HY / Matematiikan ja tilastotieteen laitos
Todennäköisyyslaskenta II, syksy 2016
Harjoitus 12 – Ratkaisuehdotuksia

Tehtäväsarja I

1. Oletetaan, että tietä pitkin havaintokohdan ohi mittauspäivinä i ajaa y_i autoa ja päivänä i havainnoidaan x_i tunnin ajan. Tässä tehtävässä käytämme mallia, jonka mukaan y_i:t ovat riippumattomien Poissonin jakaumia noudattavien satunnaismuuttujien Y_i havaittuja arvoja, joiden jakauma on $Y_i \sim \text{Poi}(\lambda x_i)$, $i = 1, \ldots, n$. Parametri λ kertoo jotain liikenteen intensiteetistä.

   a) Kirjoita uskottavuusfunktio $\lambda \mapsto f(\mathbf{y} \mid \lambda)$, missä $\mathbf{y} = (y_1, \ldots, y_n)$.

   b) sekä etsi parametrin suurimman uskottavuuden estimaatti λ^{ML}.

   Ratkaisu: Koska havainnot y_i ovat riippumattomia, uskottavuusfunktio $\lambda \mapsto f(\mathbf{y} \mid \lambda)$ saadaan seuraavasti:

$$f(\mathbf{y} \mid \lambda) = \prod_{i=1}^{n} f(y_i \mid \lambda) = \prod_{i=1}^{n} \frac{(\lambda x_i)^{y_i}}{(y_i)!} e^{-\lambda x_i},$$

   mikä kelpaa meille uskottavuusfunktioksi ja on λ:n funkio.

   Koska etsimme suurimman uskottavuuden estimaattia, etsimme siis parametrille λ sitä arvoa, mikä maksimoi yllä olevan uskottavuusfunktion. Tämä saadaan helposti tehtyä etsimällä uskottavuusfunktion derivaatan nollakohta ja tarkistamalla funktion konkaavisuudesta, että tämä kriittinen piste on todella maksimoi funktion.

   Laskujen helpottamiseksi otamme logaritmin uskottavuusfunktiosta. Näin saamme *logaritmisen uskottavuusfunktion*, joka maksimoituu samassa pisteessä kuin alkuperäinen uskottavuusfunktio mutta on helpompi derivoida ja käsitellä.

$$l(\mathbf{y} \mid \lambda) = \log f(\mathbf{y} \mid \lambda) = \sum_{i=1}^{n} \log(\frac{(\lambda x_i)^{y_i}}{(y_i)!} e^{-\lambda x_i}) \propto n\bar{y}(\log(\lambda)) - \lambda n\bar{x}, \tag{1}$$

   missä $\bar{y}$ on y:n otoskeskiarvo $\frac{1}{n}\sum_{i=1}^{n} y_i$ (x:lle vastaavasti), ja verrannollisuuskertoimella on poistettu lausekkeesta kaikki λ:sta riippumattomat termit (logaritmisessa uskottavuusfunktiossa vakiosta riippuvat termit eivät vaikuta maksimoivaan kohtaan. Miksi?). Näin saimme helposti derivoituvan funktion.

$$l'(\mathbf{y} \mid \lambda) = \frac{n\bar{y}}{\lambda} - n\bar{x} = 0 \Leftrightarrow \lambda = \frac{\bar{y}}{\bar{x}},$$

   ja koska $l''(\bar{y}) = -\frac{n\bar{y}}{\lambda^2} < 0$ kaikilla $\bar{y} > 0, \lambda > 0$, niin funktio l on konkaavi ja siten löydetty lokaali ääriarvokohta on todellinen maksimi. Siten SUE on

$$\lambda^{ML} = \frac{\bar{y}}{\bar{x}}.$$

   Huomaamme, että jos keskiarvo $\bar{y} = 0$, suurimman uskottavuuden estimaattime ei kuulu parametriavaruuteen $(0, \infty)$. Tämä tietenkin osoittautuu monella tapaa ongelmalliseksi.

Testiasetelmassa pyrimme mallintamaan havaintoja y_i, ja arvelemme niiden olevan Poisson-jakautuneiden satunnaismuuttujien realisaatioituneita arvoja. Tilanne $\bar{y} = 0$ voi syntyä, jos tilanne on seuraava: Yliopisto on saanut lisärahoituksen valtiolta, ja tutkimusryhmämme on saanut rahoituksen tutkimukseen jossa opiskelija T istuu viikon pientareella, tutkien havaintokohdan ohi ajavien autojen määrää. Jos tänä aikana havaintopistettä ei ole ohittanut yhtään autoa, tutkimusryhmä voi helposti päätellä että havainnot eivät ole Poisson-jakautuneita (tai että parametri λ on niin pieni, että testaus ei ole järkevää) ja mallimme siten viallinen.

