Tilastollinen päättely

2. Datan redusoinnin periaatteet

2.1. Tyhjentävyys

Ansillaarisuus, Basun teoreema, Datan redusointi, Faktorointiteoreema, Informaatio, Minimaalinen tyhjentävyys, Otos, Otostunnusluku, Parametri, Riippumattomuus, Sivutunnusluku, Todennäköisyysjakauma, Tyhjentävyys, Tyhjentävyysperiaate, Täydellisyys, Yhteisjakauma

2.2. Uskottavuus

Birnbaumin teoreema, Datan redusointi, Ehdollisuusperiaate, Evidenssi, Fidusiaalisuus, Formaali tyhjentävyysperiaate, Formaali uskottavuusperiaate, Otos, Otostunnusluku, Parametri, Riippumattomuus, Tyhjentävyys, Tyhjentävyysperiaate, Uskottavuus, Uskottavuusfunktio, Uskottavuusperiaate, Yhteisjakauma

2.1. Tyhjentävyys

Otos ja otostunnusluvut

Olkoon

$$X_1, X_2, \ldots , X_n$$

satunnaisotos jakaumasta, jonka *pistetodennäköisyys*- tai *tiheysfunktio* $f(x;\theta)$ riippuu parametrista θ. Tällöin havainnot $X_1, X_2, \ldots , X_n$ ovat *riippumattomia*, *identtisesti jakautuneita satunnaismuuttujia*, joilla on *sama pistetodennäköisyys*- tai *tiheysfunktio* $f(x;\theta)$:

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim f(x;\theta)\,,\; i = 1, 2, \ldots, n$$

Olkoon

$$\mathbf{X} = (X_1, X_2, \ldots , X_n)$$

satunnaismuuttujien $X_1, X_2, \ldots , X_n$ muodostama n-vektori.

Kutsumme *satunnaismuuttujien* $X_1, X_2, \ldots , X_n$ (mitallisia) *funktioita*

$$T(\mathbf{X}) = T(X_1, X_2, \ldots, X_n)$$

(**otos-**) **tunnusluvuiksi**.

Olkoot satunnaismuuttujien $X_1, X_2, \ldots , X_n$ *havaitut arvot*

$$x_1, x_2, \ldots , x_n$$

Merkitään tätä:

$$X_1 = x_1, X_2 = x_2, \ldots , X_n = x_n$$

Satunnaismuuttujien $X_1, X_2, \ldots , X_n$ havaitut arvot $x_1, x_2, \ldots , x_n$ määräävät *havaintopisteen*

$$\mathbf{x} = (x_1, x_2, \ldots , x_n)$$

Jos satunnaismuuttujat $X_1, X_2, \ldots , X_n$ ovat saaneet otannan tuloksena havaituiksi arvoikseen *havaintoarvot* $x_1, x_2, \ldots , x_n$, niin tunnusluku

$$T(\mathbf{X}) = T(X_1, X_2, \ldots, X_n)$$

saa havaituksi arvokseen t funktion $T(\cdot)$ arvon havaintopisteessä $\mathbf{x} = (x_1, x_2, \ldots , x_n)$:

$$t = T(\mathbf{x}) = T(x_1, x_2, \ldots, x_n)$$

Tyhjentävyysperiaate

Tunnusluku $T(\mathbf{X})$ on *tyhjentävä parametrille* θ, *jos se käyttää – jossakin mielessä – kaiken otoksessa olevan informaation parametrista* θ.

Tämä on tyhjentävyyden *naiivi määritelmä*, joka kuitenkin tavoittaa olennaiset piirteet tyhjentävyyden käsitteestä. Määritelmä täsmennetään seuraavassa kappaleessa.

Tyhjentävyysperiaate sanoo, että jos $T(\mathbf{X})$ on *tyhjentävä parametrille* θ, niin parametria θ koskevat johtopäätökset riippuvat otoksesta $\mathbf{X}$ vain tunnusluvun $T(\mathbf{X})$ arvojen kautta. Toisin sanoen, jos $\mathbf{x}$ ja $\mathbf{y}$ ovat kaksi havaintopistettä, joille

$$T(\mathbf{x}) = T(\mathbf{y})$$

niin ne johtavat *samoihin* johtopäätöksiin parametrista θ.

Tyhjentävyys

Tunnusluku $T(\mathbf{X})$ on **tyhjentävä parametrille** θ, jos otoksen $\mathbf{X}$ ehdollinen jakauma kiinteällä tunnusluvun $T(\mathbf{X})$ arvolla (ehdolla $T(\mathbf{X}) = t$) *ei riipu* parametrista θ.

Tarkastelemme seuraavassa (ja myös jatkossa) tyhjentävyyttä ja sen karakterisointia vain *diskreettien jakaumien* tapauksessa. *Sivuutamme tässä esityksessä jatkuvien jakaumien tapauksen käsittelyn* siihen liittyvien teknisten hankaluuksien takia.

Hankaluudet jatkuvien jakaumien tapauksessa liittyvät siihen, että soveltamamme ehdollisen todennäköisyyden määritelmä ei salli sellaisten tapahtumien ehdollisten todennäköisyyksien käsittelyn, joissa ehtotapahtuman todennäköisyys = 0 (muista, että jatkuvien jakaumien tapauksessa yhden pisteen todennäköisyys = 0). Ehdollisen todennäköisyyden määritelmä voidaan kuitenkin yleistää sellaiseen muotoon, että tästä ei ole haittaa. Siten kaikki se, mitä tässä (ja myös jatkossa) tyhjentävyydestä esitetään diskreettien jakaumien tapauksessa, pätee myös jatkuvien jakaumien tapauksessa.

Oletetaan, että tunnusluku $T(\mathbf{X})$ on *tyhjentävä* parametrille θ. Olkoon t tunnusluvun $T(\mathbf{X})$ mahdollinen arvo eli sellainen arvo, jolle

$$\Pr_\theta(T(\mathbf{X}) = t) > 0$$

Tarkastellaan ehdollista todennäköisyyttä

$$\Pr_\theta(\mathbf{X} = \mathbf{x} \mid T(\mathbf{X}) = t)$$

Jos $\mathbf{x}$ on havaintopiste, jolle

$$T(\mathbf{x}) \neq t$$

niin

$$\Pr_\theta(\mathbf{X} = \mathbf{x} \mid T(\mathbf{X}) = t) = 0$$

Tämä seuraa siitä, että (ehdollisen todennäköisyyden määritelmän mukaan)

$$\Pr_\theta(\mathbf{X} = \mathbf{x} \mid T(\mathbf{X}) = t) = \frac{\Pr_\theta(\mathbf{X} = \mathbf{x} \ ja\ T(\mathbf{X}) = t)}{\Pr_\theta(T(\mathbf{X}) = t)}$$

ja

$$\{\mathbf{X} = \mathbf{x} \ ja\ T(\mathbf{X}) = t\} \subset \{T(\mathbf{X}) = t\}$$

joten

$$\Pr_\theta(\mathbf{X} = \mathbf{x} \mid T(\mathbf{X}) = t) = 1$$

Siten voimme rajoittua tarkastelemaan ehdollisia todennäköisyyksiä

$$\Pr_\theta(\mathbf{X} = \mathbf{x} \mid T(\mathbf{X}) = T(\mathbf{x}))$$

Koska olemme olettaneet, että tunnusluku $T(\mathbf{X})$ on tyhjentävä parametrille θ, niin (tyhjentävyyden määritelmän mukaan) ehdollinen todennäköisyys

$$\Pr_\theta(\mathbf{X} = \mathbf{x} \mid T(\mathbf{X}) = T(\mathbf{x}))$$

ei riipu parametrista θ. Siten voimme jatkossa jättää indeksin θ pois tästä ehdollisesta todennäköisyydestä ja kirjoittaa

$$\Pr(\mathbf{X} = \mathbf{x} \mid T(\mathbf{X}) = T(\mathbf{x}))$$

Tarkastelemme seuraavassa sitä, missä mielessä *parametrille θ tyhjentävä tunnusluku $T(\mathbf{X})$ sisältää kaiken otoksessa olevan informaation parametrista θ.*

Oletetaan, että henkilö A havaitsee otoksen $\mathbf{X} = \mathbf{x}$ ja määrää tunnusluvun $T(\mathbf{X})$ arvon $T(\mathbf{x})$. Tehdessään päätelmiä parametrista θ, hän voi siis käyttää sekä tietoa siitä, että $\mathbf{X} = \mathbf{x}$ että tietoa siitä, että $T(\mathbf{X}) = T(\mathbf{x})$.

