Tilastolliset menetelmät:
Tilastolliset testit

Sisällys

8. Tilastollinen testaus

Tilastollisella testauksella tarkoitetaan tutkimuksen kohteena olevasta *perusjoukosta esitettyjen väitteiden tai oletuksien asettamista koetteelle havainnoista saatua informaatiota vastaan*.

Perusjoukosta esitetyt väitteet tai oletukset on tilastollisessa testauksessa puettava tutkimuksen kohteena olevan perusjoukon alkioiden ominaisuuksien vaihtelua perusjoukossa kuvaavien *todennäköisyysjakaumia* tai *niiden parametreja* koskeviksi **hypoteeseiksi**. **Tilastollinen testi** on *päätössääntö*, joka sanoo *onko hypoteesi hylättävä vai ei* havainnoista saadun informaation valossa. Testi mittaa hypoteesin ja havaintojen *yhteensopivuutta*.

Tarkastelemme tässä luvussa tilastollisen testiteorian *peruskäsitteitä* sekä tarkastelemme esitetytyn teorian havainnollistuksena *normaalijakauman* parametreja koskevien hypoteesien testaamista.

Avainsanat:

Ensimmäisen lajin virhe, Frekvenssitulkinta, Havainto, Hylkäysalue, Hylkäysvirhe, Hypoteesi, Hyväksymis-alue, Hyväksymisvirhe, Järjestysasteikko, χ^2-jakauma, χ^2-testi, Kaksisuuntainen vaihtoehtoinen hypoteesi, Kriittinen arvo, Kriittinen raja, Laatueroasteikko, Maailman tila, Merkitsevyys, Merkitsevyystaso, Mitta-asteikko, Nollahypoteesi, Normaaliarvo, Normaalijakauma, Odotusarvo, Otantamenetelmä, Otos, Parametri, *p*-arvo, Perusjoukko, Satunnaisotos, Suhdeasteikko, *t*-jakauma, *t*-testi, Testausasetelma, Testi, Testin tulos, Testisuure, Tilastollinen hypoteesi, Tilastollisten hypoteesien testaus, Todennäköisyysjakauma, Toisen lajin virhe, Vaihtoehtoinen hypoteesi, Varianssi, Voimakkuus, Välimatka-asteikko, Yksisuuntainen vaihtoehtoine hypoteesi, Yleinen hypoteesi

8.1. Tilastollisen testauksen idea

Tilastollisen testauksen lähtökohta

Lähtökohta:

> **Tutkimuksen kohteena olevasta perusjoukosta on esitetty jokin *väite* tai *oletus*.**

Kysymys:

> **Miten esitettyä väitettä tai oletusta voidaan *testata*?**

Vastaus:

> **Väitettä tai oletusta voidaan *testata tilastollisesti*, jos väite tai oletus voidaan pukea tutkimuksen kohteena olevan perusjoukon ominaisuuden vaihtelua perusjoukossa kuvaavaa *todennäköisyysjakaumaa tai sen parametreja koskevaksi oletukseksi eli hypoteesiksi*.**

Olkoon X tutkimuksen kohteena olevan perusjoukon jonkin ominaisuuden vaihtelua perusjoukossa kuvaava *satunnaismuuttuja* ja olkoon satunnaismuuttujan X *todennäköisyysjakauman piste-todennäköisyys*- tai *tiheysfunktio*

$$f(x;\theta)$$

jossa θ on funktion f muodon määräävä *tuntematon* **parametri**. Yksinkertaisissa testausasetelmissa kiinnostuksen kohteena on **hypoteesi**, jonka mukaan *parametrilla* θ *on arvo* θ_0.

Miten todennäköisyysjakauman $f(x;\theta)$ parametria θ koskevaa hypoteesia

$$\theta = \theta_0$$

voidaan **testata tilastollisesti**? Tilastollisessa testauksessa *hypoteesi* $\theta = \theta_0$ *asetetaan koetteelle havaintojen todennäköisyysjakaumasta* $f(x;\theta)$ *sisältämää informaatiota vastaan*.

