Lineaaristen mallien kurssi - kesätentti 9.8.2007

1. Määrittele yleinen (täysiasteinen) lineaarinen malli ja siihen liittyvät käsitteet (pienimmän neliösumman) sovite ja residuaali. Selvitä muutamalla virkkeellä mallin sekä sovitteen ja residuaalin tulkinta ja laske lisäksi $\mathsf{Cov}(\hat{\varepsilon})$ ja $\mathsf{Cov}(\hat{\varepsilon}, \hat{\mathbf{Y}})$, kun $\hat{\varepsilon}$ on aineistosta laskettua residuaalivektoria vastaava satunnaisvektori ja $\hat{\mathbf{Y}}$ on aineistosta laskettua sovitevektoria vastaava satunnaisvektori.

2. Olkoon $Y_1, ..., Y_5$ riippumattomia normaalisti jakautuneita satunnaismuuttujia, joille pätee $Y_i \sim \mathsf{N}(\mu_i, \sigma^2)$ ja

$$\mu_i = \begin{cases} \beta_1, & \text{kun } i = 1,2 \\ \beta_1 + 2\beta_2, & \text{kun } i = 3 \\ \beta_1 - \beta_2, & \text{kun } i = 4,5. \end{cases}$$

Muotoile tilanne lineaarisena mallina ja estimoi parametrit β_1 ja β_2 soveltamalla lineaarisen mallin estimointiteoriaa. Kuvaa lyhyesti käyttämäsi estimointiperiaate ja selvitä parametrivektorin $\boldsymbol{\beta} = [\beta_1 \;\; \beta_2]'$ estimaattorin $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1 \;\; \hat{\beta}_2]'$ jakauma. Ovatko $\hat{\beta}_1$ ja $\hat{\beta}_2$ riippumattomia?

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ $(n > 2)$ on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n-1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n-1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n-1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right\},$$

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi_k^2 / k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$. Lisäksi $\mathsf{E}(\chi_k^2) = k$, $\mathsf{Var}\left(\chi_k^2\right) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z/\sqrt{\frac{1}{k}\chi_k^2}$, josssa $Z \sim \mathsf{N}(0,1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k\left(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k\right)$ ja matriisi $\mathbf{P}$ $(k \times k)$ astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X} - \boldsymbol{\mu})' \mathbf{P} (\mathbf{X} - \boldsymbol{\mu}) / \sigma^2 \sim \chi_r^2$.

Lineaaristen mallien kurssi - yleistentti 18.12.2007

1. Määrittele yleinen (täysiasteinen) lineaarinen malli ja siihen liittyvät käsitteet (pienimmän neliösumman) sovite ja residuaali. Selvitä muutamalla virkkeellä mallin sekä sovitteen ja residuaalin tulkinta ja laske $\hat{\mathbf{y}}'\hat{\boldsymbol{\varepsilon}}$, kun $\hat{\mathbf{y}} = [\hat{y}_1 \ \cdots \ \hat{y}_n]'$ on aineistosta laskettu sovitevektori, $\hat{\boldsymbol{\varepsilon}} = [\hat{\varepsilon}_1 \ \cdots \ \hat{\varepsilon}_n]'$ on aineistosta laskettu residuaalivektori ja n on havaintojen lukumäärä.

2. Olkoon $Y_1, ..., Y_5$ riippumattomia normaalisti jakautuneita satunnaismuuttujia, joille pätee $Y_i \sim \mathsf{N}(\mu_i, \sigma^2)$ ja

$$\mu_i = \begin{cases} \beta_1, & \text{kun } i = 1,2 \\ \beta_1 + 2\beta_2, & \text{kun } i = 3 \\ \beta_1 - \beta_2, & \text{kun } i = 4,5. \end{cases}$$

Muotoile tilanne lineaarisena mallina ja estimoi parametrit β_1 ja β_2 soveltamalla lineaarisen mallin estimointiteoriaa. Kuvaa lyhyesti käyttämäsi estimointiperiaate ja selvitä parametrivektorin $\boldsymbol{\beta} = [\beta_1 \ \ \beta_2]'$ estimaattorin $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1 \ \ \hat{\beta}_2]'$ jakauma. Ovatko $\hat{\beta}_1$ ja $\hat{\beta}_2$ riippumattomia?

