



**Høgskolen i Telemark**

Avdeling for allmenne fag

## EKSAMEN

### I

## 5551 STATISTIKK II

28.11.08

Tid: 9-13 (4 timer).  
Sidetall: 3.  
Hjelpemidler: Alle trykte og skrevne samt kalkulator.

---

### OPPGAVE 1

Som et ledd i en undersøkelse om lønnsforhold i store industribedrifter er det trukket ut 24 ansatte, 12 menn og 12 kvinner, og deres årslønn er registrert. Gjennomsnittlig årslønn for de 24 uttrukne var ca 426 tusen kr. Av de 12 mennene hadde 8 årslønn over gjennomsnittet og 4 under. Tilsvarende tall for kvinnene var 5 med årslønn over gjennomsnittet og 7 under.

- a) Sett opp resultatene i en  $2 \times 2$ -tabell. Bruk en kjiqvadrattest til å undersøke om en på grunnlag av disse resultatene kan påstå at det er forskjell i årslønn mellom menn og kvinner. Angi nullhypotese og alternativ hypotese for testen og definer testvariabel. Gjennomfør testingen og angi konklusjonen når signifikansnivået velges lik 5%.

### OPPGAVE 2

Situasjonen er som i oppgave 1. Vi lar  $X_1, X_2, \dots, X_{12}$  være årslønner for de 12 uttrukne mennene og antar at disse er normalfordelte med forventning  $\mu_1$  og standardavvik  $\sigma$ .

Vi lar  $Y_1, Y_2, \dots, Y_{12}$  være årslønner for de 12 uttrukne kvinnene, og vi antar at disse er normalfordelte med forventning  $\mu_2$  og standardavvik  $\sigma$ .

$X_1, X_2, \dots, X_{12}$  og  $Y_1, Y_2, \dots, Y_{12}$  er dessuten uavhengige variabler.

Vi definerer  $\delta = \mu_1 - \mu_2$

Resultatene er gitt til slutt i oppgaven.

- a) Estimer  $\sigma$ .  
Angi et 95% konfidensintervall for  $\delta$ .
- b) Har menn høyere lønn enn kvinner i store industribedrifter? Formuler dette spørsmålet som en hypotesetest om  $\delta$ . Gjennomfør testingen og angi konklusjonen. Bruk nivå 5%.

### Resultater

Årslønn menn (i tusen kroner),  $X$ :

469 375 478 385 447 358 510 558 445 535 422 444

$$\bar{X} = 452.2 \quad \sum (X_i - \bar{X})^2 = 42\,745.7$$

Årslønn kvinner (i tusen kroner),  $Y$ :

437 327 478 370 428 380 432 434 345 328 420 415

$$\bar{Y} = 399.5 \quad \sum (Y_i - \bar{Y})^2 = 25\,877$$

### OPPGAVE 3

Situasjonen er som i oppgave 1 og 2. For hver av de 22 uttrukne personene er det i tillegg til årslønnen registrert følgende variabler:

- $X_1$  = " Antall år i bedriften (= ansiennitet)"
- $X_2$  = " Antall år med høyere utdanning"
- $X_3$  = " Antall personer den ansatte er leder for "
- $X_4$  = " Kjønn (0 betyr kvinne og 1 mann)"

Vi skal bruke en multippel regresjonsmodell til å studere sammenhengen mellom årslønnen (i kroner), som vi nå betegner  $Y$ , og de 4 forklaringsvariablene  $X_1$ ,  $X_2$ ,  $X_3$  og  $X_4$

Merk at variablene  $Y$  og  $X_1$ ,  $X_2$ ,  $X_3$  og  $X_4$  har en annen betydning enn i oppgave 2!

Observasjonene og regresjonsanalysen går fram av Minitab-utskriften til slutt i oppgaven. Studer denne før du svarer på spørsmålene nedenfor.

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

Skriv opp den estimerte regresjonsmodellen og forklar kort hva de estimerte regresjonskoeffisientene uttrykker. Prediker årslønnen til en kvinne med 5 års ansiennitet, 4 års høyere utdanning som er leder for 7 ansatte.

Forklar hva  $R^2$  ( $R$ -sq i Minitab-utskriften) uttrykker.

- b) Hvilke av de 4 forklaringsvariablene har signifikant innvirkning på årslønnen? Formuler selv aktuelle hypoteser og angi konklusjoner på testene. Bruk signifikansnivå 5%.
- c) Angi et 95% konfidensintervall for  $\beta_4$ , som er regresjonskoeffisient tilhørende  $X_4$  (Kjønn). Kommenter kort spørsmålet om menn og kvinner har forskjellig årslønn med bakgrunn i testene i oppgave 1, 2b og 3b. Kan du gi en forklaring på eventuelle forskjeller i konklusjonene på testene?

## Utskrift fra Minitab

Results for: s2-08.mtw

### Data Display

Row	Y	X1	X2	X3	X4
1	469000	18	6	5	1
2	437000	15	6	4	0
3	375000	5	1	0	1
4	327000	6	0	1	0
5	478000	22	5	7	1
6	385000	3	2	6	1
7	447000	21	3	9	1
8	478000	18	3	5	0
9	370000	0	3	4	0
10	358000	0	5	0	1
11	510000	6	2	2	1
12	428000	3	4	3	0
13	380000	2	2	5	0
14	432000	9	1	2	0
15	558000	12	5	6	1
16	445000	6	2	2	1
17	434000	9	0	5	0
18	345000	0	0	2	0
19	535000	15	4	4	1
20	328000	0	0	2	0
21	422000	19	2	6	1
22	444000	4	5	8	1
23	420000	2	4	2	0
24	415000	6	1	3	0

### Regression Analysis: Y versus X1; X2; X3; X4

The regression equation is  
 $Y = 357736 + 3604 X1 + 8796 X2 + 993 X3 + 19742 X4$

Predictor	Coef	SE Coef	T	P
Constant	357736	20527	17,43	0,000
X1	3604	1727	2,09	0,051
X2	8796	5744	1,53	0,142
X3	993	5144	0,19	0,849
X4	19742	21575	0,92	0,372

S = 47055,3    R-Sq = 50,7%    R-Sq(adj) = 40,3%