Oletetaan, että henkilölle B kerrotaan vain se, että tunnusluku $T(\mathbf{X})$ on saanut arvon $T(\mathbf{x})$. Henkilö B voi tämän tiedon perusteella määrätä todennäköisyydet

$$\Pr(\mathbf{X} = \mathbf{x} \mid T(\mathbf{X}) = T(\mathbf{x}))$$

eli joukossa

$$A_{T(\mathbf{x})} = \{\mathbf{y} \mid T(\mathbf{y}) = T(\mathbf{x})\}$$

määritellyn todennäköisyysjakauman, koska siihen liittyvät todennäköisyydet voidaan tyhjentävyyden määritelmän mukaan määrätä ilman tietoa parametrin θ todellisesta arvosta. Siten A voi generoida (esimerkiksi pseudosatunnaislukujen avulla) havainnon $\mathbf{Y}$, joka toteuttaa ehdon

$$\Pr(\mathbf{Y} = \mathbf{y} \mid T(\mathbf{X}) = T(\mathbf{x})) = \Pr(\mathbf{X} = \mathbf{y} \mid T(\mathbf{X}) = T(\mathbf{x}))$$

Kuten alla osoitetaan, satunnaismuuttujilla $\mathbf{X}$ ja $\mathbf{Y}$ on sama ei-ehdollinen todennäköisyysjakauma. Tämä merkitsee sitä, että A:lla ja B:llä on käytettävissään täsmälleen yhtä paljon informaatiota parametrista θ. Koska havainto $\mathbf{Y}$ on *generoitu*, B:n informaatio parametrista θ ei ole *aidosti* lisääntynyt. B:n *aito* informaatio parametrista θ sisältyy siihen, että tunnusluvulla $T(\mathbf{X})$ on arvo $T(\mathbf{x})$.

Edellä esitetystä seuraa, että B:llä, joka tietää vain sen, että

$$T(\mathbf{X}) = T(\mathbf{x})$$

on täsmälleen yhtä paljon informaation parametrista θ kuin A:lla, joka tuntee myös otoksen

$$\mathbf{X} = \mathbf{x}$$

Täydennetään yllä esitettyä tarkastelua lopuksi sillä, että näytetään, että satunnaismuuttujilla $\mathbf{X}$ ja $\mathbf{Y}$ on sama ei-ehdollinen todennäköisyysjakauma eli että

$$\Pr_\theta(\mathbf{X} = \mathbf{x}) = \Pr_\theta(\mathbf{Y} = \mathbf{x})$$

Huomaa, että tapahtumat $\{\mathbf{X} = \mathbf{x}\}$ ja $\{\mathbf{Y} = \mathbf{x}\}$ ovat tapahtuman $\{T(\mathbf{X}) = T(\mathbf{x})\}$ osajoukkoja. Lisäksi

$$\Pr(\mathbf{X} = \mathbf{x} \mid T(\mathbf{X}) = T(\mathbf{x})) = \Pr(\mathbf{Y} = \mathbf{x} \mid T(\mathbf{X}) = T(\mathbf{x}))$$

ja nämä ehdolliset todennäköisyydet *eivät riipu* parametrista θ.

Siten

$$\begin{aligned}
\Pr_\theta(\mathbf{X}=\mathbf{x}) &= \Pr_\theta(\mathbf{X}=\mathbf{x} \textit{ ja } T(\mathbf{X})=T(\mathbf{x})) \\
&= \Pr(\mathbf{X}=\mathbf{x} \mid T(\mathbf{X})=T(\mathbf{x}))\Pr_\theta(T(\mathbf{X})=T(\mathbf{x})) \\
&= \Pr(\mathbf{Y}=\mathbf{x} \mid T(\mathbf{X})=T(\mathbf{x}))\Pr_\theta(T(\mathbf{X})=T(\mathbf{x})) \\
&= \Pr_\theta(\mathbf{Y}=\mathbf{x} \text{ ja } T(\mathbf{X})=T(\mathbf{x})) \\
&= \Pr_\theta(\mathbf{Y}=\mathbf{x})
\end{aligned}$$

Tyhjentävyyden karakterisointi

Jotta voisimme käyttää yllä esitettyä tyhjentävyyden määritelmää todistaaksemme, että tunnusluku $T(\mathbf{X})$ on *tyhjentävä parametrille* θ, meidän on todistettava, että ehdollinen todennäköisyys

$$\Pr_\theta(\mathbf{X}=\mathbf{x} \mid T(\mathbf{X})=t)$$

ei riipu parametrin θ arvosta kaikille kiinteille $\mathbf{x}$ ja t.

Lause:

Olkoon

$$f(\mathbf{x};\theta)$$

otoksen $\mathbf{X}$ yhteisjakauman pistetodennäköisyys- tai tiheysfunktio ja

$$q(t;\theta)$$

tunnusluvun $T(\mathbf{X})$ pistetodennäköisyys- tai tiheysfunktio. Tällöin $T(\mathbf{X})$ on *tyhjentävä* parametrille θ, jos suhde

$$\frac{f(\mathbf{x};\theta)}{q(T(\mathbf{x});\theta)}$$

ei riipu parametrista θ.

Perustelu:

Todistamme lauseen vain *diskreettien jakaumien* tapauksessa.

Jos $\mathbf{x}$ on havaintopiste, jolle

$$T(\mathbf{x}) \neq t$$

niin

$$\Pr_\theta(\mathbf{X}=\mathbf{x} \mid T(\mathbf{X})=t)=0$$

kaikille parametrin θ arvoille. Siten riittää todistaa, että todennäköisyys

$$\Pr_\theta(\mathbf{X}=\mathbf{x} \mid T(\mathbf{X})=T(\mathbf{x}))$$

ei riipu parametrista θ.

Koska tapahtuma $\{\mathbf{X} = \mathbf{x}\}$ on tapahtuman $\{T(\mathbf{X}) = T(\mathbf{x})\}$ osajoukko, niin

$$\begin{aligned}\Pr_\theta(\mathbf{X}=\mathbf{x} \mid T(\mathbf{X})=T(\mathbf{x})) &= \frac{\Pr_\theta(\mathbf{X}=\mathbf{x}\ ja\ T(\mathbf{X})=T(\mathbf{x}))}{\Pr_\theta(T(\mathbf{X})=T(\mathbf{x}))} \\ &= \frac{\Pr_\theta(\mathbf{X}=\mathbf{x})}{\Pr_\theta(T(\mathbf{X})=T(\mathbf{x}))} \\ &= \frac{f(\mathbf{x};\theta)}{q(T(\mathbf{x});\theta)}\end{aligned}$$

Tässä

$$f(\mathbf{x};\theta)$$

on otoksen **X** yhteisjakauman pistetodennäköisyysfunktio ja

$$q(t;\theta)$$

on tunnusluvun $T(\mathbf{X})$ pistetodennäköisyysfunktio.

Siten tunnusluku $T(\mathbf{X})$ on tyhjentävä parametrille θ, jos ja vain jos suhde

$$\frac{f(\mathbf{x};\theta)}{q(T(\mathbf{x});\theta)}$$

ei riipu parametrista θ.

■

Esimerkki 1.1: Otos normaalijakaumasta

Satunnaismuuttuja X noudattaa *normaalijakaumaa* parametrein

$$\mu = \mathrm{E}(X)$$
$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2]$$

jos sen tiheysfunktio on muotoa

$$f(x_i;\mu) = (2\pi\sigma^{-2})^{-1/2}\exp\left[-\frac{1}{2\sigma^2}(x_i-\mu)^2\right], i = 1,2,\ldots,n$$

Oletamme tässä, että varianssi σ^2 on *tunnettu*.

Oletetaan, että *havainnot*

$$X_1, X_2, \ldots, X_n$$

muodostavat satunnaisotoksen normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$. Tällöin

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim \mathrm{N}(\mu,\sigma^2), i = 1,2,\ldots,n$$

Olkoot satunnaismuuttujien $X_1, X_2, \ldots, X_n$ *havaitut arvot*

$$x_1, x_2, \ldots, x_n$$

Merkitään tätä:

$$X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$$

Olkoon

$$\mathbf{X} = (X_1, X_2, \dots, X_n)$$

satunnaismuuttujien $X_1, X_2, \dots, X_n$ muodostama vektori ja

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

niiden havaittujen arvojen muodostama vektori.

Näytetään, että havaintojen aritmeettinen keskiarvo

$$T(\mathbf{X}) = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

on tyhjentävä tunnusluku odotusarvoparametrille μ.