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ $(n > 2)$ on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n-1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n-1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n-1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\},$$

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi_k^2 / k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$. Lisäksi $\mathsf{E}(\chi_k^2) = k$, $\mathsf{Var}(\chi_k^2) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z / \sqrt{\frac{1}{k}\chi_k^2}$, jossa $Z \sim \mathsf{N}(0,1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k)$ ja matriisi $\mathbf{P}$ $(k \times k)$ astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X} - \boldsymbol{\mu})' \mathbf{P} (\mathbf{X} - \boldsymbol{\mu}) / \sigma^2 \sim \chi_r^2$.

Lineaaristen mallien kurssi - yleistentti 04.03.2008

1. Yleisen lineaarisen mallin lähtökohdaksi voidaan ottaa malliyhtälö, jonka matriisiesitys on $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$. Mallin oletukset eivät kuitenkaan selviä tästä. Esitä kaikki mallista tehtävät oletukset ja selvitä muutamalla virkkeellä mallin tulkinta. Selvitä lisäksi mitä tarkoitetaan käsitteillä (pienimmän neliösumman) residuaali ja sovite. Laske lisäksi $\mathsf{Cov}(\hat{\varepsilon})$ ja $\mathsf{Cov}(\hat{\varepsilon}, \hat{\mathbf{Y}})$, kun $\hat{\varepsilon}$ on aineistosta laskettua residuaalivektoria vastaava satunnaisvektori ja $\hat{\mathbf{Y}}$ on aineistosta laskettua sovitevektoria vastaava satunnaisvektori.

2. Oletetaan, että tehtävässä 1 matriisin $\mathbf{X}$ sarakkeet ovat ortogonaalisia eli kohtisuorassa toisiaan vastaan. Estimoi parametrivektori $\boldsymbol{\beta}$ soveltamalla lineaarisen mallin estimointiteoriaa ja selvitä estimaattorisi todennäköisyysjakauma. Muodosta lisäksi $100(1-\alpha)\%$:n luottamusväli parametrille β_j ($= \boldsymbol{\beta}$:n j. komponentti) ja selvitä miten se muuttuu, jos matriisista $\mathbf{X}$ poistetaan k. sarake ($k \neq j$).

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ ($n > 2$) on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n-1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n-1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n-1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right\},$$

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k=1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi^2_k / k\chi^2_m$, jossa $\chi^2_k \perp\!\!\!\perp \chi^2_m$.
  Lisäksi $\mathsf{E}(\chi^2_k) = k$, $\mathsf{Var}(\chi^2_k) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z/\sqrt{\frac{1}{k}\chi^2_k}$, josssa $Z \sim \mathsf{N}(0,1)$ ja $Z \perp\!\!\!\perp \chi^2_k$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}-\boldsymbol{\mu}) \sim \chi^2_k$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k)$ ja matriisi $\mathbf{P}$ ($k \times k$) astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X}-\boldsymbol{\mu})' \mathbf{P} (\mathbf{X}-\boldsymbol{\mu}) / \sigma^2 \sim \chi^2_r$.

Lineaaristen mallien kurssi - yleistentti 03.04.2008

1. Määrittele yleinen (täysiasteinen) lineaarinen malli ja siihen liittyvät käsitteet (pienimmän neliösumman) sovite ja residuaali. Selvitä muutamalla virkkeellä mallin sekä sovitteen ja residuaalin tulkinta ja laske $\hat{\mathbf{y}}'\hat{\boldsymbol{\varepsilon}}$, kun $\hat{\mathbf{y}} = [\hat{y}_1 \ \cdots \ \hat{y}_n]'$ on aineistosta laskettu sovitevektori, $\hat{\boldsymbol{\varepsilon}} = [\hat{\varepsilon}_1 \ \cdots \ \hat{\varepsilon}_n]'$ on aineistosta laskettu residuaalivektori ja n on havaintojen lukumäärä.

