Noregs teknisk–
naturvitskaplege universitet
Institutt for matematiske fag

Fagleg kontakt under eksamen:
Turid Follestad 98066880

EKSAMEN I FAG TMA4240/TMA4245 STATISTIKK

19. august 2006

Tid: 09:00–13:00

Tillatne hjelpemiddel:
Gult A5-ark med eigne handskrivne notatar.
Tabeller og formler i statistikk (Tapir Forlag).
K. Rottmann: *Matematisk formelsamling.*
Kalkulator: HP30S.

NYNORSK

Sensur: 11. september 2006.

Oppgåve 1 Kvalitetskontroll av poteter

Ein grønsaksprodusent produserer poteter, og er blitt godkjent som brukar av kvalitetsmerket PrimaVare. Dette inneber mellom anna at potetene skal tilhøyre ein bestemt vektklasse. Produsenten tar jamnleg stikkprøver for å kontrollere at potetene held seg innanfor kvalitetsspesifikasjonen.

La X vere vekta av ei tilfeldig valt potet, og føreset at vekta er normalfordelt med forventning $\mu = 110\ g$ og varians σ^2. La kvalitetsspesifikasjonen for vekt vere at vekta på potetene ikkje skal avvike meir enn 20 g frå den forventa vekta på 110 g.

a) Føreset, berre i dette punktet, at variansen er kjent, og lik $\sigma^2 = 100\ g^2$.

Kva er sannsynet for at ei tilfeldig valt potet veg meir enn 120 g?

Kva er sannsynet for at ei tilfeldig valt potet er innanfor spesifikasjonen, dvs. $P(|X - 110| \leq 20)$?

Kva er sannsynet for at ein pose med 14 tilfeldig valde poteter veg mindre enn 1500 g?

Føreset at det blir tatt ei stikkprøve med n tilfeldig valde poteter frå produksjonen. La Y vere talet på poteter som *ikkje* følgjer spesifikasjonen, blant dei n i stikkprøva.

b) Forklar kvifor vi kan føresette at Y er binomisk fordelt med parametrar n og $p = 1 - P(|X - 110| \leq 20)$, der X er definert som tidlegare i oppgåva.

Ein estimator for p er $\hat{p} = \frac{Y}{n}$. Vis at dette er ein forventningsrett estimator for p, og utlei variansen til $\hat{p}$.

Bruk denne estimatoren til å finne eit anslag for variansen σ^2 til vekta på ei potet, når det i ei stikkprøve på storleik $n = 19$ blir observert at $y = 2$ poteter ikkje følgjer spesifikasjonen.

Oppgåve 2 Internettkafé

Torkel driv ein internettkafé der han har tre datamaskiner tilgjengelige. Iblant er det mange kundar innom, og enkelte blir misfornøgde dersom det er lenge å vente på ledig maskin. Torkel bestemmer seg for å gjennomføre ein statistisk analyse av påloggingstidene til kundar. Føreset at påloggingstida T til ein tilfeldig valt kunde har kumulativ sannsynsfordeling

$$F(t) = 1 - \exp(-\lambda t), \qquad \text{for } t > 0 \text{ og } \lambda > 0.$$

a) Føreset, kun i dette punktet, at $\lambda = 0.1$. Berekn $P(T < 15)$. Bruk definisjonen av betinga sannsyn til å berekne $P(T < 45|T > 30)$. Samanlikn dei to resultata og drøft i den sammenhengen kva for ein eigenskap som gjeld for denne sannsynsfordelinga.

For å gjennomføre analysen sin registrerer Torkel ti uavhengige påloggingar. La T_i vere påloggingstida til person $i = 1, \ldots, 10$, og føreset at desse følgjer den kumulative sannsynsfordelinga gitt over. Resultatet er følgjande påloggingstider (i minutt): 11.1, 9.4, 40.4, 19.5, 6.9, 5.4, 14.5, 2.2, 8.0 og 7.7.

b) Vis at sannsynstettleiken til T_i, $i = 1, \ldots, 10$, er

$$f(t) = \begin{cases} \lambda \exp(-\lambda t) & \text{for } t > 0 \\ 0 & \text{elles.} \end{cases}$$

Basert på talmaterialet i undersøkinga vil Torkel estimere λ. Set opp rimelighetsfunksjonen (likelihoodfunksjonen) med tanke på dette føremålet. Utlei sannsynsmaksimeringsestimatoren (maximum likelihood estimatoren) for λ og rekn ut estimatet.

Torkel vil nytte resultatet frå undersøkinga til å informere kundane om ventetida på ledig maskin.

c) Ein person kjem inn på kaféen, men alle tre maskinene er opptekne. Kva er sannsynet for at kunden får ein ledig maskin innan 15 minutt?

