Tilastollinen päättely

4. Hypoteesien testaus

4.1. Johdanto

Hylkäysalue, Hypoteesi, Hyväksymisalue, Kriittinen alue, Nollahypoteesi, Otos, Parametri, Parametriavaruus, Perusjoukko, Päätös, Testi, Testisuure, Vaihtoehtoinen hypoteesi, Väite

4.2. Testien konstruointi

Hylkäysalue, Hypoteesi, Hyväksymisalue, Nollahypoteesi, Normaalijakauma, Osamäärä-testisuure, Otos, Parametri, Parametriavaruus, Päätös, Suurimman uskottavuuden estimaattori, Suurimman uskottavuuden estimointimenetelmä, Testi, Testisuure, Tyhjentävyys, Uskottavuus-funktio, Uskottavuusosamäärä, Vaihtoehtoinen hypoteesi

4.3. Testien vertailu

1. lajin virhe, 2. lajin virhe, Harhaton testi, Hylkäysalue, Hylkäysvirhe, Hypoteesi, Hyväksymisalue, Hyväksymisvirhe, Kaksisuuntainen hypoteesi, Karlinin ja Rubinin teoreema, Kelvollinen *p*-arvo, Merkitsevyystaso, Monotoninen uskottavuusosamäärä, Neymanin ja Pearsonin lemma, Nolla-hypoteesi, Osamäärätestisuure, Otos, Parametri, Parametriavaruus, *p*-arvo, Päätös, Suurimman uskottavuuden estimaattori, Suurimman uskottavuuden estimointimenetelmä, Tasaisesti voimakkain testi, Testi, Testin koko, Testin taso, Testisuure, Tyhjentävyys, Uskottavuusfunktio, Uskottavuusosamäärä, Vaihtoehtoinen hypoteesi, Virheet testauksessa, Virhetodennäköisyys, Voimakkuus, Voimakkuusfunktio, Yhdistetty hypoteesi, Yksinkertainen hypoteesi, Yksisuuntainen hypoteesi

4.1. Johdanto

Hypoteesi

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* $f(x;\theta)$:

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

Olkoon

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

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

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*

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

Satunnaismuuttujan X *pistetodennäköisyys*- tai *tiheysfunktio* $f(x;\theta)$ määrittelemä todennäköisyysjakauma kuvaa *satunnaismuuttujan X arvojen vaihtelua perusjoukossa* ja parametri θ kuvaa jotakin *perusjoukon ominaisuutta*.

Tilastollinen **hypoteesi** on jokin parametria θ koskeva *väite*. **Hypoteesin testauksen** *tavoitteena on päättää kumpi kahdesta vastakkaisesta perusjoukon parametria koskevasta hypoteesista eli väitteestä on tosi perusjoukosta poimitun otoksen perusteella*.

Huomautus:

Tarkastelemme tässä luvussa vain todennäköisyysjakauman *parametreja koskevien hypoteesien* testausta.

Hypoteesin testauksen vastakkaisia hypoteeseja kutsutaan **nollahypoteesiksi** ja **vaihtoehtoiseksi hypoteesiksi**. Merkitään

H_0 : Nollahypoteesi

H_1 : Vaihtoehtoinen hypoteesi

Olkoon θ perusjoukkoa kuvaava parametri ja olkoon Θ *parametriavaruus* eli mahdollisten parametrin θ arvojen joukko. Nollahypoteesin H_0 yleinen muoto on

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

jossa Θ_0 on jokin parametriavaruuden Θ osajoukko. Tällöin vaihtoehtoinen hypoteesin H_1 yleinen muoto on

$$H_1 : \theta \in \Theta_0^c \subset \Theta$$

jossa Θ_0^c on nollahypoteesin H_0 määrittelemän joukon Θ_0 komplementti.

Esimerkkejä:

Olkoon *nollahypoteesi* H_0 muotoa

$$H_0 : \theta = \theta_0$$

jossa θ_0 on parametrin θ jokin mahdollinen arvo. Tällöin *vaihtoehtoinen hypoteesi* H_1 on muotoa

$$H_1 : \theta \neq \theta_0$$

Olkoon *nollahypoteesi* H_0 muotoa

$$H_0 : \theta \geq \theta_0$$

jossa θ_0 on parametrin θ jokin mahdollinen arvo. Tällöin *vaihtoehtoinen hypoteesi* H_1 on muotoa

$$H_1 : \theta < \theta_0$$

Olkoon *nollahypoteesi* H_0 muotoa

$$H_0 : \theta \leq \theta_0$$

jossa θ_0 on parametrin θ jokin mahdollinen arvo. Tällöin *vaihtoehtoinen hypoteesi* H_1 on muotoa

$$H_1 : \theta > \theta_0$$

Testi

Testauksessa pyritään päättämään *otoksesta saadun informaation perusteella* **jätetäänkö nollahypoteesi** H_0 **voimaan** (eli **hyväksytäänkö nollahypoteesi** H_0) vai **hylätäänkö nollahypoteesi** H_0 ja **hyväksytään vaihtoehtoinen hypoteesi** H_1.

Testi on *päätössääntö*, joka jakaa *mahdollisten havaintoarvojen joukon* kahteen osajoukkoon:

(i) Niiden havaintoarvojen joukko, jotka johtavat *nollahypoteesin* H_0 *hyväksymiseen*.

(ii) Niiden havaintoarvojen joukko, jotka johtavat *nollahypoteesin* H_0 *hylkäämiseen* ja *vaihtoehtoisen hypoteesin* H_1 *hyväksymiseen*.

Sitä otosavaruuden osajoukkoa, joka johtaa *nollahypoteesin* H_0 *hylkäämiseen* kutsutaan testin **hylkäysalueeksi** tai **kriittiseksi alueeksi**. Sitä otosavaruuden osajoukkoa, joka johtaa *nollahypoteesin* H_0 *jäämiseen voimaan* kutsutaan testin **hyväksymisalueeksi**.

Tilastollinen testi on siis päätössääntö, joka jakaa otosavaruuden hyväksymisalueeseen ja hylkäysalueeseen. Hylkäysalue pyritään valitsemaan siten, että nollahypoteesista H_0 voidaan pitää kiinni, ellei otoksesta saatu informaatio ole *kyllin vahvaa* nollahypoteesin H_0 hylkäämiseksi.

Olkoon

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

perusjoukosta poimittu *otos*. Testi perustetaan aina johonkin *testisuureeseen*. **Testisuure** on otoksen funktio; merkitään testisuuretta yleisesti:

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

Esimerkki 1.1: Normaalijakauma

Olkoon

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

otos *normaalijakaumasta* $\mathrm{N}(\mu,\sigma^2)$ ja olkoon *nollahypoteesina*

$$\mathrm{H}_0 : \mu \ge \mu_0$$

ja *vaihtoehtoisena hypoteesina*

$$\mathrm{H}_1 : \mu < \mu_0$$

Koska havaintojen *aritmeettinen keskiarvo*

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

on *tyhjentävä* ja *paras harhaton estimaattori* parametrille μ, niin on luontevaa perustaa testi nollahypoteesille H_0 tunnuslukuun $\bar{X}$.

On järkevää valita *hylkäysalueeksi* joukko

$$\{(x_1, x_2, \ldots, x_n) \,|\, \bar{x} < c\}$$

jolloin *hyväksymisalue* saa muodon

$$\{(x_1, x_2, \ldots, x_n) \,|\, \bar{x} \ge c\}$$

Tarkastelemme myöhemmin kappaleessa 4.3 sitä, miten *kriittinen raja* tai *arvo c* on valittava, jotta testisuureen todennäköisyys joutua hylkäysalueelle nollahypoteesin H_0 pätiessä eli todennäköisyys

$$\Pr_{\theta_0}\{(x_1, x_2, \ldots, x_n) \,|\, \bar{x} < c\}$$

olisi halutun suuruinen.