2. Olkoon $Y_1, ..., Y_5$ riippumattomia normaalisti jakautuneita satunnaismuuttujia, joille pätee $Y_i \sim \mathsf{N}(\mu_i, \sigma^2)$ ja

$$\mu_i = \begin{cases} \beta_1, & \text{kun } i = 1, 2 \\ \beta_1 + 2\beta_2, & \text{kun } i = 3 \\ \beta_1 - \beta_2, & \text{kun } i = 4, 5. \end{cases}$$

Muotoile tilanne lineaarisena mallina ja estimoi parametrit β_1 ja β_2 soveltamalla lineaarisen mallin estimointiteoriaa. Kuvaa lyhyesti käyttämäsi estimointiperiaate ja selvitä parametrivektorin $\boldsymbol{\beta} = [\beta_1 \ \ \beta_2]'$ estimaattorin $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1 \ \ \hat{\beta}_2]'$ jakauma. Ovatko $\hat{\beta}_1$ ja $\hat{\beta}_2$ riippumattomia?

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ $(n > 2)$ on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n-1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n-1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n-1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right\},$$

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi^2_k / k\chi^2_m$, jossa $\chi^2_k \perp\!\!\!\perp \chi^2_m$. Lisäksi $\mathsf{E}(\chi^2_k) = k$, $\mathsf{Var}(\chi^2_k) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z/\sqrt{\frac{1}{k}\chi^2_k}$, josssa $Z \sim \mathsf{N}(0,1)$ ja $Z \perp\!\!\!\perp \chi^2_k$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2_k$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k)$ ja matriisi $\mathbf{P}$ $(k \times k)$ astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X} - \boldsymbol{\mu})' \mathbf{P} (\mathbf{X} - \boldsymbol{\mu}) / \sigma^2 \sim \chi^2_r$.

Lineaaristen mallien kurssi - yleistentti 13.05.2008

1. Yleisen lineaarisen mallin lähtökohdaksi voidaan ottaa malliyhtälö, jonka matriisiesitys on $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Mallin oletukset eivät kuitenkaan selviä tästä. Esitä kaikki mallista tehtävät oletukset ja selvitä muutamalla virkkeellä mallin tulkinta. Selvitä lisäksi mitä tarkoitetaan käsitteillä (pienimmän neliösumman) residuaali ja sovite. Laske lisäksi $\mathsf{Cov}(\hat{\varepsilon})$ ja $\mathsf{Cov}(\hat{\varepsilon}, \hat{\mathbf{Y}})$, kun $\hat{\varepsilon}$ on aineistosta laskettua residuaalivektoria vastaava satunnaisvektori ja $\hat{\mathbf{Y}}$ on aineistosta laskettua sovitevektoria vastaava satunnaisvektori.

2. Oletetaan, että tehtävässä 1 matriisin $\mathbf{X}$ sarakkeet ovat ortogonaalisia eli kohtisuorassa toisiaan vastaan. Estimoi parametrivektori $\boldsymbol{\beta}$ soveltamalla lineaarisen mallin estimointiteoriaa ja selvitä estimaattorisi todennäköisyysjakauma. Muodosta lisäksi $100(1-\alpha)\%$:n luottamusväli parametrille β_j ($= \boldsymbol{\beta}$:n j. komponentti) ja selvitä miten se muuttuu, jos matriisista $\mathbf{X}$ poistetaan k. sarake ($k \neq j$).

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ ($n > 2$) on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n-1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n-1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n-1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on
$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\},$$
jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi_k^2 / k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$.
Lisäksi $\mathsf{E}(\chi_k^2) = k$, $\mathsf{Var}(\chi_k^2) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z/\sqrt{\frac{1}{k}\chi_k^2}$, josssa $Z \sim \mathsf{N}(0,1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k)$ ja matriisi $\mathbf{P}$ $(k \times k)$ astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X} - \boldsymbol{\mu})' \mathbf{P} (\mathbf{X} - \boldsymbol{\mu}) / \sigma^2 \sim \chi_r^2$.

