



Høgskolen i Telemark

Avdeling for allmenne fag

EKSAMEN

I

5551 STATISTIKK II

2.12.09

Tid: 9-13 (4 timer).
Sidetall: 5.
Hjelpemidler: Alle trykte og skrevne samt kalkulator.

Oppgave 1

Et reisebyrå har i lengre tid arrangert bussturer i samarbeid med en turoperatør. Byrået har jevnlig målt kundenes tilfredshet ved å be dem fylle ut et spørreskjema som bl.a. inneholder følgende spørsmål: *Hvor fornøyd er du med turen?*

Ut fra erfaring gjennom flere år med samme turoperatør har reisebyrået satt opp følgende modell for kundenes valg av svaralternativer:

SVARALTERNATIV	PROSENTVIS SVAR
Svært fornøyd	20 %
Godt fornøyd	10 %
Fornøyd	50 %
Noe misfornøyd	15 %
Svært misfornøyd	5 %

For et halvt år siden skiftet reisebyrået til en annen turoperatør. Spørsmålet om *hvor fornøyd er du med turen* er stilt til $n = 500$ kunder som har reist med den nye turoperatøren.

Svarene fordelte seg slik:

SVARALTERNATIV	ANTALL SVAR
Svært fornøyd	127
Godt fornøyd	54
Fornøyd	223
Noe misfornøyd	62
Svært misfornøyd	34

- a) Tyder svarene fra kundene som har reist med den nye turoperatøren, på at fordelingen for de forskjellige svaralternativene er annerledes enn de var for den forrige turoperatøren? Formuler dette spørsmålet som en hypotesetest. Gjennomfør testingen og angi konklusjonen. Velg signifikansnivå 5%.

Dersom en slår sammen de 3 "fornøyd"-alternativene og de 2 "misfornøyd"-alternativene til henholdsvis Fornøyd og Misfornøyd, får en følgende fordeling på svaralternativene for den forrige turoperatøren:

SVARALTERNATIV	PROSENTVIS SVAR
Fornøyd	80 %
Misfornøyd	20 %

Svarene fra de $n = 500$ kundene som har reist med den nye turoperatøren, fordeler seg slik på de to sammenslåtte svaralternativene:

SVARALTERNATIV	ANTALL SVAR
Fornøyd	404
Misfornøyd	96

- b) Gjennomfør en tilsvarende test som i a) for de sammenslåtte svaralternativene. Bruk også her signifikansnivå 5 %.

Kommenter kort resultatene fra de to testene. Forsøk å beskrive eventuell endring i kundenes tilfredshet etter skifte av turoperatør.

Oppgave 2

Det er samlet inn data for salgspris Y (i 1000 kr) og boareal X (i m^2) for 25 eneboliger. Boligene ligger i samme område nær en stor by, og tomtene har omtrent samme størrelse.

Vi skal studere sammenhengen mellom salgspris og boareal ved hjelp av en enkel lineær regresjonsmodell

$$Y = \beta_0 + \beta_1 X + U$$

Observasjoner og resultater fra regresjonsanalysen er gitt i Minitab-utskriften til slutt i oppgaven.

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

Skriv opp den estimerte sammenhengen mellom salgspris og boareal. Forklar kort hva de estimerte verdiene av β_0 og β_1 uttrykker. Angi et 95% konfidensintervall for β_1 .

- b) For en valgt verdi av boarealet X kan en angi både et konfidensintervall og et prediksjonsintervall for salgsprisen. Forklar kort forskjellen på de to intervallene. Finn prediksjonsintervallet når $X = 150 \text{ m}^2$. Bruk sikkerhet 95%.

Resultater fra Minitab

Data Display

Row	X	Y
1	125	2510
2	150	3015
3	115	2240
4	135	2570
5	145	2680
6	200	3650
7	120	2290
8	145	2820
9	170	2890
10	165	2785
11	140	2535
12	125	2370
13	175	3090
14	190	3270
15	155	2710
16	130	2270
17	155	2890
18	160	3050
19	175	3165
20	185	3250
21	160	2600
22	150	2990
23	120	2060
24	145	2750
25	170	2795

MTB > Describe c2

Descriptive Statistics: X

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
X	25	0	152,20	4,62	23,10	115,00	132,50	150,00	170,00

Variable	Maximum
X	200,00

MTB > Regress 'Y' 1 'X';

Regression Analysis: Y versus X

The regression equation is
 $Y = 516 + 14,8 X$

Predictor	Coef	SE Coef	T	P
Constant	515,9	211,0	2,45	0,023
X	14,809	1,371	10,80	0,000

S = 155,133 R-Sq = 83,5% R-Sq(adj) = 82,8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	2807900	2807900	116,67	0,000
Residual Error	23	553524	24066		
Total	24	3361424			

Oppgave 3

Det er gjennomført et forsøk for å sammenligne bensinforbruket til en bestemt bilmodell over en gitt strekning for 4 bensintyper A, B, C og D. En er interessert i å finne ut om det er forskjeller i bensinforbruket for de 4 bensintypene.

20 nye biler av samme modell ble delt tilfeldig inn i 4 grupper à 5 biler. Første gruppe fikk A-bensin, andre gruppe B-bensin osv. Bilene ble kjørt med tilnærmet samme fart over en gitt strekning, og deretter ble bensinforbruket målt.

Vi lar X_{ij} (liter) være målt bensinforbruk for bil nr j i gruppe nr i . Resultatene er gitt til slutt i oppgaven (neste side).

Vi antar at X_{ij} er normalfordelt med forventning μ_i og varians σ^2 ($i = 1, 2, 3, 4$ og $j = 1, 2, 3, 4, 5$). Alle observasjonene er uavhengige.

- Angi estimatorer for de ukjente parametrene i modellen ($\mu_1, \mu_2, \mu_3, \mu_4$ og σ^2). Finn et 95% konfidensintervall for forventet bensinforbruk for bensin av type A.
- Tyder forsøket på at det er forskjeller i bensinforbruket for de 4 bensintypene? Formuler dette spørsmålet som en hypotesetest. Gjennomfør testingen og angi konklusjonen når vi velger signifikansnivå 5%.
- La oss tenke oss at vi ønsker å sammenligne forbruket av bensin for de 4 bensintypene for 3 forskjellige bilmodeller (f.eks. VW Golf, Toyota Corolla og Honda Civic). Beskriv kort data, modell og parametre for en slik analyse.

Resultater

	A-bensin	B-bensin	C-bensin	D-bensin
Gruppe	1	2	3	4
X_{ij}	30.0	30.4	30.3	28.1
	31.5	28.7	30.3	30.4
	30.2	29.9	29.2	30.8
	31.9	29.8	27.5	30.1
	30.6	30.3	29.3	28.3
$\bar{X}_i$	30.84	29.82	29.32	29.54
$\sum_{j=1}^5 (X_{ij} - \bar{X}_i)^2$	2.73	1.83	5.25	6.25