4.2. Testien konstruointi

Uskottavuusosamäärätesti

Uskottavuusosamäärämenetelmä on yleinen menetelmä testien konstruoimiseksi ja sillä on läheinen kytkentä *suurimman uskottavuuden estimointimenetelmään*.

Satunnaisotos

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* $f(x;\theta)$:

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

Uskottavuusfunktio

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* on

$$f(x_1, x_2, \ldots, x_n;\theta) = f(x_1;\theta) f(x_2;\theta) \cdots f(x_n;\theta)$$

jossa

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

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

Otoksen $X_1, X_2, \ldots , X_n$ **uskottavuusfunktio**

$$L(\theta; x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n;\theta)$$

on havaintojen $X_1, X_2, \ldots , X_n$ *yhteisjakauman pistetodennäköisyys*- tai *tiheysfunktion* f arvo pisteessä

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

tulkittuna *parametrin* θ *arvojen funktioksi*. Uskottavuusperiaatteen mukaan uskottavuusfunktio L sisältää *kaiken* (stokastisen) *informaation otoksesta*.

Osamäärätestisuure ja osamäärätesti

Olkoon θ jakaumaa $f(x;\theta)$ kuvaava parametri ja olkoon Θ *parametriavaruus* eli mahdollisten parametrin θ arvojen joukko. Olkoon nollahypoteesi H_0 muotoa

$$H_0 : \theta \in \Theta_0 \subset \Theta$$

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.

Osamäärätestisuure hypoteesille H_0 hypoteesia H_1 vastaan on

$$\lambda(\mathbf{x}) = \frac{\sup\limits_{\theta\in\Theta_0} L(\theta;\mathbf{x})}{\sup\limits_{\theta\in\Theta} L(\theta;\mathbf{x})}$$

Koska

$$\Theta_0 \subset \Theta$$

niin

$$\sup_{\theta\in\Theta_0} L(\theta;\mathbf{x}) \le \sup_{\theta\in\Theta} L(\theta;\mathbf{x})$$

ja siten

$$0 \le \lambda(\mathbf{x}) \le 1$$

Osamäärätesti on *päätössääntö*, jonka *hylkäysalue* eli *kriittinen alue* on muotoa

$$\{\mathbf{x} \mid \lambda(\mathbf{x}) \le c\}$$

jossa c on toteuttaa ehdon

$$0 \le c \le 1$$

Tarkastelemme *kriittisen rajan* eli *arvon c* valitsemista kappaleessa 4.3.

Jos osamäärätestisuure $\lambda(\mathbf{x})$ saa *pienen* (lähellä nollaa olevan) arvon, niin vaihtoehtoisen hypoteesin H_1 rajoittamassa parametriavaruuden Θ osajoukossa Θ_1 on sellaisia parametrin θ arvoja, jotka tekevät havaitusta otoksesta *uskottavamman* kuin mikään sellainen parametrin arvo, joka kuuluu nollahypoteesin H_0 rajoittamaan parametriavaruuden Θ osajoukkoon Θ_0. Siten *pienet* (lähellä nollaa olevat) *osamäärätestisuureen* $\lambda(\mathbf{x})$ *arvot viittaavat siihen*, *että nollahypoteesi saattaa olla syytä hylätä*.

Jos osamäärätestisuure $\lambda(\mathbf{x})$ saa *suuren* (lähellä ykköstä olevan) arvon, niin nollahypoteesin H_0 rajoittamassa parametriavaruuden Θ osajoukossa Θ_0 on sellaisia parametrin θ arvoja, jotka tekevät havaitusta otoksesta lähes *yhtä uskottavan* kuin mielivaltainen parametrin θ arvo, joka kuuluu vaihtoehtoisen hypoteesin H_1 rajoittamaan parametriavaruuden Θ osajoukkoon Θ_1. Siten *suuret* (lähellä ykköstä olevat) *osamäärätestisuureen* $\lambda(\mathbf{x})$ *arvot viittaavat siihen*, *että nollahypoteesia ei ole syytä hylätä*.

Olkoon

$$\hat{\theta} = \hat{\theta}(\mathbf{x})$$

se parametrin θ arvo, joka maksimoi uskottavuusfunktion $L(\theta;\mathbf{x})$ koko parametriavaruudessa Θ ja olkoon

$$\hat{\theta}_0 = \hat{\theta}_0(\mathbf{x})$$

se parametrin θ arvo, joka maksimoi uskottavuusfunktion $L(\theta;\mathbf{x})$ nollahypoteesin H_0 rajoittamassa parametriavaruuden Θ osajoukossa Θ_0. Tällöin

$$\lambda(\mathbf{x}) = \frac{L(\hat{\theta}_0;\mathbf{x})}{L(\hat{\theta};\mathbf{x})}$$

Huomaa, että

$$\hat{\theta} = \hat{\theta}(\mathbf{X})$$

on parametrin θ **rajoittamaton suurimman uskottavuuden estimaattori**, joka saadaan maksimoimalla uskottavuusfunktio $L(\theta;\mathbf{x})$ koko parametriavaruudessa Θ ja

$$\hat{\theta}_0 = \hat{\theta}_0(\mathbf{X})$$

on parametrin θ **rajoitettu suurimman uskottavuuden estimaattori**, joka saadaan maksimoimalla uskottavuusfunktio $L(\theta;\mathbf{x})$ nollahypoteesin H_0 rajoittamassa parametriavaruuden Θ osajoukossa Θ_0.

Esimerkki 2.1: Osamäärätesti normaalijakauman odotusarvolle

Tarkastelemme tässä *uskottavuusosamäärätestiä normaalijakauman odotusarvolle*, kun jakauman *varianssi on tunnettu*; vrt. esimerkkiä 2.1. esimerkkiin 2.2.

Satunnaismuuttuja X noudattaa *normaalijakaumaa* parametrein

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

jos sen *tiheysfunktio* on

$$f(x;\mu,\sigma^2) = (2\pi\sigma^2)^{-1/2}\exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$
$$-\infty < \mu < +\infty\,,\ \sigma > 0\,,\ -\infty < x < +\infty$$

Parametri μ on normaalijakauman *odotusarvo* ja parametri σ^2 on normaalijakauman *varianssi*.

Oletetaan, että *havainnot*

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

muodostavat satunnaisotoksen normaalijakaumasta $\mathrm{N}(\mu,1)$. Tällöin

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

Olkoon *nollahypoteesi* H_0 muotoa

$$\mathrm{H}_0 : \mu = \mu_0$$

ja *vaihtoehtoinen hypoteesi* H_1 muotoa

$$\mathrm{H}_1 : \mu \neq \mu_0$$

Parametriavaruus on tässä muotoa

$$\Theta = \{\mu \mid -\infty < \mu < +\infty\}$$

Määrätään *osamäärätesti* hypoteesille H_0 hypoteesia H_1 vastaan. *Osamäärätestisuure* $\lambda(\mathbf{x})$ on tässä

$$\lambda(\mathbf{x}) = \frac{\max\limits_{\Theta_0} L(\mu;\mathbf{x})}{\max\limits_{\Theta} L(\mu;\mathbf{x})}$$

Koska nollahypoteesin H_0 rajoittama parametriavaruuden Θ osajoukko on muotoa

$$\Theta_0 = \{\mu_0\}$$

niin osamäärätestisuureen *osoittaja* on

$$L(\mu_0;\mathbf{x})$$

Osamäärätestisuureen *nimittäjä* on

$$L(\hat{\mu};\mathbf{x})$$

jossa $\hat{\mu}$ *maksimoi uskottavuusfunktion* $L(\mu;\mathbf{x})$ parametrin μ suhteen koko parametri-avaruudessa Θ.