Lineaaristen mallien kurssin kesätentti 12.06.2008

1. Määrittele yleinen lineaarinen malli kaikkine oletuksineen ja siihen liittyen käsitteet (pienimmän neliösumman) sovite ja residuaali. Laske lisäksi $\mathsf{Cov}(\hat{\varepsilon})$, $\mathsf{Cov}(\hat{\mathbf{Y}})$ ja $\mathsf{Cov}(\hat{\varepsilon}, \hat{\mathbf{Y}})$, kun $\hat{\varepsilon}$ on aineistosta laskettua residuaalivektoria vastaava satunnaisvektori ja $\hat{\mathbf{Y}}$ on aineistosta laskettua sovitevektoria vastaava satunnaisvektori.

2. Olkoon $Y_1, ..., Y_5$ riippumattomia normaalisti jakautuneita satunnaismuuttujia, joille pätee $Y_i \sim \mathsf{N}(\mu_i, \sigma^2)$ ja

$$\mu_i = \begin{cases} \beta_1, & \text{kun } i = 1, 2 \\ \beta_1 + \beta_2, & \text{kun } i = 3, 4 \\ \beta_1 - 2\beta_2, & \text{kun } i = 5. \end{cases}$$

Muotoile tilanne lineaarisena mallina ja johda parametrien β_1, β_2 ja σ^2 suurimman uskottavuuden estimaatit.

3. Olkoon $Y_1, ..., Y_4$ riippumattomia normaalisti jakautuneita satunnaismuuttujia, joille pätee $Y_i \sim \mathsf{N}(i\beta, \sigma^2)$ $(i = 1, ..., 4)$. Muotoile tilanne lineaarisena mallina ja johda testi hypoteesille $H : \beta = 1$ vaihtoehtoa $\beta \neq 1$ vastaan.

4. Oletetaan, että satunnaismuuttujat $Y_{11}, ..., Y_{1n_1}, Y_{21}, ..., Y_{2n_2}, Y_{31}, ..., Y_{3n_3}$ ovat riippumattomia ja $Y_{ji} \sim \mathsf{N}\left(\mu_j, \sigma^2\right)$, $i = 1, ..., n_j$, $j = 1, 2, 3$. Muotoile tilanne lineaarisena mallina ja johda testi hypoteesille $H : \mu_1 = \mu_2 = \mu_3$.

Perustele vastauksesi mahdollisimman hyvin.

Oheismateriaalia

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

  $$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right\},$$

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.
- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi_k^2 / k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$.
  Lisäksi $\mathsf{E}(\chi_k^2) = k$, $\mathsf{Var}\left(\chi_k^2\right) = 2k$.
- t_k-jakauman määrittelee satunnaismuuttuja $Z / \sqrt{\frac{1}{k}\chi_k^2}$, jossa $Z \sim \mathsf{N}(0, 1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.
- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_k^2$.
- Jos $\mathbf{X} \sim \mathsf{N}_k\left(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k\right)$ ja matriisi $\mathbf{P}$ $(k \times k)$ astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X} - \boldsymbol{\mu})' \mathbf{P} (\mathbf{X} - \boldsymbol{\mu}) / \sigma^2 \sim \chi_r^2$.

Lineaaristen mallien kurssi - kesätentti 14.08.2008

1. Yleisen lineaarisen mallin lähtökohdaksi voidaan ottaa malliyhtälö, jonka matriisiesitys on $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Mallin oletukset eivät kuitenkaan selviä tästä. Esitä kaikki mallista tehtävät oletukset ja selvitä muutamalla virkkeellä mallin tulkinta. Selvitä lisäksi mitä tarkoitetaan käsitteillä (pienimmän neliösumman) residuaali ja sovite. Laske lisäksi $\mathsf{Cov}(\hat{\varepsilon})$ ja $\mathsf{Cov}(\hat{\varepsilon}, \hat{\mathbf{Y}})$, kun $\hat{\varepsilon}$ on aineistosta laskettua residuaalivektoria vastaava satunnaisvektori ja $\hat{\mathbf{Y}}$ on aineistosta laskettua sovitevektoria vastaava satunnaisvektori.

2. Oletetaan, että tehtävässä 1 matriisin $\mathbf{X}$ sarakkeet ovat ortogonaalisia eli kohtisuorassa toisiaan vastaan. Estimoi parametrivektori $\boldsymbol{\beta}$ soveltamalla lineaarisen mallin estimointiteoriaa ja selvitä estimaattorisi todennäköisyysjakauma. Muodosta lisäksi $100(1-\alpha)\%$:n luottamusväli parametrille β_j ($= \boldsymbol{\beta}$:n j. komponentti) ja selvitä miten se muuttuu, jos matriisista $\mathbf{X}$ poistetaan k. sarake ($k \neq j$).

