Tilastollinen päättely

6. Bayeslaiset menetelmät

6.1. Johdanto

Bayesin kaava, Bayeslainen lähestymistapa, Ennakkotieto, Estimointi, Frekventistinen lähestymistapa, Frekvenssitulkinta, Klassinen lähestymistapa, Luottamustaso, Luottamusväli, Merkitsevyystaso, Otos, Otostieto, Parametri, Posteriorijakauma, Priorijakauma, Prioritieto, Testaus, Todennäköisyys, Todennäköisyysjakauma, Uskomus, Väliestimointi

6.2. Bayesin kaava

Bayesin kaava, Diskreetti jakauma, Ehdollinen jakauma, Ehdollinen todennäköisyys, Jatkuva jakauma, Kokonaistodennäköisyyden kaava, Ositus, Pistetodennäköisyysfunktio, Posterioritodennäköisyys, Prioritodennäköisyys, Tiheysfunktio

6.3. Priorijakaumat ja posteriorijakaumat

Bayesin kaava, Bayes-uskottavuus, Diskreetti jakauma, Ehdollinen jakauma, Epäinformatiivinen priorijakauma, Epäoleellinen priorijakauma, Jatkuva jakauma, Konjugaattiperhe, Konjugaattipriorijakauma, Otos, Otostieto, Parametri, Posteriorijakauma, Priorijakauma, Prioritieto, Uskottavuusfunktio

6.4. Bayes-estimointi

Estimaattori, Estimointi, Mediaani, Moodi, Odotusarvo, Parametri, Posteriorijakauma, Tunnusluku

6.5. Bayes-testit

Hylkäysalue, Nollahypoteesi, Parametri, Posteriorijakauma, Testaus, Testi, Testisuure, Vaihtoehtoinen hypoteesi

6.6. Bayes-luottamusvälit

Bayeslainen optimaalisuus, Luottamustaso, Luottamusväli, Parametri, Peittotodenäköisyys, Posteriorijakauma, Uskottavuusjoukko, Uskottavuustodennäköisyys, Uskottavuusväli

6.1. Johdanto

Klassisessa eli *frekventistisessä lähestymistavassa* tilastolliseen päättelyyn oletetaan, että havainnot noudattavat jakaumaa, jonka indeksoiva parametri θ on *kiinteä*, *ei-satunnainen vakio*. Klassisessa lähestymistavassa ajatellaan, että otos *sisältää kaiken parametria* θ *koskevan informaation* ja siten kaikki *parametria* θ *koskevat johtopäätökset voidaan perustaa otokseen ja sen jakaumaan*.

Klassisessa lähestymistavassa estimointiin ja testaukseen nojataan voimakkaasti *todennäköisyyden frekvenssitulkintaan* – siitä nimitys *frekventistinen lähestymistapa*:

(i) *Väliestimoinnissa* luottamusvälin *luottamustasolle* $1-\alpha$ annetaan seuraava tulkinta:

Jos otantaa toistetaan, niin luottamusväli peittää parametrin θ kiinteän, mutta tuntemattoman arvon $(1-\alpha)$ %:ssa otoksia.

(ii) *Testauksessa* testin *merkitsevyystasolle* α annetaan seuraava tulkinta:

Jos otantaa toistetaan tilanteessa, jossa nollahypoteesi H_0 koko ajan pätee, niin se joudutaan virheellisesti hylkäämään α %:ssa otoksia.

Bayeslaisessa lähestymistavassa tilastolliseen päättelyyn ajatellaan, että parametri θ on *satunnaismuuttuja*, jonka arvojen vaihtelua voidaan kuvata todennäköisyysjakaumalla, jota kutsutaan *priorijakaumaksi*.

Priorijakauma on *subjektiivinen* jakauma, joka kuvastaa tutkimuksen tekijän *uskomuksia* tai *ennakkokäsityksiä* parametrin θ käyttäytymisestä. Huomaa, että tutkimuksen tekijän pitää pystyä formuloimaan ennakkokäsityksensä priorijakauman muotoon *ilman otostietojen apua* (siis *jo ennen havaintojen keräämistä*).