Satunnaisotos

Oletamme jatkossa, että *havainnot*

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

muodostavat **satunnaisotoksen** jakaumasta, jonka pistetodennäköisyys- tai tiheysfunktio on

$$f(x;\theta)$$

Tällöin $X_1, X_2, \ldots, X_n$ ovat *riippumattomia, identtisesti jakautuneita satunnaismuuttujia,* joilla on *sama* pistetodennäköisyys- tai tiheysfunktio $f(x;\theta)$:

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

Tavoitteenamme on testata tilastollisesti muotoa

$$\theta = \theta_0$$

olevaa **parametrista hypoteesia**. Testin suorittamista varten valitaan **testisuure**, joka *mittaa* satunnaismuuttujien

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

havaittujen arvojen

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

ja hypoteesin $\theta = \boldsymbol{\theta_0}$ *yhteensopivuutta*. *Hyvä yhteensopivuus* merkitsee sitä, että havainnot ovat *sopusoinnussa* oletuksen $\theta = \boldsymbol{\theta_0}$ kanssa ja *huono yhteensopivuus* merkitsee sitä, että havainnot ja oletus $\theta = \boldsymbol{\theta_0}$ ovat *ristiriidassa* keskenään.

8.2. Tilastolliset hypoteesit

Kun todennäköisyysjakauman parametreja koskevia väitteitä tai oletuksia testataan tilastollisesti, **testausasetelman** kiinnittämiseksi on tehtävä seuraavat kolme oletusta:

(i) *Testausasetelmaa koskevat perusoletukset*, joista pidetään kiinni testauksen aikana, muodostavat testin **yleisen hypoteesin**.

(ii) *Testattavaa oletusta* kutsutaan **nollahypoteesiksi**.

(iii) **Vaihtoehtoinen hypoteesi** on *oletus, joka astuu voimaan, jos nollahypoteesi hylätään testissä*.

Yleinen hypoteesi

Yleiset testausasetelmaa koskevat oletukset muodostavat testin **yleisen hypoteesin** H. *Yleinen hypoteesi* H sisältää oletukset

- *perusjoukosta*
- *otantamenetelmästä*
- *perusjoukon jakaumasta*

Yleisen hypoteesin H oletuksista *pidetään kiinni koko testauksen ajan*, mikä merkitsee sitä, että tilastollinen testaus tehdään aina ehdollisesti yleisen hypoteesin H oletusten suhteen.

Huomautus:

- Yleisen hypoteesin sisältämiä jakaumaoletuksia voidaan ja on tavallisesti myös syytä testata erikseen; ks. esimerkiksi lukua **Yhteensopivuuden, homogeenisuuden ja riippumattomuuden testaaminen**.

Nollahypoteesi

Sitä *perusjoukon jakauman parametreja koskevaa väitettä* tai *oletusta, jota halutaan testata* kutsutaan **nollahypoteesiksi**. Nollahypoteesille käytetään tavallisesti merkintää H_0 .

Testissä nollahypoteesi H_0 *asetetaan koetteelle havaintojen perusjoukon jakaumasta sisältämää informaatiota vastaan*. Nollahypoteesista H_0 *pidetään kiinni*, elleivät havaintojen sisältämät *todisteet nollahypoteesia vastaan* ole *kyllin voimakkaita*.

Olkoon

$$f(x;\theta)$$

tutkimuksen kohteena olevaa perusjoukon ominaisuutta kuvaavan *todennäköisyysjakauman piste-todennäköisyys-* tai *tiheysfunktio*. *Yksinkertaisissa testausasetelmissa* nollahypoteesi on muotoa

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

Huomautus:

- Nollahypoteesit ovat yksinkertaisissa testausasetelmissa muotoa ”*on sama*” tai muotoa ”*ei ole eroa*”.

Vaihtoehtoinen hypoteesi

Vaihtoehtoinen hypoteesi H_1 on *oletus, joka astuu voimaan, jos nollahypoteesi* H_0 *hylätään*.

Vaihtoehtoinen hypoteesi voidaan tavallisesti muotoilla usealla eri tavalla. Jos nollahypoteesi on muotoa ”*on sama*” tai ”*ei ole eroa*”, vaihtoehtoinen hypoteesi on tavallisesti muotoa ”*ei ole sama*” tai ”*on eroa*”.

