

Tautikartoitus gaussisilla prosesseilla

Loppuseminaari: Terveysthuollon uudet
analyysimenetelmät (TERANA)

Jarno Vanhatalo

jarno.vanhatalo@tkk.fi



HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Biomedical Engineering and Computational Science

1.4.2009

- Tavoitteena visualisoida alueellisia eroja tauti-/kuolemariskissä
- Analyysin vaiheet
 - 1 poistetaan väestörakenteesta johtuva vaihtelu
 - 2 luodaan tilastollinen malli vielä selittämättä jääneelle vaihtelulle
 - 3 esitetään tulokset kartalla, josta näkee alueellisen vaihtelun riskissä
- Tulosten tulkinta
 - tuloksissa esitetään riski suhteessa koko maan keskiarvoon

Bayesilaisen tautikartoitusmallin osat:

- havaintomalli
 - esim. poisson jakauma
- alueellinen prioril
 - kertoo, kuinka samankaltaisia lähekkäin olevien alueiden oletetaan olevan

⇒ Posteriori

- Tietämys alueellisista eroista havaitun aineiston ja priorioletusten yhteisvaikutuksena ("siloitettu suhteellinen riski")

Bayesilaisen tautikartoitusmallin osat:

- havaintomalli
 - esim. poisson jakauma
 - alueellinen priorii
 - kertoo, kuinka samankaltaisia lähekkäin olevien alueiden oletetaan olevan
- ⇒ Posteriori
- Tietämys alueellisista eroista havaitun aineiston ja priorioletusten yhteisvaikutuksena ("siloitettu suhteellinen riski")

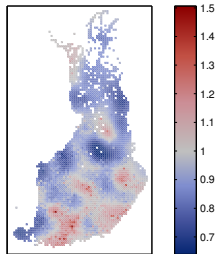
- (a) kuolemantapauksia eniten isoissa keskuksissa, joissa paljon ihmisiä
- (b) Kuolemantapausten lukumäärä suhteessa väestörakenteen mukaiseen odotusarvoon on kohinainen
- (c) Sattuman vaikutus poistettu



(a) Kuolemantapausten lukumäärä



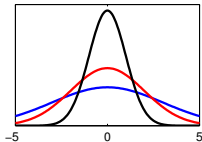
(b) Raaka suhteellinen riski



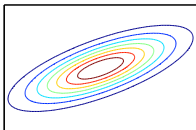
(c) Siloitettu suhteellinen riski

Gaussinen prosessi

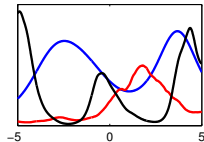
- Stokastinen prosessi. Tehokas rakennuspalikka moniin bayesilaisiin malleihin
- Yleistää gaussin jakauman
- Voidaan käyttää määrittelemään todennäköisyyksiä funktioille



(d) Yksiulotteinen gaussin jakauma



(e) kaksiulotteinen gaussin jakauma

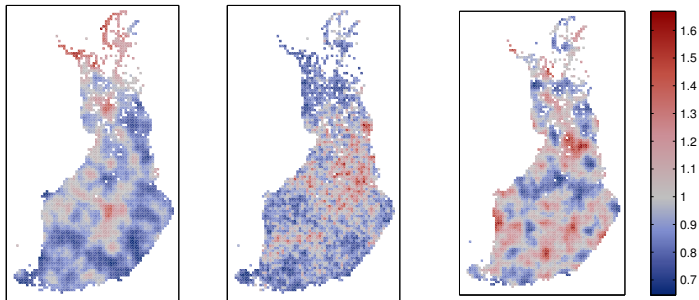


(f) Satunnaisesti näytteistettyjä funktioita gaussisesta prosessista

Gaussinen prosessi tautikartoituksessa

Priori oletukset

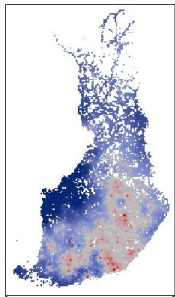
Esimerkkejä gaussinen prosessi priorista näytteistetyistä alueellisista riskipinnoista.



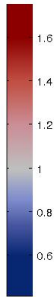
Gaussinen prosessi tautikartoituksessa

Posteriori tietämys

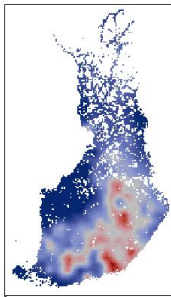
Alueellinen riski alkoholikuolemista. Priorioletuksena, että riskissä voi olla pitkän ja lyhyen korrelaation vaihtelua.



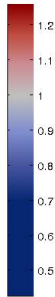
(j) kokonaisriski



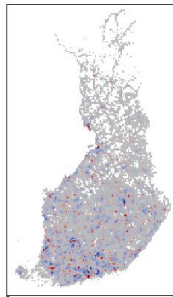
=



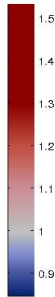
(k) Pitkän korrelaation komponentti



×



(l) Lyhyen korrelaation komponentti



Bayesilainen mallintaminen periaatteessa yksinkertaista ja suoraviivaista.

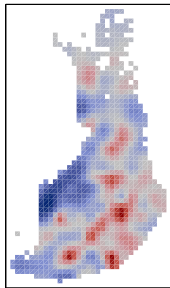
Käytännön ongelmana laskenta:

- Bayesilainen mallintamisen ongelma on integrointi
 - yksinkertaisissakin malleissa integraalit ratkaisemattomia
 - ⇒ numeeriset ja analyttiset aproksimaatiot
- Gaussiset prosessit laskennallisesti raskaita
 - lineaarialgebraa ja integraaleja hyvin isoissa ulottuvuuksissa
 - naivilla toteutuksella laskenta hidastuu nopeasti datan määrän lisääntyessä
 - ⇒ Harvat gaussiset prosessit, lineaarialgebra harvoilla matriiseilla, aproksimaatiot integraaleille

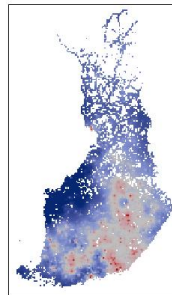
Gaussiset prosessit TERANA projektissa

Laskennan nopeutus

- 2006: 20km resoluutio (900 datapistettä), analyysi 2-3 vrk
2009: 5km resoluutio (10 500 datapistettä), analyysi 20 min



(m) Tilanne 2006.



(n) Tilanne 2009.

Kuva: Suhteellinen riski alkoholiperäisissä taudeissa.