



Høgskolen i Telemark

Avdeling for allmenne fag

EKSAMEN

I

STATISTIKK II

19.12.01

Tid:	9-13 (4 timer).
Fagkode:	037
Sidetall:	3.
Hjelpemidler:	Alle trykte og skrevne samt kalkulator.
Vedlegg:	Minitabutskrift (til oppgave 2).

OPPGAVE 1

Som et ledd i en undersøkelse om lønnsforhold i store industribedrifter er det trukket ut 22 ansatte, 12 menn og 10 kvinner, og deres årslønn er registrert. Årslønnene til de 22 uttrukne er gitt til slutt i oppgaven, ordnet i stigende rekkefølge for menn og for kvinner.

I første omgang skal vi teste om lønnsnivået for menn er høyere enn for kvinner i de store industribedriftene. Formuler dette problemet som en hypotesetest. Bruk Wilcoxons to-utvalgstest. Angi forutsetningene for testen, gjennomfør testingen og angi konklusjonen. Bruk signifikansnivå 5%.

Årslønn menn (kr):	278000	295000	305000	346000	364000	365000
	367000	389000	398000	430000	455000	478000
Årslønn kvinner (kr):	247000	248000	265000	290000	300000	346000
	352000	354000	357000	398000		

OPPGAVE 2

Situasjonen er som i oppgave 1. For de 22 uttrukne personene er det i tillegg til årslønnen Y 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 regresjonmodell til å studere sammenhengen mellom årslønnen Y og de 4 forklaringsvariablene X_1 , X_2 , X_3 og X_4 .

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%.

Kommenter spesielt hvordan innvirkningen fra X_4 (Kjønn) samsvarer med resultatet fra testen i oppgave 1. Kan du gi en forklaring på eventuell forskjell i konklusjon?

OPPGAVE 3

Det skal undersøkes om tilbakefallsprosenten for kriminelle er forskjellig for menn og kvinner.

Vi lar p_1 være sannsynligheten for at en mann som er tatt for en forbrytelse, begår en ny forbrytelse i løpet av de neste 5 årene. Tilsvarende sannsynlighet for en kvinne som er tatt for en forbrytelse, er p_2 .

Følgende hypotese er satt opp

$$H_0: p_1 = p_2 \quad \text{mot} \quad H_1: p_1 \neq p_2$$

Av 553 mannlige forbrytere som ble tatt i 1995, viste det seg at 249 av dem begikk en ny forbrytelse i løpet av de neste 5 årene. Av 97 kvinner som ble tatt i 1995, begikk 30 en ny forbrytelse i samme tidsrom.

Gjennomfør testing av hypotesen ovenfor. Definer selv variabler du bruker ved testingen, og angi hvilken fordeling de har. Forklar kort testmetoden og regn ut signifikanssannsynligheten for testen (tilnærmet). Angi konklusjonen når signifikansnivået velges lik 5%.

Angi også et tilnærmet 95% konfidensintervall for $p_1 - p_2$

Vedlegg til oppgave 2

19.11.01 08:01:37

Welcome to Minitab, press F1 for help.
Retrieving worksheet from file: M:\Minitab\s2-01.mtw
Worksheet was saved on 13.11.01 13:06:37

Results for: s2-01.mtw

MTB > print c1-c5

Data Display

Row	Y	X1	X2	X3	X4
1	389000	18	6	5	1
2	357000	15	6	4	0
3	295000	5	1	0	1
4	247000	6	0	1	0
5	398000	22	5	7	1
6	305000	3	2	6	1
7	367000	21	3	9	1
8	398000	18	3	5	0
9	290000	0	3	4	0
10	278000	0	5	0	1
11	430000	6	2	2	1
12	346000	3	4	3	0
13	300000	2	2	5	0
14	352000	9	1	2	0
15	478000	12	5	6	1
16	365000	6	2	2	1
17	354000	9	0	5	0
18	265000	0	0	2	0
19	455000	15	4	4	1
20	248000	0	0	2	0
21	346000	19	2	6	1
22	364000	4	5	8	1

MTB >

MTB >

MTB > regress c1 4 c2 c3 c4 c5

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

The regression equation is

$$Y = 272913 + 3720 X1 + 8154 X2 + 1282 X3 + 24563 X4$$

Predictor	Coef	SE Coef	T	P
Constant	272913	22375	12,20	0,000
X1	3720	1806	2,06	0,055
X2	8154	6285	1,30	0,212
X3	1282	5369	0,24	0,814
X4	24563	23319	1,05	0,307

S = 48810

R-Sq = 52,4%

R-Sq(adj) = 41,2%