Luvussa 3 on näytetty, että parametrin μ *rajoittamaton suurimman uskottavuuden estimaattori* $\hat{\mu}$ on havaintojen *aritmeettinen keskiarvo*

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

Siten *osamäärätestisuure* $\lambda(\mathbf{x})$ saa muodon

$$\begin{aligned}
\lambda(\mathbf{x}) &= \frac{L(\mu_0;\mathbf{x})}{L(\overline{x};\mathbf{x})} \\
&= \frac{(2\pi)^{-n/2}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right]}{(2\pi)^{-n/2}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i-\overline{x})^2\right]} \\
&= \exp\left[-\frac{1}{2}\left(\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\overline{x})^2\right)\right]
\end{aligned}$$

Osamäärätestisuure $\lambda(\mathbf{x})$ voidaan kirjoittaa yksinkertaisempaan muotoon, kun otetaan huomioon, että

$$\begin{aligned}
\sum_{i=1}^{n}(x_i-\mu_0)^2 &= \sum_{i=1}^{n}(x_i-\overline{x}+\overline{x}-\mu_0)^2 \\
&= \sum_{i=1}^{n}(x_i-\overline{x})^2 - 2\sum_{i=1}^{n}(x_i-\overline{x})(\overline{x}-\mu_0) + \sum_{i=1}^{n}(\overline{x}-\mu_0)^2 \\
&= \sum_{i=1}^{n}(x_i-\overline{x})^2 + n(\overline{x}-\mu_0)^2
\end{aligned}$$

koska

$$\begin{aligned}\sum_{i=1}^{n}(x_i-\bar{x})(\bar{x}-\mu_0) &= (\bar{x}-\mu_0)\sum_{i=1}^{n}(x_i-\bar{x}) \\ &= (\bar{x}-\mu_0)\left(\sum_{i=1}^{n}x_i - n\bar{x}\right) \\ &= (\bar{x}-\mu_0)\left(\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}x_i\right) \\ &= 0\end{aligned}$$

Siten osamäärätestisuure $\lambda(\mathbf{x})$ nollahypoteesille H_0 vaihtoehtoista hypoteesia H_1 vastaan voidaan esittää muodossa

$$\lambda(\mathbf{x}) = \exp\left[-\frac{1}{2}n(\bar{x}-\mu_0)^2\right]$$

Osamäärätesti on testi, joka hylkää nollahypoteesin H_0 *pienille* testisuureen $\lambda(\mathbf{x})$ arvoille.

Hylkäysalue

$$\{\mathbf{x} \mid \lambda(\mathbf{x}) \le c\}$$

voidaan esittää ekvivalentissa muodossa

$$\left\{\mathbf{x} \,\middle|\, |\bar{x}-\mu_0| \ge \sqrt{-2\log(c)/n}\right\}$$

Kun c vaihtelee välillä $[0,1]$, niin $\sqrt{-2\log(c)/n}$ vaihtelee välillä $[0,\infty)$. Siten *osamäärätesti hylkää nollahypoteesin* H_0, jos havaintoarvojen aritmeettinen keskiarvo $\bar{x}$ poikkeaa parametrin μ nollahypoteesin H_0 kiinnittämästä arvosta μ_0 *enemmän* kuin jokin ennen testin tekemistä kiinnitetty lukuarvo.

Olkoon tunnusluku $T(\mathbf{X})$ on *tyhjentävä* parametrille θ ja olkoon sen *pistetodennäköisyys-* tai *tiheysfunktio* $g(t;\theta)$. Tarkastellaan osamäärätestiä, joka perustuu tunnuslukuun T ja sen uskottavuus*funktioon*

$$L^*(\theta;\, t) = g(t;\theta)$$

Olkoon tunnuslukuun T perustuvaa osamäärätestisuure $\lambda^*(t)$.

Koska kaikki otokseen sisältyvä informaatio parametrista θ sisältyy tunnuslukuun T, tuntuisi luontevalta, että tunnuslukuun T perustuva testi on *yhtä hyvä* kuin otokseen $\mathbf{X}$ perustuva testi. Itse asiassa nämä *testit ovat ekvivalentteja*, kuten seuraava lause osoittaa.

Lause:

Oletetaan, että tunnusluku $T(\mathbf{X})$ on *tyhjentävä* parametrille θ. Olkoon $\lambda^*(t)$ tunnuslukuun T perustuva osamäärätesti ja olkoon $\lambda(\mathbf{x})$ otokseen $\mathbf{X}$ perustuva osamäärätesti. Tällöin

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

jokaiselle havaintopisteelle $\mathbf{x}$.

Todistus:

Koska tunnusluku T on tyhjentävä parametrille θ, niin *faktorointiteoreemasta* (ks. lukua 2) seuraa, että otoksen $\mathbf{X}$ pistetodennäköisyys- tai tiheysfunktio $f(\mathbf{x};\theta)$ voidaan esittää muodossa

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

jossa $g(t;\theta)$ on tunnusluvun T pistetodennäköisyys- tai tiheysfunktio ja $h(\mathbf{x})$ *ei riipu* parametrista θ.

Siten

$$\begin{aligned}\lambda(\mathbf{x}) &= \frac{\sup_{\theta\in\Theta_0} L(\theta;\mathbf{x})}{\sup_{\theta\in\Theta} L(\theta;\mathbf{x})} \\ &= \frac{\sup_{\theta\in\Theta_0} f(\mathbf{x};\theta)}{\sup_{\theta\in\Theta} f(\mathbf{x};\theta)} \\ &= \frac{\sup_{\theta\in\Theta_0} g(T(\mathbf{x});\theta)h(\mathbf{x})}{\sup_{\theta\in\Theta} g(T(\mathbf{x});\theta)h(\mathbf{x})} \\ &= \frac{\sup_{\theta\in\Theta_0} g(T(\mathbf{x});\theta)}{\sup_{\theta\in\Theta} g(T(\mathbf{x});\theta)} \\ &= \frac{\sup_{\theta\in\Theta_0} L^*(\theta;T(\mathbf{x}))}{\sup_{\theta\in\Theta} L^*(\theta;T(\mathbf{x}))} \\ &= \lambda^*(T(\mathbf{x}))\end{aligned}$$

■

Esimerkki 2.2: Osamäärätesti normaalijakauman odotusarvolle

Tarkastelemme tässä *uskottavuusosamäärätestiä normaalijakauman odotusarvolle*, kun jakauman *varianssi ei ole tunnettu*; vrt. esimerkkiä 2.2. esimerkkiin 2.1.

Satunnaismuuttuja X noudattaa *normaalijakaumaa* parametrein

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

jos sen *tiheysfunktio* on muotoa

$$f(x;\mu,\sigma^2) = (2\pi\sigma^2)^{-1/2}\exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$
$$-\infty < \mu < +\infty\,,\ \sigma > 0\,,\ -\infty < x < +\infty$$

Parametri μ on normaalijakauman *odotusarvo* ja parametri σ^2 on normaalijakauman *varianssi*.

Oletetaan, että *havainnot*

$X_1, X_2, \ldots, X_n$

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

Tällöin

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

Olkoon *nollahypoteesi* H_0 muotoa

$$\mathrm{H}_0 : \mu \le \mu_0$$

ja *vaihtoehtoinen hypoteesi* H_1 muotoa

$$\mathrm{H}_1 : \mu > \mu_0$$

Parametriavaruutena on tässä

$$\Theta = \{\mu, \sigma^2 \mid -\infty < \mu < +\infty\,,\ \sigma^2 \ge 0\}$$

Nollahypoteesin H_0 rajoittama parametriavaruuden Θ osajoukko on

$$\Theta_0 = \{\mu, \sigma^2 \mid \mu \le \mu_0\,,\ \sigma^2 \ge 0\}$$

Määrätään *osamäärätestisuure* hypoteesille H_0 hypoteesia H_1 vastaan. Parametri σ^2 on tässä *kiusaparametri* (*engl*. nuisance parameter) σ^2 on kyllä malliin liittyvä tuntematon parametri, mutta sen arvosta ei sinänsä olla kiinnostuneita.