Kun tilastollista testiä tehdään, toivotaan usein, että *nollahypoteesi voidaan hylätä* ja *vaihtoehtoinen hypoteesi hyväksyä*. Vaihtoehtoisen hypoteesin *hyväksyminen* merkitsee yleensä *informaation lisääntymistä*.

Jos nollahypoteesi on *yksinkertaista muotoa*

$$H_0 : \theta = \theta_0$$

vaihtoehtoinen hypoteesi voidaan muotoilla seuraavilla kolmella tavalla:

(i) $H_1 : \theta > \theta_0$

(ii) $H_1 : \theta < \theta_0$

(iii) $H_1 : \theta \neq \theta_0$

Tapauksissa (i) ja (ii) sanomme, että vaihtoehtoinen hypoteesi on **yksisuuntainen**. Tapauksessa (iii) sanomme, että vaihtoehtoinen hypoteesi on **kaksisuuntainen**.

Huomautus:

- Vaihtoehtoisen hypoteesin muoto vaikuttaa tavallisesti siihen tapaan, jolla testi suoritetaan.

8.3. Tilastolliset testit ja testisuureet

Testi

Tilastollinen testi on *päätössääntö*, joka kertoo jokaisessa yksittäisessä testaustilanteessa eli jokaiselle otokselle, onko nollahypoteesi H_0 hylättävä vai ei.

Testisuure

Tilastollinen testi perustuu aina **testisuureeseen**, joka *mittaa havaintojen ja nollahypoteesin* H_0 *yhteensopivuutta*. Testisuure on *satunnaismuuttuja*, jonka arvo riippuu havainnoista ja nollahypoteesista H_0. Havaintojen ja nollahypoteesin H_0 *yhteensopivuuden mittaaminen* tarkoittaa sitä, että *tutkitaan kuinka todennäköistä on saada sellaisia testisuureen arvoja kuin on saatu*. Tämä vaatii testisuureen jakauman tuntemista.

Jos havaintojen ja nollahypoteesin H_0 *yhteensopivuus on testisuureella mitattuna hyvä, nollahypoteesi* H_0 *jätetään voimaan*. Jos havaintojen ja nollahypoteesin H_0 *yhteensopivuus on testisuureella mitattuna huono, nollahypoteesi* H_0 *hylätään ja vaihtoehtoinen hypoteesi* H_1 *hyväksytään*.

Testisuureen odotusarvoa nollahypoteesin H_0 pätiessä kutsutaan testisuureen **normaaliarvoksi**. Jos testisuureen havaittu arvo *on lähellä testisuureen normaaliarvoa, havainnot ovat sopusoinnussa* *nollahypoteesin* H_0 *kanssa*. Jos testisuureen otoksesta määrätty arvo *poikkeaa merkitsevästi testisuureen normaaliarvosta, havainnot sisältävät todisteita nollahypoteesia* H_0 *vastaan*.

8.4. Virheet testauksessa

Hylkäysvirhe

Jos nollahypoteesi H_0 *hylätään silloin, kun se on tosi*, tehdään **hylkäysvirhe**. *Hylkäysvirheen todennäköisyys* α on ehdollinen todennäköisyys

$$\Pr(H_0 \text{ hylätään} \mid H_0 \text{ on tosi}) = \alpha$$

Hylkäysvirheen todennäköisyyden α *komplementtitodennäköisyys*

$$\Pr(H_0 \text{ hyväksytään} \mid H_0 \text{ on tosi}) = 1 - \alpha$$

on todennäköisyys *hyväksyä* nollahypoteesi *silloin, kun se on tosi*.

Tilastollisessa tutkimuksessa noudatetaan tieteen yleistä *varovaisuusperiaatetta*:

Hypoteeseja ei saa hylätä ilman riittäviä syitä.

Siksi nollahypoteesin H_0 *virheellisen hylkäyksen todennäköisyys halutaan tehdä* tilastollisessa testauksessa *mahdollisimman pieneksi*. Siksi *havainnoilta vaaditaan vahvoja todisteita nolla-hypoteesia* H_0 *vastaan* ennen kuin se suostutaan hylkäämään.