Otoksen $X_1, X_2, \dots, X_n$ *yhteisjakauman tiheysfunktio* voidaan kirjoittaa satunnaismuuttujien $X_1, X_2, \dots, X_n$ riippumattomuuden takia muotoon

$$\begin{aligned}
f(\mathbf{x};\mu) &= \prod_{i=1}^{n} f(x_i;\mu) \\
&= \prod_{i=1}^{n} (2\pi)^{-2}\sigma^{-1}\exp\left[-\frac{1}{2\sigma^2}(x_i-\mu)^2\right] \\
&= (2\pi)^{-n/2}\sigma^{-n}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right]
\end{aligned}$$

Nyt

$$\begin{aligned}
\sum_{i=1}^{n}(x_i-\mu)^2 &= \sum_{i=1}^{n}(x_i-\overline{x}+\overline{x}-\mu)^2 \\
&= \sum_{i=1}^{n}(x_i-\overline{x})^2 + 2\sum_{i=1}^{n}(x_i-\overline{x})(\overline{x}-\mu) + \sum_{i=1}^{n}(\overline{x}-\mu)^2 \\
&= \sum_{i=1}^{n}(x_i-\overline{x})^2 + n(\overline{x}-\mu)^2
\end{aligned}$$

jossa

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = T(\mathbf{x})$$

Tämä seuraa siitä, että

$$\begin{aligned}
\sum_{i=1}^{n}(x_i-\overline{x})(\overline{x}-\mu) &= (\overline{x}-\mu)\sum_{i=1}^{n}(x_i-\overline{x}) \\
&= (\overline{x}-\mu)\left[\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\overline{x}\right] \\
&= (\overline{x}-\mu)\left[\sum_{i=1}^{n}x_i - n\overline{x}\right] = 0
\end{aligned}$$

Siten

$$f(\mathbf{x};\mu) = (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2\right]\right\}$$

Otosjakaumia koskevassa luvussa on todistettu, että yllä esitettyjen oletuksien pätiessä *aritmeettinen keskiarvo* $T(\mathbf{X}) = \bar{X}$ *noudattaa normaalijakaumaa* parametrein μ ja σ^2/n:

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

Siten sen tiheysfunktio on muotoa

$$q(T(\mathbf{x});\mu) = (2\pi)^{-1/2}\sigma^{-1}n^{1/2}\exp\left[-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right]$$

Koska tiheysfunktioiden

$$\begin{aligned}\frac{f(\mathbf{x};\mu)}{q(t(\mathbf{x});\mu)} &= \frac{(2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2\right]\right\}}{(2\pi)^{-1/2}\sigma^{-1}n^{1/2}\exp\left[-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right]} \\ &= n^{-1/2}(2\pi)^{-(n-1)/2}\sigma^{-(n-1)}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]\right\}\end{aligned}$$

suhde ei riipu parametrista μ, tunnusluku $T(\mathbf{X}) = \bar{X}$ on *tyhjentävä* parametrille μ.

Tyhjentävyyden todistaminen tyhjentävyyden määritelmään nojaten on usein hankalaa. Todistaminen tapahtuu tavallisesti paljon helpommin vetoamalla seuraavassa esitettävään *faktorointiteoreemaan*.

Faktorointiteoreema:

Olkoon

$$f(\mathbf{x};\theta)$$

otoksen $\mathbf{X}$ yhteisjakauman pistetodennäköisyys- tai tiheysfunktio. Tunnusluku $T(\mathbf{X})$ on *tyhjentävä* parametrille θ, jos ja vain jos on olemassa funktiot $g(t;\theta)$ ja $h(\mathbf{x})$ siten, että

$$f(\mathbf{x};\theta) = g(T(\mathbf{x});\theta)h(\mathbf{x})$$

kaikille havaintopisteille $\mathbf{x}$ ja parametrin θ mahdollisille arvoille ja funktio g riippuu otoksesta $\mathbf{X} = \mathbf{x}$ vain tunnusluvun $T(\mathbf{X})$ kautta ja funktio h ei riipu parametrista θ.

Perustelu:

Todistamme lauseen vain *diskreettien jakaumien* tapauksessa.

(i) Olkoon $T(\mathbf{X})$ *tyhjentävä*.

Valitaan

$$g(t;\theta) = \Pr_\theta(T(\mathbf{X}) = t)$$

ja

$$h(\mathbf{x}) = \Pr(\mathbf{X} = \mathbf{x} \,|\, T(\mathbf{X}) = T(\mathbf{x}))$$

Funktio $g(t;\theta)$ *riippuu* parametrista θ. Mutta koska $T(\mathbf{X})$ on tyhjentävä, funktion $h(\mathbf{x})$ määrittelevä ehdollinen todennäköisyys *ei riipu* parametrista θ.

Käyttämällä hyväksi yllä esitettyjä määritelmiä, otoksen **X** yhteisjakauman pistetodennäköisyysfunktio $f(\mathbf{x};\theta)$ voidaan kirjoittaa seuraaviin muotoihin:

$$\begin{aligned} f(\mathbf{x};\theta) &= \Pr_\theta(\mathbf{X} = \mathbf{x}) \\ &= \Pr_\theta(\mathbf{X} = \mathbf{x} \textit{ ja } T(\mathbf{X}) = T(\mathbf{x})) \\ &= \Pr_\theta(T(\mathbf{X}) = T(\mathbf{x}))\Pr(\mathbf{X} = \mathbf{x} \,|\, T(\mathbf{X}) = T(\mathbf{x})) \\ &= g(T(\mathbf{x});\theta)h(\mathbf{x}) \end{aligned}$$

Siten otoksen **X** yhteisjakauman pistetodennäköisyysfunktio $f(\mathbf{x};\theta)$ *voidaan faktoroida* vaaditulla tavalla.

Lisäksi näemme, että

$$\Pr_\theta(T(\mathbf{X}) = T(\mathbf{x})) = g(T(\mathbf{x});\theta)$$

joten $g(T(\mathbf{x});\theta)$ on tunnusluvun $T(\mathbf{X})$ pistetodennäköisyysfunktio.

(ii) Oletetaan, että otoksen **X** yhteisjakauman pistetodennäköisyysfunktio $f(\mathbf{x};\theta)$ *voidaan faktoroida* seuraavalla tavalla:

$$f(\mathbf{x};\theta) = g(T(\mathbf{X});\theta)h(\mathbf{x})$$

Olkoon $q(t;\theta)$ tunnusluvun $T(\mathbf{X})$ pistetodennäköisyysfunktio. Näyttääksemme, että $T(\mathbf{X})$ on tyhjentävä, tarkastelemme suhdetta

$$\frac{f(\mathbf{x};\theta)}{q(T(\mathbf{x});\theta)}$$

Määritellään joukko

$$A_{T(\mathbf{x})} = \{\mathbf{y} \,|\, T(\mathbf{y}) = T(\mathbf{x})\}$$

Siten

$$\begin{aligned} \frac{f(\mathbf{x};\theta)}{q(T(\mathbf{x});\theta)} &= \frac{g(T(\mathbf{x});\theta)h(\mathbf{x})}{q(T(\mathbf{x});\theta)} \\ &= \frac{g(T(\mathbf{x});\theta)h(\mathbf{x})}{\sum_{A_{T(\mathbf{x})}} g(T(\mathbf{x});\theta)h(\mathbf{y})} \\ &= \frac{g(T(\mathbf{x});\theta)h(\mathbf{x})}{g(T(\mathbf{x});\theta)\sum_{A_{T(\mathbf{x})}} h(\mathbf{y})} \\ &= \frac{h(\mathbf{x})}{\sum_{A_{T(\mathbf{x})}} h(\mathbf{y})} \end{aligned}$$

Koska olemme siis todistaneet, että jos suhde

$$\frac{f(\mathbf{x};\theta)}{q(T(\mathbf{x});\theta)}$$

ei riipu parametrista θ, niin *tyhjentävyyden karakterisointilauseesta* seuraa, että tunnusluku $T(\mathbf{X})$ on *tyhjentävä* parametrille θ.

■

Esimerkki 1.2: Otos normaalijakaumasta (jatkoa esimerkille 1.1)

Sovelletaan *faktorointiteoreemaa* esimerkin 1.1 tilanteeseen.

Esimerkissä 1.1 todettiin, että *otoksen* $X_1, X_2, \ldots, X_n$ *yhteisjakauman tiheysfunktio* voidaan kirjoittaa muotoon

$$f(\mathbf{x};\mu) = (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2\right]\right\}$$

Siten tiheysfunktio $f(\mathbf{x};\theta)$ voidaan faktoroida seuraavalla tavalla:

$$f(\mathbf{x};\mu) = (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]\right\}\exp\left\{-\frac{1}{2\sigma^2}n(\bar{x}-\mu)^2\right\}$$

Olkoon

$$g(t;\mu) = \exp\left\{-\frac{1}{2\sigma^2}n(\bar{x}-\mu)^2\right\}$$

$$h(\mathbf{x}) = (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]\right\}$$

Funktio $g(t;\mu)$ riippuu havaintoarvoista $\mathbf{x}$ vain funktion

$$T(\mathbf{x}) = \bar{x}$$

kautta ja funktio $h(\mathbf{x})$ ei riipu tuntemattomasta parametrista μ.

