EKSAMENSOPPGÅVE

Eksamen i: STA-1001, Statistikk og sannsyn 1

Dato : Måndag 11. juni 2007

Tid : 0900 - 1300

Stad : AudMax

Tillatne hjelpemiddel: - 4 sider handskrivne notat

- Godkjent kalkulator

- «Tabeller og formler i statistikk» av Kvaløy og Tjelmeland

Oppgåvesettet er på 4 sider inkl. forside

Kontaktperson: Georg Elvebakk, 77646532

DET MATEMATISK-NATURVITSKAPLEGE FAKULTET

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

Deloppgåvene vil telje likt ved vurderinga.

Oppgåve 1

Ein forskar har som oppgåve å overvake pH-verdien i ferskvatn i Agder-fylka (det er fleire 1000 av dei). Dette blir gjort ved å måle pH-verdiar i nokre vatn kvar sommar. Ho går ut frå at pH-verdien i vatna er normalfordelt med forventning μ og standardavvik σ. Frå tidlegare år veit ho at $\mu = 6.0$ og $\sigma = 0.25$. Desse verdiane skal brukast i punktet a).

a) Om forskaren vel tilfeldig ut eit vatn, kva blir sannsynet for at
i) pH-verdien ligg under 5.85? ii) pH-verdien ligg mellom 5.9 og 6.1?
Dersom ho vel tilfeldig 4 vatn, kva er sannsynet for at gjennomsnittleg pH-verdi ligg mellom 5.9 og 6.1?

For å få nye estimat på μ (forventa pH-verdi i Agder-fylka) og σ vel forskaren ein sommar å måle pH-verdien i 6 tilfeldig valde vatn. Ho får følgjande 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 frå data.
Bruk desse til å lage eit 95%-konfidensintervall for μ. Kva fortel dette intervallet deg?

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

c) Kan vi ut frå data konkludere med at forventa pH-verdi i Agder-fylka har sokke i forhold til verdien $\mu = 6.0$? Formuler og utfør ein passande test for dette med signifikansnivå 5%.
Vil dette vere ein fornuftig test dersom forskaren gjerne vil ha stort sannsyn for å oppdage ein eventuell nedgang til 5.8? Om ikkje, kva endringar vil du tilrå?

Neste sommar vil forskaren gjerne finne ut om forventa pH-verdi har endra seg sidan dei forrige målingane blei tatt. Ho gjer derfor 6 nye målingar i dei same vatna 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 frå tidlegare kjent at kovariansen mellom målingar i same vatn 2 år på rad er 0.05.

d) Lag eit intervall som med 95% sannsyn inneholder forskjellen i forventa pH-verdi fra i fjor til i år.
Bør forskaren konkludere med at forventa pH-verdi i Agder-fylka har endra seg?

Oppgåve 2

Ein liten dyrepark er open kvar søndag og vil gjerne lage ein modell for kor mange som vil vitje dei. Dei veit av erfaring at dei har eit omland med 5000 innbyggjarar. Frå tidlegare år trur parken at kvar av desse har eit sannsyn på $\frac{1}{100}$ for å komme innom ein tilfeldig søndag.

a) Kva slags fordeling vil du foreslå for X, talet på personar som vitjar parken ein tilfeldig søndag? Presiser dei føresetnadene du må gjere, virkar dei realistiske?
Kva blir forventning og standardavvik for X?
Kva blir sannsynet for at det skal komme over 40 personar ein tilfeldig søndag?

Etter å ha sett over tala frå tidlegare år finn parken ut at det er noko galt med modellen. Gjennomsnittet ser ut til å stemme, men berre i litt over 75% av dagane får dei over 40 personar. Dei undrast på om dette kan ha noko med veret å gjere. Frå meteorologisk institutt finn dei ut at sannsynet for at det skal regne ein tilfeldig dag er $\frac{1}{3}$. Parken gjer òg følgjande føresetnader:

- Alle innbyggjarar har eit sannsyn på p_R for å vitje parken på regndagar.
- Dette sannsynet aukar med 50% dersom det ikkje regnar.

b) Vis at $p_R = \frac{3}{400}$ dersom sannsynet for at ein tilfeldig person skal vitje parken ein tilfeldig søndag framleis skal vere $\frac{1}{100}$?
Kva blir no sannsynet for at det skal komme over 40 personar ein tilfeldig søndag?

Det kostar parken 5000 kroner i veka å drive. For kvar person som vitjar den får dei inn 120 kroner.

c) Dersom parken ei veke hadde eit overskot, kor sannsynleg er det at det var opphaldsver den søndagen?

Oppgåve 3

I denne oppgåva skal vi sjå på modellering av trafikktettleik. På ei vegstrekning har vegvesenet sett opp apparatur som automatisk registrerer bilar som passerer mellom klokka 9:00 og 10:00 på kvardagar. Dei meiner at tidsavstanden mellom to påfølgjande bilar følgjer ei fordeling som liknar eksponensialfordelinga, men der avstanden ikkje kan komme under ein viss minste avstand. Altså for X, tida i sekund mellom to tilfeldig valde påfølgjande bilar, har vi sannsynstettleik

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

der d er parameteren for minimumsavstanden mellom 2 bilar.

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

No har vegvesenet registrert n slike tidsavstandar, $X_1, X_2, \ldots, X_n$. Under føresetnaden at desse er uavhengige skal dei nyttast til å estimere parametrane β og d.

b) Skriv opp sannsynsmaksimeringsfunksjonen (likelihoodfunksjonen) for denne situasjonen.
Argumenter ut frå sannsynsmaksimeringsfunksjonen at estimatoren for d må bli minimum av observasjonane:

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

Finn deretter (på vanleg måte) ein estimator for β.

Avstandane 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 visast (skal ikkje gjerast) at fordelinga til $\hat{d}$ er av samme type som fordelinga til X, einaste forskjell er at parameteren β er erstatta av $\frac{\beta}{n}$.

c) Bruk opplysningane du har til å finne forventninga til estimatorane for d og β.
Er dei forventningsrette? Om ikkje, konstruer forventningsrette estimatorar basert på dei.

Delvis integrasjon:

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