Hyväksymisvirhe

Jos nollahypoteesi H_0 *jätetään voimaan silloin, kun se ei ole tosi*, tehdään **hyväksymisvirhe**. *Hyväksymisvirheen todennäköisyys* β on ehdollinen todennäköisyys

$$\Pr(H_0 \text{ jätetään voimaan} \mid H_0 \text{ ei ole tosi}) = \beta$$

Huomautus:

- Hylkäysvirheen todennäköisyys α ja hyväksymisvirheen todennäköisyys β *eivät ole* toistensa komplementtitodennäköisyyksiä.

Testin voimakkuus

Hyväksymisvirheen todennäköisyyden β *komplementtitodennäköisyyttä*

$$\Pr(H_0 \text{ hylätään} \mid H_0 \text{ ei ole tosi}) = 1 - \beta$$

kutsutaan testin **voimakkuudeksi**. Hyvä testi on *voimakas*, koska voimakkaalla testillä on pieni hyväksymisvirheen todennäköisyys β. Testin voimakkuus $(1 - \beta)$ riippuu tavallisesti testattavan parametrin *todellisesta arvosta*.

Testin voimakkuutta testattavan parametrin arvojen funktiona kutsutaan **voimakkuusfunktioksi**; esimerkki: ks. kappaletta **Yhden otoksen *t*-testi** luvussa **Testejä suhdeasteikollisille muuttujille**.

1. ja 2. lajin virheet

Koska testiä tehtäessä pyritään *ensisijaisesti* varomaan sitä, että nollahypoteesi H_0 *hylätään* silloin, kun se *on tosi*, *hylkäysvirhettä* kutsutaan usein **ensimmäisen lajin virheeksi**. Tällöin *hyväksymis-virhettä* eli sitä, että nollahypoteesi H_0 *hyväksytään* silloin, kun se *ei ole tosi*, kutsutaan **toisen lajin virheeksi**.

Testin tulos ja maailman tilat

Maailman tilat ja *testin tulokset* voidaan ryhmitellä seuraavaksi nelikentäksi:

<table>
<tr><td colspan="2"></td><td colspan="2">Maailman tila</td></tr>
<tr><td colspan="2"></td><td>Nollahypoteesi pätee</td><td>Nollahypoteesi ei päde</td></tr>
<tr><td rowspan="2">Testin tulos</td><td>Nollahypoteesi jää voimaan</td><td>Oikea johtopäätös</td><td>Hyväksymis-virhe</td></tr>
<tr><td>Nollahypoteesi hylätään</td><td>Hylkäysvirhe</td><td>Oikea johtopäätös</td></tr>
</table>

Testin hylkäys- ja hyväksymisalueet

Kun testi formuloidaan *päätössääntönä*, testiä varten konstruoidun testisuureen *mahdollisten arvojen joukko* jaetaan kahteen osaan, *hylkäysalueeseen* ja *hyväksymis-alueeseen*:

(i) Jos testisuureen havainnoista määrätty arvo joutuu **hylkäysalueelle**, *nollahypoteesi* H_0 *hylätään*.

(ii) Jos testisuureen havainnoista määrätty arvo joutuu **hyväksymisalueelle**, *nollahypoteesi* H_0 *jätetään voimaan*.

Huomautus:

- Testisuureen mahdollisten arvojen joukon jako hylkäys- ja hyväksymisalueisiin *ei saa riippua havainnoista*.

8.5. Merkitsevyystaso ja testin hylkäysalue

Merkitsevyystaso

Testin merkitsevyystaso α on *todennäköisyys, että testisuureen havainnoista määrätty arvo joutuu hylkäysalueelle, jos nollahypoteesi* H_0 *pätee*.

Jos testisuureen havainnoista määrätty arvo joutuu nollahypoteesin H_0 pätiessä hylkäysalueelle, nollahypoteesi H_0 hylätään virheellisesti ja seurauksena on hylkäysvirhe, jonka todennäköisyys on α.