Siten *faktorointiteoreemasta* seuraa, että tunnusluku $T(\mathbf{X}) = \bar{X}$ on *tyhjentävä* parametrille μ.

Esimerkeissä 1.1 ja 1.2 tyhjentävänä tunnuslukuna on ollut otoksen reaaliarvoinen funktio ja kiinnostuksen kohteena olevaa parametria θ koskeva informaatio otoksesta on tiivistetty *yhteen* tunnuslukuun $T(\mathbf{x})$.

Tilastotieteessä kohdataan kuitenkin usein tilanteita, joissa parametria koskevaa informaatiota ei voida tiivistää yhteen tunnuslukuun. Tällöin tyhjentävänä tunnuslukuna $\mathbf{T}(\mathbf{X})$ on jokin vektori:

$$\mathbf{T}(\mathbf{X}) = (T_1(\mathbf{X}), \ldots, T_r(\mathbf{X}))$$

Tällainen on tilanne usein silloin, kun myös parametrina on vektori:

$$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_s)$$

Tavallisesti $r = s$, mutta näin ei tarvitse olla.

Myös vektoriarvoiset tyhjentävät tunnusluvut löydetään tavallisesti helpoiten *faktorointiteoreemaa* soveltamalla.

Esimerkki 1.3: Otos normaalijakaumasta (jatkoa esimerkille 1.1)

Oletetaan, että *havainnot*

$$X_1, X_2, \ldots , X_n$$

muodostavat satunnaisotoksen normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$ kuten esimerkissä 1.1, mutta oletamme nyt, että sekä odotusarvoparametri μ että varianssiparametri σ^2 ovat *tuntemattomia*.

Siten parametrivektorina on

$$\boldsymbol{\theta} = (\mu, \sigma^2)$$

Esimerkin 1.1 mukaan *otoksen* $X_1, X_2, \ldots , X_n$ *yhteisjakauman tiheysfunktio* voidaan kirjoittaa muotoon

$$f(\mathbf{x};\mu,\sigma^2) = (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\overline{x})^2 + n(\overline{x}-\mu)^2\right]\right\}$$

Määritellään otosvarianssi S^2 kaavalla

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

jossa

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n}X_i$$

Olkoon

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2$$

jossa

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n}x_i$$

tunnusluvun S^2 havaittu arvo. Siten otoksen normaalijakautuneen otoksen yhteisjakauman tiheysfunktio voidaan kirjoittaa muotoon

$$f(\mathbf{x};\mu,\sigma^2) = (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)t_2 + n(t_1-\mu)^2\right]\right\}$$

jossa

$$t_1 = T_1(\mathbf{x}) = \overline{x}$$
$$t_2 = T_2(\mathbf{x}) = s^2$$

Määritellään vektori

$$\mathbf{t} = (t_1, t_2)$$

Olkoon

$$g(\mathbf{t};\boldsymbol{\theta}) = f(\mathbf{x};\mu,\sigma^2)$$
$$h(\mathbf{x}) \equiv 1$$

Olemme siis näyttäneet, että

$$f(\mathbf{x};\mu,\sigma^2) = g(t_1,t_2;\mu,\sigma^2)h(\mathbf{x}) = g(T_1(\mathbf{x}),T_2(\mathbf{x});\mu,\sigma^2)h(\mathbf{x})$$

jossa funktio g riippuu otoksesta $\mathbf{X} = \mathbf{x}$ vain tunnusluvun $\mathbf{T}(\mathbf{X}) = (T_1(\mathbf{X}),T_2(\mathbf{X})) = (\bar{X},S^2)$ kautta ja funktio h ei riipu parametrista $\boldsymbol{\theta} = (\mu,\sigma^2)$.

Siten *faktorointiteoreemasta* seuraa, että tunnusluku

$$\mathbf{T}(\mathbf{X}) = (T_1(\mathbf{X}),T_2(\mathbf{X})) = (\bar{X},S^2)$$

on *tyhjentävä* parametrille $\boldsymbol{\theta} = (\mu,\sigma^2)$.

Esimerkki 1.4: Otos eksponenttiperheestä

Oletetaan, että *havainnot*

$$X_1, X_2, \ldots, X_n$$

muodostavat satunnaisotoksen **eksponenttiperheestä**, jonka pistetodennäköisyys- tai tiheysfunktio on muotoa

$$f(x;\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})\exp\left[\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right]$$

jossa $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_d)$, $d \le k$. Tällöin tunnusluku

$$\mathbf{T}(\mathbf{X}) = \left(\sum_{j=1}^{n} t_1(X_j),\ldots,\sum_{j=1}^{n} t_k(X_j)\right)$$

on *tyhjentävä* parametrille $\boldsymbol{\theta}$.

Huomautus:

Huomattava osa tilastotieteen tavanomaisista todennäköisyysjakaumista kuuluu eksponenttiperheeseen. Tällaisia jakaumia ovat esimerkiksi sellaiset diskreetit jakaumat kuten *Bernoulli-jakauma*, *binomijakauma*, *geometrinen jakauma*, *negatiivinen binomijakauma* ja *Poisson-jakauma* sekä sellaiset jatkuvat jakaumat kuten *eksponenttijakauma*, *normaalijakauma*, *gammajakauma*, χ^2*-jakauma* ja *betajakauma*.

Tyhjentävän tunnusluvun funktioiden tyhjentävyys

Lause:

Jokainen tyhjentävän tunnusluvun *bijektio* on tyhjentävä.

Perustelu:

Olkoon funktio r bijektio, jonka käänteisfunktio on r^{-1}.

Oletetaan, että $T(\mathbf{X})$ on tyhjentävä tunnusluku ja

$$T^*(\mathbf{x}) = r(T(\mathbf{x}))$$

kaikille **x**. Faktorointiteoreeman mukaan on olemassa funktiot g ja h siten, että

$$f(\mathbf{x};\theta) = g(T(\mathbf{X});\theta)h(\mathbf{x}) = g(r^{-1}(T^*(\mathbf{X}));\theta)h(\mathbf{x})$$

Määritellään

$$g^*(t;\theta) = g(r^{-1}(t);\theta)$$

Tällöin

$$f(\mathbf{x};\theta) = g^*(T^*(\mathbf{X});\theta)h(\mathbf{x})$$

joten *faktorointiteoreeman* mukaan tunnusluku $T^*(\mathbf{X})$ on *tyhjentävä*.

Esimerkki 1.5: Otos normaalijakaumasta (jatkoa esimerkille 1.3)

Oletetaan, että *havainnot*

$$X_1, X_2, \ldots , X_n$$

muodostavat satunnaisotoksen normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$ kuten esimerkissä 1.1, mutta oletamme nyt, että sekä odotusarvoparametri μ että varianssiparametri σ^2 ovat *tuntemattomia*.

Esimerkissä 1.3 todettiin, että tunnusluku

$$\mathbf{T}(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X})) = (\bar{X}, S^2)$$

on *tyhjentävä* parametrille $\theta = (\mu,\sigma^2)$.

Koska kuvaus

$$\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right) \to (\bar{X}, S^2)$$

on *bijektio*, niin myös havaintojen $X_1, X_2, \ldots , X_n$ *summa*

$$\sum_{i=1}^n X_i$$

ja *neliösumma*

$$\sum_{i=1}^n X_i^2$$

ovat *yhdessä tyhjentäviä* parametreille μ ja σ^2.

Minimaalinen tyhjentävyys

Tyhjentävä tunnusluku $T(\mathbf{X})$ on **minimaalisesti tyhjentävä**, jos $T(\mathbf{X})$ on jokaisen (muun) tyhjentävän tunnusluvun funktio.

Tällä tarkoitetaan seuraavaa: Jos $T^*(\mathbf{X})$ on mielivaltainen *tyhjentävä* tunnusluku ja

$$T^*(\mathbf{x}) = T^*(\mathbf{y})$$

niin tällöin $T(\mathbf{X})$ on *minimaalisesti tyhjentävä*, jos

$$T(\mathbf{x}) = T(\mathbf{y})$$

Minimaalisen tyhjentävyyden karakterisointi

Lause:

Olkoon

$$f(\mathbf{x};\theta)$$

otoksen **X** yhteisjakauman pistetodennäköisyys- tai tiheysfunktio. Oletetaan, että funktiolla $T(\mathbf{X})$ on seuraava ominaisuus: Suhde

$$\frac{f(\mathbf{x};\theta)}{f(\mathbf{y};\theta)}$$

ei riipu parametrista θ (eli on *vakio* parametrin θ funktiona), jos ja vain jos

$$T(\mathbf{x}) = T(\mathbf{y})$$

Tällöin tunnusluku $T(\mathbf{X})$ on *minimaalisesti tyhjentävä*.