Osamäärätestisuure $\lambda(\mathbf{x})$ on tässä muotoa

$$\lambda(\mathbf{x}) = \frac{\max\limits_{\Theta_0} L(\mu,\sigma^2;\mathbf{x})}{\max\limits_{\Theta} L(\mu,\sigma^2;\mathbf{x})}$$

Testisuureen *nimittäjä* on

$$L(\hat{\mu},\hat{\sigma}^2;\mathbf{x})$$

jossa $\hat{\mu}$ ja $\hat{\sigma}^2$ maksimoivat uskottavuusfunktion $L(\mu,\sigma^2;\mathbf{x})$ koko parametriavaruudessa Θ parametrien μ ja σ^2 suhteen.

Luvussa 3 on näytetty, että parametrin μ *rajoittamaton suurimman uskottavuuden estimaattori* $\hat{\mu}$ on havaintojen *aritmeettinen keskiarvo*

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

ja parametrin σ^2 *rajoittamaton suurimman uskottavuuden estimaattori* $\hat{\sigma}^2$ on havaintojen (harhainen) *otosvarianssi*

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2$$

Jos $\hat{\mu} = \bar{x} \le \mu_0$, uskottavuusfunktion $L(\mu, \sigma^2; \mathbf{x})$ rajoitettu maksimi on sama kuin sen rajoittamaton maksimi, mutta jos $\hat{\mu} = \bar{x} > \mu_0$, niin sen rajoitettu maksimi on

$$L(\mu_0, \hat{\sigma}_0^2; \mathbf{x})$$

jossa

$$\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2$$

Siten osamäärätestisuure $\lambda(\mathbf{x})$ saa muodon

$$\lambda(\mathbf{x}) = \begin{cases} 1 & , \text{ jos } \bar{x} \le \mu_0 \\ \dfrac{L(\mu_0, \hat{\sigma}_0^2; \mathbf{x})}{L(\bar{x}, \hat{\sigma}^2; \mathbf{x})} & , \text{ jos } \bar{x} > \mu_0 \end{cases}$$

Tarkastellaan osamäärätestisuuretta $\lambda(\mathbf{x})$ lähemmin tapauksessa $\bar{x} > \mu_0$.

Tällöin

$$\begin{aligned} \lambda(\mathbf{x}) &= \frac{L(\mu_0, \hat{\sigma}_0^2; \mathbf{x})}{L(\bar{x}, \hat{\sigma}^2; \mathbf{x})} \\ &= \frac{(2\pi\hat{\sigma}_0^2)^{-n/2} \exp\left[-\dfrac{1}{2\hat{\sigma}_0^2}\sum_{i=1}^{n}(x_i - \mu_0)^2\right]}{(2\pi\hat{\sigma}^2)^{-n/2} \exp\left[-\dfrac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]} \\ &= \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)^{-n/2} \end{aligned}$$

koska

$$\sum_{i=1}^{n}(x_i - \mu_0)^2 = n\hat{\sigma}_0^2$$

ja

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = n\hat{\sigma}^2$$

Siten osamäärätestisuure $\lambda(\mathbf{x})$ saa muodon

$$\begin{aligned} \lambda(\mathbf{x}) &= \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)^{-n/2} \\ &= \left(\frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{-n/2} \end{aligned}$$

Ottamalla huomioon sen, että (ks. esimerkkiä 2.1)

$$\sum_{i=1}^{n}(x_i-\mu_0)^2=\sum_{i=1}^{n}(x_i-\bar{x})^2+n(\bar{x}-\mu_0)^2$$

sekä lisäksi sen, että

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

osamäärätestisuure $\lambda(\mathbf{x})$ voidaan kirjoittaa muotoon

$$\begin{aligned}
\lambda(\mathbf{x}) &= \left(\frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{-n/2} \\
&= \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2+n(\bar{x}-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{-n/2} \\
&= \left(1+\frac{n(\bar{x}-\mu_0)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{-n/2} \\
&= \left(1+\frac{(\bar{x}-\mu_0)^2}{(n-1)s^2/n}\right)^{-n/2} \\
&= \left(1+\frac{t^2}{n-1}\right)^{-n/2}
\end{aligned}$$

jossa

$$t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}$$

on tavanomainen *t-testisuure* nollahypoteesille H_0 vaihtoehtoista hypoteesia H_1 vastaan.

Luvussa 1 on näytetty, että

$$t \sim t(n-1)$$

jos nollahypoteesi H_0 pätee.

Osamäärätesti hylkää nollahypoteesin H_0 *pienille* testisuureen $\lambda(\mathbf{x})$ arvoille.

Hylkäysalue

$$\{\mathbf{x} \mid \lambda(\mathbf{x}) \le c\}$$

voidaan kirjoittaa ekvivalenttiin muotoon

$$\left\{\mathbf{x} \,\middle|\, t \ge \sqrt{(n-1)(c^{-n/2}-1)}\right\}$$

Kun c vaihtelee välillä [0,1], niin lauseke

$$\sqrt{(n-1)(c^{-n/2}-1)}$$

vaihtelee välillä $[0,\infty)$. Siten *osamäärätesti hylkää nollahypoteesin* H_0, jos tavanomainen t-testisuure saa *suuremman* arvon kuin jokin ennen testin tekemistä kiinnitetty lukuarvo.

4.3. Testien vertailu

Hylätessään tai hyväksyessään nollahypoteesin testin tekijä voi tehdä *virheellisen* päätöksen. Tavallisesti testejä vertaillaan vertailemalla virheiden todennäköisyyksiä.

Virheet testauksessa, virheiden todennäköisyydet ja testin voimakkuus

Hylkäysvirhe ja hyväksymisvirhe

Olkoon θ perusjoukkoa kuvaava parametri ja olkoon Θ *parametriavaruus* eli mahdollisten parametrin θ arvojen joukko. Olkoon nollahypoteesi H_0 muotoa

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

jossa Θ_0 on jokin parametriavaruuden Θ osajoukko. Tällöin vaihtoehtoinen hypoteesin H_1 voidaan esittää muodossa

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

jossa Θ_0^c on joukon Θ_0 komplementti.

Jos nollahypoteesi H_0 *hylätään* silloin, kun *se on tosi*, tehdään 1. *lajin virhe* eli *hylkäysvirhe*. Jos nollahypoteesi H_0 *hyväksytään* silloin, kun *se on epätosi*, tehdään 2. *lajin virhe* eli *hyväksymisvirhe*.

