Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Faglig kontakt under eksamen:
Jo Eidsvik 901 27 472

EKSAMEN I FAG SIF5075 LEVETIDSANALYSE

Torsdag 22. mai 2003

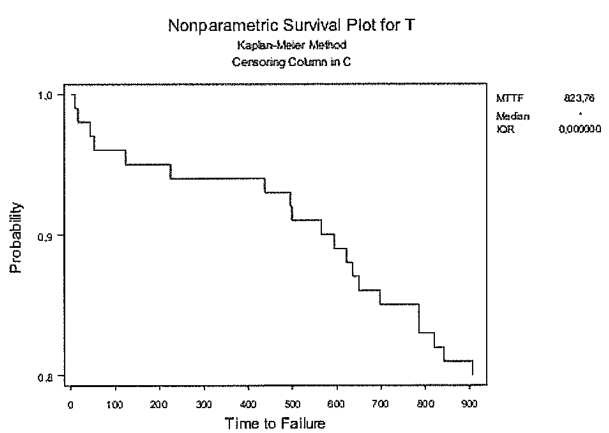
Tid: 09:00–14:00

Tillatte hjelpemidler:
Alle trykte og håndskrevne hjelpemidler, samt kalkulatorene HP30S og Casio fx-350W.

Sensur: 12. juni 2003

(MAI 2003: 1)

Oppgave 1



(MAI 2003: 1)

Levetiden T til en bestemt komponent er studert gjennom et laboratorieforsøk. Et tilfeldig utvalg på n=100 komponenter ble satt på test ved tidspunkt 0. Forsøket løp til tidspunktet $t_0 = 1000$ (timer). Da hadde $R = 20$ enheter feilet. Feiltidene $t_1, \ldots, t_{20}$ (i timer) for de feilende enhetene ble notert.

a) Hva kaller vi den type sensurering som ble brukt i forsøket?

Figuren på Side 1 av oppgavesettet viser et Kaplan-Meier plott (KM-plott) basert på de observerte feiltidene.

Hva er det som estimeres ved dette plottet, og hvilke forutsetninger bygger det på? Gjør kort rede for at disse er oppfylt i dette tilfellet.

Bruk plottet til å anslå $P(T \leq 400)$.

Definer p-kvantilen t_p ved $P(T \leq t_p) = p$ for $0 < p < 1$. Anslå $t_{0.10}$ fra KM-plottet.

Hvorfor kan ikke medianen $t_{0.50}$ estimeres ut fra KM-plottet?

Basert på fysisk kunnskap om feilmekanismene brukte ingeniørene en eksponensiell fordeling som modell for komponentens levetid. Levetiden T ble dermed antatt å ha tetthet

$$f(t;\theta) = \frac{1}{\theta} e^{-t/\theta};\ t > 0$$

der $\theta > 0$ er en ukjent parameter. Denne modellen skal brukes i resten av oppgaven.

b) Hvilken fortolkning har parameteren θ i den gitte modellen?

Finn likelihood-funksjonen $L(\theta)$ basert på forsøket beskrevet i begynnelsen av oppgaven. Begrunn utledningen av likelihood-uttrykket ut fra den type sensurering som er brukt.

Vis at log-likelihood funksjonen kan skrives

$$l(\theta) = -20 \log \theta - (1/\theta) \left(\sum_{i=1}^{20} t_i + 80 t_0 \right)$$

c) Finn et uttrykk for maksimum likelihood estimatoren $\hat{\theta}$ for θ.

Finn en estimator for standardavviket til $\hat{\theta}$.

Regn ut estimatet for θ når det er oppgitt at $\sum_{i=1}^{20} t_i = 9816$.

Hva blir da 95%-standardintervallet for θ?

(MAI 2003:1)

d) Finn en estimator for p-kvantilen t_p (definert i punkt a) for $0 < p < 1$.

Hva blir estimatet for $t_{0.10}$ når du bruker den oppgitte $\sum_{i=1}^{20} t_i$ fra forrige punkt? Sammenlign med svaret fra punkt (a).

Fra tidligere erfaring hadde man antatt at $t_{0.10} = 700$. For å undersøke eventuelle endringer ønsket man å teste

$$H_0 : t_{0.10} = 700 \text{ mot } H_1 : t_{0.10} \neq 700$$

Utled en test for dette ved å bruke log-likelihood funksjonen $l(\theta)$. Velg signifikansnivå tilnærmet 5%. Hva blir konklusjonen når du bruker den oppgitte $\sum_{i=1}^{20} t_i$ fra forrige punkt?