Perustelu:

Todistuksen yksinkertaistamiseksi oletetaan, että

$$f(\mathbf{x};\theta) > 0$$

kaikille havaintopisteille **x** ja parametrin arvoille θ.

(i) Näytetään ensin, että tunnusluku $T(\mathbf{X})$ on *tyhjentävä*.

Olkoon

$$T = \{t \mid t = T(\mathbf{x}) \text{ jollekin } \mathbf{x} \in X\}$$

jossa X on **kaikkien mahdollisten havaintopisteiden x** *joukko*. Siten T on joukon X kuva kuvauksessa $T(\mathbf{x})$.

Olkoon

$$A_t = \{\mathbf{x} \mid T(\mathbf{x}) = t\}$$

kuvauksen $T(\mathbf{x})$ määrittelemä *ositus* joukossa X. Valitaan jokaisesta joukosta A_t *yksi* mielivaltainen alkio $\mathbf{x}_t \in A_t$. Tällöin $\mathbf{x}_{T(\mathbf{x})} \in A_t$ jokaiselle $\mathbf{x} \in X$.

Koska $\mathbf{x}_t$ ja $\mathbf{x}_{T(\mathbf{x})}$ kuuluvat aina samaan joukkoon A_t, niin

$$T(\mathbf{x}) = T(\mathbf{x}_{T(\mathbf{x})})$$

ja suhde

$$\frac{f(\mathbf{x};\theta)}{f(\mathbf{x}_{T(\mathbf{x})};\theta)}$$

on *vakio* parametrin θ funktiona. Siten voimme määritellä funktion

$$h(\mathbf{x}) = \frac{f(\mathbf{x};\theta)}{f(\mathbf{x}_{T(\mathbf{x})};\theta)}$$

joukossa X ja funktio $h(\mathbf{x})$ *ei riipu* parametrista θ.

Määritellään vielä funktio

$$g(t;\theta) = f(\mathbf{x}_t;\theta)$$

joukossa T.

Yllä esitetystä seuraa, että

$$f(\mathbf{x};\theta) = h(\mathbf{x}) = \frac{f(\mathbf{x}_{T(\mathbf{x})};\theta)f(\mathbf{x};\theta)}{f(\mathbf{x}_{T(\mathbf{x})};\theta)} = g(T(\mathbf{x});\theta)h(\mathbf{x})$$

joten tunnusluku $T(\mathbf{X})$ on faktorointiteoreeman mukaan *tyhjentävä* parametrille θ.

(ii) Näytetään, että tunnusluku $T(\mathbf{X})$ on *minimaalisesti tyhjentävä*.

Olkoon $T'(\mathbf{X})$ *mielivaltainen* toinen tyhjentävä tunnusluku. *Faktorointiteoreeman* mukaan on olemassa funktiot g' ja h' siten, että

$$f(\mathbf{x};\theta) = g'(T'(\mathbf{x});\theta)h'(\mathbf{x})$$

Olkoot $\mathbf{x}$ ja $\mathbf{y}$ kaksi havaintopistettä, joille $T'(\mathbf{x}) = T'(\mathbf{y})$.

Tällöin

$$\frac{f(\mathbf{x};\theta)}{f(\mathbf{y};\theta)} = \frac{g'(T'(\mathbf{x});\theta)h'(\mathbf{x})}{g'(T'(\mathbf{y});\theta)h'(\mathbf{y})} = \frac{h'(\mathbf{x})}{h'(\mathbf{y})}$$

Koska tämä suhde *ei riipu* parametrista θ, oletuksesta seuraa, että

$$T(\mathbf{x}) = T(\mathbf{y})$$

joten tunnusluku $T(\mathbf{x})$ on tunnusluvun $T'(\mathbf{x})$ funktio. Koska $T'(\mathbf{X})$ oli valittu mielivaltaisesti, tunnusluku $T(\mathbf{X})$ on *minimaalisesti tyhjentävä*.

■

Esimerkki 1.6: Otos normaalijakaumasta (jatkoa esimerkeille 1.1-1.3)

Oletetaan, että *havainnot*

$$X_1, X_2, \ldots , X_n$$

muodostavat satunnaisotoksen normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$ kuten esimerkissä 1.1, jossa oletettiin, että varianssiparametri σ^2 on *tunnettu*.

Esimerkeissä 1.1 ja 1.2 todettiin, että tunnusluku $T(\mathbf{X}) = \overline{X}$ on *tyhjentävä* parametrille μ.

Esimerkin 1.3 mukaan *otoksen* $X_1, X_2, \ldots, X_n$ *yhteisjakauman tiheysfunktio* voidaan kirjoittaa muotoon

$$\begin{aligned} f(\mathbf{x};\mu) &= (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2\right]\right\} \\ &= (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)t_2 + n(t_1-\mu)^2\right]\right\} \end{aligned}$$

jossa

$$t_1 = T_1(\mathbf{x}) = \bar{x}$$
$$t_2 = T_2(\mathbf{x}) = s^2$$

Siten *faktorointiteoreemasta* seuraa, että myös tunnusluku

$$\mathbf{T}'(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X})) = (\bar{X}, S^2)$$

on *tyhjentävä* parametrille μ.

Tunnusluku

$$T(\mathbf{X}) = \bar{X}$$

selvästi redusoi havaintoaineiston voimakkaammin kuin tunnusluku

$$\mathbf{T}'(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X})) = (\bar{X}, S^2)$$

koska emme vielä tunne otosvarianssin arvoa, jos tunnemme tunnusluvun $T(\mathbf{X}) = \bar{X}$ arvon.

Tunnusluku $T(\mathbf{X})$ on tunnusluvun $\mathbf{T}'(\mathbf{X})$ funktio, mikä nähdään määrittelemällä funktio

$$r(a,b) = a$$

jolloin

$$r(\mathbf{T}'(\mathbf{x})) = r(T_1(\mathbf{x}), T_2(\mathbf{x})) = r(\bar{x}, s^2) = \bar{x} = T(\mathbf{x})$$

Koska tunnusluvut $T(\mathbf{X})$ ovat $\mathbf{T}'(\mathbf{X})$ tyhjentäviä parametrille μ, ne sisältävät saman informaation parametrista μ.

Siten otosvarianssiin S^2 sisältyvä lisäinformaatio ei lisää tietoamme parametrista μ, kun varianssi σ^2 on *tunnettu*. Jos varianssi σ^2 on *tuntematon*, tunnusluku

$$T(\mathbf{X}) = \bar{X}$$

ei ole tyhjentävä ja tunnusluku

$$\mathbf{T}'(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X})) = (\bar{X}, S^2)$$

sisältää tunnuslukua $T(\mathbf{X}) = \bar{X}$ enemmän informaatiota parametrista $\boldsymbol{\theta} = (\mu, \sigma^2)$.

Esimerkki 1.7: Otos normaalijakaumasta (jatkoa esimerkeille 1.1-1.3 ja 1.6)

Oletetaan, että *havainnot*

$X_1, X_2, \ldots, X_n$

muodostavat satunnaisotoksen normaalijakaumasta $N(\mu,\sigma^2)$ kuten esimerkissä 1.1, mutta oletamme nyt, että sekä odotusarvoparametri μ että varianssiparametri σ^2 ovat *tuntemattomia*.

Olkoot **x** ja **y** kaksi mielivaltaista havaintopistettä ja olkoot

$$\bar{x}, \bar{y}, s_x^2, s_y^2$$

havaintoarvoista $x_1, x_2, \ldots, x_n$ ja $y_1, y_2, \ldots, y_n$ määrätyt aritmeettiset keskiarvot ja otosvarianssit.

Esimerkin 1.3 mukaan voimme kirjoittaa

$$\begin{aligned}\frac{f(\mathbf{x};\mu,\sigma^2)}{f(\mathbf{y};\mu,\sigma^2)} &= \frac{(2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\dfrac{1}{2\sigma^2}\left[(n-1)s_x^2+n(\bar{x}-\mu)^2\right]\right\}}{(2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\dfrac{1}{2\sigma^2}\left[(n-1)s_y^2+n(\bar{y}-\mu)^2\right]\right\}} \\ &= \exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)(s_x^2-s_y^2)+n(\bar{x}^2-\bar{y}^2)-2n\mu(\bar{x}-\bar{y})\right]\right\}\end{aligned}$$

Siten suhde

$$\frac{f(\mathbf{x};\mu,\sigma^2)}{f(\mathbf{y};\mu,\sigma^2)}$$

on riippumaton parametreista μ ja σ^2, jos ja vain jos

$$\bar{x} = \bar{y}$$

$$s_x^2 = s_y^2$$

Siten *minimaalisen tyhjentävyyden karakterisointilauseesta* seuraa, että tunnusluku

$$(\bar{X}, S^2)$$

on minimaalisesti tyhjentävä parametrille (μ,σ^2) .