Otoksen poimimisen jälkeen *parametrin* θ *priorijakauma päivitetään* ns. *posteriorijakaumaksi käyttämällä Bayesin kaavaa*. Bayesin kaava yhdistää parametria koskevan prioritiedon ja otostiedon toisiinsa. Bayeslaisessa lähestymistavassa kaikki *parametria* θ *koskevat johtopäätökset perustetaan sen posteriorijakaumaan*.

Matemaattisissa tieteissä ei yleensä ole koulukuntia. Tilastotiede on kuitenkin jakautunut *frekventisteihin* ja *bayeslaisiin*.

Frekventistisestä näkökulmasta prioritiedon liittäminen malleihin tuo mukaan epätoivottavan subjektiivisen elementin tutkimukseen. Lisäksi frekventistit ovat huomauttaneet, että vaikka prioritietoa parametrista olisikin käytettävissä, sen formulointi priorijakauman muotoon saattaa olla hyvin vaikeata.

Bayeslaiset tilastotieteilijät ovat tehneet tähän sen vastahuomautuksen, että tutkijalla on joka tapauksessa joitakin ennakkokäsityksiä tutkimustilanteesta ja on parempi pyrkiä formuloimaan ennakkokäsitykset priorijakauman muotoon kuin yrittää piilottaa ne maton alle. Jos tutkijalla ei ole kunnollista prioritietoa tutkimustilanteesta, hän voi pyrkiä formuloimaan epätietoisuutensa ns. *epäinformatiivisen* priorijakauman muotoon; ks. kappaletta 6.3. alla).

Lisäksi bayeslaiset tilastotieteilijät ovat kiinnittäneet huomiota siihen, että todennäköisyyden frekvenssitulkinnan käyttöön estimoinnissa ja testauksessa liittyy vaikeasti tulkittavaa puhetta siitä, *mitä tapahtuu*, *jos otantaa toistetaan*, *kun todellisuudessa käytettävissä on vain yksi ainoa otos – se otos*, *joka on poimittu*.

6.2. Bayesin kaava

Ositus

Joukon S osajoukot

$$B_1, B_2, \ldots , B_n$$

muodostavat joukon S **osituksen**, jos seuraavat ehdot pätevät:

(i) $B_i \neq \varnothing$, $i = 1, 2, \ldots, n$

(ii) $B_i \cap B_j = \varnothing$, $i \neq j$

(iii) $S = B_1 \cup B_2 \cup \cdots \cup B_n$

Kokonaistodennäköisyyden kaava

Olkoon A epätyhjä otosavaruuden S osajoukko:

$$A \subset S \text{ , } A \neq \varnothing$$

Oletetaan, että joukot

$$B_1, B_2, \ldots , B_n$$

muodostavat otosavaruuden S *osituksen*. Tällöin pätee **kokonaistodennäköisyyden kaava**

$$\Pr(A) = \sum_{i=1}^{n} \Pr(B_i)\Pr(A \mid B_i)$$

Bayesin kaava

Olkoon A epätyhjä otosavaruuden S osajoukko:

$$A \subset S \text{ , } A \neq \varnothing$$

Oletetaan, että joukot

$$B_1, B_2, \ldots , B_n$$

muodostavat otosavaruuden S *osituksen*. Tällöin pätee **Bayesin kaava**

$$\Pr(B_i \mid A) = \frac{\Pr(B_i)\Pr(A \mid B_i)}{\sum_{i=1}^{n} \Pr(B_i)\Pr(A \mid B_i)} \text{ , } i = 1, 2, \ldots, n$$

Bayesin kaava kertoo miten *prioritodennäköisyydet* $\Pr(B_i)$ (lat. *prior* = edeltävä) voidaan tapahtuman A havaitsemisen jälkeen päivittää *posterioritodennäköisyyksiksi* $\Pr(B_i|A)$ (lat. *posterior* = jälkeen tuleva).

Bayesin kaava diskreeteille todennäköisyysjakaumille

Olkoon $f_{XY}(x,y)$ *diskreettien* satunnaismuuttujien X ja Y *yhteisjakauman pistetodennäköisyysfunktio*. Olkoon lisäksi $f_Y(y)$ satunnaismuuttujan Y *reunajakauman pistetodennäköisyysfunktio* ja $f_{X|Y}(x|y)$ *satunnaismuuttujan x ehdollisen jakauman* (ehdolla y) *pistetodennäköisyysfunktio*.