(Vink: Finn en verdi θ_0 av θ slik at H_0 er ekvivalent med hypotesen $\theta = \theta_0$).

e) Kan man estimere θ dersom man ikke kjenner de enkelte feiltidene $t_1, \ldots, t_{20}$, men bare vet at $R = 20$?

Foreslå i så fall en estimator $\tilde{\theta}$ som er en funksjon av bare R og t_0.

Oppgave 2

(MAI 2003:2)

Nedenfor er gitt tidspunkter for utskifting av ventilseter i fire dieselmotorer. For hver motor er angitt antall dager motoren har vært under observasjon (fra den var ny), samt tidspunktene (målt i antall dager siden observasjonsstart) for alle utskiftinger i denne tiden. Merk at for motor nr. 2 har det ikke vært utskiftinger.

Motor nr.	Dager under observasjon	Utskifting	
1	389	206	348
2	485		
3	606	408	604
4	648	61	

Det antas at tidspunktene for utskifting av ventilsete for en motor av denne typen kan modelleres ved en ikke-homogen Poisson-prosess (NHPP) med ROCOF $w(t)$ for $t \geq 0$. Videre antas at utskiftingstidspunkter for ulike motorer er stokastisk uavhengige.

a) Tegn Nelson-Aalen plottet for disse dataene. Hva er det som estimeres ved dette plottet? Forklar hvordan du beregner koordinatene til punktene i diagrammet.

Les av (omtrentlig!) fra plottet et anslag for forventet antall utskiftinger i løpet av ett år for en motor av denne typen.

(MAI 2003:2)

b) Regn ut estimerte standardavvik for hvert av punktene i plottet. Tegn inn tilnærmede 95% konfidensgrenser for hvert punkt.

c) Tyder plottet i punkt (a) på at det er noen sammenheng mellom alder og utskiftingshyppighet?

Man vil teste nullhypotesen at det ikke er noen slik sammenheng, mot den alternative hypotese at utskiftingshyppigheten er voksende. Formuler dette som hypoteser om ROCOF-funksjonen $w(t)$.

Gjør kort rede for en eller flere metoder for å teste dette basert på de gitte dataene. Du trenger ikke gjøre alle involverte beregninger.

Oppgave 3

En komponent antas under normalstress å ha overlevelsesfunksjon (reliability function) $R_0(t)$ og hasardfunksjon (sviktintensitet) $z_0(t)$.

Man ønsker å estimere påliteligheten av denne komponenttypen ved hjelp av akselerert levetidstesting. Dette gjøres ved å utsette komponenter for stress s, $0 \leq s < \infty$, og måle levetiden (eventuelt sensurerte levetider). Normalstress svarer til $s = 0$.

To modeller betraktes:

Modell 1: Proporsjonal hasardmodell. Under stress s har komponenten hasardfunksjon

$$z_s^{PH}(t) = z_0(t)g(s)$$

for en funksjon $g(s)$ med g(0)=1.

Modell 2: Akselerert levetidsmodell. Under stress s har komponenten overlevelsesfunksjon

$$R_s^{AL}(t) = R_0(\phi(s)t)$$

for en funksjon $\phi(s)$ med $\phi(0) = 1$.

a) Forklar kort hva som er hensikten med akselerert levetidstesting. Hva er ideen bak de to modellene? Hva uttrykker de to funksjonene $g(s)$ og $\phi(s)$?

Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Faglig kontakt under eksamen:
Bo Lindqvist 975 89 418

EKSAMEN I FAG TMA4275 LEVETIDSANALYSE

Lørdag 4. juni 2005

Tid: 09:00–13:00

Tillatte hjelpemidler:
Alle trykte og håndskrevne hjelpemidler. Alle kalkulatorer tillatt.

Sensur: 25. juni 2005.

