Høgskolen i Telemark
Avdeling for allmennvitenskaplige fag

EKSAMEN

I

STATISTIKK II

24.11.06

Tid:	12-16 (4 timer).
Fagkode:	5551.
Sidetall:	3.
Hjelpemidler:	Alle trykte og skrevne samt kalkulator.

OPPGAVE 1

Som et ledd i en undersøkelse om lønnsforhold blant nyutdannede kandidater er det trukket ut 10 nyutdannede økonomer med bachelorgrad. Vi lar $X_1, X_2, \ldots, X_{10}$ være årslønnene til økonomene som er trukket ut, og antar at disse er uavhengige kontinuerlige variabler. Sannsynlighetstettheten $f(x)$ er symmetrisk, og medianen i fordelingen er m.

a) I første omgang skal vi teste om medianen i lønnsfordelingen til de nyutdannede økonomene er over 300 000 kr. Formuler dette spørsmålet som en hypotesetest om m.

Hvilke to tester fra pensum kan brukes i denne situasjonen? Forklar kort hvordan de to aktuelle testmetodene virker. Hvilken av de to metodene ville du valgt når antakelser og observasjoner er som gitt i denne oppgaven?

b) Gjennomfør testingen med én av de to aktuelle testmetodene. Regn ut signifikans-sannsynligheten og angi konklusjonen når nivået velges lik 5 %. Observasjonene er gitt til slutt i oppgaven. Alle beløp er i hundre tusen kroner.

10 nyutdannede realister med bachelorgrad er også trukket ut, og deres årslønner er $Y_1, Y_2, \ldots, Y_{10}$. Observasjonene er gitt til slutt i oppgaven (i hundre tusen kroner).

c) Vi skal bruke Wilcoxons to-utvalgstest til å teste om lønnsnivået for nyutdannede økonomer er høyere enn for nyutdannede realister. Formuler dette spørsmålet som en hypotesetest. Angi forutsetningene for testen, gjennomfør testingen og angi konklusjonen. Bruk signifikansnivå 5%.

d) I dette punktet antar vi at årslønner for uttrukne nyutdannede kandidater er uavhengige og normalfordelte variabler. Bruk to-utvalgs t-test til å teste samme hypotese som i c). Angi konklusjonen på testen når signifikansnivået velges lik 5%.

Til slutt i oppgaven er det gitt en utskrift fra Minitab der det er gjennomført en én-faktor variansanalyse for å teste om det er forskjell i lønn mellom de to gruppene økonomer og realister. Sammenlign denne testen med t-testen ovenfor ved å peke på eventuelle likheter og forskjeller. (Du trenger ikke å redegjøre for selve variansanalysen).

Observasjoner

Årslønn økonomer i hundre tusen kroner (ordnet i stigende rekkefølge), X:

278 295 305 346 364 365 367 389 398 410

$\overline{X} = 351.7 \qquad \sum (X_i - \overline{X})^2 = 18356.1$

Årslønn realister i hundre tusen kroner (ordnet i stigende rekkefølge), Y:

247 248 265 290 300 347 352 354 357 382

$\overline{Y} = 314.2 \qquad \sum (Y_i - \overline{Y})^2 = 22623.6$

Variansanalyse (oppgave 1 d))

One-way ANOVA: Økonomer; Realister

```
Analysis of Variance
Source     DF        SS        MS        F        P
Factor      1      7031      7031     3,09    0,096
Error      18     40980      2277
Total      19     48011

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
Økonomer   10    351,70     45,16                (---------*----------)
Realiste   10    314,20     50,14  (----------*---------)
                                    ------+---------+---------+---------+
Pooled StDev =    47,71              300       330       360       390
```

OPPGAVE 2

En bilimportør vil undersøke i hvilken grad reklameinnsats i TV, aviser og radio virker inn på omsetningen. For 10 perioder har importøren satt opp sammenhørende verdier for omsetning (Y) og reklameinnsats i perioden i henholdsvis TV (X_1), aviser (X_2) og radio (X_3). Alle beløp er i hundre tusen kroner.

Vi skal bruke en multippel regresjonsmodell til å studere sammenhengen mellom omsetningen for bilimportøren og satsing på de 3 reklamemediene. Observasjoner og analyse går fram av Minitab-utskriften til slutt i oppgaven.

a) Gjør kort greie for antakelsene i modellen. Anta at disse er oppfylt i resten av oppgaven.

Skriv opp den estimerte sammenhengen mellom Y og de 3 forklaringsvariablene. Forklar kort hva de estimerte regresjonskoeffisientene uttrykker. Angi et 95% konfidensintervall for regresjonskoeffisienten for TV-reklame.

b) For hvilke medier gir økt satsing på reklame signifikant økning i omsetningen? Formuler selv aktuelle hypoteser og angi konklusjoner på testene. Bruk signifikansnivå 5 %.

Hvordan kan det henge sammen at estimert uttelling i omsetningen pr investert hundre tusen kroner er større for reklame i radio enn i aviser samtidig som avisreklame gir en mer signifikant (lavere signifikanssannsynlighet) økning i omsetningen enn radioreklame?

Utskrift fra Minitab

```
MTB > print c1-c4
```

Data Display

Row	Y	X1	X2	X3
1	3611	83	44	61
2	3440	63	42	49
3	3779	99	59	63
4	3715	94	33	61
5	3654	104	27	52
6	3645	90	35	51
7	3729	92	41	60
8	3794	106	48	64
9	3626	93	42	55
10	3875	105	60	59

```
MTB > Regress 'Y' 3 'X1'-'X3';
SUBC>   Constant;
SUBC>   Brief 2.
```

Regression Analysis: Y versus X1; X2; X3

```
The regression equation is
Y = 2662 + 6,73 X1 + 3,26 X2 + 4,51 X3
```

Predictor	Coef	SE Coef	T	P
Constant	2662,3	163,4	16,29	0,000
X1	6,727	1,344	5,01	0,002
X2	3,257	1,642	1,98	0,095
X3	4,507	3,703	1,22	0,269

```
S = 44,18       R-Sq = 91,1%     R-Sq(adj) = 86,6%
```

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	119453	39818	20,40	0,002
Residual Error	6	11711	1952		
Total	9	131164			