Tavallisesti testin hylkäysalue määrätään kiinnittämällä testissä käytettävä merkitsevyystaso α etukäteen so. jo ennen havaintojen keräämistä; ks. kappaletta **Esimerkki: Normaalijakauman parametrien testaaminen**.

Merkitsevyystason frekvenssitulkinta

Oletetaan, että yleisen hypoteesin H lisäksi myös nollahypoteesi H_0 pätee testausasetelmassa ja että olemme valinneet testin merkitsevyystasoksi luvun α. *Toistetaan otantaa* ja *sovelletaan jokaiseen* *otokseen samaa testiä*. **Tällöin joudumme virheellisesti hylkäämään nollahypoteesin** H_0 **keskimäärin**

$$\alpha\ \%\text{:ssa}$$

otoksia, **vaikka nollahypoteesi** H_0 **koko ajan pätee**.

Tavanomaiset merkitsevyystasot

Koska testeissä halutaan ensisijaisesti suojautua hylkäysvirhettä vastaan, testin merkitsevyystasoksi α on tapana valita pieniä lukuja.

Ns. **tavanomaiset merkitsevyystasot** ovat

$$\alpha = 0.05$$

$$\alpha = 0.01$$

$$\alpha = 0.001$$

(i) Jos nollahypoteesi H_0 voidaan hylätä merkitsevyystasolla $\alpha = 0.05$, sanotaan: Testisuureen arvo (tai testin tulos) on **melkein merkitsevä**.

(ii) Jos nollahypoteesi H_0 voidaan hylätä merkitsevyystasolla $\alpha = 0.01$, sanotaan: Testisuureen arvo (tai testin tulos) on **merkitsevä**.

(iii) Jos nollahypoteesi H_0 voidaan hylätä merkitsevyystasolla $\alpha = 0.001$, sanotaan: Testisuureen arvo (tai testin tulos) on **erittäin merkitsevä**.

Merkitsevyystasoa α valittaessa on aina syytä ottaa huomioon väärän päätöksen seuraukset.

Hylkäysalueen määrääminen yksikertaisissa testausasetelmissa

Testin hylkäysalue riippuu yksinkertaisissa testausasetelmissa – paitsi valitusta merkitsevyystasosta α – myös vaihtoehtoisen hypoteesin muodosta.

Olkoon parametria θ koskeva nollahypoteesi yksinkertaista muotoa

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

Valitaan testin merkitsevyystasoksi α ja oletetaan, että testisuureena on (jatkuva) satunnaismuuttuja Z. Tehdään testisuureesta Z seuraavat oletukset:

(1) Testisuureen Z mahdolliset arvot kuuluvat väliin (a,b), jossa voi olla $a = -\infty$ ja/tai $b = +\infty$.

(2a) Testisuureella Z on taipumus saada *suuria* arvoja, jos

$$\theta > \theta_0$$

(2b) Testisuureella Z on taipumus saada *pieniä* arvoja, jos

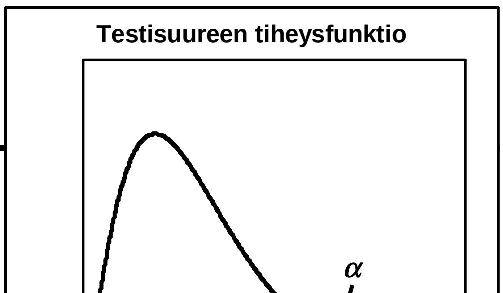
$$\theta < \theta_0$$

Huomautus:

- Oletukset 2a-b pätevät kaikille testisuureille tässä esityksessä.

Tarkastellaan testin hylkäysalueen määräämistä erilaisten vaihtoehtoisten hypoteesien tapauksissa.

(i) Jos vaihtoehtoisena hypoteesina on *yksisuuntainen vaihtoehto*



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

niin testin **hylkäysalue** on muotoa

$$(u,b)$$

jossa **kriittinen arvo** tai **raja** u määrätään siten, että

$$\Pr(Z \geq u \mid \mathrm{H}_0) = \alpha$$

Ks. kuvaa oikealla.