Tällöin

$$f_{XY}(x,y) = f_{X|Y}(x\,|\,y)f_Y(y)$$

ja satunnaismuuttujan y *ehdollisen jakauman* (ehdolla x) *pistetodennäköisyysfunktio* $f(y|x)$ saadaan kaavasta

$$f_{Y|X}(y\,|\,x) = \frac{f_{X|Y}(x\,|\,y)f_Y(y)}{f_X(x)}$$

jossa

$$f_X(x) = \sum_y f_{X|Y}(x\,|\,y)f_Y(y)$$

satunnaismuuttujan X *reunajakauman pistetodennäköisyysfunktio*. Summa lasketaan yli kaikkien satunnaismuuttujan Y mahdollisten arvojen.

Bayesin kaava jatkuville todennäköisyysjakaumille

Olkoon $f_{XY}(x,y)$ *jatkuvien* satunnaismuuttujien X ja Y *yhteisjakauman tiheysfunktio*. Olkoon lisäksi $f_Y(y)$ satunnaismuuttujan Y *reunajakauman tiheysfunktio* ja $f_{X|Y}(x|y)$ *satunnaismuuttujan x ehdollisen jakauman* (ehdolla y) *tiheysfunktio*.

Tällöin

$$f_{XY}(x,y) = f_{X|Y}(x\,|\,y)f_Y(y)$$

ja satunnaismuuttujan y *ehdollisen jakauman* (ehdolla x) *tiheysfunktio* $f(y|x)$ saadaan kaavasta

$$f_{Y|X}(y\,|\,x) = \frac{f_{X|Y}(x\,|\,y)f_Y(y)}{f_X(x)}$$

jossa

$$f_X(x) = \int f_{X|Y}(x\,|\,y)f_Y(y)dy$$

satunnaismuuttujan X *reunajakauman tiheysfunktio*.

Huomautus:

Bayesin kaava todennäköisyysjakaumille voidaan ilmeisellä tavalla formuloida myös *sekamuodoille*, joissa toinen satunnaismuuttujista X ja Y on *diskreetti* ja toinen on *jatkuva*.

6.3. Priorijakaumat ja posteriorijakaumat

Otosinformaatio

Olkoon

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

satunnaisotos satunnaismuuttujan X 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* $f(x;\theta)$:

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

Koska parametria θ pidetään Bayeslaisessa lähestymistavassa *satunnaismuuttujana*, havaintojen jakauma $f(x;\theta)$ tulkitaan *satunnaismuuttujan X ehdolliseksi jakaumaksi ehdolla* θ. Kirjoitamme siksi jatkossa

$$X \sim f(x \mid \theta)$$

Koska havainnot $X_1, X_2, \ldots, X_n$ muodostavat satunnaisotoksen jakaumasta $f(x|\theta)$, niin otoksen

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

yhteisjakauman pistetodennäköisyys- tai *tiheysfunktio* voidaan esittää muodossa

$$f(\mathbf{x} \mid \theta) = f(x_1, x_2, \ldots, x_n \mid \theta) = f(x_1 \mid \theta) f(x_2 \mid \theta) \cdots f(x_n \mid \theta)$$

jossa

$$f(x_i \mid \theta)\,,\; i = 1, 2, \ldots, n$$

on yksittäiseen havaintoon X_i , $i = 1, 2, \ldots, n$ liittyvä pistetodennäköisyys- tai tiheysfunktio ja

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

Otoksen **X uskottavuusfunktio**

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

on otoksen *yhteisjakauman pistetodennäköisyys-* tai *tiheysfunktion f* arvo pisteessä **x** tulkittuna *parametrin* θ *arvojen funktioksi*. *Uskottavuusperiaatteen* mukaan uskottavuusfunktio $L(\theta)$ sisältää kaiken *otosinformaation* parametrista θ.

Priorijakauma

Oletetaan nyt, että parametri θ on *satunnaismuuttuja*, jonka *pistetodennäköisyys-* tai *tiheysfunktio* on $\pi(\theta)$:

$$\theta \sim \pi(\theta)$$

Kutsumme todennäköisyysjakaumaa $\pi(\theta)$ parametrin θ **priorijakaumaksi** ja se sisältää *priori-informaation* eli *ennakkotiedon* parametrista θ.