2. Käsittelemme nyt edellisen tehtävän Bayes-versiota. Parametri $\Lambda > 0$ on jatkuvasti jakautunut satunnaismuuttuja. Ehdolla Λ sm:t $Y_1, \dots, Y_n$ ovat ehdollisesti riippumattomia, ja $Y_i \mid (\Lambda = \lambda) \sim \text{Poi}(\lambda x_i)$, $i = 1, \dots, n$.

   Parametrin Λ priorijakauma on $\text{Gam}(\alpha, \beta)$, jossa $\alpha, \beta > 0$ ovat vakioita. Johda posteriorijakauma $\Lambda \mid (\mathbf{Y} = \mathbf{y})$, jossa $\mathbf{Y} = (Y_1, \dots, Y_n)$ on havaintoja vastaava satunnaisvektori ja $\mathbf{y} = (y_1, \dots, y_n)$, jossa luvut $y_i \geq 0$ ovat havaittuja lukumääriä.

   Ratkaisu: Merkitään prioria $f_\Lambda(\lambda)$, uskottavuusfunktiota $f_{\mathbf{Y}|\Lambda}(\mathbf{y} \mid \lambda)$ ja $\mathbf{y}$:n yhteistodennäköisyyttä $f_\mathbf{Y}(\mathbf{Y})$ Kaivettuamme ulkomuistista Gammajakauman tiheysfunktion, saamme unnetusti Bayesin kaavan mukaan posteriorijakauma voidaan esittää muodossa:

$$f_{\Lambda|\mathbf{Y}}(\lambda \mid \mathbf{y}) = \frac{f_\Lambda(\lambda) f_{\mathbf{Y}|\Lambda}(\mathbf{y} \mid \lambda)}{f_\mathbf{Y}(\mathbf{Y})} \propto f_\Lambda(\lambda) f_{\mathbf{Y}|\Lambda}(\mathbf{y} \mid \lambda) \propto \lambda^{n\bar{y}} e^{-\lambda n \bar{x}} \lambda^{\alpha - 1} e^{-\beta\lambda}$$
$$= \lambda^{n\bar{y} + \alpha - 1} e^{-\lambda(n\bar{x} + \beta)},$$

   mikä tunnistetaan vakiota vaille jakauman $\text{Gam}(n\bar{y} + \alpha, n\bar{x} + \beta)$ tiheysfunktioksi. Ensimmäisen verrannollisuuskertoimen kohdalla pystyimme jättämään Γ:sta riippumattoman nimittäjän pois, ja toisen kohdalla nappasimme priori-ja uskottavuusfunktiosta λ:sta riippumattomat vakiot pois. Teknisesti ottaen tämä posteriorijakauma ei ole *todennäköisyysjakauma*, sillä se ei integroidu ykköseksi yli $\mathbb{R}$:n mutta *stokastisilta ominaisuuksiltaan* se on hyvinkin passeli. Esimerkiksi saadun "posteriorijakauman" moodi on samassa pisteessä kun vakiot sisältävän posteriorijakauman moodi, mikä esimerkiksi kiinnostaa usein Bayesiläisen tilastotieteen harjoittajia.

3. Tietystä populaatiosta satunnaisesti valitun henkilön ikä A (vuosia), pituus L (cm) ja paino W (kg) mallinnetaan kolmiulotteisella normaalijakaumalla siten, että sv:lle $\mathbf{V} = (L, A, W)$

$$\mathbb{E}\mathbf{V} = \begin{pmatrix} 170 \\ 23 \\ 68 \end{pmatrix} \quad \text{Cov}(\mathbf{V}) = \begin{pmatrix} 49 & 0 & 63 \\ 0 & 81 & 42 \\ 63 & 42 & 121 \end{pmatrix}$$

   a) Mikä on pituuden L (reuna-)jakauma?

   b) Johda ehdollinen jakauma $W|(A = a, L = l)$. Mikä on 32 vuotta vanhojen ja 160 cm pitkien henkiöiden painojakauma?

   c) Ovatko henkilöiden ikä ja pituus riippumattomia?