(ii) Jos vaihtoehtoisena hypoteesina on *yksisuuntainen vaihtoehto*

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

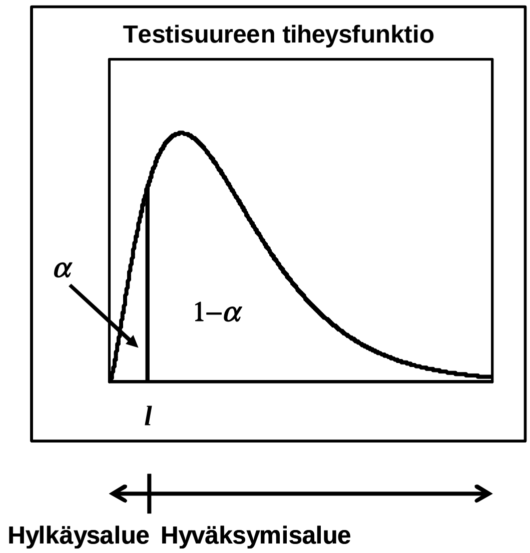
niin testi **hylkäysalue** on muotoa

$$(a,l)$$

jossa **kriittinen arvo** tai **raja** l määrätään siten, että

$$\Pr(Z \leq l \mid \mathrm{H}_0) = \alpha$$

Ks. kuvaa oikealla.



(iii) Jos vaihtoehtoisena hypoteesina on *kaksisuuntainen vaihtoehto*

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

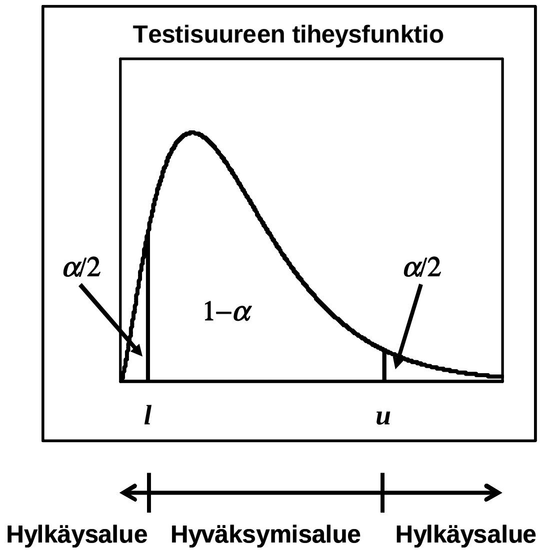
niin testin **hylkäysalue** on muotoa

$$(a,l) \cup (u,b)$$

jossa **kriittiset arvot** tai **rajat** l ja u määrätään siten, että

$$\Pr(Z \geq u \mid \mathrm{H}_0) = \Pr(Z \leq l \mid \mathrm{H}_0) = \alpha/2$$

Ks. kuvaa oikealla.



Jos testisuureen Z jakauma on *symmetrinen* origon suhteen, niin edellä esitetyille kriittisille arvoille pätee:

$l = -u$

Huomautus:

- Todennäköisyydet kohdissa (i)-(iii) ovat ehdollisia todennäköisyyksiä, joissa ehtotapahtumana on se, että nollahypoteesi H_0 pätee. Siten todennäköisyydet määrätään testisuureen *Z* jakaumasta, kun jakaumaa määrättäessä oletetaan, että nollahypoteesi H_0 pätee.

8.6. Testin *p*-arvo

p-arvo

Nollahypoteesin hylkääminen voidaan perustaa etukäteen valitun merkitsevyystason ja sitä vastaavan hylkäysalueen määräämisen sijasta testin *p-arvoon*.

Testin *p*-arvo on *pienin merkitsevyystaso*, jolla nollahypoteesi H_0 voidaan *hylätä*.

Tilastolliset ohjelmistot tulostavat nykyään lähes aina sovellettavien testien *p*-arvot ja siksi *p*-arvojen käyttö *on lähes kokonaan syrjäyttänyt* etukäteen valittujen kiinteiden merkitsevyystasojen käytön.

Testin ***p*-arvo** määrätään seuraavalla tavalla:

(i) Lasketaan valitun *testisuureen arvo* havainnoista.