JUNI 2005 : 1

Oppgave 1

En studie er gjennomført for å finne ut hvordan levetiden til en spesiell type bremseklosser avhenger av type kjøring. Seks biler av samme merke og årsmodell ble valgt ut, og typen av kjøring for hver av dem ble klassifisert ved variabelen w med verdier 0 = *hovedsakelig bykjøring*, 1 = *blandet kjøring*, eller 2 = *hovedsakelig motorveikjøring.*

Følgende data ble innhentet:

y_i	d_i	w_i	x_{i1}	x_{i2}
22.7	1	0	0	0
41.0	0	1	1	0
54.2	1	2	0	1
59.8	0	0	0	0
62.4	1	1	1	0
73.1	0	2	0	1

Her er y_i (med enhet 1000 km) den observerte bremseklosslevetid og d_i er sensureringsstatus for bil nummer i ($d_i = 1$ betyr ikke-sensurert, $d_i = 0$ betyr høyresensurert). Kovariaten w_i er verdien på variabelen w for type kjøring. I analysen vil w bli representert ved kovariatene x_1 og x_2, hvor $x_1 = 1$ hvis $w = 1$ og $x_1 = 0$ ellers; og $x_2 = 1$ hvis $w = 2$ og $x_2 = 0$ ellers. Verdiene for x_1 og x_2 for hver bil er for enkelhets skyld også gitt i tabellen.

(JUNI 2005 : 1)

I studien bruker man en Cox-modell for dataene.

a) Sett opp et uttrykk for hasardraten (feilraten) for levetiden til en bremsekloss som funksjon av tiden t og kovariatene x_1 og x_2. Hvilke antagelser ligger til grunn når en Cox-modell brukes for denne situasjonen?

Finn et uttrykk for pålitelighetsfunksjonen (overlevelsesfunksjonen) for en bremsekloss på en bil som brukes hovedsakelig på motorveier.

b) Beregn et uttrykk for Cox' partielle likelihood for de gitte dataene.

c) Det viser seg at den maksimale verdi for log partiell likelihood er −3.67. Bruk dette til å regne ut verdien for en testobservator for å teste nullhypotesen at type kjøring ikke har betydning for levetiden til bremseklossene.

Hva er konklusjonen på denne testen? Bruk signifikansnivå 5%.

JUNI 2005 : 2

Oppgave 2

Man har registrert tidspunktene for kritiske feil på en enkeltstående kompressor opp til $\tau = 650$ dager etter start. Det antas at reparasjoner er minimale og tar en neglisjerbar tid. Vi antar derfor at feiltidspunktene følger en ikke-homogen Poisson-prosess med intensitetsfunksjon $w(t)$ og kumulativ intensitetsfunksjon $W(t)$.

Følgende data ble observert (i dager etter start):

i	s_i
1	120
2	347
3	420
4	512
5	595

a) Hvordan ville du estimere den kumulative intensitet $W(t)$ når det ikke gjøres noen parametriske antagelser? Tegn den estimerte kurven for $W(t)$.

Tyder plottet på en trend i feilintensiteten?

Finn et tilnærmet 95% konfidensintervall for verdien til $W(595)$.

b) Gjennomfør Laplace-testen for å undersøke om det er trend i feiltidspunktene. Hva blir konklusjonen hvis du velger signifikansnivå 5%?

Hva kan fortegnet på denne testobservatoren si oss om eventuell voksende eller avtagende trend?

(JUNI 2005, 2)

I resten av oppgaven antar vi at ROCOF $w(t)$ for feilprosessen er av såkalt *loglineær* type, dvs.

$$w(t) = e^{\alpha+\beta t} \text{ for } t > 0,$$

der $\alpha > 0$ og $-\infty < \beta < \infty$ er ukjente parametre.

