

ELINAIKA- JA TAPAHTUMAHISTORIAMALLIEN TEORIAA

Matematiikan ja tilastotieteen laitos

Tentti 6.3.2007

Vastaa neljään seuraavista kysymyksistä

1. Vastaa lyhyesti mutta täsmällisesti, mielellään kaavoilla, seuraaviin kysymyksiin. Mitä tarkoittavat

- (a) Kiihdytetyn ajan regressiomalli?
- (b) Verrannollisten hasardien malli?
- (c) Epäinformatiivinen sensurointi?

2. Mallin diagnostiikka: Miten tutkit

- (a) Coxin mallin verrannollisuusoletuksen pätevyyttä?
- (b) Parametroidun, esim. eksponentiaalisen mallin jakaumaoletusta?
- (c) Mallin kykyä ennustaa tapahtumien lukumäärää?

3. Tavalliselle lineaariselle mallille pätee, että $X'(y - X\hat{\beta}) = 0$ eli mallin systemaattinen (selitetty) osa on korreloimaton satunnaisen osan (residuaalin) kanssa. Kun Coxin mallin log-osittaisuskottavuusfunktio kirjoitetaan tapahtumaindikaattorin d_i avulla muotoon $\sum_{i=1}^n d_i(z'_i\beta - \log \sum_{i=1}^n \exp(z'_i\beta))$ niin osoita, että sen score funktiolle pätee

$$\sum_{i=1}^n z_i(d_i - \exp(z'_i\hat{\beta})\hat{\Lambda}_0(t_i)) = 0,$$

jossa $\hat{\Lambda}_0(t_i)$ on Breslow'n estimaattori perushasardille ja $d_i = 1$ kun tapahtuma toteutuu ja 0 muuten. Mitkä kurssilla esitellyistä residuaaleista siis vastaavat tavallisen lineaarisen mallin residuaaleja $y - X\hat{\beta}$?

4. Oletetaan, että aikuisiän diabetespotilailla tutkitaan riskiä sairastua joko sydän- ja verisuonitauteihin, silmänpohjan rappeumatautiin tai munuaisten vajaatoimintaan. Miten tässä kilpailevien riskien tutkimusasetelmassa ilmaistaan (hasardifunktioita käyttäen) todennäköisyys sairastua sydän- ja verisuonitauteihin 70. ikävuoteen mennessä diabeteksen diagnostisoinnista?

5. Hepatiittipotilaiden hoidossa testattiin steroiditerapian hyödyllisyyttä siten, että satunnaistettiin potilaat steroidi- ja kontrolliryhmään. Seuraavat havainnot ovat steroidiryhmän potilaiden kuolinaikoja (+ sensuroitu). Seurannan pituus on 12 viikkoa.

1,1,1,2+,4,5+,7,8,10,10+,11+,12+,12+,12+

Sovitetaan aineistoon eksponenttimallia $f(t) = \lambda \exp(-\lambda t)$.

a) Estimoi kuolinintensiteetti λ suurimman uskottavuuden menetelmällä ja muodosta asymptoottinen 95%-luottamusväli λ :lle (Opastus: käytä ensin log-muunnosta $\log \hat{\lambda} \sim N(\log \lambda, 1/d)$, jossa d on tapausten lukumäärä, ja käännä sitten tulos λ :aa koskevaksi).

b) Estimoi eloonjäämistodennäköisyys seurannan päättyessä ja muodosta sille edellisen kohdan perusteella asymptoottinen 95%-luottamusväli.

ELINAIKA- JA TAPAHTUMAHISTORIAMALLIEN TEORIAA

Matematiikan ja tilastotieteen laitos

Uusintatentti 12.4.2007

Vastaa neljään seuraavista kysymyksistä

1. Vastaa lyhyesti mutta täsmällisesti, mielellään kaavoilla, seuraaviin kysymyksiin.

Mitä tarkoittavat

- (a) Hasardifunktio?
- (b) Kiihdytetyn ajan malli (accelerated life time model)?
- (c) Martingaaliresiduaali?

2. Sensurointi ja sen vaikutukset päättelyyn elinaika-analyysissä. Kirjoita uskottavuusfunktio sensuroidulle aineistolle tilanteessa, jossa sensurointi on tapahtumisen suhteen informatiivista ja tilanteessa, jossa sen voidaan olettaa olevan epäinformatiivista (tavallinen oletus). Opastus: Määrittele informatiivisessa tilanteessa 'sensurointitapahtuman' tiheysfunktio $g(t)$ ja käytä sitä uskottavuusfunktion osana.