To personer kjem samtidig inn på kaféen, men alle tre maskinene er opptekne. Kva er forventa tid til den siste av dei to får ein ledig maskin?

Uttrykk løysingane først som funksjon av λ, og set deretter inn den estimerte verdien din for λ frå forrige deloppgåve for å gi talsvar.

Oppgåve 3 Kneoperasjon

Ved operasjon av kneleddet er det viktig at vinkelutslaget mellom leggbeinet og lårbeinet blir nær 0 grader. Vinkelen blir rekna som negativ dersom personen er kalvbeint og positiv dersom personen er hjulbeint.

I ein medisinsk studie vart $n = 25$ pasientar kneopererte med konvensjonell kirurgi, og $m = 27$ pasientar vart opererte med ein ny teknikk som blir kalla computer-assistert kirurgi.

La X vere vinkelutslaget mellom leggbeinet og lårbeinet etter operasjon ved konvensjonell kirurgi, og føreset at X er normalfordelt med forventningsverdi $\mathrm{E}(X) = \mu$ og standardavvik $\mathrm{SD}(X) = \sigma$. Vi har $n = 25$ uavhengige målingar av vinkelutslag med konvensjonell kirurgi, og desse blir noterte $x_1, ..., x_{25}$. Det blir opplyst at $\bar{x} = \frac{1}{25}\sum_{i=1}^{25} x_i = 2.3$ og $\sum_{i=1}^{25}(x_i - \bar{x})^2 = 294$.

Tilsvarande, la Y vere vinkelutslaget mellom leggbeinet og lårbeinet etter operasjon ved computer-assisert kirurgi, og føreset at Y er normalfordelt med forventningsverdi $\mathrm{E}(Y) = \eta$ og standardavvik $\mathrm{SD}(Y) = \sigma$. Merk at standardavviket til dei to operasjonsmetodane blir føresett å vere det same. Vi har $m = 27$ uavhengige målingar av vinkelutslag med computer-assistert kirurgi, og desse blir noterte $y_1, ..., y_{27}$. Det blir opplyst at $\bar{y} = \frac{1}{27}\sum_{j=1}^{27} y_j = 1.5$ og $\sum_{j=1}^{27}(y_j - \bar{y})^2 = 115$.

Vi føreset at målingane som er gjorde frå dei to operasjonsmetodane er uavhengige av kvarandre.

a) Føreslå ein god estimator for variansen σ^2 basert på alle målingane, dvs. frå både konvensjonell og computer-assistert kirurgi. Finn forventningsverdien til estimatoren du har føreslått. Er estimatoren forventningsrett?

Utlei eit 95% konfidensintervall for variansen σ^2 basert på målingane frå både konvensjonell og computer-assistert kirurgi. Rekn også ut intervallet numerisk.

b) Vi ønskjer å undersøke om forventa vinkelutslag mellom leggbeinet og lårbeinet er forskjellig for dei to operasjonsmetodane.

Formulér dette som ein hypotesetest ved å definere nullhypotese og alternativ hypotese.

Set opp ein testobservator og finn forkastningsområdet. Kva blir konklusjonen på testen, når signifikansnivået er $\alpha = 0.05$ og du bruker dataene som er gitt i oppgåva?

Rekn ut p-verdien ved å bruke tabell 1.

t	1.05	1.03	1.01	0.99	0.97	0.95	0.75	0.55	0.35	0.15
$\nu = 25$	0.848	0.844	0.839	0.834	0.829	0.824	0.770	0.706	0.635	0.559
$\nu = 27$	0.848	0.844	0.839	0.835	0.830	0.825	0.770	0.707	0.635	0.559
$\nu = 50$	0.851	0.846	0.841	0.837	0.832	0.827	0.772	0.708	0.636	0.559
$\nu = 52$	0.851	0.846	0.841	0.837	0.832	0.827	0.772	0.708	0.636	0.559

Tabell 1: Kumulativt sannsyn i t-fordelinga. For T t-fordelt med ν fridomsgrader, så viser tabellen $P(T \leq t)$ for ulike verdiar av t.

Oppgåve 4 Simultanfordeling

Simultanfordelinga, $f(x, y)$, til dei to diskrete stokastiske variablane X og Y er gitt i følgjande tabell:

	y=0	y=1	y=2
x=-1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$
x=0	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
x=1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$

Finn marginalfordelinga til X og til Y, og rekn ut forventning og varians til X og til Y.

Rekn ut kovariansen mellom X og Y, $\text{Cov}(X, Y)$. Er X og Y uavhengige? Svaret skal grunngjevast.