Posteriorijakauma

Parametrin θ **posteriorijakauma**

$$\pi(\theta \mid \mathbf{x})$$

jossa

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

saadaan *otoksen* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ *yhteisjakaumasta* ja *parametrin* θ *priorijakaumasta Bayesin kaavalla*. Parametrin θ posteriorijakauma on parametrin θ ehdollinen jakauma ehdolla $\mathbf{X} = \mathbf{x}$. Ehdon muodostaa siis poimittu *otos* eli ne havainnot, jotka todella on saatu.

Diskreettien jakaumien tapauksessa *posteriorijakauma* saadaan kaavasta

$$\begin{aligned}\pi(\theta\,|\,\mathbf{x}) &= \frac{f(\mathbf{x}\,|\,\theta)\pi(\theta)}{f(\mathbf{x})} \\ &= \frac{f(\mathbf{x}\,|\,\theta)\pi(\theta)}{\sum_{\theta} f(\mathbf{x}\,|\,\theta)\pi(\theta)}\end{aligned}$$

jossa

$$f(\mathbf{x}) = \sum_{\theta} f(\mathbf{x}\,|\,\theta)\pi(\theta) = \sum_{\theta} f(\mathbf{x},\theta)$$

on otoksen **X** *reunajakauma*, ns. **Bayes-uskottavuus**. Summa lasketaan yli kaikkien satunnaismuuttujan θ mahdollisten arvojen. Huomaa, että

$$\sum_{\theta} f(\mathbf{x}\,|\,\theta)\pi(\theta) = \mathrm{E}_{\theta}[f(\mathbf{x}\,|\,\theta)]$$

Jatkuvien jakaumien tapauksessa *posteriorijakauma* saadaan kaavasta

$$\begin{aligned}\pi(\theta\,|\,\mathbf{x}) &= \frac{f(\mathbf{x}\,|\,\theta)\pi(\theta)}{f(\mathbf{x})} \\ &= \frac{f(\mathbf{x}\,|\,\theta)\pi(\theta)}{\int f(\mathbf{x}\,|\,\theta)\pi(\theta)d\theta}\end{aligned}$$

jossa

$$f(\mathbf{x}) = \int f(\mathbf{x}\,|\,\theta)\pi(\theta)d\theta = \int f(\mathbf{x},\theta)d\theta$$

on otoksen **X** *reunajakauma*, ns. **Bayes-uskottavuus**. Huomaa, että

$$\int f(\mathbf{x}\,|\,\theta)\pi(\theta)d\theta = \mathrm{E}_{\theta}[f(\mathbf{x}\,|\,\theta)]$$

Huomautus:

Posteriorijakauman kaava voidaan ilmeisellä tavalla formuloida myös *sekamuodoille*, joissa toinen satunnaismuuttujista X ja θ on *diskreetti* ja toinen on *jatkuva*.

Bayeslaisessa lähestymistavassa *kaikki parametria θ koskeva tilastollinen päättely perustetaan sen posteriorijakaumaan*.

Koska

$$\pi(\theta\,|\,\mathbf{x}) \propto f(\mathbf{x}\,|\,\theta)\pi(\theta)$$

niin posteriorijakaumaa johdettaessa voidaan toimia siten, että muodostetaan otoksen yhteisjakauman $f(\mathbf{x}|\theta)$ ja priorijakauman $\pi(\theta)$ tulo ja *normeerataan* tulo todennäköisyysjakaumaksi.

Itse asiassa posteriorijakaumaa muodostettaessa riittää käsitellä otoksen yhteisjakauman ja priorijakauman *ytimiä* eli niitä jakaumien pistetodennäköisyys- tai tiheysfunktioiden osia, jotka jäävät jäljelle, kun funktioiden lausekkeista jätetään pois kaikki (kerrannaiset) *vakio-osat*, so. ne (kerrannaiset) osat, jotka eivät riipu ko. funktioiden argumenteista.

