

Kurskod: TAMS65

Provkod: TEN1

MATEMATISK STATISTIK I FORTSÄTTNINGSKURS

Tentamen torsdagen den 21 mars 2019 kl 14–18

Hjälpmedel: Formelsamling i matematisk statistik utgiven av matematiska institutionen och/eller formelsamling ”Formel- och tabellsamling i matematisk statistik TAMS65 (Martin Singull)”. Inga anteckningar i formelsamlingarna är tillåtet. Miniräknare med tömda minnen. Språklexikon.

Betygsgränser: 8-11 poäng ger betyg 3, 11.5-14.5 betyg 4 och 15-18 poäng betyg 5.

Examinator: Martin Singull, Matematisk statistik, MAI

Resultatet meddelas *normalt* via LADOK inom 12 arbetsdagar.

Tydliga svar och motiveringar krävs till varje uppgift.

1. I en Amerikansk studie har man för en viss kurs undersökt kopplingen mellan studenter närvaro på undervisningen och tentamensresultatet och fått följande data

Antal missade undervisningstillfällen	Resultat på tentamen		
	≥ 80	70–79	< 70
Inget	18	11	6
1–5	14	12	6
Mer än 5	3	9	20

För varje frånvarokategori antar vi att studenterna har valts slumpmässigt ur en mycket stor population.

Undersök ed hjälp av ett lämpligt test på nivån 1% om närvaro på undervisningen verkar ha betydelse för tentamensresultatet. (2p)

2. I en maskin monteras ett verktyg i en hållare och man ”drar åt”. I ett försök utnyttjade man tre olika åtdragningsmoment mellan hållare och verktyg och i varje försök bestämda man x -koordinaterna för skärspetsen i verktyget. För varje åtdragningsmoment gjordes 21 mätningar. Resultat för x -koordinaterna (enhet: 10^{-4} mm):

Moment	\bar{x}_i	s_i
60	-20.167	4.1363
70	4.3000	3.1310
80	9.7667	4.1662

Modell: Vi har oberoende observationer från $N(\mu_i, \sigma_i)$, där $i = 1, 2, 3$.

- (a) Pröva på nivån 10% hypotesen

$$H_0 : \sigma_2 = \sigma_3 \quad \text{mot} \quad H_1 : \sigma_2 \neq \sigma_3.$$

(Egentligen ska även de andra standardavvikelseerna jämföras, men du behöver inte genomföra de testen.) (1p)

- (b) Anta nu att $\sigma_1 = \sigma_2 = \sigma_3 = \sigma$. Gör parvisa jämförelser mellan väntevärdena genom att konstruera lämpliga konfidensintervall, vart och ett med konfidensgraden 99%. Två åtdragningsmoment anses likvärdiga om $|\mu_i - \mu_k| < 10$. Är detta uppfyllt för något par? (2p)

3. Man har gjort oberoende mätningar av exekveringstider x_1, \dots, x_n (enhet: *ms*) för $n = 40$ körningar på en centraldator. Resultat:

10	19	90	40	15	11	32	17	4	152
23	13	36	101	2	14	2	23	34	15
27	1	57	17	3	30	50	4	62	48
9	11	20	13	38	54	46	12	5	26

där $\bar{x} = 29.65$.

Modell: De s.v. X_i är exponentialfördelade med väntevärde μ , dvs. med täthetsfunktionen

$$f(x) = \frac{1}{\mu} e^{-x/\mu}, \quad \text{för } x \geq 0.$$

- (a) Härled maximum-likelihood-skattningen av μ . (2p)
 (b) Visa att maximum-likelihood-skattningen är en väntevärdesriktig skattning av μ och beräkna dess varians. (1p)

4. För sjutton olika kärnkraftverk har man under ett år noterat antalet säkerhetsrelaterade fel som uppstått då reaktorn inte varit i drift (motsvarande värden för reaktorer i drift finns också). Reaktorerna är av två typer BWR = boiling water reactor, och PWR = pressurized water reactor. Observationer:

BWR	1	5	4	7	10	8	7	10	4
PWR	8	7	3	0	1	2	3	6	

Modell: Vi har två oberoende stickprov från $Po(\mu_1)$ respektive $Po(\mu_2)$.

Finns det någon skillnad mellan de båda reaktortyperna i fråga om felintensitet under ett år? Konstruera ett lämpligt konfidensintervall med konfidensgraden approximativt 95% och redovisa din slutsats. (3p)

5. Antag att vi har en observation x från den s.v. X med täthetsfunktionen

$$f(x) = \theta^2 e^{-\theta^2 x}, \quad 0 \leq x < \infty.$$

Vi vill testa hypotesen

$$H_0 : \theta = 1 \quad \text{mot} \quad H_1 : \theta > 1,$$

på nivån 5%. Beräkna det kritiska området. (2p)

Tips: Vill man förkasta H_0 till förmån för H_1 för stor eller liten observation x ?

6. Financial Times Stock Exchange 100 (FTSE-100) är ett index som omfattar de 100 största aktierna, mätt till marknadsvärde, på Londonbörsen. Man vill nu undersöka om man kan förklara avkastningen för FTSE-100 (y) med förklaringsvariablerna

- $x_1 =$ avkastningen på den brittiska obligationsmarknaden,
 $x_2 =$ avkastningen på S&P 500 (index inkluderar ett representativt urval av större ledande bolag på den amerikanska marknaden),
 $x_3 =$ avkastningen på växlingskursen mellan amerikanska dollar och brittiska pund.

Data har först analyserats enligt följande modell

$$\text{Modell 1: } Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon,$$

där $\varepsilon \sim N(0, \sigma)$ med följande resultat.

Skattad regressionslinje: $y = -0.34 - 0.06x_1 + 0.94x_2 - 0.03x_3$

i	$\hat{\beta}_i$	$d(\hat{\beta}_i)$
0	-0.3420	0.6810
1	-0.0613	0.2644
2	0.9390	0.1504
3	-0.0328	0.1908

Variansanalys:

	Frihetsgrader	Kvadratsumma
REGR	3	420.09
RES	24	252.14
TOT	27	672.23

- a) Undersök med lämpliga test eller konfidensintervall om var och en av förklaringsvariablerna gör nytta. Varje test ska vara på signifikansnivån 5%. (1p)

Data har sedan analyserats enligt följande modell

$$\text{Modell 2: } Y = \beta_0 + \beta_2 x_2 + \varepsilon,$$

där $\varepsilon \sim N(0, \sigma)$ med följande resultat.

Skattad regressionslinje: $y = -0.43 + 0.94x_2$

i	$\hat{\beta}_i$	$d(\hat{\beta}_i)$
0	-0.4258	0.5956
2	0.9359	0.1426

Variansanalys:

	Frihetsgrader	Kvadratsumma
REGR	1	419.14
RES	26	253.09
TOT	27	672.23

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 0.0364 & -0.0012 \\ -0.0012 & 0.0021 \end{pmatrix}$$

- b) Beräkna R^2 för de två modellerna. Hur stor är skillnaden? Varför är den så liten? (1p)
- c) Vilken avkastning kan vi förvänta oss för en genomsnittsdag för FTSE-100 när avkastningen för S&P 500 är 3%, dvs $x_2 = 3$. Bilda ett 95% konfidensintervall för lämplig storhet. (2p)
- d) Punktskatta skillnaden i avkastning för FTSE-100 för två dagar då avkastningen för S&P 500 är 2% och 4%. (1p)