3. Tarkastellaan riippumattomia, mahdollisesti sensuroituja, havaintoja kahdesta populaatiosta, joista n_1 on realisaatioita satunnaismuuttujasta hasardifunktiolla λ_1 ja n_2 hasardifunktiolla λ_2 . Johda suurimman uskottavuuden estimaattorit hasardifunktiolle ja uskottavuusosamäärätesti (likelihood ratio test) hypoteesille $\lambda_1 = \lambda_2$.

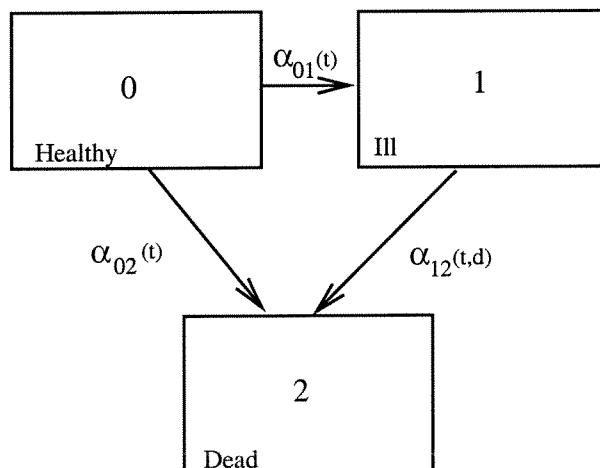
4. Coxin malli, osittaisuskottavuusfunktio ja sen tulkinta

5. Oletetaan, että aikuisiän diabetespotilailla tutkitaan riskiä sairastua joko sydän- ja verisuonitauteihin, silmänpohjan rappeumatautiin tai munuaisten vajaatoimintaan. Kirjoita tässä kilpailevien riskien tutkimusasetelmassa (hasardifunktioita käyttäen) todennäköisyys sairastua sydän- ja verisuonitauteihin 70. ikävuoteen mennessä diabeteksen diagnostisoinnista?

Tapahtumahistoriamallien sovelluksia
Loppukoe 22.5.2007

- Selitä (mahdollisesti esimerkkien avulla), mitä tarkoittavat: (a) suljettu kohortti ja avoin kohortti; (b) vasemmalta sensurointi (*left censoring*); (c) vasemmalta katkaisu (*left truncation*).
- Millainen suhde on seuraavien epidemiologiassa usein käytettyjen suureiden välillä: ilmaantuvuus, vallitsevuus ja tilassa vietetty keskimääräinen aika? Anna pari esimerkkiä.
- Tarkastellaan kolmen kilpailevan kuolinsyyn mallia. Yhteensä $n:n$ yksilön riippumattomiksi oletettavat havainnot ovat (T_i, D_i) , $i = 1, \dots, n$, jossa $D_i = j$, jos hetkellä T_i yksilö i siirtyy (lähtötilasta "0") absorboivaan tilaan j , ja $D_i = 0$, jos kyseessä on oikealta sensurointi. (a) Merkitään siirtymäkohtaisia hasardeja $\alpha_j(t)$, $j = 1, 2, 3$ (hasardit siirtymille tilasta "0" tiloihin $j = 1, 2, 3$). Kirjoita hasardien uskottavuusfunktio. (b) Jos hasardit $\alpha_j(t)$, $j = 1, 2, 3$, on annettu, mikä on todennäköisyys, että yksilö on siirtynyt tilaan j aikaan t mennessä?
- Oheisen kuvan mukaisessa tilasiirtymämallissa on tehty seuraavat havainnot (seurantavälillä hetkestä 0 hetkeen T):

Siirtymä	Siirtymiä N
$0 \rightarrow 1$	4
$0 \rightarrow 2$	6
$1 \rightarrow 2$	4



Tiloissa 0 ja 1 kertyneet henkilöajat ovat $Y_0 = 60$ ja $Y_1 = 40$ (aikayksikköä). Oletetaan, että kuoleman vakiohasardi ei riipu lähtötilasta, ajasta t eikä kestosta d eli $\alpha_{02}(t) = \alpha_{12}(t, d) = \alpha$. Lisäksi siirtymän $0 \rightarrow 1$ hasardi oletetaan vakioksi $\alpha_{01}(t) = \beta$. Laske hasardien α ja β piste-estimaatit ja keskivirheet.