Ansillaarisuus

Tunnusluku $S(\mathbf{X})$ on **ansillaarinen tunnusluku** eli **sivutunnusluku**, jos sen jakauma ei riipu parametrista θ.

Esimerkki 1.8: Paikkaparametriperhe ja havaintojen vaihteluväli

Olkoon $f(x)$ tiheysfunktio ja olkoon

$$-\infty < \mu < +\infty$$

parametri. Tällöin parametrin μ indeksoimaa tiheysfunktioiden perhettä

$$f(x-\mu)$$

kutsutaan **paikkaparametriperheeksi**, jonka *standarditiheysfunktio* on $f(x)$ ja, jonka **paikkaparametrina** on μ.

Oletetaan, että havainnot

$$X_1, X_2, \ldots, X_n$$

muodostavat satunnaisotoksen paikkaparametriperheestä, jonka kertymäfunktio on

$$F(x-\mu)\,,\ -\infty < \mu < +\infty$$

Olkoot

$$X_{(1)}, X_{(2)}, \ldots, X_{(n)}$$

otokseen $X_1, X_2, \ldots, X_n$ liittyvät *järjestystunnusluvut*. Näytämme, että *vaihteluvälin pituus*

$$R = X_{(n)} - X_{(1)}$$

on ansillaarinen tunnusluku.

Olkoon

$$Z_1, Z_2, \ldots, Z_n$$

otos jakaumasta $F(x)$, jossa siis $\mu = 0$ ja olkoot

$$X_1 = Z_1 + \mu, X_2 = Z_2 + \mu, \ldots, X_n = Z_n + \mu$$

Tällöin tunnusluvun R *kertymäfunktio* on

$$\begin{aligned} F_R(r;\mu) &= \Pr(R \le r) \\ &= \Pr(\max_i X_i - \min_i X_i \le r) \\ &= \Pr(\max_i (Z_i + \mu) - \min_i (Z_i + \mu) \le r) \\ &= \Pr(\max_i (Z_i) - \min_i (Z_i) + \mu - \mu \le r) \\ &= \Pr(\max_i (Z_i) - \min_i (Z_i) \le r) \end{aligned}$$

mikä *ei riipu* paikkaparametrista μ, koska satunnaismuuttujien $Z_1, Z_2, \ldots, Z_n$ jakauma ei riipu paikkaparametrista μ. Siten vaihteluväli R on *ansillaarinen tunnusluku*.

Esimerkki 1.9: Skaalaparametriperhe ja havaintojen osamäärät

Olkoon $f(x)$ tiheysfunktio ja olkoon

$$\sigma > 0$$

parametri. Tällöin parametrin σ indeksoimaa tiheysfunktioiden perhettä

$$f(x/\sigma)$$

kutsutaan **skaalaparametriperheeksi**, jonka *standarditiheysfunktio* on $f(x)$ ja, jonka **skaalaparametrina** on σ.

Oletetaan, että havainnot

$$X_1, X_2, \ldots, X_n$$

muodostavat satunnaisotoksen paikkaparametriperheestä, jonka kertymäfunktio on

$$F(x/\sigma)\,,\ \sigma > 0$$

Näytämme, että kaikki tunnusluvut, jotka riippuvat otoksesta ainoastaan suhteiden

$X_1/X_n, X_2/X_n, \ldots , X_{n-1}/X_n$

kautta ovat ansillaarisia.

Olkoon

$Z_1, Z_2, \ldots , Z_n$

otos jakaumasta $F(x)$, jossa siis $\sigma = 1$ ja olkoot

$X_1 = \sigma Z_1, X_2 = \sigma Z_2, \ldots , X_n = \sigma Z_n$

Tällöin satunnaismuuttujien $X_1/X_n, X_2/X_n, \ldots , X_{n-1}/X_n$ yhteisjakauman *kertymäfunktio* on

$$\begin{aligned} F(y_1,\ldots,y_{n-1};\sigma) &= \Pr(X_1 / X_n \le y_1,\ldots,X_{n-1} / X_n \le y_{n-1}) \\ &= \Pr(\sigma Z_1 /(\sigma Z_n) \le y_1,\ldots,\sigma Z_{n-1} /(\sigma Z_n) \le y_{n-1}) \\ &= \Pr(Z_1 / Z_n \le y_1,\ldots,Z_{n-1} / Z_n \le y_{n-1}) \end{aligned}$$

mikä *ei riipu* skaalaparametrista σ, koska satunnaismuuttujien $Z_1, Z_2, \ldots , Z_n$ jakauma ei riipu skaalaparametrista σ. Siten tunnuslukujen $X_1/X_n, X_2/X_n, \ldots , X_{n-1}/X_n$ jakauma on riippumaton skaalaparametrista σ ja niin on myös mikä tahansa tunnuslukujen $X_1/X_n, X_2/X_n, \ldots , X_{n-1}/X_n$ funktio.

Täydellisyys

Olkoon

$g(t;\theta)$

pistetodennäköisyys- tai tiheysfunktioiden perhe tunnusluvulle $T(\mathbf{X})$. Jakaumaperhe on **täydellinen**, jos siitä, että

$\mathrm{E}[g(t;\theta)] = 0$

kaikille θ seuraa, että

$\Pr(g(t;\theta) = 0) = 1$

kaikille θ. Tällöin myös tunnuslukua $T(\mathbf{X})$ kutsutaan **täydelliseksi**.

Täydellisyys on *minimaalisesti tyhjentävän tunnusluvun* ominaisuus, joka takaa sen, että ko. *tunnusluku on riippumaton kaikista ansillaarisista tunnusluvuista* eli *sivutunnusluvuista*.

Basun teoreema:

Jos tunnusluku $T(\mathbf{X})$ on *täydellinen* ja *minimaalisesti tyhjentävä*, niin $T(\mathbf{X})$ on riippumaton kaikista ansillaarisista tunnusluvuista.

Perustelu:

Todistamme lauseen vain *diskreettien jakaumien* tapauksessa.

Olkoon $S(\mathbf{X})$ mielivaltainen *ansillaarinen* tunnusluku. Tällöin todennäköisyys

$\Pr(S(\mathbf{X}) = s)$

ei riipu parametrista θ.

Myöskään ehdollinen todennäköisyys

$$\Pr(S(\mathbf{X}) = s \mid T(\mathbf{X}) = t) = \Pr(\mathbf{X} \in \{\mathbf{x} \mid S(\mathbf{x}) = s\} \mid T(\mathbf{X}) = t)$$

ei riipu parametrista θ, koska tunnusluku $T(\mathbf{X})$ on *tyhjentävä*.

Siten sen todistamiseksi, että tunnusluvut $T(\mathbf{X})$ ja $S(\mathbf{X})$ ovat riippumattomia, riittää osoittaa, että

$$\Pr(S(\mathbf{X}) = s \mid T(\mathbf{X}) = t) = \Pr(S(\mathbf{x}) = s)$$

kaikille $t \in T$.

Kokonaistodennäköisyyden kaavan mukaan

$$\Pr(S(\mathbf{x}) = s) = \sum_{t \in T} \Pr(S(\mathbf{X}) = s \mid T(\mathbf{X}) = t) \Pr_\theta(T(\mathbf{X}) = t)$$

Edelleen, koska

$$\sum_{t \in T} \Pr_\theta(T(\mathbf{X}) = t) = 1$$

niin

$$\Pr(S(\mathbf{x}) = s) = \sum_{t \in T} \Pr(S(\mathbf{X}) = s) \Pr_\theta(T(\mathbf{X}) = t)$$

Määritellään tunnusluku

$$g(t) = \Pr(S(\mathbf{X}) = s \mid T(\mathbf{X}) = t) - \Pr(S(\mathbf{X}) = s)$$

Yllä esitetystä seuraa, että

$$\mathrm{E}_\theta[g(t)] = \sum_{t \in T} g(t) \Pr_\theta(T(\mathbf{X}) = t) = 0$$

kaikille θ. Koska $T(\mathbf{X})$ on täydellinen, niin tästä seuraa, että

$$g(t) = 0$$

kaikille $t \in T$, jolloin

$$\Pr(S(\mathbf{X}) = s \mid T(\mathbf{X}) = t) = \Pr(S(\mathbf{X}) = s)$$

■

Esimerkki 1.10: Otos eksponenttiperheestä

Oletetaan, että *havainnot*

$$X_1, X_2, \ldots, X_n$$

muodostavat satunnaisotoksen eksponenttiperheestä, jonka pistetodennäköisyys- tai tiheysfunktio on

$$f(x;\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})\exp\left[\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right]$$

jossa $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$.

