Noregs teknisk–
naturvitskaplege universitet
Institutt for matematiske fag

Nynorsk

Fagleg kontakt under eksamen:
Håkon Tjelmeland 4822 1896

EKSAMEN I EMNE ST1201/ST6201 STATISTISKE METODAR

Laurdag 3. desember 2011
Tid: 09:00–13:00

Hjelpemiddel: Kalkulator CITIZEN SR-270X eller HP30S med tomt minne.
Statistiske tabeller og formler, Tapir forlag.
Eitt gult ark (A5 med stempel) med eigne formar og notat.

Sensur er ferdig: 24. desember 2011.

Oppgåve 1

Eit anleggsføretak har undersøkt korleis forventa opptak av fukt varierer mellom fire typar betong. I undersøkinga nytta føretaket seks prøver av kvar av dei fire betongtypane. Kvar av dei totalt 24 prøvene vart utsette for fukt i 48 timar og det vart målt kor mykje fukt som vart teke opp i prøvene. Føretaket fekk følgjande resultat.

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
$S_j^2 = \frac{1}{5}\sum_{i=1}^{6}(Y_{ij} - \bar{Y}_{\cdot j})^2$:	12133.87	2302.67	3593.50	3704.27

Ein delvis utfylt variansanalysetabell (ANOVA-tabell) for desse målingane er som følgjer.

Kilde	df	SS	MS	F
Betong	$\star$	$\star$	15734.38	$\star$
Error	$\star$	108671.50	$\star$	
Total	$\star$	$\star$		

a) Skriv opp den fullstendige ANOVA-tabellen. Vis korleis du reknar ut verda der det står $\star$ i den oppgitte tabellen.

I ANOVA-tabellen inngår det ein testobservator F. Spesifiser kva for hypotesar H_0 og H_1 denne testobservatoren relaterer seg til. Forklar spesielt kva eventuelle parametrar du nyttar i spesifikasjonen av H_0 og H_1 representerer i situasjonen skildra over.

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

b) Angje modellen som to-utval t-test baserer seg på.

Utfør ein to-utval t-test for å teste om det er grunnlag for å påstå at forventningsverda for oppteken fukt i betong av type 3 og 4 er ulike.

Samanlikn konklusjonane på dei to hypotesetestene du har utført og kommenter.

Oppgåve 2

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

a) Utlei sannsynsmaksimeringsestimator (SME) og momentestimator for p og vis at begge er gjeve som

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

Cramér-Raos ulikhet seier (med notasjon frå læreboka) at

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

b) Kva er definisjonen av ein *beste* estimator? Forklar korleis ein kan nytte Cramér-Raos teorem til å vise at ein gjeven estimator $\widehat{\theta}$ er ein beste estimator for θ.

Vis at $\widehat{p}$ er ein beste estimator for p.

Vidare i oppgåva skal vi gå ut ifrå at det ikkje er parameteren p, men parameteren $\theta = \mathrm{Var}[Y] = mp(1-p)$, vi er interesserte i å estimere. Ein mogleg estimator for θ får vi ved å erstatte p med $\widehat{p}$ i uttrykket for θ, dvs.

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

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

Kan du foreslå ein korrigert forventningsrett estimator for θ?

Oppgåve 3

I denne oppgåva skal vi sjå på ein regresjonsmodell som er noko modifisert i tilhøve til han som er handsama i læreboka. Gå ut ifrå at vi har variabelpar $(x_1, Y_1), \ldots, (x_n, Y_n)$ der $x_1, \ldots, x_n$ ikkje betraktes som stokastiske, medan $Y_1, \ldots, Y_n$ blir anteke å vere uavhengige stokastiske normalfordelte variablar med

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

Her er $\bar{x} = (1/n)\sum_{i=1}^n x_i$, verda til dei to parametrane α og β blir anteke ukjende, medan variansen σ_0^2 blir anteke å ha eit kjent verd.

a) Utlei sannsynsmaksimeringsestimatorane (SME) for α og β og vis spesielt at estimatoren for β kan skrivast på forma

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

Vis at variansen til $\widehat{\beta}$ kan skrivast på forma

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

Vidare i oppgåva kan du (utan å utleie det) nytte at $\widehat{\alpha}$ og $\widehat{\beta}$ begge er forventningsrette og at $\mathrm{Var}[\widehat{\alpha}] = \sigma_0^2/n$.

b) Kva for ein sannsynsfordeling har $\widehat{\beta}$? Gje grunn for svaret og angje eventuelle kjende eigenskapar du nyttar og forklar kvifor dei gjeld i den aktuelle situasjonen.

Utlei eit $(1-a)\cdot 100\%$ konfidensintervall for β.

c) Utlei eit $(1-a)\cdot 100\%$ prediksjonsintervall for ein ny observasjon Y_0 som skal observerast for $x = x_0$. Angje spesielt eventuelle kjende eigenskapar du nyttar og forklar kvifor dei gjeld i den aktuelle situasjonen.

For kva for eit verd av x_0 blir prediksjonsintervallet kortast?