

Johdatus tilastolliseen päättelyyn
Harjoitus 6 (2.–6. 5. 2011)

1. Fyysikko A on kiinnostunut tietyistä fysikaalisista suureista θ . Kun A pohtii ennakkokäsitystään suureesta θ , niin hän päätyy siihen, että hyvä approksimaatio on normaalijakauma, jonka keskikohta on 800. Lisäksi A:n mielestä 50%:n varmuudella θ on välillä (700, 900). (Tämän tulkitsemme tarkoittavan sitä, että priorijakauman mukaan $P(700 < \theta < 900) = 0.5$.) Auta fyysikkoo A laskemalla hänen priorijakaumansa parametrit. Varianssiparametrille riittää kahden merkitsevän numeron tarkkuus.

2. Myös fyysikko B on kiinnostunut suureesta θ . B:n ennakkokäsitystä θ :sta kuvaa normaalijakauma $N(800, 60^2)$, eli normaalijakauma odotusarvolla 800 ja varianssilla 60^2 . Kyseisestä fysikaalisesta suureesta tehtiin mittauksia X_i , jotka ovat ehdollisesti riippumattomia ehdolla θ ja joille $X_i | \theta \sim N(\theta, 40^2)$. Mittauksia tehtiin $n = 10$ kertaa, ja niiden keskiarvoksi saatiin $\bar{x} = 850$.

Johda fyysikon B posteriorijakauma suureelle θ . Johda myös sellainen väli, joka sisältää 95% fyysikon B posteriorijakaumasta, ja jonka kumpaankin häntään jää siitä 2,5%. (Posteriorijakauma on tietty normaalijakauma, minkä takia välin määrittäminen on helppoa.)

3. Fyysikolla C on paljon hatarampi käsitys suureesta θ kuin fyysikolla A tai B. Fyysikko C päättää esittää ennakkokäsityksensä ns. epäaidolla tasajakaumalla koko reaaliakselilla. Voimme esittää tämän ennakkotiedon formaalisti "tiheysfunktioilla", jonka arvo $p(\theta) = 1$, kun θ on mikä tahansa reaaliarvo. C saa käyttöönsä saman uskottavuusfunktion kuin fyysikko B.

Johda fyysikon C posteriorijakauma. Laske sellainen väli, joka sisältää 95% fyysikon C posteriorijakaumasta, ja jonka kumpaankin häntään jää siitä 2,5%. Vertaa tulosta frekventistiseen luottamusväliin samassa tilanteessa (eli laske normaalijakauman odotusarvon luottamusväli, kun varianssi on tunnettu).

KÄÄNNÄ!

4. Mallinnamme viimeisessä kahdessa tehtävässä katastrofaalisten ydinreaktorionnettomuuksien sattumista maailmassa. Mallin parametrina θ on katastrofien sattumistiheys yhtä reaktorivuotta kohta. Mallimme mukaan katastrofeja sattuu kussakin reaktorissa Poissonin prosessin mukaan intensiteetillä θ , ja eri reaktorit ovat (ehdollisesti) riippumattomia toisistaan. (Varoitus: malli olettaa, että kaikki reaktorit ovat yhtä herkkiä tuottamaan onnettomuuksia ja että ne ovat keskenään riippumattomia. Molemmat oletukset ovat epärealistisia, minkä takia mallin perusteella saataviin johtopäätöksiin pitää suhtautua varauksella. Malli perustuu saksalaisten tilastotieteen professoreiden Kauermann ja Küchenhoff sanomalehdessä *Frankfurter Allgemeine Zeitung* julkaisemaan yleistajuiseen artikkeliin.)

Tämän mallin seuraus on se, että jos tarkastellaan N :ä reaktoria, joita kutakin käytetään T vuotta, niin katastrofien lukumäärään X jakauma on

$$X \mid \theta \sim \text{Poisson}(NT\theta),$$

jossa NT on reaktorivuosien lukumäärä.

Viimeisen $T = 30$ vuoden aikana maailmassa on sattunut kaksi ydinreaktorikatastrofia, joiksi laskemme Tsernobylin nelosreaktorin onnettomuuden 26.4.1986 ja Fukushima I - ydinvoimalan onnettomuuden maaliskuussa 2011. (Fukushiman kohdalla laskemme onnettomuuden katastrofiksi vain yhdelle kompleksin kuudesta voimalasta.) Tänä 30 vuoden ajanjakson aikana maailmassa on ollut käytössä noin $N = 420$ ydinreaktoria. (Arvo NT on yliarvio todellisille reaktorivuosille; tämä johtaa riskin θ aliarvioimiseen.) Käytämme priorina gammajakaumaa $\text{Gamma}(\alpha_0, \lambda_0)$.

Osoita, että posteriorijakauma on tietty gammajakauma, ja laske sen parametrit. Mikä on posteriorijakauman odotusarvo?

5. Eräässä maassa S on $M = 4$ ydinreaktoria, eikä niiden lukumäärä muutu seuraavien 50 vuoden aikana. Laske edellisen tehtävän mallin mukainen ennustetodennäköisyys sille, että tulevien 50 vuoden aikana maassa S tapahtuu yksi tai useampia ydinreaktorikatastrofeja. Käytämme tässä tehtävässä ensin priorina gammajakaumaa $\text{Gamma}(1, 250000)$ eli eksponenttijakaumaa, jonka odotusarvo on $1/250000$ (mikä vastaa Kauermannin ja Küchenhoffin artikkelissa annettua ydinvoimateollisuuden arvioita onnettomuuden riskistä: yksi katastrofi 250000 reaktorivuotta kohti) ja sen jälkeen epäaitoa tasajakaumaa positiivisella reaaliakselilla, joka voidaan formaalisti esittää gammajakaumalla $\text{Gamma}(1, 0)$.

Opastus: jos Y on maassa S sattuvien ydinreaktorikatastrofien lukumäärä, niin ehdolla θ havaintoa 2 vastaava satunnaismuuttuja X ja tulevien katastrofien lukumäärä Y ovat riippumattomia, ja

$$Y \mid \theta \sim \text{Poisson}(MT_1\theta),$$

jossa $T_1 = 50$. Laske ennustetodennäköisyys $P(Y = 0 \mid X = 2)$ sille, että maassa S ei tapahdu lainkaan ydinreaktorikatastrofeja, ja vähennä tulos luvusta yksi.