<table>
<tr><th colspan="2" rowspan="2">Virheet testauksessa</th><th colspan="2">Maailman tila</th></tr>
<tr><th>Nollahypoteesi H₀ on tosi</th><th>Nollahypoteesi H₀ on epätosi</th></tr>
<tr><td rowspan="2">Testin tulos</td><td>Nollahypoteesi H₀ hylätään</td><td>1. lajin virhe eli hylkäysvirhe</td><td>Oikea päätös</td></tr>
<tr><td>Nollahypoteesi H₀ hyväksytään</td><td>Oikea päätös</td><td>2. lajin virhe eli hyväksymisvirhe</td></tr>
</table>

Olkoon testin *hylkäysalue* eli se otosavaruuden osajoukko, joka johtaa nollahypoteesin hylkäämiseen R. Olkoon $\mathbf{x}$ *havaintopiste*. Oletetaan, että nollahypoteesi

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

pätee. Tällöin testi johtaa nollahypoteesin H_0 virheelliseen hylkäämiseen eli 1.lajin virheeseen, jos

$$\mathbf{x} \in R$$

Siten 1. lajin virheen eli *hylkäysvirheen todennäköisyys* on

$$\Pr_\theta(\mathbf{X} \in R)$$

Vastaavasti 2. lajin virheen eli *hyväksymisvirheen todennäköisyys* on

$$\Pr_\theta(\mathbf{X} \in R^c)$$

Huomaa, että

$$\Pr_\theta(\mathbf{X} \in R^c) = 1 - \Pr_\theta(\mathbf{X} \in R)$$

Siten todennäköisyys

$$\Pr_\theta(\mathbf{X} \in R)$$

sisältää parametrin θ funktiona kaiken informaation testistä, jonka hylkäysalue on R.

Nyt

$$\Pr_\theta(\mathbf{X} \in R) = \begin{cases} \text{1. lajin virheen todennäköisyys,} & \text{jos } \theta \in \Theta_0 \\ 1 - (\text{2. lajin virheen todennäköisyys}), & \text{jos } \theta \in \Theta_0^c \end{cases}$$

Testin voimakkuus

Olkoon testin hylkäysalue R. Tällöin todennäköisyyttä

$$\Pr_\theta(\mathbf{X} \in R)$$

kutsutaan parametrin θ funktiona testin **voimakkuusfunktioksi**.

Merkitään:

$$\beta(\theta) = \Pr_\theta(\mathbf{X} \in R)$$

Hyvä testi on *voimakas*, koska sen 2. lajin virheen todennäköisyys on *pieni*.

Ideaalisen testin voimakkuusfunktio olisi muotoa

$$\beta(\theta) = \begin{cases} 0\,, & \theta \in \Theta_0 \\ 1\,, & \theta \in \Theta_0^c \end{cases}$$

Tällaista ideaalitilannetta ei kuitenkaan saavuteta kuin triviaaleissa erikoistapauksissa. Voimme kuitenkin todeta, että hyvällä testillä $\beta(\theta) = 0$ *useimmille* $\theta \in \Theta_0$ ja $\beta(\theta) = 1$ *useimmille* $\theta \in \Theta_0^c$.

Osamäärätestin tapauksessa testin *voimakkuusfunktio* on muotoa

$$\beta(\theta) = \Pr_\theta(\lambda(\mathbf{X}) \le c)$$

Testin koko ja testin taso

Kiinteälle otoskoolle 1. ja 2. lajin virheiden todennäköisyyksiä *ei voida* yleensä *tehdä samanaikaisesti mielivaltaisen pieniksi*. Siksi testi konstruoidaan tavallisesti niin, että ensin *kiinnitetään* 1. *lajin virheen todennäköisyys* ja sitten *niiden testien joukosta, joilla sama* 1. *lajin virheen todennäköisyys valitaan testi, jolla* 2. *lajin virheen todennäköisyys on mahdollisimman pieni*.

Olkoon

$$0 \le \alpha \le 1$$

ja olkoon $\beta(\theta)$ testin voimakkuusfunktio.

Sanomme, että testin **koko** on α, jos

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$$

Sanomme, että testin **taso** on α, jos

$$\sup_{\theta \in \Theta_0} \beta(\theta) \le \alpha$$

Kaikki eivät tee eroa termien *koko* ja *taso* välillä. Lisäksi testin *tasoa* kutsutaan usein **merkitsevyystasoksi**.

On syytä huomata, että kokoa α olevat testit muodostavat *osajoukon* tasoa α olevien testien joukossa. *Kokoa α olevaa testiä ei välttämättä ole olemassa*, mutta vaikka näin olisikin asianlaita, *saattaa silti olla mahdollista löytää jokin tasoa α oleva testi*.

Sovelluksissa *testin taso* valitaan usein etukäteen so. ennen testin tekemistä. Tavanomaisia valintoja ovat

$$\alpha = 0.1, 0.05, 0.01, 0.001$$

On syytä huomata, että valitsemalla testin taso etukäteen voidaan *kontrolloida* vain 1. lajin virheen eli hylkäysvirheen todennäköisyyttä.

Osamäärätestin tapauksessa testin *tasoksi* tulee α, jos vakio c valitaan siten, että

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \sup_{\theta \in \Theta_0} \Pr_\theta(\lambda(\mathbf{X}) \le c) = \alpha$$

Harhattomat testit

Testin tason lisäksi saatamme olla kiinnostuneita myös muista testin ominaisuuksista.

On järkevää toivoa, että nollahypoteesi H_0 tulee todennäköisemmin hylätyksi silloin, kun $\theta \in \Theta_0^c$, kuin silloin, kun $\theta \in \Theta_0$. Jos testillä on tämä ominaisuus, kutsutaan testiä *harhattomaksi*.

Sanomme, että testi on **harhaton**, jos

$$\beta(\theta') \ge \beta(\theta'')$$

kaikille $\theta' \in \Theta_0^c$ ja kaikille $\theta'' \in \Theta_0$.

Tasaisesti voimakkaimmat testit

Tarkastellaan sellaisten testien luokkaa, joiden *taso* on α. Edellisessä kappaleessa todettiin, että kaikilla tähän luokkaan kuuluvilla testeillä 1. *lajin virheen* eli *hylkäysvirheen* todennäköisyys on *korkeintaan α kaikille* $\theta \in \Theta_0$.

Optimaalisena testinä niiden testien joukossa, joiden taso on α, voidaan pitää sellaista testiä, jolla on *mahdollisimman pieni* 2. *lajin virheen* eli *hyväksymisvirheen todennäköisyys* eli *mahdollisimman suuri voimakkuus* kaikille $\theta \in \Theta_0^c$.

Olkoon C jokin sellaisten testien luokka, joilla testataan *nollahypoteesia*

$$H_0 : \theta \in \Theta_0$$

vaihtoehtoista hypoteesia

$$H_1 : \theta \in \Theta_0^c$$

vastaan.

Luokkaan C kuuluva testi, jonka voimakkuusfunktio on $\beta(\theta)$, on **tasaisesti voimakkain** (*engl.* uniformly most powerful, UMP) luokassa C, jos

$$\beta(\theta) \geq \beta'(\theta)$$

kaikille $\theta \in \Theta_0^c$ ja kaikille luokkaan C kuuluvien testien *voimakkuusfunktioille* $\beta'(\theta)$.

Tässä kappaleessa tarkastelun kohteena oleva testien luokka on *kaikkien* niiden testien luokka C, joiden *taso* on α. Tällöin voimme puhua **tasaisesti voimakkaimmasta tasoa α olevasta testistä**. Tällaista testiä ei ole aina olemassa, mutta jos sellainen on olemassa, sitä voidaan pitää *parhaimpana* tasoa α olevien testien luokassa.

Neymanin ja Pearsonin lemma:

Olkoon nollahypoteesina

$$\mathrm{H}_0 : \theta = \theta_0$$

ja vaihtoehtoisena hypoteesina

$$\mathrm{H}_1 : \theta = \theta_1$$

ja olkoon

$$f(\mathbf{x};\theta_i)\,,\; i = 0,1$$

otoksen $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ yhteisjakauman pistetodennäköisyys- tai tiheysfunktio, kun parametri θ saa arvon θ_i, $i = 0,1$.