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ ($n > 2$) on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n-1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n-1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n-1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right\},$$

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi_k^2 / k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$.
  Lisäksi $\mathsf{E}(\chi_k^2) = k$, $\mathsf{Var}(\chi_k^2) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z/\sqrt{\frac{1}{k}\chi_k^2}$, josssa $Z \sim \mathsf{N}(0,1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}-\boldsymbol{\mu}) \sim \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k)$ ja matriisi $\mathbf{P}$ ($k \times k$) astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X}-\boldsymbol{\mu})' \mathbf{P} (\mathbf{X}-\boldsymbol{\mu}) / \sigma^2 \sim \chi_r^2$.

Lineaaristen mallien kurssi - yleistentti 21.10.2008

1. Yleisen lineaarisen mallin lähtökohdaksi voidaan ottaa malliyhtälö, jonka matriisiesitys on $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Mallin oletukset eivät kuitenkaan selviä tästä. Esitä kaikki mallista tehtävät oletukset ja selvitä muutamalla virkkeellä mallin tulkinta. Selvitä lisäksi mitä tarkoitetaan käsitteillä (pienimmän neliösumman) residuaali ja sovite. Laske lisäksi $\mathsf{Cov}(\hat{\boldsymbol{\varepsilon}})$ ja $\mathsf{Cov}(\hat{\boldsymbol{\varepsilon}}, \hat{\mathbf{Y}})$, kun $\hat{\boldsymbol{\varepsilon}}$ on aineistosta laskettua residuaalivektoria vastaava satunnaisvektori ja $\hat{\mathbf{Y}}$ on aineistosta laskettua sovitevektoria vastaava satunnaisvektori.

2. Oletetaan, että tehtävässä 1 matriisin $\mathbf{X}$ sarakkeet ovat ortogonaalisia eli kohtisuorassa toisiaan vastaan. Estimoi parametrivektori $\boldsymbol{\beta}$ soveltamalla lineaarisen mallin estimointiteoriaa ja selvitä estimaattorisi todennäköisyysjakauma. Muodosta lisäksi $100(1-\alpha)\%$:n luottamusväli parametrille β_j ($= \boldsymbol{\beta}$:n j. komponentti) ja selvitä miten se muuttuu, jos matriisista $\mathbf{X}$ poistetaan k. sarake ($k \neq j$).

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ ($n > 2$) on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n-1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n-1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n-1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on
$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\},$$
jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi^2_k / k\chi^2_m$, jossa $\chi^2_k \perp\!\!\!\perp \chi^2_m$. Lisäksi $\mathsf{E}(\chi^2_k) = k$, $\mathsf{Var}(\chi^2_k) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z/\sqrt{\frac{1}{k}\chi^2_k}$, josssa $Z \sim \mathsf{N}(0,1)$ ja $Z \perp\!\!\!\perp \chi^2_k$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2_k$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \sigma^2\mathbf{I}_k)$ ja matriisi $\mathbf{P}$ ($k \times k$) astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X} - \boldsymbol{\mu})' \mathbf{P} (\mathbf{X} - \boldsymbol{\mu}) / \sigma^2 \sim \chi^2_r$.

Lineaaristen mallien kurssi 12.11.2008

1. Yleisen lineaarisen mallin lähtökohdaksi voidaan ottaa malliyhtälö, jonka matriisiesitys on $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Mallin oletukset eivät kuitenkaan selviä tästä. Esitä kaikki mallista tehtävät oletukset ja selvitä muutamalla virkkeellä mallin tulkinta. Selvitä lisäksi mitä tarkoitetaan käsitteillä (pienimmän neliösumman) residuaali ja sovite. Laske lisäksi $\mathsf{Cov}(\hat{\boldsymbol{\varepsilon}})$ ja $\mathsf{Cov}(\hat{\boldsymbol{\varepsilon}}, \hat{\mathbf{Y}})$, kun $\hat{\boldsymbol{\varepsilon}}$ on aineistosta laskettua residuaalivektoria vastaava satunnaisvektori ja $\hat{\mathbf{Y}}$ on aineistosta laskettua sovitevektoria vastaava satunnaisvektori.