On syytä huomata, että ei ole edes välttämätöntä, että priorijakauma *aito todennäköisyysjakauma*. Silti posteriorijakaumaksi saadaan aito todennäköisyysjakauma. Tällä tarkoitetaan sitä, että ehdon

$$\int_{-\infty}^{+\infty} \pi(\theta) d\theta = 1$$

ei tarvitse päteä, kunhan

$$\pi(\theta) > 0$$

kaikille θ. Bayesin kaava antaa silti (sopivan normeerauksen jälkeen) posteriorijakaumaksi $f(\theta \mid \mathbf{x})$ aidon todennäköisyysjakauman.

Jos $\pi(\theta) > 0$ ja

$$\int_{-\infty}^{+\infty} \pi(\theta) d\theta = \infty$$

funktiota $\pi(\theta)$ on tapana kutsua *epäoleelliseksi priorijakaumaksi*.

Bayeslaisessa lähestymistavassa on tapana käyttää tällaisia epäoleellisia priorijakaumia esimerkiksi silloin, kun halutaan kuvata sitä, että prioritieto parametrin käyttäytymisestä on niin *heikkoa*, että *kaikkia vaihtoehtoja voidaan pitää yhtä epämääräisinä*; ks. kohtaa epäinformatiivisista priorijakaumista alla.

Konjugaattiperheet

Posteriorijakauman ei tarvitse olla mitään tunnettua tyyppiä, jolloin sen hallitseminen saattaa olla hyvin vaikeata. Jos posteriorijakauma on samaa tyyppiä kuin priorijakauma, niin tällaista ongelmaa ei ole. Sanomme, että tällaisessa tapauksessa priorijakauma kuuluu havaintojen jakauman *konjugaattiperheeseen*.

Olkoon $\mathcal{F}$ parametrin θ indeksoima tiheys- tai pistetodennäköisyysfunktioiden $f(x|\theta)$ perhe. Priorijakaumien perhe Π on perheen $\mathcal{F}$ **konjugaattiperhe**, jos parametrin θ posteriorijakauma kuuluu jokaiselle $f \in \mathcal{F}$ ja jokaiselle priorijakaumalle $\pi \in \Pi$ perheeseen Π.

Alla olevaan taulukkoon on kerätty eräiden keskeisten jakaumien parametrien konjugaattipriorijakaumat.

Havaintojen jakauma	**Parametri**	**Konjugaattipriorijakauma**
Bernoulli-jakauma	$p = \mathrm{E}(X)$	Beta-jakauma
Binomijakauma	$p = \mathrm{E}(X)$	Beta-jakauma
Poisson-jakauma	$\lambda = \mathrm{E}(X)$	Gamma-jakauma
Eksponenttijakauma	$\lambda = 1/\mathrm{E}(X)$	Gamma-jakauma
Normaalijakauma (σ^2 on tunnettu)	$\mu = \mathrm{E}(X)/n$	Normaalijakauma
Normaalijakauma (μ on tunnettu)	$\sigma^2 = \mathrm{E}[(X-\mu)^2]$	Invertoitu gamma-jakauma

Epäinformatiiviset priorijakaumat

Jos parametrin käyttäytymisestä ei ole käytettävissä sellaista ennakkotietoa, joka mahdollistaa priorijakauman tarkan formuloinnin, niin Bayeslaisessa lähestymistavassa pyritään kuvaamaan kaikkien vaihtoehtojen epämääräisyyttä **epäinformatiivisen priorijakauman** avulla. Tämä voidaan tehdä määrittelemällä parametrille (tai jollekin sen muunnokselle) *lokaalisti tasainen jakauma*.

Jos parametriavaruutena Θ on reaalilukujen joukko $\mathbb{R}$, niin parametrille θ voidaan määritellä *epäinformatiivinen priorijakauma* kaavalla

$$\pi(\theta)d\theta \propto d\theta$$

jolloin

$$\pi(\theta) \propto k$$

jossa k on *vakio*.