Ratkaisu: a) Pituuden reunajakauman tiedetään olevan normaalijakautunut, ja multinormaalijakautuneen $\mathbf{V}$:n odotusarvosta ja kovarianssimatriisista poimimme suoraan arvot:

$$\mathbb{E}(L) = 170, \text{ ja } \text{Cov}(L) = \text{var}(L) = 49. \tag{2}$$

b)-kohdassa käytämme lausetta 10.6, mikä antaa meille näppärästi ehdollistetun multinormaalijakauman jakauman. Merkitään $U = (A, L)$,ja nyt halutuksi jakaumaksi saamme

$$W \mid (U = u) \sim N(\mu_w + \Sigma_{wu}\Sigma_{uu}^{-1}(u - \mu_u), \Sigma_{ww} - \Sigma_{wu}\Sigma_{uu}^{-1}\Sigma_{uw}),$$

missä

$$\mu_w = 68, \mu_u = \begin{pmatrix} 23 & 170 \end{pmatrix}^T, \Sigma_{wu} = \begin{pmatrix} 42 & 63 \end{pmatrix}, \Sigma_{uw} = \begin{pmatrix} 42 & 63 \end{pmatrix}^T \text{ ja } \Sigma_{uu} = \begin{pmatrix} 81 & 0 \\ 0 & 49 \end{pmatrix}.$$

Tästä voimme laskea auki varianssin ja odotusarvon annetuille pituuden ja iän arvoille:

$$\begin{aligned} Var(W \mid (A = 32, L = 160)) = \Sigma_{ww} - \Sigma_{wu}\Sigma_{uu}^{-1}\Sigma_{uw} &= 121 - \begin{pmatrix} 42 & 63 \end{pmatrix} \begin{pmatrix} 1/81 & 0 \\ 0 & 1/49 \end{pmatrix} \begin{pmatrix} 42 & 63 \end{pmatrix}^T \\ &= 121 - \left(\frac{42^2}{81} + \frac{63^2}{49}\right) = 121 - \frac{196}{9} - 81 \\ &\approx 18,22, \end{aligned}$$

ja vastaavasti odotusarvoksi saadaan

$$\begin{aligned} \mathbb{E}(W \mid (A = 32, L = 160)) = \mu_w + \Sigma_{wu}\Sigma_{uu}^{-1}(u - \mu_u) = \\ 68 + \begin{pmatrix} 42 & 63 \end{pmatrix} \begin{pmatrix} 1/81 & 0 \\ 0 & 1/49 \end{pmatrix} \begin{pmatrix} 32 - 23 & 160 - 170 \end{pmatrix} = \\ 68 + \frac{14}{3} - \frac{90}{7} \approx 59.80. \end{aligned}$$

Tehtävänannon kovarianssimatriisista huomaamme, että $\text{cov}(A, L) = 0$. Nyt lauseen 10.4 perusteella A ja L ovat riippumattomia, sillä (A, L) on normaalijakautunut satunnaisvektori.

4. Olkoon $Q \in \mathbb{R}^{n\times n}$ ortogonaalivakiomatriisi. Tämä tarkoittaa, että Q on kääntyvä ja $Q^{-1} = Q^\top$. Määritellään $\mathbf{X} = 4Q\mathbf{U}$, missä $\mathbf{U} \sim N_n(0, I_n)$ on standardinormaalijakautunut satunnaisvektori. Oletetaan, että $n \geq 10$ ja muodostetaan satunnaisvektori $\mathbf{Y} = (X_1, \dots, X_6)$.

   a) Mitä jakaumaa $\mathbf{X}$ noudattaa?

   b) Määrää satunnaismuuttujan $W = \mathbf{Y}^\top\mathbf{Y}/16$ jakauma?

Ratkaisu: a)-kohdassa huomaamme, että $\mathbf{X}$ on affiini muunnos multinormaalijakautunelle satunnaismuuttujalle $\mathbf{U}$, joten luentojen lauseen 10.2 mukaan

$$\mathbf{U} \sim N_n(0, 4QI_n(4Q)^T) = N_n(0, 16I_n),$$

sillä Q on ortogonaalivakiomatriisi, eli $QQ^T = QQ^{-1} = I_n$. Nyt b-kohdassa huomaamme, että $\frac{1}{4}X_i \sim N(0, 1)$, ja merkitään $\frac{1}{4}X_i =: Z_i$.

Nyt huomaamme, että

$$W = \mathbf{Y}^T\mathbf{Y}/16 = \frac{1}{4}(X_1 \cdots X_6)^T \frac{1}{4}(X_1 \cdots X_6) = (Z_1 \cdots Z_6)^T (Z_1 \cdots Z_6) = Z_1^2 + \cdots + Z_6^2,$$

ja edelleen muistamme, että standardinormaalijakautuneiden satunnaismuuttujien neliöiden summa on Khi-toiseen jakatunut n-vapausasteella, joten

$$W \sim \chi_6^2.$$

5. Hieman epäsymmetristä yksinkertaista jatkuvatilaista satunnaiskulun tilaa X_n ajan hetkellä $n = 1, 2, 3, \dots, 10000$ kuvaa sen siirtymien summa

$$X_n = \mu n + \sum_{k=1}^{n} W_k$$

missä $W_k \sim N_1(0, 1/1000)$ ja satunnaismuuttujat $W_1, W_2, \dots, W_{10000}$ ovat riippumattomia.