2. Oletetaan, että tehtävässä 1 matriisin $\mathbf{X}$ sarakkeet ovat ortogonaalisia eli kohtisuorassa toisiaan vastaan. Estimoi parametrivektori $\boldsymbol{\beta}$ soveltamalla lineaarisen mallin estimointiteoriaa ja selvitä estimaattorisi todennäköisyysjakauma. Muodosta lisäksi $100(1-\alpha)\%$:n luottamusväli parametrille β_j ($= \boldsymbol{\beta}$:n j. komponentti) ja selvitä miten se muuttuu, jos matriisista $\mathbf{X}$ poistetaan k. sarake ($k \neq j$).

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ ($n > 2$) on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n-1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n-1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n-1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right\},$$

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi_k^2 / k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$.
  Lisäksi $\mathsf{E}(\chi_k^2) = k$, $\mathsf{Var}(\chi_k^2) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z / \sqrt{\frac{1}{k}\chi_k^2}$, josssa $Z \sim \mathsf{N}(0,1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}-\boldsymbol{\mu}) \sim \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k)$ ja matriisi $\mathbf{P}$ $(k \times k)$ astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X}-\boldsymbol{\mu})' \mathbf{P} (\mathbf{X}-\boldsymbol{\mu}) / \sigma^2 \sim \chi_r^2$.

Lineaaristen mallien kurssi - tentti 16.12.2008

1. Yleisen lineaarisen mallin lähtökohdaksi voidaan ottaa malliyhtälö, jonka matriisiesitys on $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Mallin oletukset eivät kuitenkaan selviä tästä. Esitä kaikki mallista tehtävät oletukset ja selvitä muutamalla virkkeellä mallin tulkinta. Selvitä lisäksi mitä tarkoitetaan käsitteillä (pienimmän neliösumman) residuaali ja sovite. Laske lisäksi $\mathsf{Cov}(\hat{\boldsymbol{\varepsilon}})$ ja $\mathsf{Cov}(\hat{\boldsymbol{\varepsilon}}, \hat{\mathbf{Y}})$, kun $\hat{\boldsymbol{\varepsilon}}$ on aineistosta laskettua residuaalivektoria vastaava satunnaisvektori ja $\hat{\mathbf{Y}}$ on aineistosta laskettua sovitevektoria vastaava satunnaisvektori.

2. Oletetaan, että tehtävässä 1 matriisin $\mathbf{X}$ sarakkeet ovat ortogonaalisia eli kohtisuorassa toisiaan vastaan. Estimoi parametrivektori $\boldsymbol{\beta}$ soveltamalla lineaarisen mallin estimointiteoriaa ja selvitä estimaattorisi todennäköisyysjakauma. Muodosta lisäksi $100(1-\alpha)\%$:n luottamusväli parametrille β_j ($= \boldsymbol{\beta}$:n j. komponentti) ja selvitä miten se muuttuu, jos matriisista $\mathbf{X}$ poistetaan k. sarake ($k \neq j$).

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ ($n > 2$) on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n-1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n-1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n-1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right\},$$

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi_k^2 / k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$.
  Lisäksi $\mathsf{E}(\chi_k^2) = k$, $\mathsf{Var}(\chi_k^2) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z / \sqrt{\frac{1}{k}\chi_k^2}$, josssa $Z \sim \mathsf{N}(0,1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}-\boldsymbol{\mu}) \sim \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k)$ ja matriisi $\mathbf{P}$ ($k \times k$) astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X}-\boldsymbol{\mu})' \mathbf{P} (\mathbf{X}-\boldsymbol{\mu}) / \sigma^2 \sim \chi_r^2$.