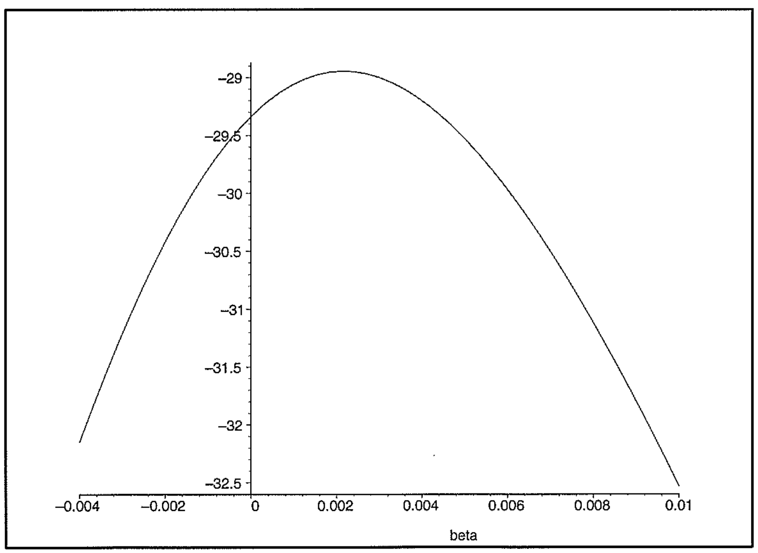
c) Vis at

$$W(t) = \frac{e^{\alpha}(e^{\beta t} - 1)}{\beta}.$$

Vis at log likelihood-funksjonen for dataene kan skrives

$$l(\alpha, \beta) = 5\alpha + 1994\beta - \frac{e^{\alpha}(e^{650\beta} - 1)}{\beta}.$$

d) Beregn profil log likelihood-funksjonen $\tilde{l}(\beta)$ for β.
En graf av $\tilde{l}(\beta)$ er gitt under.



JUNI 2005 : 2

Bruke grafen (grovt)

1. til å finne maksimum likelihood estimatet for β,
2. til å finne et tilnærmet 95% konfidensintervall for β,
3. til å teste, med signifikansnivå 5%, nullhypotesen at det ikke er noen trend i feilprosessen.

e) Anta at dataene isteden bare var gitt i intervaller, dvs. som følger:

Intervall	Antall feil
$(0, 100]$	0
$(100, 200]$	1
$(200, 300]$	0
$(300, 400]$	1
$(400, 500]$	1
$(500, 600]$	2
$(600, 650]$	0

Finn et uttrykk for log likelihood-funksjonen i dette tilfellet.

Oppgave 3

La $\Phi_0(w)$ og $\phi_0(w)$ være, henholdsvis, den kumulative fordelingsfunksjon og sannsynlighetstettheten for en stokastisk variabel W med mulige verdier i $(-\infty, \infty)$. Anta videre at $\phi_0(w) > 0$ for $w \in (-\infty, \infty)$.

a) Forklar hvordan $\Phi_0(w)$ og $\phi_0(w)$ brukes til å definere en *log-lokasjon-skala* familie av fordelinger for en levetid T, indeksert av parametre $-\infty < \mu < \infty$ og $\sigma > 0$.

Sett spesielt opp den kumulative fordelingsfunksjon og sannsynlighetstettheten for T, uttrykt ved $\Phi_0(w)$, $\phi_0(w)$, μ og σ.

La W ha fordeling gitt ved den standard *logistiske* fordeling, dvs. la

$$\Phi_0(w) = \frac{e^w}{1 + e^w} \text{ for } -\infty < w < \infty. \tag{1}$$

Det er kjent at den tilsvarende familie av fordelinger for T da kalles den *log-logistiske* familie.

b) Vis at den log-logistiske familie av fordelinger for T kan reparametriseres med $\alpha > 0$, $\gamma > 0$ slik at pålitelighetsfunksjonen (overlevelsesfunksjonen) kan skrives

$$R_T(t) = \frac{1}{1 + \gamma t^{\alpha}} \text{ for } t > 0.$$

Forklar hvordan dette kan brukes til å konstruere et sannsynlighetsplott for å sjekke om et høyresensurert datasett kan antas å komme fra den log-logistiske familie.

Problem 3

Let the lifetime T have the cumulative distribution function

$$F(t) = P(T \leq t) = e^{-\left(\frac{\theta}{t}\right)^{\alpha}} \quad \text{for } t > 0, \tag{2}$$

where $\alpha > 0$, $\theta > 0$ are parameters.

This distribution is known as the Frechet distribution, and is also called the inverse Weibull distribution because of the form of the cumulative distribution function.

a) Which requirements does a cumulative distribution function of a lifetime T have to satisfy? Check that these requirements are satisfied for $F(t)$ in (2).

Find the density $f(t)$ of T and write down an expression for the hazard rate $z(t)$ for T.

b) Compute $\ln(-\ln F(t))$.

Explain how this can be used to construct a probability plot for checking whether a right censored dataset can be assumed to come from the distribution (2).