(ii) Määrätään – *olettaen, että nollahypoteesi* H_0 *pätee* – todennäköisyys sille, että *testisuure saa* (normaaliarvoonsa verrattuna) *niin poikkeuksellisen arvon kuin se on saanut tai vielä poikkeuksellisempia arvoja*.

Jos testin *p*-arvoksi saadaan *pieni* luku, *testisuure on saanut arvon*, *joka kuuluu – nollahypoteesin* H_0 *pätiessä – epätodennäköisten testisuureen arvojen joukkoon*. Siten nollahypoteesi voidaan *hylätä*, jos testin *p*-arvo on *kyllin pieni*. Mitä *pienempi* on testin *p*-arvo, sitä *vahvempia* todisteita havainnot sisältävät nollahypoteesia H_0 *vastaan*.

Huomautus:

- Testin *p*-arvo määrätään testisuureen *Z* jakaumasta, kun jakauma on määrätty olettaen, että nollahypoteesi H_0 pätee.

p-arvon frekvenssitulkinta

Oletetaan, että yleisen hypoteesin H lisäksi myös nollahypoteesi H_0 pätee testausasetelmassa. *Toistetaan otantaa* ja *sovelletaan jokaiseen otokseen samaa testiä*. **Tällöin havaitsemme keskimäärin**

p %:ssa

poimittuja otoksia havaittua testisuureen arvoa poikkeavamman testisuureen arvon.

p-arvo ja testi päätössääntönä

Tilastollinen testi eli päätössääntö, joka kertoo jokaisessa yksittäisessä tilanteessa eli jokaiselle otokselle, onko nollahypoteesi H_0 hylättävä vai ei, voidaan perustaa seuraavalla tavalla testin *p*-arvoon:

(i) Valitaan *pieni* todennäköisyys p_0 .

(ii) Määrätään testin p-arvo.

Jos $p < p_0$, *hylätään* nollahypoteesi H_0 ja *hyväksytään* vaihtoehtoinen hypoteesi H_1.

Jos $p \geq p_0$, *jätetään* nollahypoteesi H_0 *voimaan*.

Todennäköisyyttä p_0 valittaessa on aina syytä ottaa huomioon väärän päätöksen seuraukset.

p-arvon määrääminen yksinkertaisissa testausasetelmissa

Testin *p-arvo* riippuu yksinkertaisissa testausasetelmissa vaihtoehtoisen hypoteesin muodosta.

Olkoon parametria θ koskeva nollahypoteesi yksinkertaista muotoa

$$H_0 : \theta = \theta_0$$

Oletetaan, että testisuureena on (jatkuva) satunnaismuuttuja Z. Oletetaan yksinkertaisuuden vuoksi, että testisuureen jakauma on *symmetrinen* origon suhteen. Tehdään testisuureesta Z lisäksi seuraavat oletukset:

(1) Testisuureen Z mahdolliset arvot kuuluvat väliin (a,b), jossa voi olla $a = -\infty$ ja/tai $b = +\infty$.

(2a) Testisuureella Z on taipumus saada *suuria* arvoja, jos

$$\theta > \theta_0$$

(2b) Testisuureella Z on taipumus saada *pieniä* arvoja, jos

$$\theta < \theta_0$$

Huomautus:

- Oletukset 2a-b pätevät kaikille testisuureille tässä esityksessä.

Oletetaan, että testisuureen Z *havainnoista määrätty arvo* on z.

Tarkastellaan testin p-arvon määräämistä erilaisten vaihtoehtoisten hypoteesien tapauksissa.

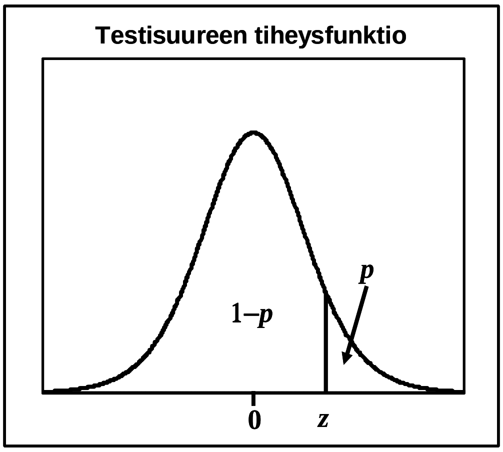
(i) Jos vaihtoehtoisena hypoteesina on *yksisuuntainen vaihtoehto*

$$H_1 : \theta > \theta_0$$

niin testin *p-arvo* on

$$p = \Pr(Z \geq z \mid H_0)$$

Ks. kuvaa oikealla.