Lineaaristen mallien kurssi - yleistentti 3.3.2009

1. Määrittele yleinen (täysiasteinen) lineaarinen malli kaikkine oletuksineen ja siihen liittyvät käsitteet (pienimmän neliösumman) sovite ja residuaali. Selvitä muutamalla virkkeellä mallin sekä sovitteen ja residuaalin tulkinta ja laske $\hat{\mathbf{y}}'\hat{\boldsymbol{\varepsilon}}$, kun $\hat{\mathbf{y}} = [\hat{y}_1 \ \cdots \ \hat{y}_n]'$ on aineistosta laskettu sovitevektori, $\hat{\boldsymbol{\varepsilon}} = [\hat{\varepsilon}_1 \ \cdots \ \hat{\varepsilon}_n]'$ on aineistosta laskettu residuaalivektori ja n on havaintojen lukumäärä.

2. Olkoon $Y_1, ..., Y_5$ riippumattomia normaalisti jakautuneita satunnaismuuttujia, joille pätee $Y_i \sim \mathsf{N}(\mu_i, \sigma^2)$ ja

$$\mu_i = \begin{cases} \beta_1, & \text{kun } i = 1, 2 \\ \beta_1 + 2\beta_2, & \text{kun } i = 3 \\ \beta_1 - \beta_2, & \text{kun } i = 4, 5. \end{cases}$$

Muotoile tilanne lineaarisena mallina ja estimoi parametrit β_1 ja β_2 soveltamalla lineaarisen mallin estimointiteoriaa. Kuvaa lyhyesti käyttämäsi estimointiperiaate ja selvitä parametrivektorin $\boldsymbol{\beta} = [\beta_1 \ \ \beta_2]'$ estimaattorin $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1 \ \ \hat{\beta}_2]'$ jakauma. Ovatko $\hat{\beta}_1$ ja $\hat{\beta}_2$ riippumattomia?

3. Oletetaan, että n riippumatonta havaintoa $Y_1, ..., Y_n$ $(n > 2)$ on poimittu normaalijakaumasta, jonka varianssi on tuntematon σ^2. Ensimmäisen $n - 1$ havainnon tiedetään olevan peräisin samasta normaalijakaumasta, mutta ennen viimeisen havainnon poimimista epäillään, että havaintojen odotusarvo on muuttunut äkillisesti ja saattaa poiketa ensimmäisen $n - 1$ havainnon odotusarvosta. Muotoile tilanne lineaarisen mallin avulla ja johda testi hypoteesille, jonka mukaan n. havainto on peräisin samasta normaalijakaumasta kuin $n - 1$ ensimmäistä havaintoa.

4. Yksisuuntainen varianssianalyysi: Ongelmanasettelu, tilastollinen malli, tavallisimmin testattava hypoteesi ja sen testaaminen.

Muistin tueksi

- Satunnaisvektorin $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ tiheysfunktio on

$$f(x) = (2\pi)^{-k/2} \det(\boldsymbol{\Sigma})^{-1/2} \exp\left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\},$$

  jossa $\det(\boldsymbol{\Sigma})$ on kovarianssimatriisin $\boldsymbol{\Sigma}$ determinantti ja yksiulotteinen tapaus saadaan asettamalle $k = 1$.

- $F_{k,m}$-jakauman määrittelee satunnaismuuttuja $m\chi_k^2 / k\chi_m^2$, jossa $\chi_k^2 \perp\!\!\!\perp \chi_m^2$. Lisäksi $\mathsf{E}(\chi_k^2) = k$, $\mathsf{Var}(\chi_k^2) = 2k$.

- t_k-jakauman määrittelee satunnaismuuttuja $Z/\sqrt{\frac{1}{k}\chi_k^2}$, jossa $Z \sim \mathsf{N}(0, 1)$ ja $Z \perp\!\!\!\perp \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, niin $(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_k^2$.

- Jos $\mathbf{X} \sim \mathsf{N}_k(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_k)$ ja matriisi $\mathbf{P}$ $(k \times k)$ astetta r oleva ortogonaalinen projektio, niin $(\mathbf{X} - \boldsymbol{\mu})' \mathbf{P} (\mathbf{X} - \boldsymbol{\mu}) / \sigma^2 \sim \chi_r^2$.

LINEAARISET MALLIT, 5 OP (aineopinnot; Pentti Saikkonen: Lineaarinen malli. Kuulustelija: yliopistonlehtori Pekka Pere.

