Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Faglig kontakt under eksamen:
Stian Lydersen 73 59 70 53

EKSAMEN I FAG SIF 5505/SIF 5506 STATISTIKK

Mandag 9. august 1999
Tid: 09:00–14:00

Hjelpemidler:
Godkjent lommekalkulator, utdelt ordliste, Statistiske tabeller og formler (Tapir forlag).

Oppgave 1 Hodekål

Ved et gartneri dyrkes hodekål. Vekten av den modne kålen, X, antas å være normalfordelt med forventningsverdi $\mu = 2$ kg og standardavvik $\sigma = 0.5$ kg.

a) Hvis en velger et kålhode tilfeldig, hva er sannsynligheten for at dette skal:
1) veie mindre enn 1.5 kg?
2) veie mellom 2 og 2.5 kg?

Hva er sannsynligheten for at vektforskjellen mellom to tilfeldige valgte kålhoder skal være mer enn 1 kg?

Kålhoder som veier mindre enn 1.5 kg oppfyller ikke kravet til klasse 1-kål.

b) Gitt at et kålhode oppfyller kravet (veier mer enn 1.5 kg), hva er sannsynligheten for at det veier mellom 2 og 2.5 kg?

Kan du sette opp fordelingen (kumulativ fordeling eller sannsynlighetstetthet) til vekten av et kålhode som oppfyller kravet til klasse 1?

Gartneriet bestemmer seg for å prøve ut en ny kåltype. Det påstås at hodene av denne nye typen gjennomgående veier mer enn hodene av den gamle typen. En liten prøveavling med 10 planter dyrkes. Vi antar først at de 10 vektmålingene $Y_1, Y_2, \cdots, Y_{10}$ er uavhengige og normalfordelte med ukjent forventning, μ_Y, og kjent standardavvik, $\sigma_Y = 0.5$ kg.

Resultatet av målingene er gitt i tabellen under:

kål i	1	2	3	4	5	6	7	8	9	10
y_i	1.23	1.95	3.30	2.44	2.56	2.23	1.73	2.35	2.78	2.72

c) Hvilken estimator vil du bruke for å estimere μ_Y? Skriv opp estimatoren og regn ut estimatet fra dataene i tabellen.

Utled et 90%-konfidensintervall for μ_Y. Hva blir intervallet med de oppgitte dataene?
Hva blir lengden av konfidensintervallet?
Hvor mange planter må vi minst ha for at lengden av konfidensintervallet skal bli mindre enn 0.2 kg?

Vi antar fra nå av at både μ_Y og σ_Y^2 er ukjente.

d) Hvilken estimator vil du bruke for å estimere σ_Y^2? Skriv opp estimatoren og regn ut estimatet fra dataene i tabellen.

Gartneriet vil vite om forventet vekt av den nye typen er gjennomgående høyere enn for den gamle. Vi skal altså teste:

$$H_0 : \mu_Y = 2 \text{ kg mot } H_1 : \mu_Y > 2 \text{ kg}$$

Hvilken testobservator vil du bruke for å utføre testen ovenfor?
Hvilken fordeling har denne testobservatoren under H_0?
Velg signifikansnivå α lik 0.10, og utfør testen. Hva blir konklusjonen?

Oppgave 2 Spørreundersøkelsen

Et politisk spørsmål blir tatt opp i en TV-debatt. Et stykke ut i debatten blir det samme spørsmålet stilt til seerne. Vi ser heretter bare på de seerne som har en oppfatning av spørsmålet. De som mener *ja*, oppfordres til å ringe et bestemt telefonnummer og de som mener *nei*, blir bedt om å ringe et annet nummer. Vi antar i denne oppgaven at 80% av seerne mener *ja*, og 20% mener *nei*. Vi antar videre at sannsynligheten for at en tilfeldig "*ja*-seer" ringer inn er 0.02. Tilsvarende sannsynlighet for en "*nei*-seer" er 0.05. Vi lar J være hendelsen at en seer mener *ja*, og R være hendelsen at seeren ringer.

Uttrykk de fire opplysningene i oppgaven som sannsynligheter (betingede eller ubetingede) for J og R (eller de komplementære hendelsene).

Hvor stor andel av innringerene mener *ja*? Gir resultatet av innringingen et riktig bilde av seernes oppfatning?

Oppgave 3 Redningshelikopteret

Ved en redningssentral ankommer alarmer som krever utrykning med redningshelikopter som en Poisson prosess. Forventet antall alarmer mottatt i løpet av ett døgn er λ. Vi vet da fra pensum at antall alarmer mottatt i løpet av ett døgn er Poisson-fordelt med forventning λ, og at tiden det går mellom to alarmer er eksponensialfordelt med forventning $1/\lambda$.

a) Anta i dette punktet at det er kjent at $\lambda = 1.5$.

Beregn sannsynligheten for at det kommer to eller flere alarmer i løpet av ett døgn.

Beregn sannsynligheten for at det kommer to eller flere alarmer i løpet av ett døgn dersom vi vet at det kommer minst én alarm.

Anta at det går 2 timer fra en alarm mottas til redningshelikopteret er klar til ny utrykning. Hva er sannsynligheten for at en eller flere nye alarmer mottas i løpet av denne tiden?

I praksis vil λ være ukjent, men kan estimeres fra observerte data. En mulig fremgangsmåte er å registrere data inntil en har observert et forhåndsbestemt antall, n, alarmer. Observasjonene vil da være $T_1, T_2, \ldots, T_n$, hvor T_1 er tida frem til første alarm, T_2 er tida fra første til andre alarm osv. Disse tidene vil være uavhengige, identisk eksponensialfordelte.

b) Utled den momentgenererende funksjonen til $Z = 2\lambda \sum_{i=1}^{n} T_i$. Vis at Z er χ^2-fordelt med $2n$ frihetsgrader.

c) Bruk resultatet i punkt **b)** til å utlede et eksakt $(1-\alpha)100\%$ konfidensintervall for λ. Regn ut et 90% konfidensintervall for λ hvis $n = 10$ og de observerte tidene (døgn) er:

tid i	1	2	3	4	5	6	7	8	9	10
t_i	1.8	0.2	0.1	1.0	0.3	1.7	0.4	0.8	0.1	1.1

d) Forklar hvorfor det er rimelig å anta at $\sum_{i=1}^{n} T_i = T_1 + T_2 + \cdots + T_n$ er tilnærmet normalfordelt. Bruk dette til å regne ut et tilnærmet 90% konfidensintervall for λ når observasjonene er som gitt ovenfor. Sammenlign med konfidensintervallet i punkt **c)** og kommenter.

e) Utled sannsynlighetsmaksimeringsestimatoren (SME) $\hat{\lambda}$ for λ basert på $T_1, T_2, \ldots, T_n$. Hva blir estimatet for λ når observasjonene er som gitt ovenfor?

Regn ut forventningsverdien til estimatoren $\hat{\lambda}$ (Hint: Du kan få bruk for resultatet i punkt **b)**). Dersom estimatoren ikke er forventningsrett, foreslå en forventningsrett estimator $\tilde{\lambda}$ basert på $\hat{\lambda}$.