Tarkastellaan testiä, jonka hylkäysalueen R määrittelee seuraavat ehdot:

On olemassa $k \geq 0$ siten, että

(1)
$$\begin{cases} \mathbf{x} \in R\ , \text{ jos } f(\mathbf{x};\theta_1) > kf(\mathbf{x};\theta_0) \\ \mathbf{x} \in R^c, \text{ jos } f(\mathbf{x};\theta_1) < kf(\mathbf{x};\theta_0) \end{cases}$$

ja

(2)
$$\Pr_{\theta_0}(\mathbf{X} \in R) = \alpha$$

Tällöin pätee seuraava:

(i) Ehtojen (1) ja (2) *riittävyys*:

Jos testi toteuttaa ehdot (1) ja (2), niin se on *tasaisesti voimakkain tasoa α oleva testi*.

(ii) Ehtojen (1) ja (2) *välttämättömyys*:

Jos on olemassa testi, joka toteuttaa ehdot (1) ja (2) jollekin $k > 0$, niin jokainen *tasaisesti voimakkain tasoa α oleva testi* on *kokoa α* eli toteuttaa ehdon (2).

Lisäksi jokainen *tasaisesti voimakkain tasoa α oleva testi* toteuttaa ehdon (1) mahdollisesti lukuun ottamatta *nollamittaista* joukkoa A eli joukko A toteuttaa ehdot

$$\Pr_{\theta_0}(\mathbf{X} \in A) = \Pr_{\theta_1}(\mathbf{X} \in A) = 0$$

Todistus:

Todistamme lauseen vain jatkuvien jakaumien tapauksessa. Todistus diskreettien jakaumien tapauksessa saadaan korvaamalla tässä esitetyn todistuksen integraalit summilla.

Toteamme ensin, että jokainen testi, joka toteuttaa ehdon

(2) $$\Pr_{\theta_0}(\mathbf{X} \in R) = \alpha$$

on *kokoa* α ja siten myös *tasoa* α, koska nollahypoteesin H_0 määrittelemässä parametriavaruuden Θ osajoukossa Θ_0 on vain yksi piste θ_0, jolloin

$$\sup_{\theta \in \Theta_0} \Pr_\theta(\mathbf{X} \in R) = \Pr_{\theta_0}(\mathbf{X} \in R) = \alpha$$

Määritellään merkintöjen yksinkertaistamiseksi **testifunktio** $\phi(\mathbf{x})$ seuraavalla tavalla:

$$\phi(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in R \\ 0, & \mathbf{x} \in R^c \end{cases}$$

Testifunktio $\phi(\mathbf{x})$ on *indikaattorifunktio* hylkäysalueelle R.

Olkoon $\phi(\mathbf{x})$ testifunktio testille, joka toteuttaa ehdot (1) ja (2) ja olkoon $\phi'(\mathbf{x})$ testifunktio mielivaltaiselle *tasoa* α olevalle testille. Olkoot $\beta(\theta)$ ja $\beta'(\theta)$ vastaavat *voimakkuusfunktiot*.

Koska

$$0 \le \phi'(\mathbf{x}) \le 1$$

niin ehdosta (1) ja siitä, että

$$\phi(\mathbf{x}) = 1, \text{ jos } f(\mathbf{x};\theta_1) > kf(\mathbf{x};\theta_0)$$

ja

$$\phi(\mathbf{x}) = 0, \text{ jos } f(\mathbf{x};\theta_1) < kf(\mathbf{x};\theta_0)$$

seuraa, että

$$[\phi(\mathbf{x}) - \phi'(\mathbf{x})][f(\mathbf{x};\theta_1) - kf(\mathbf{x};\theta_0)] \ge 0$$

Siten

(∗) $$0 \le \int [\phi(\mathbf{x}) - \phi'(\mathbf{x})][f(\mathbf{x};\theta_1) - kf(\mathbf{x};\theta_0)]\,d\mathbf{x} = \beta(\theta_1) - \beta'(\theta_1) - k[\beta(\theta_0) - \beta'(\theta_0)]$$

(i) Ehtojen (1) ja (2) *riittävyys*.

Todetaan ensin, että

$$\beta(\theta_0) - \beta'(\theta_0) = \alpha - \beta'(\theta_0) \ge 0$$

koska ϕ' on *tasoa* α oleva testi ja ϕ on *kokoa* α oleva testi.

Siten epäyhtälöstä (∗) ja siitä, että $k \ge 0$ että seuraa, että

$$0 \le \beta(\theta_1) - \beta'(\theta_1) - k[\beta(\theta_0) - \beta'(\theta_0)] \le \beta(\theta_1) - \beta'(\theta_1)$$

joten

$$\beta(\theta_1) \ge \beta'(\theta_1)$$

Siten testi ϕ on *voimakkaampi* kuin testi ϕ'. Koska ϕ' oli mielivaltainen *tasoa* α oleva testi ja vaihtoehtoisen hypoteesin H_1 määrittelemässä parametriavaruuden Θ osajoukossa Θ_1 on vain yksi piste θ_1, niin ϕ on *tasaisesti voimakkain tasoa* α *oleva testi*.

(ii) Ehtojen (1) ja (2) *välttämättömyys*.

Olkoon $\phi'(\mathbf{x})$ testifunktio mielivaltaiselle *tasaisesti voimakkaimmalle tasoa* α *olevalle testille*. Kohdan (i) mukaan jokainen testi ϕ, joka toteuttaa ehdot (1) ja (2) on myös *tasaisesti voimakkain tasoa* α *oleva testi*. Siten

$$\beta(\theta_1) = \beta'(\theta_1)$$

Koska $k > 0$, niin epäyhtälöstä (∗) seuraa, että

$$\alpha - \beta'(\theta_0) = \beta(\theta_0) - \beta'(\theta_0) \leq 0$$

Koska ϕ' on *tasoa* α oleva testi,

$$\beta'(\theta_0) \leq \alpha$$

ja edelleen

$$\beta'(\theta_0) = \alpha$$

ja siten ϕ' on *kokoa* α oleva testi. Lisäksi epäyhtälössä (∗) on voimassa yhtäsuuruus.

Epäyhtälön (∗) ei-negatiivisen integroitavan

$$[\phi(\mathbf{x}) - \phi'(\mathbf{x})][f(\mathbf{x};\theta_1) - kf(\mathbf{x};\theta_0)]$$

integraali voi olla nolla vain, jos testi ϕ' toteuttaa ehdon (1) mahdollisesti lukuun ottamatta nollamittaista joukkoa A, jolle

$$\int_A f(\mathbf{x};\theta_i)d\mathbf{x} = 0 \text{ , } i = 0,1$$

Siten myös kohdan (ii) väite on todistettu.

■

Seuraava korollaari liittää tunnusluvun *tyhjentävyyden* Neymanin ja Pearsonin lemmassa tarkasteltuun testaus-asetelmaan.

Korollaari:

Olkoon testausasetelma sama kuin Neymanin ja Pearsonin lemmassa. Olkoon $T(\mathbf{X})$ parametrin θ *tyhjentävä tunnusluku* ja olkoon

$$g(t;\theta_i) \text{ , } i = 0,1$$

tunnusluvun T pistetodennäköisyys- tai tiheysfunktio, kun parametri θ saa arvon θ_i , $i = 0,1$.

Tarkastellaan tunnuslukuun T perustuvaa testiä, jonka hylkäysalue S saadaan seuraavista ehdoista:

On olemassa $k \geq 0$ siten, että

(1) $$\begin{cases} \mathbf{x} \in S \text{ , jos } g(t;\theta_1) > kg(t;\theta_0) \\ \mathbf{x} \in S^c\text{, jos } g(t;\theta_1) < kg(t;\theta_0) \end{cases}$$

ja

(2) $\Pr_{\theta_0}(T \in S) = \alpha$

Tällöin pätee seuraava: Jos testi toteuttaa ehdot (1) ja (2), niin se on *tasaisesti voimakkain tasoa α oleva testi*.