Jos parametriavaruutena Θ on positiivisten reaalilukujen joukko $\mathbb{R}^+$, niin parametrin θ logaritmimuunnokselle $\phi = \log(\theta)$ voidaan määritellä *epäinformatiivinen priorijakauma* kaavalla

$$\pi(\phi)d\phi \propto d\phi$$

jolloin

$$\pi(\theta) \propto \frac{1}{\theta}$$

koska

$$d\phi = \frac{d\theta}{\theta}$$

Kummassakin tapauksessa priorijakauma on *epäoleellinen* eli

$$\int_{\Theta} \pi(\theta)d\theta = \infty$$

Molemmat priorijakaumat kuvaavat *vaihtoehtojen epämääräisyyttä* seuraavassa mielessä:

(i) Tapauksessa $\Theta = \mathbb{R}$ suhde

$$\frac{\Pr\left[\theta \in (-\infty, a)\right]}{\Pr\left[\theta \in (a, +\infty)\right]}$$

jossa

$$\Pr\left[\theta \in (c, d)\right] = \int_c^d \pi(\theta)d\theta \text{ , } c, d \in \mathbb{R} \cup \{-\infty, +\infty\} \text{ , } c < d$$

on *määrittelemätön* eli muotoa ∞/∞ kaikille a.

(ii) Tapauksessa $\Theta = \mathbb{R}^+$ suhde

$$\frac{\Pr\left[\theta \in (0, a)\right]}{\Pr\left[\theta \in (a, +\infty)\right]}$$

jossa

$$\Pr\left[\theta \in (c,d)\right] = \int_c^d \pi(\theta)d\theta \text{ , } c,d \in \mathbb{R} \cup \{+\infty\} \text{ , } c < d$$

on *määrittelemätön* eli muotoa ∞/∞ kaikille a.

6.4. Bayes-estimointi

”Puhdasveriset” Bayeslaiset tilastotieteilijät eivät yritä estimoida todennäköisyysjakaumien parametreja samassa mielessä kuin frekventistit. Frekventisti on tyytyväinen, kun hän on kyennyt tuottamaan parametrille *piste-estimaatin*. Sen sijaan Bayeslainen tilastotieteilijä haluaa konstruoida parametrille kokonaisen *posteriorijakauman*.

Parametrin *Bayes-estimaattorina* käytetään kuitenkin yleisesti jotakin *posteriorijakauman sijaintia kuvaavaa tunnuslukua*. Tällöin kyseeseen tulevat posteriorijakauman **odotusarvo**, **mediaani** tai **moodi**. Varsinkin posteriorijakauman *moodi* on suosittu valinta Bayeslaisessa lähestymistavassa estimointiin, koska se *asettuu* ”*todennäköisimpien*” *parametrin arvojen kohdalle*.

6.5. Bayes-testit

Olkoon

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

otoksen

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

yhteisjakauman tiheysfunktio,

$$\pi(\theta)$$

parametrin θ *priorijakauma* ja

$$\pi(\theta \mid \mathbf{x})$$

parametrin θ *posteriorijakauma*.

Olkoon testauksen kohteena oleva nollahypoteesi H_0 muotoa

$$\mathrm{H}_0 : \theta \in \Theta_0 \subset \Theta$$

jossa Θ_0 on jokin parametriavaruuden Θ osajoukko ja olkoon vaihtoehtoinen hypoteesi H_1 muotoa

$$\mathrm{H}_1 : \theta \in \Theta_0^c \subset \Theta$$

jossa Θ_0^c on joukon Θ_0 komplementti.

Määritellään todennäköisyydet

$$\Pr(\theta \in \Theta_0 \mid \mathbf{X} = \mathbf{x}) = \Pr(\mathrm{H}_0 \text{ pätee} \mid \mathbf{X} = \mathbf{x})$$

ja

$$\Pr(\theta \in \Theta_0^c \mid \mathbf{X} = \mathbf{x}) = \Pr(\mathrm{H}_1 \text{ pätee} \mid \mathbf{X} = \mathbf{x})$$

Huomaa, että nämä todennäköisyydet eivät ole mielekkäitä frekventistille, koska klassisessa lähestymistavassa parametri θ ei ole satunnaismuuttuja vaan kiinteä luku ja lisäksi hypoteesit ovat väitteitä, jotka ovat tosia tai epätosia ja siten niihin ei voida liittää nollasta ja ykkösestä poikkeavia todennäköisyyksiä.

