Bokmål

EKSAMENSOPPGAVE I STA-0001

Eksamen i	:	Sta-0001 Brukerkurs i statistikk 1
Dato	:	Mandag 11. juni 2007
Tid	:	Kl. 1500 – 1900
Sted	:	Åsgårdveien 9
Tillatte hjelpemidler	:	Alle trykte og skrevne hjelpemidler samt kalkulator.

Oppgavesettet er på 3 sider inklusive forside.

Kontaktperson: Sigrunn H. Sørbye, tlf. 92 43 97 22.

DET MATEMATISK-NATURVITENSKAPELIGE FAKULTET
Universitetet i Tromsø, N-9037 Tromsø, Telefon 77 64 40 01, Telefaks 77 64 47 65

Bokmål

Oppgave 1

Vi antar at potetavlingen X (kg) på et mål av en åker er normalfordelt med forventning lik 5000 kg og standardavvik lik 300 kg.

a) Hva er sannsynligheten for at potetavlingen på et mål er over 5500 kg? Hva er sannsynligheten for at avlingen på et mål ligger mellom 4500 kg og 5000 kg? Tegn en skisse av sannsynlighetstettheten til X og marker de sannsynlighetene du har funnet som arealer under kurven.

Bonden som dyrker poteter regner med at dersom avlingen er X kg pr. mål, så er inntekten Z (i kroner) av disse potetene gitt ved

$$Z = 2.20 \cdot X - 1500.$$

b) Hva er forventet inntekt fra potetavlingen på et mål? Hva er sannsynligheten for at inntekten av potetavlingen på et mål blir over 10 000 kroner?

Bonden ønsker å undersøke om en ny potetsort gir større avling enn den som hittil har vært dyrket. Han foretar en prøvedyrking av den nye potetsorten på totalt 14 mål. Avlingsresultatene pr. mål, gitt ved $Y_1, Y_2, \ldots, Y_{14}$, antas å være uavhengige og normalfordelte variabler med ukjent forventning μ, men kjent standardavvik lik 300 kg.

c) Formuler hypoteser for den gitte problemstillingen. Utfør testen og angi konklusjonen dersom gjennomsnittlig avling for de 14 målene var 5184 kg. Benytt signifikansnivå lik 5%.

d) Hva blir sannsynligheten for at nullhypotesen i punkt c) forkastes dersom sann forventningsverdi er $\mu = 5100$ kg? For hvilken verdi av μ vil teststyrken bli lik 0.90? Skisser styrkefunksjonen for testen.

Oppgave 2

Et strømforsyningssystem består av et hovedaggregat og et nødaggregat. Dersom hovedaggregatet svikter vil nødaggregatet overta. La hendelsen A angi at hovedaggregatet svikter en gitt dag og la hendelsen B angi at nødaggregatet svikter den gitte dagen. Vi antar at

$$P(A) = \frac{1}{20}, \quad P(B \mid A) = \frac{1}{15}, \quad P(B \mid \bar{A}) = 0.$$

a) Forklar kort med ord hva det vil si at $P(B \mid \bar{A}) = 0$ i denne situasjonen.

Er A og B disjunkte hendelser? Er A og B uavhengige hendelser? Begrunn svarene ved å sette opp aktuelle sannsynligheter.

b) Finn sannsynligheten for at strømforsyningssystemet ikke svikter en gitt dag.

Gitt at strømforsyningssystemet fungerer, hva er da sannsynligheten for at hovedaggregatet har sviktet denne dagen?

Bokmål

Oppgave 3

Erfaring viser at en mye brukt influensavaksine gir beskyttelse gjennom vinteren for 90% av de som vaksinerer seg om høsten. 10% får influensa til tross for at de er vaksinert. En ny og forbedret utgave av vaksinen prøves ut på 300 personer. Vi ønsker å undersøke om den nye utgaven av vaksinen gir bedre beskyttelse mot influensa enn den gamle. La p angi sannsynligheten for at en person som har tatt den nye vaksinen får influensa.

a) Formuler hypoteser for den gitte problemstillingen. Beskriv hvilke antakelser du gjør og sett opp testobservatoren du vil bruke for å utføre hypotesetesten. Hva blir p-verdien for testen dersom det viser seg at 23 av de 300 forsøkspersonene fikk influensa? Tolk svaret og konkluder hvorvidt nullhypotesen bør forkastes eller ikke.

Oppgave 4

Statistisk Sentralbyrå innhenter årlig opplysninger om forbruk av tobakk i Norge ved å spørre et stort og representativt utvalg av personer om deres røykevaner. Følgende tabell viser observert prosentandel dagligrøykere i aldersgruppen 16 - 74 år i Norge i perioden 1997-2006, der $x = 1$ svarer til år 1997, $x = 2$ svarer til 1998 osv.

Antall år etter 1996 (x)	1	2	3	4	5	6	7	8	9	10
Andel dagligrøykere (Y)	33.6	33.0	32.0	31.9	29.8	29.4	26.3	26.0	25.0	23.9

For å beskrive utviklingen av prosentandel dagligrøykere som funksjon av x benyttes en lineær regresjonsmodell

$$Y_i = \alpha + \beta x_i + e_i, \quad i = 1, \ldots, 10.$$

Vi antar at feilleddene $e_1, \ldots, e_{10}$ er uavhengige og normalfordelte variabler med forventning lik 0 og ukjent standardavvik σ. En lineær regresjonsanalyse av det gitte datasettet i Excel gav følgende resultater:

SAMMENDRAG (UTDATA)

Variansanalyse

	fg	SK	GK	F	Signifikans-F
Regresjon	1	109.970	109.970	243.126	0.000
Residualer	8	3.619	0.452		
Totalt	9	113.589			

	Koeffisienter	Standardfeil	t-Stat	P-verdi
Skjæringspunkt	35.440	0.459	77.138	0.000
X-variabel 1	-1.155	0.074	-15.593	0.000

a) Hva er de estimerte verdiene for parametrene α og β? Tegn observasjonene samt den estimerte regresjonslinjen inn i et spredningsplott. Forklar kort hva den estimerte verdien for β angir og finn et 95% konfidensintervall for β.

b) Finn den empiriske korrelasjonskoeffisienten r for de 10 observasjonsparene samt størrelsen r^2. Beskriv hvilken informasjon disse to verdiene gir.

c) Har det vært en signifikant nedgang i prosentandelen dagligrøykere i løpet av perioden 1997-2006? Formuler dette som en hypotesetest og konkluder utfra et signifikansnivå på 1%.
Hva blir predikert prosentandel dagligrøykere for år 2007?