

Kurskod: TAMS65
Provkod: TEN1/TEN2

MATEMATISK STATISTIK I FORTSÄTTNINGSKURS

Tentamen onsdagen den 21 augusti 2019 kl 8–12

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.

TEN1 Betygsgränser: 8-11 poäng ger betyg 3, 11.5-14.5 betyg 4 och 15-18 poäng betyg 5.

TEN2 Betygsgränser: 7-9 poäng ger betyg 3, 9.5-12 betyg 4 och 12.5-15 poäng betyg 5.

Examinator: Martin Singull, Matematisk statistik, MAI

Resultatet meddelas *normalt* via LADOK inom 15 arbetsdagar.

Tydliga svar och motiveringar krävs till varje uppgift.

1. För att undersöka om det finns en koppling mellan resultaten i ett traine-program och hur väl de anställda (som har genomfört traine-programmet) utför sina arbetsuppgifter gjorde ett storföretag följande undersökning.

		Betyg från traine-program		
		3	4	5
Betyg på arbetet	Bra	23	60	29
	Mycket bra	28	79	60

Pröva på 5%-nivån om det finns någon koppling mellan betyg på det utförda arbetet och betyget från traine-programmet. (2p)

2. En studie gjordes för att undersöka tiden det tar att köra en lastbil från A till B via tre olika vägar ($i = 1, 2, 3$) i en stor stad. Den observerade tiden ges nedan i minuter. Man bör påpeka att transportererna skedde ungefär under samma tidpunkt, dag som tid, och under samma väderförhållanden.

i	x_{ij}											\bar{x}_i	s_i
1	18.6	20.5	20.3	18.1	18.1	16.8	17.5	19.9	16.4	18.7	18.1	18.37	0.41
2	18.3	19.9	21.9	22.3	21.0	18.3	18.8	21.0	22.3	21.0	21.9	20.61	0.47
3	20.5	20.9	17.4	18.2	19.2	16.1	18.3	17.6	16.4	19.7	18.2	18.41	0.47

- (a) Har de tre olika stickproven samma standardavvikelse? Genomför lämpliga test på nivån 5%. Det räcker att du gör ett test men motivera hur du drar dina slutsatser. (1p)
 - (b) Bilda lämpliga konfidensintervall för att jämföra de olika rutterna. Den simultana konfidensgraden ska vara minst 94%. Finns det skillnader mellan rutterna? (2p)
3. Låt x_1, \dots, x_n vara n oberoende observationer från de stokastiska variablerna $X_k \sim \text{Exp}(k\theta)$, $k = 1, \dots, n$, där $E(X_k) = k\theta$.

- (a) Bestäm maximum-likelihood-skattningen av θ . (2p)
- (b) (Endast TEN2) Är skattningen i (a) en väntevärdesriktig skattning av θ ? Motivera ditt svar. (1p)
- (c) (Endast TEN2) Beräkna variansen för skattningen i (a). Är skattning en konsistent skattning av θ ? Motivera ditt svar. (1p)

4. Antag att x_1, \dots, x_{16} är ett stickprov på X som antas vara $N(\mu, \sigma)$ med $\sigma = 2$.

- (a) Man vill testa $H_0 : \mu = 1$ mot $H_1 : \mu \neq 1$ på nivån $\alpha = 5\%$. Beräkna de kritiska området, dvs, beräkna när du ska förkasta H_0 . (1p)
- (b) Beräkna och skissa styrkefunktionen $h(\mu)$ för testet i (a). Funktionen ska anges i fördelningsfunktionen $\Phi(\cdot)$. (2p)

5. I en provproduktion med $n = 200$ komponenter fann man 24 felaktiga komponenter.

- (a) Beräkna ett approximativt 95% konfidensintervall för felandelen p . (1.5p)
- (b) Man planerar att göra en ny provproduktion och vill då få ett approximativt 95% konfidensintervall med högst längden 8%. Hur många komponenter ska man då undersöka? (1.5p)

6. (Endast TEN1) I en serie försök studerades samband mellan bullerexponering och bullerkänslighet. Man uppmätte blodtrycksförhöjning y (enhet: mm Hg) i samband med olika bullernivåer x (enhet: dB). Resultat

Nivåer (dB)	x	60	63	65	70	70	70	80	80	80	80	85	88	90	90	90	90	94	100	100	100
Höjning (mm Hg)	y	1	0	1	1	2	5	2	3	4	6	5	4	4	5	6	8	7	7	8	9

Följande regressionsmodell

$$\text{Modell: } Y = \beta_0 + \beta_1 x + \varepsilon,$$

där $\varepsilon \sim N(0, \sigma)$ analyserades.

Skattad regressionslinje: $y = -10.7 + 0.18x$ och variansanalysen

i	$\hat{\beta}_i$	$d(\hat{\beta}_i)$
0	-10.7454	2.0613
1	0.1841	0.0248

Variansanalys:

	Frihetsgrader	Kvadratsumma
REGR	?	101.64
RES	?	33.16
TOT	19	134.80

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 2.30671 & -0.02744 \\ -0.02744 & 0.00033 \end{pmatrix}$$

- (a) Hur många frihetsgrader har kvadratsummorna för REGR och RES? (1p)
- (b) Plotta observationerna och skissa den skattade regressionslinjen. (1p)
- (c) Undersök med lämpliga test på eller konfidensintervall om bullernivåerna verkar påverka blodtrycksförhöjningen (dvs bullerkänsligheten). Använd signifikansnivån 5%. (1p)
- (d) Ett riktvärde för att god miljö kvalitet ska nås utanför bostäder är gränsvärdet 70 dB. Vilken blodtrycksförhöjningen kan vi förvänta oss i genomsnitt för en normalt frisk person? Bilda ett lämpligt 95% intervall som svar. (2p)