Yleistentti 11.6.

Kaikissa tehtävissä oletetaan, että monisteessa esitetyn yleisen lineaarisen mallin

$$
\begin{aligned}
\mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\
&= \begin{bmatrix} \mathbf{x}_1'\boldsymbol{\beta} \\ \vdots \\ \mathbf{x}_n'\boldsymbol{\beta} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{x}_{(1)} & \cdots & \mathbf{x}_{(p)} \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\varepsilon}.
\end{aligned} \tag{1}
$$

oletukset pätevät.

1.

a) Selitä yksityiskohtaisesti (monisteessa esitetyt) yleisen lineaarisen mallin (1) merkinnät, oletukset ja kaikkien suureiden tulkinta kunkin yhtäsuuruuden kohdalla.

b) Määrittele kaavoin ja sanoin pienimmän neliösumman (PNS) estimaattori ($\hat{\boldsymbol{\beta}}$) parametrivektorille $\boldsymbol{\beta}$, ja selitä intuitiivisesti PNS-estimaattorin idea. Selitä, miksi PNS-estimaattori on lineaarinen estimaattori.

2.

a) Todista, että PNS-estimaattori $\hat{\boldsymbol{\beta}}$ on $\boldsymbol{\beta}$:n harhaton estimaattori. Perustele todistuksesi jokainen vaihe huolellisesti.

b) Todista, että suurimman uskottavuuden estimaattori ($\hat{\sigma}^2$) on jäännösvarianssin σ^2 harhainen estimaattori. Perustele todistuksesi jokainen vaihe huolellisesti.

c) Johda PNS-estimaattorin $\hat{\boldsymbol{\beta}}$ kovarianssimatriisi. Miten se liittyy parametrivektorin $\boldsymbol{\beta}$ yksittäisiä elementtejä testaaviin t-testisuureisiin?

3. Olkoon $Y_1, \ldots, Y_n \perp\!\!\!\perp$, $Y_i \sim \mathsf{N}(\beta, \sigma^2)$ eli olkoon lineaarisessa mallissa vakio ainoa selittäjä.

a) Laske $\mathbf{P}$ ja $\mathbf{Py}$ (monisteen merkinnöin) ja tulkitse ne sanoin.

b) Yleisesti pätee, että $\mathsf{R}^2 \in [0, 1]$. Todista tulos tehtävän erityistilanteessa.

4. Tarkastellaan yhden selittäjän lineaarista regressiomallia kahden ryhmän tapauksessa:

$$Y_1, \ldots, Y_n \perp\!\!\!\perp, \quad Y_i \sim \begin{cases} \mathsf{N}\left(\beta_1 + \beta_2 x_i, \sigma^2\right), & \text{kun } i = 1, \ldots, n_1 \\ \mathsf{N}\left(\beta_3 + \beta_4 x_i, \sigma^2\right), & \text{kun } i = n_1 + 1, \ldots, n_1 + n_2 = n, \end{cases}$$

jossa $n_1, n_2 > 2$.

a) Osoita, että kysymyksessä on lineaarinen malli käyttäen sen matriisiesitystä ($\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$).

b) Osoita, että parametrien β_1 ja β_2 (vastaavasti β_3 ja β_4) PNS-estimaatit saadaan käyttäen vain havaintoyksiköiden $i = 1, \ldots, n_1$ (vastaavasti $i = n_1 + 1, \ldots, n$) havaintoja.

c) Miten tulkitset ehtoja $\beta_1 = \beta_3$ ja $\beta_2 = \beta_4$, kun y on lääkärien kuukausipalkka, x on työvuosien määrä ja sukupuoli määrää ryhmäjaon?

APUTULOKSIA

- $\mathsf{E}(Z) = k$ ja $\mathsf{Var}(Z) = 2k$, jos $Z \sim \chi^2_k$.
- $0 \leq \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2$ eli $\sum_{i=1}^n y_i^2 \geq n\bar{y}^2$.

Tenttien korjaaminen saattaa viedä tavallista pidemmän ajan kesällä. Oikein hyvää ja aurinkoista kesää tilastotieteen ja muun kesäpuuhailun merkeissä!