



Høgskolen i Telemark
Fakultet for allmennvitenskapelige fag

EKSAMEN

6019
Statistikk II

9.12.2015

Tid: 9-13 (4 timer)

Målform: Bokmål

Sidetal: 4 (inkludert denne)

Hjelpemiddel: Alle trykte og skrevne samt kalkulator.

Merknader: Ingen

Vedlegg: Ingen.

Sensuren finner du på StudentWeb.

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
Godt fornøyd	30 %
Fornøyd	50 %
Misfornøyd	20 %

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

Svarene fordelte seg slik:

Svaralternativ	Antall svar
Godt fornøyd	181
Fornøyd	223
Misfornøyd	96

- 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%. Kommenter kort eventuelle forskjeller i kundenes tilfredshet etter overgang til ny turoperatør.
- Reisebyrået arrangerer bussturer til 4 forskjellige storbyer. Blant de 500 som har svart på spørreskjemaet, er kunder fra alle 4 reisemål representert. Reisebyrået vil undersøke om det er forskjeller mellom de 4 reisemålene når det gjelder kundenes tilfredshet.

Gjør rede for modell og testmetode for å analysere denne problemstillingen. Sett opp nullhypotese og alternativ hypotese og forklar kort hvordan testen gjennomføres.

Oppgave 2

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, \dots, 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 .

Observasjonene er gitt til slutt i oppgaven. Alle beløp er i tusen kroner.

- a) I første omgang skal vi teste om medianen i lønnsfordelingen til de nyutdannede økonomene er over 450 000 kr. Formuler dette spørsmålet som en hypotesetest om m .

Hvilke to tester fra pensum kan brukes i denne situasjonen?

Gjennomfør testingen med den av de to aktuelle testene som du mener er best i denne situasjonen. Regn ut signifikanssannsynligheten og angi konklusjonen når nivået velges lik 5 %.

Vi skal dessuten sammenligne lønn for økonomer med lønn for ingeniører og sykepleiere. Disse yrkesgruppene har alle fullført en treårig bachelor-utdanning, og vi vil undersøke om det er forskjeller i lønnen som de tre gruppene tilbys etter endt utdanning.

I tillegg til årslønn for 10 nyutdannede økonomer er det nedenfor gitt årslønn for 8 tilfeldig utvalgte nyutdannede ingeniører og 8 tilfeldig utvalgte nyutdannede sykepleiere. Alle beløp er i tusen kroner.

Vi antar at lønnen innen hver yrkesgruppe følger en kontinuerlig sannsynlighetsfordeling, og at alle observasjonene kan oppfattes som uavhengige variabler.

- b) Tyder observasjonene nedenfor på at det er forskjeller mellom økonomer, ingeniører og sykepleiere når det gjelder lønnen disse tre yrkesgruppene tilbys etter endt utdanning? Formuler dette spørsmålet som en hypotesetest, gjennomfør testingen og angi konklusjonen. Bruk signifikansnivå 5%.

Hvilken modell kunne en ha brukt til å analysere denne problemstillingen dersom en i tillegg til antakelsene ovenfor hadde antatt at lønnen innen hver av de tre yrkesgruppene var normalfordelt, og at variansen var felles for de tre gruppene?

Årslønn for nyutdannede kandidater (ordnet i stigende rekkefølge innen hver yrkesgruppe)

Økonomer

410 425 441 456 464 465 467 489 498 510

Ingeniører

433 448 458 476 478 489 512 523

Sykepleiere

395 402 415 432 444 453 459 478

Oppgave 3

På en forsøksgård har en gjennomført dyrkingsforsøk med 3 forskjellige gjødseltyper (G_1 , G_2 , og G_3) og 4 forskjellige kornsorter (K_1 , K_2 , K_3 og K_4).

Vi lar X_{ij} være målt avlingsmengde for gjødseltype G_i og kornsort K_j ($i = 1, 2, 3$ og $j = 1, 2, 3, 4$).

Vi antar at målingene er uavhengige og normalfordelte med varians σ^2 .

Forventningene er gitt ved

$$E(X_{ij}) = \mu + \alpha_i + \beta_j \quad i = 1, 2, 3 \quad j = 1, 2, 3, 4$$

$$\text{der } \alpha_1 + \alpha_2 + \alpha_3 = 0 \quad \text{og} \quad \beta_1 + \beta_2 + \beta_3 + \beta_4 = 0$$

Til slutt i oppgaven fins en tabell med resultatene av forsøket og deler av utskrift fra en Minitab-kjøring der det er gjennomført to-faktor variansanalyse. Resultatene i utskriften kan brukes i svarene på spørsmålene nedenfor.

- a) Forklar kort hva parametrene μ , α_i og β_j uttrykker i modellen.
Regn ut estimater for μ , α_1 , β_1 og σ^2 .
- Estimer forventet avlingsmengde for kombinasjonen G_1 og K_1 .
- b) Tyder resultatene på at det er forskjeller i avlingsmengdene mellom de 3 gjødseltypene?
Formuler dette spørsmålet som en hypotesetest. Gjennomfør testingen og angi konklusjonen når signifikansnivået velges lik 5%.

Resultater

Gjødseltype	Kornsort				Gjennomsnitt
	K ₁	K ₂	K ₃	K ₄	
G ₁	86	88	77	84	83.75
G ₂	92	91	81	93	89.25
G ₃	75	80	83	79	79.25
Gjennomsnitt	84.33	86.33	80.33	85.33	84.08

Fra Minitab

ANOVA: Avling versus Gjødsel; Korn

Analysis of Variance for Avling

Source	DF	SS	MS
Gjødsel	2	200,67	100,33
Korn	3	62,25	20,75
Error	6	132,00	22,00
Total	11	394,92	