EKSAMENSOPPGAVE

Eksamen i: STA-1001, Statistikk og sannsynlighet 1

Dato : Mandag 11. juni 2007

Tid : 0900 - 1300

Sted : AudMax

Tillatte hjelpemidler:
- **4 sider handskrevne notat**
- **Godkjent kalkulator**
- **«Tabeller og formler i statistikk» av Kvaløy og Tjelmeland**

Oppgavesettet er på 4 sider inkl. forside

Kontaktperson: Georg Elvebakk, 77646532

DET MATEMATISK-NATURVITENSKAPELIGE FAKULTET

Universitetet i Tromsø, N-9037 Tromsø, Telefon 77 64 40 01, Telefaks 77 64 47 65

Deloppgavene vil telle likt ved vurderinga.

Oppgave 1

En forsker har som oppgave å overvåke pH-verdien i ferskvann i Agder-fylka (det er fleire 1000 av dem). Dette blir gjort ved å måle pH-verdier i noen vann hver sommer. Ho går ut fra at pH-verdien i vanna er normalfordelt med forventning μ og standardavvik σ. Fra tidlegere år veit ho at $\mu = 6.0$ og $\sigma = 0.25$. Disse verdiene skal brukes i punktet a).

a) Om forskeren velger tilfeldig ut et vann, hva blir sannsynligheten for at i) pH-verdien ligger under 5.85? ii) pH-verdien ligger mellom 5.9 og 6.1? Dersom ho velger tilfeldig 4 vann, hva er sannsynligheten for at gjennomsnittlig pH-verdi ligger mellom 5.9 og 6.1?

For å få nye estimat på μ (forventa pH-verdi i Agder-fylka) og σ velger forskeren en sommer å måle pH-verdien i 6 tilfeldig valgte vann. Ho får følgende data:

vatn nr.	1	2	3	4	5	6
pH-verdi:	5.90	6.17	5.81	5.35	5.70	5.58

b) Estimer μ og σ fra data.
Bruk disse til å lage et 95%-konfidensintervall for μ. Hva forteller dette intervallet deg?

I resten av oppgaven skal du igjen gå ut fra at standardavviket er **kjent**, $\sigma = 0.25$.

c) Kan vi ut fra data konkludere med at forventa pH-verdi i Agder-fylka har sunket i forhold til verdien $\mu = 6.0$? Formuler og utfør en passende test for dette med signifikansnivå 5%.
Vil dette være en fornuftig test dersom forskeren gjerne vil ha stor sannsynlighet for å oppdage en nedgang så liten som 0.2. Om ikke, hvilke endringer vil du anbefale?

Neste sommer vil forskeren gjerne finne ut om forventa pH-verdi har endra seg siden de forrige målingene blei tatt. Ho gjør derfor 6 nye målinger i de samme vanna som året før:

vatn nr.	1	2	3	4	5	6
pH-verdi:	5.79	6.00	5.66	5.43	5.70	5.39

Det er fra tidlegere kjent at kovariansen mellom målinger i samme vann 2 år på rad er 0.05.

d) Lag et 95%-konfidensintervall for forskjellen i forventa pH-verdi fra i fjor til i år.
Bør forskeren konkludere med at forventa pH-verdi i Agder-fylka har endra seg?

Oppgave 2

En liten dyrepark er åpen hver søndag og vil gjerne lage en modell for hvor mange som vil besøke dem. De veit av erfaring at de har et omland med 5000 innbyggere. Fra tidlegere år trur parken at hver av disse har en sannsynlighet på $\frac{1}{100}$ for å komme innom en tilfeldig søndag.

a) Hva slags fordeling vil du foreslå for X, antallet personar som besøker parken en tilfeldig søndag? Presiser de forutsetningene du må gjøre, virker de realistiske?
Hva blir forventning og standardavvik for X?
Hva blir sannsynligheten for at det skal komme over 40 personer en tilfeldig søndag?

Etter å ha sett over talla fra tidlegere år finner parken ut at det er noe galt med modellen. Gjennomsnittet ser ut til å stemme, men bare i litt over 75% av dagene får de over 40 personer. De undrer på om dette kan ha noe med været å gjøre. Og fra meteorologisk institutt finner de ut at sannsynligheten for at det skal regne en tilfeldig dag er $\frac{1}{3}$. Parken gjør òg følgende forutsetninger:

- Alle innbyggere har en sannsynlighet på p_R for å besøke parken på regndager.
- Denne sannsynligheten øker med 50% dersom det ikke regner.

b) Vis at $p_R = \frac{3}{400}$ dersom sannsynligheten for at en tilfeldig person skal besøke parken en tilfeldig søndag fremdeles skal være $\frac{1}{100}$?
Hva blir nå sannsynligheten for at det skal komme over 40 personer en tilfeldig søndag?

Det koster parken 5000 kroner i uka å drive. For hver person som besøker den får de inn 120 kroner.

c) Dersom parken ei uke hadde et overskudd, hvor sannsynlig er det at det var oppholdsvær den søndagen?

Oppgave 3

I denne oppgava skal vi se på modellering av trafikktetthet. På ei vegstrekning har vegvesenet satt opp apparatur som automatisk registrerer biler som passerer mellom klokka 9:00 og 10:00 på hverdager. De meiner at tidsavstanden mellom to påfølgende biler følger ei fordeling som likner eksponensialfordelinga, men der avstanden ikke kan komme under en viss minste avstand. Altså for X, tida i sekund mellom to tilfeldig valgte påfølgende biler, har vi sannsynlighetstetthet

$$f(x;d,\beta) = \begin{cases} \frac{1}{\beta}e^{-\frac{(x-d)}{\beta}}, & x > d \\ 0 & \text{ellers.} \end{cases}$$

der d er parameteren for minimumsavstanden mellom 2 biler.

a) Lag ei grov skisse av fordelinga.
Vis ved utrekning at forventninga til X blir $d + \beta$.
Dersom $\beta = 3$ og $d = 2$ finn sansynligheten for at en avstand skal være under 4 sekund.

Nå har vegvesenet registrert n slike tidsavstandar, $X_1, X_2, \ldots, X_n$. Under forutsetningen at disse er uavhengige skal de brukes til å estimere parametrene β og d.

b) Skriv opp sannsynlighetsmaksimeringsfunksjonen (likelihoodfunksjonen) for denne situasjonen.
Argumenter ut fra sannsynlighetsmaksimeringsfunksjonen at estimatoren for d må bli minimum av observasjonene:

$$\hat{d} = \min_{i=1,\ldots,n} X_i$$

Finn deretter (på vanlig måte) en estimator for β.

Avstandene vegvesenet registrerte var

Nr.	1	2	3	4	5	6	7	8	9	10
Avstand:	2.86	2.16	5.76	3.53	11.60	4.57	2.24	5.43	8.02	2.64

Bruk observerte data til å finne estimat for d og β.

Det kan vises (skal ikke gjøres) at fordelinga til $\hat{d}$ er av samme type som fordelinga til X, eneste forskjell er at parameteren β er erstatta av $\frac{\beta}{n}$.

c) Bruk opplysningene du har til å finne forventninga til estimatorene for d og β.
Er de forventningsrette? Om ikke, konstruer forventningsrette estimatorer basert på dem.

Delvis integrasjon:

$$\int f(x)G(x)dx = F(x)G(x) - \int F(x)g(x)dx$$