Tällöin tunnusluku

$$\mathbf{T}(\mathbf{X}) = \left(\sum_{j=1}^{n} t_1(X_j), \ldots, \sum_{j=1}^{n} t_k(X_j) \right)$$

on *täydellinen*, jos joukko

$$\{(w_1(\boldsymbol{\theta}), \ldots, w_k(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\}$$

(Θ on parametrin $\boldsymbol{\theta}$ mahdollisten arvojen muodostama *parametriavaruus*) sisältää avaruuden $\mathbb{R}^k$ avoimen joukon.

Basun teoreeman todistuksessa *ei ole käytetty hyväksi* tyhjentävän tunnusluvun *minimaalisuutta*. Itse asiassa voidaan todistaa, että seuraava lause pätee:

Lause:

Jos minimaalisesti tyhjentävä tunnusluku *on olemassa*, niin jokainen täydellinen tunnusluku on minimaalisesti tyhjentävä.

Siten vaatimus tyhjentävän tunnusluvun minimaalisuudesta Basun teoreemassa on redundantti.

2.2. Uskottavuus

Uskottavuusfunktio

Olkoon

$$f(\mathbf{x};\theta)$$

otoksen

$$\mathbf{X} = (X_1, X_2, \ldots, X_n)$$

yhteisjakauman pistetodennäköisyys- tai *tiheysfunktio*. Oletetaan, että

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

on otoksen $\mathbf{X}$ havaittu arvo eli havaintopiste. Tällöin parametrin θ funktiota

$$L(\theta;\mathbf{x}) = f(\mathbf{x};\theta)$$

kutsutaan (otoksen $\mathbf{X}$) **uskottavuusfunktioksi**.

Erityisesti, jos $\mathbf{X}$ on *diskreetti* satunnaismuuttuja, niin

$$L(\theta;\mathbf{x}) = \Pr_\theta(\mathbf{X} = \mathbf{x})$$

Jos vertaamme uskottavuusfunktion arvoa kahdessa eri parametriavaruuden pisteessä ja havaitsemme, että

$$\Pr_{\theta_1}(\mathbf{X} = \mathbf{x}) = L(\theta_1;\mathbf{x}) > L(\theta_2;\mathbf{x}) = \Pr_{\theta_2}(\mathbf{X} = \mathbf{x})$$

niin otos, jonka olemme havainneet, on *uskottavampi*, jos

$$\theta = \theta_1$$

kuin silloin, kun

$$\theta = \theta_2$$

Uskottavuusperiaate

Olkoot $\mathbf{x}$ ja $\mathbf{y}$ ovat kaksi havaintopistettä, joille

$$L(\theta;\mathbf{x}) \propto L(\theta;\mathbf{y})$$

eli joille on olemassa parametrista θ riippumaton vakio $C(\mathbf{x}, \mathbf{y})$ siten, että

$$L(\theta;\mathbf{x}) = L(\theta;\mathbf{y})C(\mathbf{x},\mathbf{y})$$

Uskottavuusperiaatteen mukaan havaintopisteistä $\mathbf{x}$ ja $\mathbf{y}$ pitää tällöin tehdä *samat* parametria θ johtopäätökset. Erikoistapauksessa

$$C(\mathbf{x}, \mathbf{y}) = 1$$

uskottavuusperiaate sanoo, että jos havaintopisteet $\mathbf{x}$ ja $\mathbf{y}$ tuottavat *saman* uskottavuusfunktion, niin havaintopisteet $\mathbf{x}$ ja $\mathbf{y}$ sisältävät *saman* informaation parametrista θ.

Fidusiaalisuus

Fidusiaalisuusperiaatteen mukaan uskottavuudet voidaan tulkita *todennäköisyyksiksi*. Tämä merkitsee sitä, että jos uskottavuusfunktio

$$L(\theta;\mathbf{x})$$

jaetaan *normeeraustekijällä*

$$N(\mathbf{x}) = \int_{-\infty}^{+\infty} L(\theta;\mathbf{x})d\theta$$

(jos parametriavaruus on äärellinen tai numeroituva, niin integraali on korvattava summalla) niin

$$\frac{L(\theta;\mathbf{x})}{N(\mathbf{x})}$$

voidaan tulkita parametrin θ todennäköisyysjakaumaksi (olettaen, että $N(\mathbf{x}) < \infty$).

On syytä ottaa huomioon, että houkuttelevuudestaan huolimatta monet tilastotieteen teorian keskeisistä kehittäjistä *eivät ole hyväksyneet* fidusiaalisuusperiaatetta.

Esimerkki 2.1: Normaalinen fidusiaalinen jakauma

Oletetaan, että *havainnot*

$$X_1, X_2, \ldots, X_n$$

muodostavat satunnaisotoksen normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$ kuten esimerkissä 1.1, jossa oletettiin, että varianssiparametri σ^2 on *tunnettu*.

Esimerkissä 1.2 todettiin, että *otoksen yhteisjakauman tiheysfunktio* $f(\mathbf{x};\mu)$ voidaan faktoroida seuraavalla tavalla:

$$f(\mathbf{x};\mu) = (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]\right\}\exp\left\{-\frac{1}{2\sigma^2}n(\bar{x}-\mu)^2\right\}$$

Siten otoksen $X_1, X_2, \ldots, X_n$ uskottavuusfunktio on muotoa

$$L(\mu;\mathbf{x}) = (2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]\right\}\exp\left\{-\frac{1}{2\sigma^2}n(\bar{x}-\mu)^2\right\}$$

Olkoot **x** ja **y** kaksi havaintopistettä. Tällöin

$$L(\mu;\mathbf{x}) = L(\mu;\mathbf{y})C(\mathbf{x},\mathbf{y})$$

jos ja vain jos

$$\bar{x} = \bar{y}$$

jolloin

$$C(\mathbf{x},\mathbf{y}) = \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 - \sum_{i=1}^{n}(y_i-\bar{y})^2\right]\right\}$$

Siten *uskottavuusperiaatteesta* seuraa, että havaintopisteistä $\mathbf{x}$ ja $\mathbf{y}$ pitää tehdä *samat johtopäätökset parametrista* μ, jos

$$\bar{x} = \bar{y}$$

Fidusiaalinen todennäköisyysjakauma parametrille μ saadaan jakamalla uskottavuusfunktio $L(\theta;\mathbf{x})$ *normeeraustekijällä*

$$\begin{aligned} N(\mathbf{x}) &= \int_{-\infty}^{+\infty} L(\mu;\mathbf{x})d\mu \\ &= (2\pi)^{-(n-1)/2}\sigma^{-(n-1)}n^{-1}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right] \end{aligned}$$

Tulokseksi saadaan

$$\frac{L(\mu;\mathbf{x})}{N(\mathbf{x})} = (2\pi)^{-1/2}\sigma^{-1}n\exp\left\{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\right\}$$

Siten parametrin μ *fidusiaalisena jakaumana* on normaalijakauma $\mathrm{N}(\bar{x},\sigma^2/n)$.

Evidenssi

Olkoon $\mathbf{X}$ satunnaisvektori, jonka pistetodennäköisyys- tai tiheysfunktio on $f(\mathbf{x};\theta)$ ja olkoon θ parametriavaruuden Θ piste. Kutsutaan kolmikkoa

$$E = (\mathbf{X}, \theta, \{f(\mathbf{x};\theta)\})$$

tilastolliseksi kokeeksi.

Oletetaan, että on tehty tilastollinen koe E, jolloin on havaittu otos $\mathbf{X} = \mathbf{x}$ ja haluamme tehdä otoksen $\mathbf{X} = \mathbf{x}$ perusteella jonkin parametria θ koskevan johtopäätöksen. Olkoon tämä johtopäätös

$$\mathrm{Ev}(E,\mathbf{x})$$

millä tarkoitetaan *kokeeseen* E ja *havaintoihin* $\mathbf{x}$ *sisältyvää* **evidenssiä** *parametrista* θ.

Formaali tyhjentävyysperiaate

Formuloidaan tyhjentävyysperiaate uudelleen.