Also explain how the parameters α and θ can be estimated graphically from this plot.

c) Show that (2) defines a log-location-scale family, i.e. that $Y = \ln T$ has a cumulative distribution function of the form

$$F_Y(y) = P(Y \leq y) = \Phi_0\left(\frac{y-\mu}{\sigma}\right)$$

for a cumulative distribution function Φ_0 and constants μ (location) og σ (scale).

Find the function Φ_0, and find the constants μ and σ expressed in terms of α and θ.

Faglig kontakt under eksamen:
Navn: Per Hokstad
Tlf.: (735)9 27 54,
90 58 41 32.

EKSAMEN I FAG SIF5075 LEVETIDSANALYSE

Lørdag 19. mai 2001
Tid: kl 0900 - 1400

Hjelpemidler: Godkjent lommekalkulator.
Statistiske tabeller (Tapir).
Tabeller og formler i statistikk (Tapir).
Se også vedlagt formelark (s.4).

SENSURDATO: Tirsdag 5. JUNI

Oppgave 1.

a) En har registrert tid til svikt (i måneder) for en ventil:

18, 32, 45, 55*, 60, 73, 87, 90*.

Måling nr. 4 og 8 som er merket med stjerne er sensurerte. Levetidene oppfattes som uavhengige identisk fordelte observasjoner.

Formuler sensureringsmodellen og finn Nelsons estimator for kumulativ sviktintensitet, $Z(t)=\int_0^t z(u)\,du$ i fordelingen. Gir plottet noen *indikasjon* på voksende/avtakende sviktintensitet?

b) Start med definisjonen av sviktintensitet, $z(t)$, og utled sammenhengen en har mellom $Z(t)$ og overlevelsesfunksjonen, $R(t)$. Finn Nelson-estimatoren for overlevelsesfunksjonen til dataene i a) og lag et plott.

MAI 2001 : 2

Oppgave 2.

En setter n komponenter i funksjon ved tid 0, og levetidene til disse komponentene, $T_1, T_2, \ldots, T_n$ antas uavhengige og identisk fordelte.

a) Anta at observasjonene har følgende sannsynlighetstetthet

$$f(t) = \tfrac{1}{2}\lambda^3 t^2 \exp(-\lambda t)$$

Hvilken fordelingsklasse hører denne fordelingen til? Hvilken sammenheng har denne til eksponensialfordelingen. Bruk sammenhengen til å påvise at $E(T) = 3/\lambda$.

(MAI 2001 : 2)

b) Utled og beregn sannsynlighetsmaksimeringsestimatoren (SME) for λ, når en har følgende 5 levetider (i år): 2.2, 2.9, 3.5, 3.8, 4.6.

c) Hva er de asymptotiske egenskapene til SME for λ? Bruk dette til å finne et tilnærmet 90% konfidensintervall for λ.

d) Gjennomfør sannsynlighetskvotetesten for H_0: $\lambda = 1.0$.

Oppgave 3

I denne oppgaven kan en bruke at en stokastisk variabel, X som er invers gamma-fordelt med parametre a og b har sannsynlighetstetthet

$$f(x)=\frac{b^a}{\Gamma(a)}\cdot\left(\frac{1}{x}\right)^{a+1}e^{-b/x}$$

og forventningen er $E(X) = b/(a-1)$; $a>1$.

a) En har observert tid til svikt/sensurering for n komponenter av samme type. Anta at tid til svikt, T, har sannsynlighetstetthet

$$f(t)=\tfrac{1}{\theta}\exp(-t/\theta)$$

Hva er her $E(T)$? Anta at observasjonene er fremkommet ved sensurering av type IV, og sett opp uttrykket for rimelighets-funksjonen (*likelihood*).

En vil estimere θ ved hjelp av Bayes-metodikken. Som a priori fordeling vil en bruke en invers gamma-fordeling. Hva blir uttrykket for a posteriori fordelingen for θ?

b) Utled uttrykket for Bayes-estimatoren for θ (når en bruker kvadratisk tapsfunksjon). Fra tidligere erfaring med denne type ventiler antas at $\theta \approx 5$ år (=60 mnd). Hvilken a priori fordeling bør statistikeren bruke hvis han også mener at denne informasjonen er tilnærmet "likeverdig" med å observere feil på komponenter med en samlet operasjonstid på 10 år?