Bayeslaisessa lähestymistavassa testaukseen voidaan toimia esimerkiksi seuraavalla tavalla:

Hyväksytään nollahypoteesi H_0, jos

$$\Pr(\theta \in \Theta_0 \mid \mathbf{X} = \mathbf{x}) \geq \Pr(\theta \in \Theta_0^c \mid \mathbf{X} = \mathbf{x})$$

ja *hylätään* nollahypoteesi H_0, jos

$$\Pr(\theta \in \Theta_0 \mid \mathbf{X} = \mathbf{x}) < \Pr(\theta \in \Theta_0^c \mid \mathbf{X} = \mathbf{x})$$

Tämä merkitsee sitä, että **testisuureena** käytetään otoksen funktiota

$$\kappa(\mathbf{X}) = \Pr(\theta \in \Theta_0^c \mid \mathbf{X})$$

ja testin **hylkäysalue** on muotoa

$$\left\{ \mathbf{x} \,\middle|\, \Pr(\theta \in \Theta_0^c \mid \mathbf{X} = \mathbf{x}) > \tfrac{1}{2} \right\}$$

jolloin testin **hyväksymisalue** on muotoa

$$\left\{ \mathbf{x} \,\middle|\, \Pr(\theta \in \Theta_0^c \mid \mathbf{X} = \mathbf{x}) \leq \tfrac{1}{2} \right\}$$

Jos haluamme erityisesti suojautua nollahypoteesin H_0 virheellistä hylkäystä vastaan, voimme ottaa hylkäysalueeksi joukon

$$\left\{ \mathbf{x} \,\middle|\, \Pr(\theta \in \Theta_0^c \mid \mathbf{X} = \mathbf{x}) > \gamma \right\}$$

jossa γ on jokin ykköstä lähellä oleva luku kuten 0.95 tai 0.99.

6.6. Bayes-luottamusvälit

Kun frekventistisessä lähestymistavassa tilastotieteeseen puhutaan parametrin ja sen luottamusvälin suhteesta, olisi aina hyvä käyttää sanontaa:

> ”*Otoksesta konstruoitu luottamusväli peittää parametrin arvon luottamustason ilmaisemalla todennäköisyydellä*.”

Näin tulisi korostetuksi sitä, että frekventistille *luottamisväli on satunnaissuure*, *jonka päätepisteet vaihtelevat satunnaisesti otoksesta toiseen* ja *parametri on kiinteä*, *ei-satunnainen vakio*.

Näin ei kuitenkaan aina sanota, vaan sanotaan, että

> ”*Parametrin arvo on otoksesta konstruoidun luottamusvälin sisällä luottamustason ilmaisemalla todennäköisyydellä*.”

Tähän sanontaan liittyy se epätäsmällisyys, että luottamusväli on *satunnaissuure*, jolle on realisoitunut poimitussa otoksessa yksi sen mahdollisista arvoista ja koska parametri on *kiinteä*, *ei-satunnainen vakio*, sen arvo on realisoituneen välin sisällä joko todennäköisyydellä 0 tai todennäköisyydellä 1.

Kun sanomme, että realisoitunut luottamusväli peittää parametrin arvon luottamustasolla $(1-\alpha)$, niin tarkasti ottaen ilmaisu tarkoittaa sitä, että $(1-\alpha)$ % kaikista mahdollisista havaintopisteistä johtaa luottamusväliin, joka peittää parametrin ja α % kaikista mahdollisista havaintopisteistä johtaa luottamusväliin, joka ei sitä tee.

Koska erilaisia käsitteitä ei haluta sekoittaa toisiinsa, Bayeslaisessa lähestymistavassa tilastotieteeseen on tapana puhua luottamusjoukkojen tai -välien sijasta *uskottavuusjoukoista* tai *-väleistä*.

Olkoon

$$\pi(\theta \mid \mathbf{x})$$

parametrin θ *posteriorijakauma* ehdolla $\mathbf{X} = \mathbf{x}$ (so. kun otostietona on $\mathbf{X} = \mathbf{x}$).

