

EKSAMEN

6019 Statistikk II

6.12.2016

Tid: 4 timer (9-13)

Målform: Bokmål

Sidetall: 4 (inkludert denne)

Hjelpemiddel: Alle trykte og skrevne samt kalkulator

Sensuren finner du på StudentWeb.

Oppgave 1

På en høgskole er det gjennomført en undersøkelse om snusbruk blant studenter. Vi lar θ_1 være andelen av kvinnelige studenter som bruker snus, og θ_2 er tilsvarende andel blant mannlige studenter.

500 studenter ble trukket ut og spurt om de bruker snus eller ikke. Av dem som ble trukket ut, var det 270 kvinner, og 41 av dem svarte at de bruker snus. Av de 230 mannlige studentene som ble trukket ut, svarte 27 at de bruker snus.

- Angi et tilnærmet 95% konfidensintervall for $\delta = \theta_1 - \theta_2$
- Er det forskjell mellom kjønnene når det gjelder andelen av studenter som bruker snus? Formuler dette spørsmålet som en hypotesetest. Gjennomfør testingen og angi konklusjonen når signifikansnivået velges lik 5 %. Finn signifikanssannsynligheten (tilnærmet) for testen.
- Man kan også benytte en kjiqvadrattest til å undersøke om det er forskjell mellom kvinnelige og mannlige studenter når det gjelder bruk av snus. Sett opp resultatene av undersøkelsen i en 2x2-tabell. Formuler nullhypotese og alternativ hypotese og definer testvariabel. Gjennomfør testingen og angi konklusjonen når signifikansnivået velges lik 5 %.

Oppgave 2

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

Vi lar ρ være korrelasjonskoeffisienten mellom X og Y .

- Hva kan du si om korrelasjonen mellom X og Y ut fra den estimerte verdien for ρ som er gitt i Minitab-utskriften til slutt i oppgaven?

En test fra pensum kan brukes til å teste om ρ er større enn 0. Sett opp nullhypotese, alternativ hypotese, gjennomfør testingen og vis at vi med denne testen kan konkludere med at ρ er større enn 0 når signifikansnivået velges lik 1 %.

Hvor stor må den estimerte korrelasjonskoeffisienten minst være for at vi kan påstå at ρ er større enn 0 med 1 % signifikansnivå?

Vi skal også 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.

- 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 . Hvor stor prosent av variasjonen i salgsprisen er forklart ved boligenes boareal?

- c) En har inntil nå regnet med at gjennomsnittlig økning i salgsprisen pr. økt kvadratmeter i boareal er 12.5 tusen kroner for boliger i dette området. Tyder resultatene fra salgsprisene for de 25 boligene som er solgt i løpet av de tre siste månedene, på at gjennomsnittlig økning pr. kvadratmeter nå er høyere? Formuler dette spørsmålet som en hypotesetest. Gjennomfør testingen og angi konklusjonen når vi velger signifikansnivå 5 %.

Resultater fra Minitab

Data Display

Row	X	Y
1	125	3310,00
2	150	3815,00
3	115	3040,00
4	135	3370,00
5	145	3480,00
6	200	4450,00
7	120	3090,00
8	145	3620,00
9	170	3690,00
10	165	3585,00
11	140	3335,00
12	125	3170,00
13	175	3890,00
14	190	4070,00
15	155	3510,00
16	130	3070,00
17	155	3690,00
18	160	3850,00
19	175	3965,00
20	185	4050,00
21	160	3400,00
22	150	3790,00
23	120	2860,00
24	145	3550,00
25	170	3595,00

Correlation: X; Y

Pearson correlation of X and Y = 0,914
P-Value = 0,000

Regression Analysis: Y versus X

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
155,133	83,53%	82,82%	80,88%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1316	211	6,24	0,000	
X	14,81	1,37	10,80	0,000	1,00

Regression Equation

$Y = 1316 + 14,81 X$

Oppgave 3

Det er gjennomført et forsøk for å sammenligne bensinforbruket til en bestemt bilmodell for 3 bensintyper A, B, og C. En er interessert i å finne ut om det er forskjeller i bensinforbruket for de 3 bensintypene.

18 nye biler av samme modell ble delt tilfeldig inn i 3 grupper à 6 biler. Første gruppe fikk A-bensin, andre gruppe B-bensin og tredje gruppe C-bensin. 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 . Vi antar at X_{ij} er normalfordelt med forventning μ_i og varians σ^2 ($i = 1, 2, 3$ og $j = 1, 2, 3, 4, 5, 6$). Alle observasjonene er uavhengige.

Resultatene er gitt til slutt i oppgaven.

- a) Tyder forsøket på at det er forskjeller i bensinforbruket for de 3 bensintypene? Formuler dette spørsmålet som en hypotesetest. Gjennomfør testingen og angi konklusjonen når vi velger signifikansnivå 5%.

Målinger

Målt bensinforbruk i liter over en gitt strekning.

X_{1j} (A-bensin):	30.0	31.9	30.6	31.1	31.5	30.2
	$\bar{X}_1 = 30.88 \quad \sum (X_{1j} - \bar{X}_1)^2 = 2.788$					
X_{2j} (B-bensin):	30.3	28.7	29.8	30.6	29.2	29.9
	$\bar{X}_2 = 29.75 \quad \sum (X_{2j} - \bar{X}_2)^2 = 2.455$					
X_{3j} (C-bensin):	28.1	30.0	30.6	29.7	29.4	28.3
	$\bar{X}_3 = 29.35 \quad \sum (X_{3j} - \bar{X}_3)^2 = 4.775$					
Totalt gjennomsnitt:	$\bar{X} = 29.99$					