tabellsamling i matematisk statistik (from MAI);
- Formel- och tabellsamling i matematisk statistik, TAMS65;
- TAMS24: Notations and Formulas (by 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. Uppgifterna är i nummerordning.

Lösningar dyker upp på hemsidan någon timme efter skrivningens sluttid.

1. En fullt rimlig fråga är vilket av Morbid Angels första 7 studioalbum som är bäst (vi ignorerar live-album och Abominations of Desolation-plattan[1]). Följande data samlades in från internet.

Albumtitel	Websida	
	Nuclear War Now!	Metalstorm.net
Altars of Madness	67	82
Blessed Are The Sick	18	34
Covenant	11	32
Domination	1	21
Formulas Fatal To The Flesh	6	4
Gateways to Annihilation	1	10
Heretic	0	2

Använd ett lämpligt test med signifikansnivå 0.01 för att undersöka om det finns en skillnad mellan åsikterna på de två sidorna. (2p)

2. Lina experimenterar med vattenkylning till sina datorer. Hon har mätt temperaturerna i processorerna i 5 olika datorer (de som överlevde experimentet). Först skedde mätningen med vanlig kylning och sedan efter byte till vattenkylning.

	Temperatur				
Dator:	C-1	C-2	C-3	C-4	C-5
Vanlig kylning	55	36	55	64	53
Vattenkylning	50	38	39	50	44

[1] Eftersom den givetvis är bäst.

Vi antar att temperaturer är normalfördelade och att olika datorer är oberoende.

(a) Finn ett 95%:igt konfidensintervall för de förväntade temperaturerna med vanlig kylning respektive med vattenkylning. (2p)

(b) Utför ett lämpligt test för att se om det finns en skillnad mellan de förväntade temperaturerna med vanlig kylning respektive vattenkylning. Använd nivån 5%. (2p)

(c) Hitta ett konfidensintervall $I_{\sigma^2} = [0, a)$ för variansen hos temperaturen när man använder vattenkylning. Använd konfidensgraden 90%. (1p)

3. Inom statistiken pratar man ofta om *stokastiska processer*. Ett exempel på en diskret sådan skulle kunna skrivas $X(n)$ för $n = 1, 2, 3, \ldots$, där $X(n)$ är en stokastisk variabel för varje n. Låt vår process $X(n)$ uppfylla följande: den har väntevärde 0 för alla n (d v s $\mathrm{E}(X(n)) = 0$) och om $Y(n)$ är den stokastiska vektorn $Y(n) = (\,X(n),\ X(n-1),\ X(n-2)\,)^T$, så ges kovariansmatrisen av

$$C_{Y(n)} = \mathrm{E}(Y(n)Y(n)^T) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

för alla $n = 3, 4, 5, \ldots$.

Finn en linjär prediktor $\widehat{X}(n) = aX(n-1) + bX(n-2)$ för $X(n)$ som minimerar kvadratfelet. Med andra ord, finn a och b så att $\mathrm{E}((\widehat{X}(n) - X(n))^2)$ blir minimal. (2p)

4. Crawford Tillinghast har byggt en maskin som möjliggör för personer att se och interagera med alternativa dimensioner. Den fungerar genom att stimulera tallkottkörteln i hjärnan med hjälp av resonansvågor. Under byggnationen så mätte Crawford frekvensen hos vågorna som funktion av spänningen och han noterade även om mätningen skett när det var natt (representerat av 0) eller dag (representerat av 1).

Frekvens (f)	1.86	2.41	3.26	3.88	4.64	5.50	6.40
Spänning (v)	1	2	3	4	5	6	7
Dag/Natt (u)	0	0	1	0	0	1	1

Crawford tror att frekvensen beror linjärt på spänningen, men han är inte säker på att tidpunkten för mätningen spelar någon roll (även om hans mest spektakulära resultat brukar uppstå på natten). Han betraktar två olika modeller:

$$\text{Modell 1:} \quad F = \beta_0 + \beta_1 v + \epsilon$$

and

$$\text{Modell 2:} \quad F = \beta_0 + \beta_1 v + \beta_2 u + \epsilon,$$

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

Modell 1:

i	$\widehat{\beta}_i$	$d(\widehat{\beta}_i)$
0	0.9671	0.0984
1	0.7564	0.0220

Variansanalys

	Frihetsgrader	Kvadratsumma
REGR	1	16.0212
RES	5	0.0678
TOT	6	16.0889

Modell 2:

i	$\widehat{\beta}_i$	$d(\widehat{\beta}_i)$
0	0.9866	0.0923
1	0.7370	0.0250
2	0.1363	0.1009

Variansanalys

	Frihetsgrader	Kvadratsumma
REGR	2	16.0424
RES	4	0.0466
TOT	6	16.0889

(a) Har termen i modell 2 som motsvarar dag/natt någon mening? Utför ett test på 1%-nivån. Vad är tolkningen av ditt resultat? (2p)

(b) Hitta ett konfidensintervall med konfidensgrad 95% för β_1 (använd modell 2). (1p)

5. Spelet dödsormsroulette (vanligare i australiensiska ödemarkerna) går ut på att personer i tur och ordning försöker klappa en giftig orm (typiskt en *death adder*) på huvudet. Spelet fortlöper tills dess att någon blir biten. Antag att sannolikheten att bli biten är konstant.

(a) Antag att en person blir biten vid omgång n. Hitta en rimlig punktskattning för p (där du använder n) och beräkna det förväntade värdet för din punktskattningsvariabel. Är din skattning väntevärdesriktig? (2p)

(b) En deltagare hävdar att ormen i fråga är riktigt irriterad, så $p = 0.4$. Vid ett spel så blir den femte personen som klappar ormen den första att bli biten. Testa hypotesen $H_0 : p = 0.4$ mot $H_1 : p < 0.4$ på signifikansnivån 5%. (1p)

(c) Vad blir styrkan för testet i $p = 0.2$? (1p)

6. Antag att $\boldsymbol{Y} = (Y_1\ Y_2\ \cdots\ Y_k)^T \sim N(\mathbf{0}, I_k)$, där I_k är identitetsmatrisen av ordning $k \times k$. Om $A \in \mathbf{R}^{k\times k}$ uppfyller att A är symmetriskt, A har rang l $(0 < l \leq k)$, och $A^2 = A$, härled fördelningen för $\boldsymbol{Y}^T A \boldsymbol{Y}$. (2p)