Todistus:

Otoksen **X** suhteen, tunnuslukuun T perustuvan testin hylkäysalue on

$$R = \{\mathbf{x} \mid T(\mathbf{x}) \in S\}$$

Koska tunnusluku T on tyhjentävä parametrille θ, niin *faktorointiteoreemasta* (ks. lukua 2) seuraa, että otoksen **X** yhteisjakauman pistetodennäköisyys- tai tiheysfunktio $f(\mathbf{x};\theta_i)$, $i = 0,1$ voidaan esittää muodossa

$$f(\mathbf{x};\theta_i) = g(T(\mathbf{x});\theta_i)h(\mathbf{x})\,,\ i = 0,1$$

jollekin ei-negatiiviselle funktiolle $h(\mathbf{x})$.

Kertomalla ehdon (1) epäyhtälöt funktiolla $h(\mathbf{x})$ nähdään, että

$$\begin{cases} \mathbf{x} \in R\ \text{, jos } f(\mathbf{x};\theta_1) = g(T(\mathbf{x});\theta_1)h(\mathbf{x}) > kg(T(\mathbf{x});\theta_0)h(\mathbf{x}) = kf(\mathbf{x};\theta_0) \\ \mathbf{x} \in R^c\text{, jos } f(\mathbf{x};\theta_1) = g(T(\mathbf{x});\theta_1)h(\mathbf{x}) < kg(T(\mathbf{x});\theta_0)h(\mathbf{x}) = kf(\mathbf{x};\theta_0) \end{cases}$$

Lisäksi ehdosta (2) seuraa, että

$$\Pr_{\theta_0}(\mathbf{X} \in R) = \Pr_{\theta_0}(T(\mathbf{X}) \in S) = \alpha$$

Siten Neymanin ja Pearsonin lauseen kohdasta (i) seuraa, että tunnuslukuun T perustuva testi on *tasaisesti voimakkain tasoa α oleva testi*.

■

Kutsumme Neymanin ja Pearsonin lemman hypoteeseja H_0 ja H_1 **yksinkertaisiksi**, koska ne kumpikin *kiinnittävät* vain yhden todennäköisyysjakauman otokselle **X**. Tavallisesti haluamme kuitenkin antaa kiinnostuksen kohteena olevien hypoteesien kiinnittää useampia todennäköisyys-jakaumia otokselle. Tällaisia hypoteeseja kutsutaan **yhdistetyiksi**.

Esimerkkejä yhdistetyistä hypoteeseista ovat **yksisuuntaiset hypoteesit**

$$\mathrm{H}: \theta \geq \theta_0 \text{ tai } \mathrm{H}: \theta > \theta_0$$
$$\mathrm{H}: \theta \leq \theta_0 \text{ tai } \mathrm{H}: \theta < \theta_0$$

sekä **kaksisuuntainen hypoteesi**

$$\mathrm{H}: \theta \neq \theta_0$$

Koska tasaisesti voimakkaimman testin määritelmässä vaaditaan, että testin on oltava voimakkain kaikille $\theta \in \Theta_0^c$, niin Neymanin ja Pearsonin lemmaa voidaan soveltaa usein soveltaa myös tilanteissa, joissa hypoteesit H_0 ja H_1 ovat *yhdistettyjä*.

Sellainen testien luokka, jossa *yksisuuntaiselle hypoteesille* voidaan konstruoida *tasaisesti voimakkain tasoa α oleva testi*, on niiden testien luokka, jotka perustuvat *monotoniseen uskottavuusosamäärään*.

Olkoon

$$\{g(t;\theta);\theta \in \Theta\}$$

yksiulotteisen satunnaismuuttujan T pistetodennäköisyys- tai tiheysfunktioiden perhe, joka riippuu parametrista θ. Perheellä on **monotoninen uskottavuusosamäärä**, jos osamäärä

$$\frac{g(t;\theta_2)}{g(t;\theta_1)}$$

on kaikille $\theta_2 > \theta_1$ muuttujan t *monotoninen* (*ei-kasvava* tai *ei-vähenevä*) *funktio* joukossa

$$\{t \mid g(t;\theta_1) > 0 \text{ tai } g(t;\theta_2) > 0\}$$

Tehdään sopimus, että

$$c/0 = \infty, \text{ jos } c > 0$$

Monilla tavallisilla jakaumilla on monotoninen uskottavuusosamäärä. Tällaisia jakaumia ovat esimerkiksi normaalijakauma, jossa odotusarvoparametri on tuntematon ja varianssi on tunnettu, Poisson-jakauma ja binomijakauma. Voidaan itse asiassa osoittaa, että kaikilla *säännölliseen eksponenttiperheeseen*

$$g(t;\theta) = h(t)\exp[w(\theta)t]$$

kuuluvilla jakaumilla on monotoninen uskottavuusosamäärä, jos $w(\theta)$ on parametrin θ *ei-vähenevä* funktio.

Karlinin ja Rubinin teoreema:

Olkoon nollahypoteesina

$$\mathrm{H}_0 : \theta \le \theta_0$$

ja vaihtoehtoisena hypoteesina

$$\mathrm{H}_1 : \theta > \theta_0$$

Olkoon tunnusluku T *tyhjentävä* parametrille θ ja oletetaan lisäksi, että tunnusluvun T jakaumien perheellä

$$\{g(t;\theta) \mid \theta \in \Theta\}$$

on *monotoninen uskottavuusosamäärä*. Valitaan t_0. Jokainen testi, joka hylkää nollahypoteesin H_0, jos ja vain jos $T > t_0$ on *tasaisesti voimakkain tasoa* α *oleva testi*, jossa

$$\alpha = \Pr_{\theta_0}(T > t_0)$$

Todistus:

Olkoon

$$\beta(\theta) = \Pr_{\theta}(T > t_0)$$

testin *voimakkuusfunktio*. Valitaan $\theta' > \theta_0$ ja tarkastellaan testiä, jossa testataan nollahypoteesia

$$\mathrm{H}_0' : \theta = \theta_0$$

vaihtoehtoista hypoteesia

$$\mathrm{H}_1' : \theta = \theta'$$

vastaan.

Koska tunnusluvun T jakaumien perheellä on *monotoninen uskottavuusosamäärä*, *voimakkuusfunktio* $\beta(\theta)$ on *ei-vähenevä*.

Tästä seuraa:

(i) $$\sup_{\theta \le \theta_0} \beta(\theta) = \beta(\theta_0) = \alpha$$

ja testin taso on α.

(ii) Jos määrittelemme suureen

$$k' = \inf_{t \in \Im} \frac{g(t;\theta')}{g(t;\theta_0)}$$

jossa

$$\Im = \{t \mid t > t_0 \text{ ja joko } g(t;\theta') > 0 \text{ tai } g(t;\theta_0) > 0\}$$

niin

$$T = t_0 \quad \Leftrightarrow \quad \frac{g(t;\theta')}{g(t;\theta_0)} > k'$$

Siten kohdista (i) ja (ii) seuraa Neymanin ja Pearsonin lauseen korollaarin nojalla, että

$$\beta(\theta') \ge \beta^*(\theta')$$

jossa $\beta^*(\theta)$ on mielivaltaisen toisen *tasoa* α olevan nollahypoteesin H_0' testin voimakkuusfunktio eli sellaisen testin voimakkuusfunktio, jolle

$$\beta(\theta_0) \le \alpha$$

Toisaalta mielivaltaisen *tasoa* α olevan nollahypoteesin H_0 testin voimakkuusfunktio toteuttaa ehdon

$$\beta^*(\theta_0) = \sup_{\theta \in \Theta_0} \beta^*(\theta) \le \alpha$$

Siten

$$\beta(\theta') \ge \beta^*(\theta')$$

mielivaltaiselle *tasoa* α olevan nollahypoteesin H_0 testille. Koska θ' oli mielivaltainen, testi on *tasaisesti voimakkain tasoa* α *oleva testi*.