Olkoon A jokin parametriavaruuden Θ osajoukko:

$$A \subset \Theta$$

Tällöin joukon A **uskottavuustodennäköisyys** on

$$\Pr(\theta \in A \mid \mathbf{x}) = \int_A \pi(\theta \mid \mathbf{x}) d\theta$$

ja A on **uskottavuusjoukko** parametrille θ. Jos posteriorijakauma $\pi(\theta \mid \mathbf{x})$ on jonkin diskreetin jakauman pistetodennäköisyysfunktio, niin yo. kaavan integraali on korvattava summalla.

Uskottavuusjoukkojen *uskottavuustodennäköisyydet* määrätään *posteriorijakaumasta*. Posteriorijakauman todennäköisyysmekanismi taas perustuu *priorijakauman* todennäköisyysmekanismiin. Tutkija on kuvannut priorijakauman avulla parametria koskevien *ennakkokäsitystensä* varmuutta. Tämän kuvauksen pitää tapahtua jo ennen kuin tutkija on saanut käyttöönsä otostietoa. *Uskottavuustodennäköisyydet* kuvaavat tutkijan parametria koskevien käsitysten varmuutta sen jälkeen, kun hän on *päivittänyt* parametria koskevat ennakkokäsityksensä otostiedolla.

Kun *bayeslainen tilastotieteilijä* sanoo, että parametri kuuluu 90 %:n todennäköisyydellä konstruoituun uskottavuusjoukkoon, hän tarkoittaa seuraavaa: Sen jälkeen kun otostieto on yhdistetty prioritietoon, hän on 90 %:sen varma siitä, että parametri kuuluu uskottavuusjoukkoon.

Sen sijaan luottamusjoukkojen *peittotodennäkäisyydet* kuvaavat *otantaan liittyvää epävarmuutta* ja siihen liittyvä todennäköisyysmekanismi perustuu siihen, että ajatellaan, että otantaa voidaan ainakin hypoteettisesti toistaa.

Kun *frekventisti* sanoo, että luottamusjoukko peittää 90 %:n todennäköisyydellä parametrin arvon, hän tarkoittaa seuraavaa: Jos otantaa toistetaan, niin 90 % konstruoiduista luottamusjoukoista peittää parametrin todellisen arvon.

Bayeslainen optimaalisuus

Klassisisen lähestymistavan *luottamusjoukkoja* konstruoitaessa haluttiin niistä joukoista, joilla on sama *peittotodennäköisyys*, valita mahdollisimman ”*pieni*”. Sama ajatus voidaan liittää myös bayeslaisiin *uskottavuusjoukkoihin*.

Olkoon $\pi(\theta \mid \mathbf{x})$ parametrin θ *posteriorijakauma* ehdolla $\mathbf{X} = \mathbf{x}$ (so. kun otostietona on $\mathbf{X} = \mathbf{x}$). Haluamme siis löytää joukon $C(\mathbf{x})$, joka toteuttaa seuraavat kaksi ehtoa:

(i) $$\int_{C(\mathbf{x})} \pi(\theta \mid \mathbf{x}) d\theta = 1 - \alpha$$

(ii) Joukon $C(\mathbf{x})$ koko on korkeintaan yhtä suuri kuin joukon $C^*(\mathbf{x})$ koko kaikille joukoille $C^*(\mathbf{x})$, jotka toteuttavat ehdon

$$\int_{C^*(\mathbf{x})} \pi(\theta \mid \mathbf{x}) d\theta \geq 1-\alpha$$

Jos uskottavuusjoukkona on *väli* ja joukon kokoa mitataan *välin pituudella*, pätee seuraava lause (vrt. luvun 5. kappaleen 5.3. kohdassa **Peittotodennäköisyys ja luottamusjoukon koko** esitettyyn lauseeseen):

Lause:

Olkoon posteriorijakauma $\pi(\theta \mid \mathbf{x})$ *unimodaalinen*. Valitaan luku $\alpha \in [0,1]$. Tällöin *lyhyin* uskottavuusväli parametrille θ toteuttaa ehdon

$$\{\theta \mid \pi(\theta \mid \mathbf{x}) \geq k\}$$

jossa

$$\int_{\{\theta \mid \pi(\theta \mid \mathbf{x}) \geq k\}} \pi(\theta \mid \mathbf{x}) d\theta = 1-\alpha$$

Lauseen määrittelemää uskottavuusjoukkoa kutsutaan **korkeimman posterioritiheyden alueeksi**.