a) Esitä satunnaismuuttujat $X_1, \dots, X_{10000}$ ja $W_1, \dots, W_{10000}$ satunnaisvektoreina $\mathbf{X}$ ja $\mathbf{W}$ ja määrää sitten matriisi A sekä vektori λ jolla $\mathbf{X} = A\mathbf{W} + \lambda$. (Matriisihan on valtaisa, joten riittää hyvin, että käyttää $\dots$ merkintöjä tai sitten voi kertoa, mikä on A_{ij} kullakin rivillä i ja sarakkeella j. Sama pätee vektoriin λ).

b) Määritä satunnaisvektorin X jakauma.

Ratkaisu: Alkuun toteamme, että $\mathbf{W}$ noudattaa moniulotteista monijakaumaa, $W \sim N_{10000}(0, \frac{1}{1000} I_{10000})$ missä I_{10000} on 10000×10000 kokoinen diagonaalimatriisi. Nyt etsimme sellaiset A, λ, että $\mathbf{X} = A\mathbf{W} + \lambda$ pätee. Huomaamme, että

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & \ddots & 0 & 0 \\ \vdots & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} W_1 \\ \vdots \\ W_{10000} \end{pmatrix} + \begin{pmatrix} \mu \\ \vdots \\ 10000\mu \end{pmatrix} = \begin{pmatrix} \Sigma W_1 \\ \vdots \\ \Sigma W_{10000} \end{pmatrix} + \begin{pmatrix} \mu \\ \vdots \\ 10000\mu \end{pmatrix} = \mathbf{X}$$

eli

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & \ddots & 0 & 0 \\ \vdots & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \text{ ja } \lambda = \begin{pmatrix} \mu \\ \vdots \\ 10000\mu \end{pmatrix},$$

missä A on 10000×10000 matriisi, jolla on ”yksikköalakolmio” ja λ on 10000×1-vektori. Kohdassa b)-meidän pitää ratkaista satunnaismatriisin $\mathbf{X}$ jakauma. Lauseen 10.2 mukaan $\mathbf{X}$ noudattaa multinormaalijakaumaa, joka saadaan matriisin $\mathbf{W}$ muunnoksesta. Nähdään, että

$$\mathbb{E}X = A\mathbb{E}\mathbf{W} + \mathbb{E}\lambda = \lambda,$$

ja

$$\mathrm{Cov}(\mathbf{X}) = A\mathrm{Cov}(\mathbf{W})A^T = \frac{1}{1000} AA^T = \frac{1}{1000} \begin{pmatrix} 1 & 1 & \cdots & & 1 \\ 1 & 2 & \cdots & & 2 \\ 1 & 2 & 3\cdots & & 3 \\ \vdots & & & \ddots & \\ 1 & 2 & 3 & \cdots & 10000 \end{pmatrix} =: B,$$

eli $X \sim N_10000(\lambda, B)$.

6. Jatkoa edelliseen tehtävään. Laske sm:n X_{1005}:n ehdollinen tiheys, ehdolla, että $X_5 = a$. (opastus: voit käyttää Luentojen lausetta 10.6 tai voit etsiä satunnaismuuttujan V, jolle $X_{1005} = V + X_5$ ja jolle $X_5 \perp\!\!\!\perp V$, tällöin X_{1005}:n jakauma ehdolla $X_5 = a$ on sama kuin sm:n $V + a$ jakauma)

Ratkaisu: Käytämme tässä tehtävässä käytämme tehtävän 3 tapaan lausetta 10.6. Etenemme sen merkinnöin, ja pidemmittä puheitta merkitään $Y = X_{1005}$, $Z = X_5$ ja edelleen

$$\mu_y = 1005\mu, \mu_z = 5\mu, \Sigma_{yy} = \frac{1}{1000}1005, \Sigma_{zz} = \frac{1}{1000}5, \Sigma_{yz} = \frac{1}{1000}1005.$$

Saamme siis odotusarvoksi

$$\mathbb{E}(X_{1005} \mid X_5 = a) = \mu_y + \Sigma_{yz}\Sigma_{zz}^{-1}(a - \mu_z) = a + 1000\mu$$

ja varianssiksi

$$\operatorname{var}(X_{1005} \mid X_5 = a) = \Sigma_{yy} - \Sigma_{yz}\Sigma_{zz}^{-1}\Sigma_{zy} = 1,$$

joten viimein

$$X_{1005} \mid X_5 = a \sim N(a + 1000\mu, 1).$$

Kiitos kurssista ja ahkerasta tehtävien teosta, nähdään III-periodin kurssilla Tilastollinen päättely II!