



Bokmål

Faglig kontakt under eksamen:
Håkon Tjelmeland 4822 1896

EKSAMEN I EMNE ST1201/ST6201 STATISTISKE METODER

Lørdag 3. desember 2011

Tid: 09:00–13:00

Hjelpemidler: Kalkulator CITIZEN SR-270X eller HP30S med tomt minne.
Statistiske tabeller og formler, Tapir forlag.
Ett gult ark (A5 med stempel) med egne håndskrevne formler og notater.

Sensur senest ferdig: 24. desember 2011.

Oppgave 1

En anleggsbedrift har undersøkt hvordan forventet opptak av fuktighet varierer mellom fire typer betong. I undersøkelsen benyttet bedriften seks prøver av hver av de fire betongtypene. Hver av de totalt 24 prøvene ble utsatt for fuktighet i 48 timer og det ble målt hvor mye fuktighet som ble tatt opp i prøvene. Dataene ble som følger:

Betongtype:	1	2	3	4
Y_{ij}	551	595	639	550
	457	580	615	449
	450	508	511	517
	731	583	573	438
	499	633	648	415
	632	517	677	555
$\sum_{i=1}^6 Y_{ij}$:	3320	3416	3663	2924
$S_j^2 = \frac{1}{5} \sum_{i=1}^6 (Y_{ij} - \bar{Y}_{.j})^2$:	12133.87	2302.67	3593.50	3704.27

En delvis utfylt variansanalysetabell (ANOVA-tabell) for disse målingene er som følger.

Kilde	df	SS	MS	F
Betong	*	*	15734.38	*
Error	*	108671.50	*	
Total	*	*		

- a) Skriv opp den fullstendige ANOVA-tabellen. Vis hvordan du beregner verdiene der det står * i den oppgitte tabellen.

I ANOVA-tabellen inngår det en testobservator F . Spesifiser hvilke hypoteser H_0 og H_1 denne testobservatoren relaterer seg til. Forklar spesielt hva eventuelle parametre du benytter i spesifikasjonen av H_0 og H_1 representerer i situasjonen beskrevet over.

Utfør hypotesetesten for signifikansnivå $\alpha = 5\%$ og konkluder.

- b) Angi modellen som to-utvalg t -test baserer seg på.

Utfør en to-utvalg t -test for å teste om det er grunnlag for å påstå at forventningsverdiene for opptatt fuktighet for betong av type 3 og 4 er ulike.

Sammenlign konklusjonene på de to hypotesetestene du har utført og kommenter.

Oppgave 2

La $Y \sim \text{bin}(m, p)$, der verdien til parameteren p er ukjent.

- a) Utled sannsynlighetsmaksimeringsestimator (SME) og momentestimator for p og vis at begge er gitt som

$$\hat{p} = \frac{Y}{m}.$$

Cramér-Raos ulikhet sier (med notasjon fra læreboka) at

$$\text{Var}[\hat{\theta}] \geq \left\{ -n \text{E} \left[\frac{\partial^2 \ln f_Y(Y; \theta)}{\partial \theta^2} \right] \right\}^{-1}.$$

- b) Hva er definisjonen av en *beste* estimator? Beskriv hvordan man kan benytte Cramér-Raos teorem til å vise at en gitt estimator $\hat{\theta}$ er en beste estimator for θ .

Vis at \hat{p} er en beste estimator for p .

Videre i oppgaven skal vi anta at det ikke er parameteren p , men parameteren $\theta = \text{Var}[Y] = mp(1-p)$, vi er interessert i å estimere. En mulig estimator for θ får vi ved å erstatte p med \hat{p} i uttrykket for θ , dvs.

$$\hat{\theta} = m\hat{p}(1 - \hat{p}) = Y \left(1 - \frac{Y}{m}\right).$$

c) Vis at $\hat{\theta}$ er forventningsskjev? Er $\hat{\theta}$ asymptotisk forventningsrett?

Kan du foreslå en korrigeret forventningsrett estimator for θ ?

Oppgave 3

I denne oppgaven skal vi se på en regresjonsmodell som er noe modifisert i forhold til den som er behandlet i læreboka. Anta at vi har variabelpar $(x_1, Y_1), \dots, (x_n, Y_n)$ der x_1, \dots, x_n ikke betraktes som stokastiske, mens Y_1, \dots, Y_n antas å være uavhengige stokastiske normalfordelte variabler med

$$E[Y_i] = \alpha + \beta(x_i - \bar{x}) \quad \text{og} \quad \text{Var}[Y_i] = \sigma_0^2.$$

Her er $\bar{x} = (1/n) \sum_{i=1}^n x_i$, verdiene til de to parametrene α og β antas ukjente, mens variansen σ_0^2 antas å ha en kjent verdi.

a) Utled sannsynlighetsmaksimeringsestimatorene (SME) for α og β og vis spesielt at estimatoren for β kan skrives på formen

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Vis at variansen til $\hat{\beta}$ kan skrives på formen

$$\text{Var}[\hat{\beta}] = \frac{\sigma_0^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Videre i oppgaven kan du (uten å utlede det) benytte at $\hat{\alpha}$ og $\hat{\beta}$ begge er forventningsrette og at $\text{Var}[\hat{\alpha}] = \sigma_0^2/n$.

b) Hvilken sannsynlighetsfordeling har $\hat{\beta}$? Begrunn svaret og angi eventuelle kjente egenskaper du benytter og forklar hvorfor de(n) gjelder i den aktuelle situasjonen.

Utled et $(1-a) \cdot 100\%$ konfidensintervall for β .

c) Utled et $(1-a) \cdot 100\%$ prediksjonsintervall for en ny observasjon Y_0 som skal observeres for $x = x_0$. Angi eventuelle kjente egenskaper du benytter og forklar hvorfor de(n) gjelder i den aktuelle situasjonen.

For hvilken verdi av x_0 blir prediksjonsintervallet kortest?