■

Vastaavasti, jos nollahypoteesi on muotoa

$$\mathrm{H}_0 : \theta \le \theta_0$$

ja vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : \theta > \theta_0$$

ja Karlinin ja Rubinin lauseen ehdot pätevät, niin jokainen testi, joka hylkää nollahypoteesin H_0, jos ja vain jos $T < t_0$ on *tasaisesti voimakkain tasoa* α *oleva testi*, jossa

$$\alpha = \Pr_{\theta_0}(T < t_0)$$

Testin *p*-arvo

Testin tekemisen jälkeen testin tuloksesta pitää kertoa jollakin tilastollisesti merkitsevällä tavalla.

Eräs mahdollisista tavoista on kertoa testin *koko* α ja se hylättiinkö vai hyväksyttiinkö nollahypoteesi H_0. Jos testin koko α oli *pieni*, *päätös hylätä nollahypoteesi* H_0 *perustuu suhteellisen vahvoihin todisteisiin nollahypoteesia vastaan*. Sen sijaan, jos testin koko α oli *suuri*, *päätös hylätä nollahypoteesi* H_0 *ei perustu kovin vahvoihin todisteisiin nollahypoteesia vastaan*.

Toinen mahdollinen tapa kertoa testin tuloksesta on ilmoittaa tunnusluku, jota kutsutaan testin *p-arvoksi*.

Testin ***p*-arvo** $p(\mathbf{X})$ on tunnusluku, joka toteuttaa ehdon

$$0 \le p(\mathbf{x}) \le 1$$

jokaiselle havaintopisteelle $\mathbf{x}$. *Pieni p*-arvo viittaa siihen, että *vaihtoehtoinen hypoteesi* H_1 *on tosi*.

Sanomme, että *p*-arvo on **kelvollinen**, jos

$$\Pr_\theta(p(\mathbf{X}) \le \alpha) \le \alpha$$

kaikille $\theta \in \Theta_0$ ja kaikille $\alpha, 0 \le \alpha \le 1$.

Jos $p(\mathbf{X})$ on kelvollinen *p*-arvo, niin voidaan helposti konstruoida *tasoa* α oleva testi. Testi, joka hylkää nollahypoteesin H_0, jos ja vain jos

$$p(\mathbf{x}) \le \alpha$$

on *tasoa* α.

Testin *p*-arvon $p(\mathbf{X})$ kertominen sisältää aina *enemmän informaatiota* kuin se, että kerrotaan vain testin *taso* α ja se hylättiinkö vai hyväksyttiinkö nollahypoteesi H_0. (Lähes) *kaikki tilastolliset ohjelmistot kertovatkin nykyään testin p-arvon*.

Seuraavaan lauseeseen sisältyy *yleisin tapa määritellä* **kelvollinen *p*-arvo**.

Lause:

Olkoon $W(\mathbf{X})$ testisuure, jonka *suuret* arvot viittaavat siihen, että vaihtoehtoinen hypoteesi H_1 pätee. Olkoon

$$p(\mathbf{x}) = \sup_{\theta \in \Theta_0} \Pr_\theta(W(\mathbf{X}) \ge W(\mathbf{x}))$$

Tällöin $p(\mathbf{X})$ on *kelvollinen p-arvo*.

Todistus:

Valitaan $\theta \in \Theta_0$. Olkoon $F_\theta(w)$ tunnusluvun $-W(\mathbf{X})$ kertymäfunktio ja olkoon

$$p_\theta(\mathbf{x}) = \Pr_\theta(W(\mathbf{X}) \ge W(\mathbf{x})) = \Pr_\theta(-W(\mathbf{X}) \le -W(\mathbf{x})) = F_\theta(-W(\mathbf{x}))$$

Siten satunnaismuuttujan $p_\theta(\mathbf{x})$ jakauma on *stokastisesti suurempi tai yhtä suuri* kuin välillä (0,1) määritelty jatkuva tasainen jakauma Uniform(0,1).

Jos satunnaismuuttuja $p_\theta(\mathbf{x})$ on *jatkuva* tämä seuraa siitä, että aina pätee seuraava:

Olkoon X jatkuva satunnaismuuttuja, jonka kertymäfunktio on F_X ja olkoon

$$Y = F_X(X)$$

Tällöin

$$Y \sim \text{Uniform}(0,1)$$

ts. $\Pr(Y \le y) = y$ kaikille y, $0 < y < 1$.

Väite voidaan perustella myös silloin, kun satunnaismuuttuja $p_\theta(\mathbf{x})$ on *diskreetti*.

Stokastinen suuremmuus:

Olkoon $X \sim F_X$ ja $Y \sim F_Y$. Satunnaismuuttuja X on *stokastisesti suurempi tai yhtä suuri* kuin satunnaismuuttuja Y, jos

$$F_X(t) \le F_Y(t)$$

kaikille t. Tällöin

$$\Pr(X > t) \ge \Pr(Y > t)$$

kaikille t. Jos X on stokastisesti suurempi kuin Y, niin satunnaismuuttujalla X on taipumus saada suurempia arvoja kuin satunnaismuuttujalla Y.

Edellä todetusta seuraa, että

$$\Pr_\theta(p_\theta(\mathbf{X}) \le \alpha) \le \alpha$$

Koska

$$p(\mathbf{x}) = \sup_{\theta' \in \Theta_0} p_{\theta'}(\mathbf{x}) \ge p_\theta(\mathbf{x})$$

kaikille θ, niin

$$\Pr_\theta(p(\mathbf{X}) \le \alpha) \le \Pr_\theta(p_\theta(\mathbf{X}) \le \alpha) \le \alpha$$

mikä pätee kaikille $\theta \in \Theta_0$ ja kaikille $\alpha, 0 \le \alpha \le 1$.

Siten $p(\mathbf{X})$ on *kelvollinen p-arvo*.

■

Toinen tapa määritellä **kelvollinen *p*-arvo** perustuu *ehdollistamiseen tyhjentävän tunnusluvun suhteen*. Olkoon tunnusluku $S(\mathbf{X})$ *tyhjentävä* nollahypoteesin H_0 määrittelemälle mallille

$$\{f(\mathbf{x};\theta) \mid \theta \in \Theta_0\}$$

Jos nollahypoteesi H_0 on tosi, niin otoksen $\mathbf{X}$ ehdollinen jakauma ehdolla $S = s$ *ei riipu* parametrista θ. Olkoon $W(\mathbf{X})$ testisuure, jonka suuret arvot viittaavat siihen, että vaihtoehtoinen hypoteesi H_1 pätee.

Olkoon

$$p(\mathbf{x}) = \Pr(W(\mathbf{X}) \geq W(\mathbf{x}) \mid S = S(\mathbf{x}))$$

Samantapaisella argumentilla kuin edellisessä lauseessa voidaan todistaa, että

$$\Pr(p(\mathbf{X}) \leq \alpha \mid S = s) \leq \alpha$$

kaikille α, $0 \leq \alpha \leq 1$. Jos S on diskreetti satunnaismuuttuja, niin

$$\Pr_\theta(p(\mathbf{X}) \leq \alpha) = \sum_s \Pr(p(\mathbf{X}) \leq \alpha \mid S = s)\Pr_\theta(S = s) \leq \alpha \sum_s \Pr_\theta(S = s) \leq \alpha$$

joten $p(\mathbf{X})$ on *kelvollinen p-arvo*. Jos S on jatkuva satunnaismuuttuja, niin summalausekkeet yllä olevissa kaavoissa on korvattava integraaleilla.