Beregn så Bayes-estimatoren når observasjonene er som gitt i oppgave 1a).

c) Vis at når X er invers gammafordelt (a, b) vil $Z = 2b/X$ bli χ^2-fordelt med $2a$ frihetsgrader. Finn et 90% "troverdighetsintervall" (*credibility interval*) for θ.

c) Finn sannsynlighetsmaksimeringsestimatoren (SME) for λ og beregn estimatet du får med de gitte dataene.

Hvordan vil du estimere sannsynligheten for at en komponent av denne typen overlever 200 dager?

d) Man er usikker på om eksponensialfordelingen gir en tilfredsstillende tilpasning for disse dataene. Man betrakter derfor en Weibull-modell med parametre α og λ og tetthet

$$\alpha\lambda(\lambda t)^{\alpha-1}e^{-(\lambda t)^\alpha}, \quad t>0$$

som en alternativ modell. Når man setter SME for α og λ inn i log-likelihoodfunksjonen for Weibull-modellen (dvs. ln til rimelighetsfunksjonen), får man verdien -28.90.

Formuler hypoteser og gjennomfør en test for å undersøke om eksponensialmodellen bør forkastes som modell for dataene. Hva blir konklusjonen?

MAI 2002; 2

Oppgave 2

Betrakt en komponenttype der levetiden T for en ny komponent har tetthet

$$f(t) = te^{-t}, \quad t>0$$

a) Vis at sviktintensiteten for komponenten er gitt ved

$$z(t) = \frac{t}{1+t}, \quad t>0$$

Finn også overlevelsesfunksjonen $R(t)$ og vis at MTTF for komponenten er 2.

Komponenten repareres når den feiler og settes deretter i drift igjen. Anta at reparasjonstidene er neglisjerbare og derfor settes til 0. La $N(t)$ være antall feil for komponenten i tidsrommet $(0,t]$, definert for alle $t>0$.

b) Anta i dette punktet at komponenten etter hver reparasjon er så god som ny.

Hva slags prosess er $N(t)$ i dette tilfellet?

Hva er (tilnærmet) forventet antall feil i løpet av de 20 første tidsenheter? (Du kan her bruke et resultat som gjelder når tiden går mot uendelig).

c) Anta nå at det utføres bare en minimal reparasjon ved feil.

MAI 2002 : 2

Forklar kort hva dette betyr og forklar hvorfor vi da får en Poisson-prosess med ROCOF gitt ved

$$w(t) = \frac{t}{1+t}, \quad t > 0$$

Finn forventet antall feil for komponenten i løpet av de 20 første tidsenheter også under denne antagelsen.

Oppgave 3

En bestemt type ventiler er observert på ulike installasjoner i et oljefelt. La X være totalt antall svikt for denne ventiltypen i løpet av en samlet operasjonstid t. Det antas at X er Poisson-fordelt med parameter λt.

Parameteren λ skal estimeres ved hjelp av Bayes-metoden. Man oppfatter da λ som en realisasjon av en stokastisk variabel Λ. Som apriorifordeling for Λ brukes en gamma-fordeling med parametre $\alpha > 0$, $\beta > 0$ og tetthet

$$\frac{\beta^\alpha \lambda^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta\lambda}, \quad \lambda > 0$$

Du kan bruke at denne fordelingen har forventning α/β.

a) Hvilken fortolkning har parameteren λ for denne ventiltypen?

Hva blir aposteriorifordelingen for Λ når X er observert?

b) Finn Bayes-estimatoren for λ (basert på kvadratisk tapsfunksjon).

Forklar hvordan man i praksis kan bruke forhåndsinformasjon om ventilen til å velge verdier for α og β i apriorifordelingen.

Anta at man ender opp med å velge $\alpha = 4$, $\beta = 20$. Kan du gi dette valget en praktisk fortolkning?

c) Beregn Bayes-estimatet når $X = 11$, $t = 56$ år og apriorifordelingen har parametre som angitt til slutt i forrige punkt.

Sammenlign med den vanlige SME for λ.

Finn også et 90% "troverdighetsintervall" (credibility interval) for λ. Bruk at for gitt $X = x$ vil

$$Z = 2(\beta + t)\Lambda$$

være tilnærmet kjikvadratfordelt med $2(\alpha + x)$ frihetsgrader.