Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Bokmål

Faglig kontakt under eksamen:
Henning Omre 73 59 35 31
Hugo Hammer 452 10 184

EKSAMEN I EMNE TMA4240 STATISTIKK

Torsdag 8. desember 2005
Tid: 09:00–13:00

Hjelpemidler: *Tabeller og formler i statistikk*, Tapir Forlag
K. Rottmann: *Matematisk formelsamling*
Gult A5 ark (IMF stemplet) med egne håndskrevne notater
Kalkulator HP30S

Sensurfrist: 5. januar 2003

Oppgave 1 Søtsaker i matvareforretning

Det er kjent at matvareforretninger strategisk plasserer søtsaker i nærheten av kassene i forretningen for å hjelpe på salget. En butikkeier har til nå hatt søtsakene plassert tilfeldig rundt i butikken og erfart at kundene uavhengig av hverandre kjøper søtsaker med sannsynlighet $p_0 = 0.25$.

a) I løpet av en time handlet 20 personer i butikken hans. La X være antallet av disse kundene som handlet søtsaker. Under hvilke forutsetninger er X binomisk fordelt? Anta at X er binomisk fordelt, hva er da $P(X \leq 5)$ og $P(5 \leq X < 10)$?

Butikkeieren ønsker å teste om strategien med å plassere søtsakene i nærheten av kassene i butikken kan hjelpe på salget også i hans butikk, dvs. at personer handler søtsaker med større sannsynlighet enn $p_0 = 0.25$. Han plasserer derfor nå søtsakene i nærheten av kassene og observerer i en undersøkelse at av n tilfeldig valgte kunder handlet nå X kunder søtsaker. Anta at $X \sim b(x; n, p)$, dvs. binomisk fordelt med parametre n og p og at p er ukjent.

b) Vis at sannsynlighetsmaksimeringsestimatoren (maximum likelihood estimatoren) for p er

$$\hat{p} = \frac{X}{n}.$$

Regn ut forventning og varians til estimatoren $\hat{p}$.

Anta at $n = 18$ i undersøkelsen over og at 8 av disse kundene handlet søtsaker.

c) Selger butikkeieren mer søtsaker ved å plassere søtsakene i nærheten av kassene? Formuler problemet som en hypotesetest. Hva blir beslutningen i testen med et signifikansnivå 0.05? (Merk at vi *ikke* har tilstrekkelig antall observasjoner til at sentralgrenseteoremet kan benyttes i testen).

Oppgave 2 Kulepenn-forhandleren

En forhandler kjøper inn et stort parti kulepenner og selger det videre til kunder. Han får klage på alle kulepenner som ikke virker. Anta først at antallet klager, X, er poissonfordelt med parameter $\lambda > 0$:

$$p(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda} \quad ; x = 0, 1, 2, \ldots$$

a) Finn et uttrykk for sannsynligheten for at forhandleren får en eller flere klager. Utled et uttrykk for sannsynligheten for at forhandleren får mindre enn tre klager, gitt at han får en eller flere klager.

Leverandøren av kulepenner har tre fabrikker, benevnt A, B og C, som produserer med ulik kvalitet. Forhandleren mottar varepartiet fra en av fabrikkene men han vet ikke fra hvilken. Dersom varepartiet kommer fra fabrikk A vil antall klager være poissonfordelt med parameter $\lambda_A = 5$, hvis partiet er fra fabrikk B og C er klagefordelingsparameteren henholdsvis $\lambda_B = 15$ og $\lambda_C = 20$. Forhandleren antar i utgangspunktet at det er sannsynlighet $p_A = 0.5$ for at varepartiet kommer fra fabrikk A og sannsynlighet $p_B = p_C = 0.25$ for hver av fabrikkene B og C.

b) Utled et uttrykk for den marginale sannsynlighetsfordelingen for antall klager forhandleren kan regne med å få.

Utled uttrykk for både forventet antall klager og variansen i antall klager han mottar når han selger varepartiet med kulepenner. Sett inn verdiene for λ_A, λ_B og λ_C samt for p_A, p_B og p_C, og regn ut tallsvarene.

Etter å ha solgt hele varepartiet registrerer forhandleren at han har mottatt $x = 13$ klager.

c) Utled et uttrykk for sannsynligheten for at varepartiet kommer fra fabrikk B gitt all informasjonen spesifisert i oppgaven. Sett inn verdiene for λ_A, λ_B og λ_C, p_A, p_B og p_C samt x, og regn ut tallsvaret.

Oppgave 3 Slitasje på bildekk

En bildekkfabrikant ønsker å undersøke slitasjeraten på en viss type dekk. Han setter opp følgende modell som er gyldig for forholdsvis nye dekk:

$$Y = d_0 - \beta x + \varepsilon,$$

hvor Y representerer mønsterdybden på dekkene, x er antall kilometre kjørt, d_0 er en kjent, konstant mønsterdybde i nye dekk og β er den ukjente slitasjeraten som skal bestemmes. Feilleddet ε antas å være normalfordelt (Gaussisk) $n(\varepsilon; 0, \sigma)$ hvor variansen σ^2 er ukjent.

Fabrikanten planlegger å leie inn n testkjørere som skal bruke dekkene i forskjellige antall kilometre, x, og deretter måle mønsterdybde, Y. Observasjonene blir da $Y_1, Y_2, \dots, Y_n$ som er uavhengige og fordelte slik at $Y_i \sim n(y_i; d_0 - \beta x_i, \sigma)\,;\;\; i = 1, \dots, n$.

Fabrikanten bruker følgende estimatorer for de ukjente β og σ^2:

$$B = \frac{d_0 \sum_{i=1}^n x_i - \sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2},$$
$$S^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - d_0 + Bx_i)^2.$$

a) Vis at B er en forventningsrett estimator for β. Utled variansen til estimatoren B.

Estimatoren B er normalfordelt (Gaussisk), forklar hvorfor.

b) Utled et 95% konfidensintervall for slitasjeraten β.

Bildekk må ikke være for 'bløte' eller for 'harde'. Standarden tilsier at slitasjeraten skal være β_0. Skissér en testprosedyre for $H_0 : \beta = \beta_0$ mot $H_1 : \beta \neq \beta_0$ med signifikansnivå 0.05.