Olkoon

$$E = (\mathbf{X}, \theta, \{f(\mathbf{x};\theta)\})$$

tilastollinen koe ja olkoon tunnusluku $T(\mathbf{X})$ *tyhjentävä* parametrille θ. Jos $\mathbf{x}$ ja $\mathbf{y}$ ovat kaksi havaintopistettä, joille

$$T(\mathbf{X}) = T(\mathbf{Y})$$

niin

$$\mathrm{Ev}(E,\mathbf{x}) = \mathrm{Ev}(E,\mathbf{y})$$

Ehdollisuusperiaate

Olkoot

$$E_1 = (\mathbf{X}_1, \theta, \{f_1(\mathbf{x}_1;\theta)\})$$

ja

$$E_2 = (\mathbf{X}_2, \theta, \{f_2(\mathbf{x}_2;\theta)\})$$

kaksi tilastollista koetta, joilla ei välttämättä ole muita yhteisiä elementtejä kuin parametri θ.

Olkoon J satunnaismuuttuja, jolle (riippumatta parametrista θ ja satunnaismuuttujista $\mathbf{X}_1$ ja $\mathbf{X}_2$)

$$\Pr(J = 1) = \Pr(J = 2) = 0.5$$

Tarkastellaan *sekoitettua* koetta, jossa havaitaan *ensin* satunnaismuuttujan J arvo ja tehdään *sen jälkeen* koe E_J. Tämä merkitsee sitä, että sekoitettu koe on muotoa

$$E^* = (\mathbf{X}^*, \theta, \{f^*(\mathbf{x}^*;\theta)\})$$

jossa

$$\mathbf{X}^* = (j, \mathbf{X}_j)\ ,\ j = 1, 2$$

ja

$$f^*(\mathbf{x}^*;\theta) = f^*((j, \mathbf{x}_j);\theta) = \frac{1}{2} f_j(\mathbf{x}_j;\theta)\ ,\ j = 1, 2$$

Ehdollisuusperiaatteen mukaan

$$\mathrm{Ev}(E^*, (j, \mathbf{x}_j)) = \mathrm{Ev}(E_j, \mathbf{x}_j)\ ,\ j = 1, 2$$

Ehdollisuusperiaate sanoo, että jos valitsemme kahdesta tilastollisesta kokeesta toisen satunnaisesti ja havaitsemme kokeen tuloksena havaintopisteen $\mathbf{x}$, *parametrista θ saatava informaatio riippuu ainoastaan tehdystä kokeesta*.

Formaali uskottavuusperiaate

Olkoot

$$E_1 = (\mathbf{X}_1, \theta, \{f_1(\mathbf{x}_1;\theta)\})$$

ja

$$E_2 = (\mathbf{X}_2, \theta, \{f_2(\mathbf{x}_2;\theta)\})$$

kaksi tilastollista koetta, joilla ei välttämättä ole muita yhteisiä elementtejä kuin parametri θ.

Oletetaan, että $\mathbf{x}_1^*$ on havaintopiste kokeesta E_1 ja $\mathbf{x}_2^*$ on havaintopiste kokeesta E_2. Oletetaan edelleen, että

$$L(\theta;\mathbf{x}_2^*) = CL(\theta;\mathbf{x}_1^*)$$

kaikille θ. Vakio C saa riippua havaintopisteistä $\mathbf{x}_1^*$ ja $\mathbf{x}_2^*$, mutta ei saa riippua parametrista θ.

Formaalin uskottavuusperiaatteen mukaan

$$\mathrm{Ev}(E_1, \mathbf{x}_1^*) = \mathrm{Ev}(E_2, \mathbf{x}_2^*)$$

Formaalista uskottavuusperiaatteesta seuraa, että jos

$$E = (\mathbf{X}, \theta, \{f(\mathbf{x};\theta)\})$$

on *tilastollinen koe*, niin

$$\mathrm{Ev}(E,\mathbf{x})$$

saa riippua kokeesta E ja havaintopisteestä $\mathbf{x}$ vain *uskottavuusfunktion*

$$L(\theta; \mathbf{x})$$

kautta.

Birnbaumin teoreema:

Formaali tyhjentävyysperiaate ja *ehdollisuusperiaate* implikoivat *formaalin uskottavuus-periaatteen*, ja kääntäen, *formaali uskottavuusperiaate* implikoi *formaalin tyhjentävyys-periaatteen* ja *ehdollisuusperiaatteen*.

Perustelu (luonnos):

(i) Todistetaan, että *formaali tyhjentävyysperiaate* ja *ehdollisuusperiaate* implikoivat *formaalin uskottavuusperiaatteen*.

Olkoot (kuten *formaalissa uskottavuusperiaatteessa*)

$$E_1 = (\mathbf{X}_1, \theta, \{f_1(\mathbf{x}_1;\theta)\})$$

ja

$$E_2 = (\mathbf{X}_2, \theta, \{f_2(\mathbf{x}_2;\theta)\})$$

kaksi tilastollista koetta, joilla ei välttämättä ole muita yhteisiä elementtejä kuin parametri θ ja olkoot lisäksi $\mathbf{x}_1^*$ on havaintopiste kokeesta E_1 ja $\mathbf{x}_2^*$ on havaintopiste kokeesta E_2.

Olkoon (kuten *ehdollisuusperiaatteessa*)

$$E^* = (\mathbf{X}^*, \theta, \{f^*(\mathbf{x}^*;\theta)\})$$

sekoitettu koe, jossa havaitaan *ensin* satunnaismuuttujan J arvo ja tehdään *sen jälkeen* koe E_J, jossa

$$\mathbf{X}^* = (j, \mathbf{X}_j)\ ,\ j = 1,2$$

ja

$$f^*(\mathbf{x}^*;\theta) = f^*((j,\mathbf{x}_j);\theta) = \frac{1}{2} f_j(\mathbf{x}_j;\theta)\ ,\ j = 1,2$$

Määritellään kokeen E^* otosavaruudessa tunnusluku

$$T(j,\mathbf{x}_j) = \begin{cases} (1, x_1^*) & \text{, jos } j = 1 \text{ ja } \mathbf{x}_1 = \mathbf{x}_1^* \text{ tai } j = 2 \text{ ja } \mathbf{x}_2 = \mathbf{x}_2^* \\ (j, x_j) & \text{, muulloin} \end{cases}$$

Olkoon lisäksi

$$g(t;\theta) = g((j,\mathbf{x}_j);\theta) = f^*((j,\mathbf{x}_j);\theta_j)\ ,\ j = 1,2$$

ja

$$h(j,\mathbf{x}_j) = \begin{cases} C\ ,\ \text{jos } (j,\mathbf{x}_j) = (2,\mathbf{x}_2^*) \\ 1\ \ ,\ \text{muulloin} \end{cases}$$

Koska

$$g(T(j,\mathbf{x}_j);\theta)h(j,\mathbf{x}_j) = f^*((j,\mathbf{x}_j);\theta_j)\ ,\ j = 1,2$$

kaikille (j, $\mathbf{x}_j$), niin *faktorointiteoreemasta* seuraa, että

$$T(J,\mathbf{x}_J)$$

on tyhjentävä tunnusluku kokeessa E^*.

Edelleen *formaalista tyhjentävyysperiaatteesta* seuraa, että

$$\text{Ev}(E^*,(1,\mathbf{x}_1^*)) = \text{Ev}(E^*,(2,\mathbf{x}_2^*))$$

ja *ehdollisuusperiaatteesta* seuraa, että

$$\text{Ev}(E^*,(1,\mathbf{x}_1^*)) = \text{Ev}(E_1,\mathbf{x}_1^*))$$
$$\text{Ev}(E^*,(2,\mathbf{x}_2^*)) = \text{Ev}(E_2,\mathbf{x}_2^*))$$

Siten

$$\text{Ev}(E_1,\mathbf{x}_1^*)) = \text{Ev}(E_2,\mathbf{x}_2^*))$$

mikä merkitsee sitä, että *formaali uskottavuusperiaate on tosi*.

(ii) Todistetaan, että *formaali uskottavuusperiaate* implikoi *formaalin tyhjentävyysperiaatteen* ja *ehdollisuusperiaatteen*.

Tarkastellaan kokeita E^* ja E_j, jotka on määritelty kuten kohdassa (i).

Voidaan osoittaa, että

$$\text{Ev}(E^*,(j,\mathbf{x}_j^*)) = \text{Ev}(E_j,\mathbf{x}_j)\ ,\ j = 1,2$$

mikä merkitsee sitä, että *ehdollisuusperiaate on tosi*.

Edelleen, jos $T(\mathbf{X})$ on tyhjentävä ja

$$T(\mathbf{x}) = T(\mathbf{y})$$

niin uskottavuusfunktiot ovat proportionaalisia ja formaalista uskottavuusperiaatteesta seuraa, että

$$\text{Ev}(E,\mathbf{x}) = \text{Ev}(E,\mathbf{y})$$

mikä merkitsee sitä, että *ehdollisuusperiaate on tosi*.

■