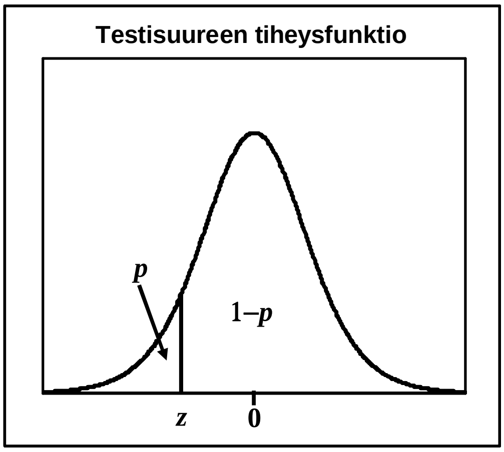
(ii) Jos vaihtoehtoisena hypoteesina on *yksisuuntainen vaihtoehto*

$$H_1 : \theta < \theta_0$$

niin testin *p-arvo* on

$$p = \Pr(Z \le z \mid H_0)$$

Ks. kuvaa oikealla.

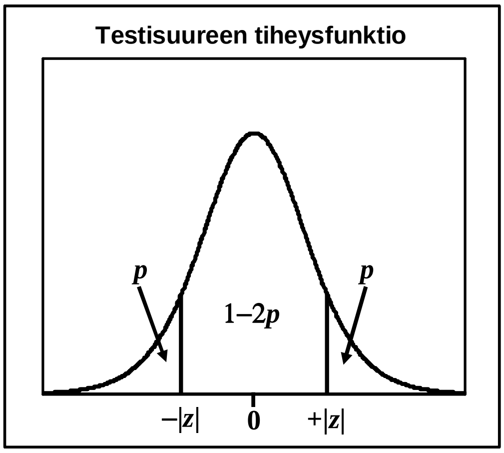


(iii) Jos vaihtoehtoisena hypoteesina on *kaksisuuntainen vaihtoehto*

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

niin testin *p-arvo* on

$$2p = 2\times\Pr(Z \ge |z| \mid H_0)$$



Huomautus:

- Todennäköisyydet kohdissa (i)-(iii) ovat ehdollisia todennäköisyyksiä, joissa ehtotapahtumana on se, että nollahypoteesi H_0 pätee. Siten todennäköisyydet määrätään testisuureen Z jakaumasta, kun jakaumaa määrättäessä on oletettu, että nollahypoteesi H_0 pätee.

8.7. Testin suorittaminen

Testin suorittaminen merkitsevyystason valintaan perustuvissa testausasetelmissa

Jos testi perustetaan *merkitsevyystason valintaan*, testin suorittamisessa on seuraavat vaiheet:

(1) Asetetaan testin **hypoteesit**:

- *Yleinen hypoteesi* H
- Testauksen kohteena oleva *nollahypoteesi* H_0
- *Vaihtoehtoinen hypoteesi* H_1

(2) Valitaan testiä varten **testisuure**.

- Testisuureen tehtävänä on mitata havaintojen ja nollahypoteesin H_0 *yhteensopivuutta*.

(3) Valitaan **merkitsevyystaso** α ja konstruoidaan testille sitä vastaava **hylkäysalue**.

(4) Poimitaan **otos** niin, että *yleisen hypoteesin* H *oletukset pätevät*.

- Jos havaintojen sisältämät todisteet nollahypoteesia H_0 vastaan ovat testisuureella mitattuna *kyllin vahvoja*, nollahypoteesi H_0 *hylätään* ja vaihtoehtoinen hypoteesi H_1 *hyväksytään*.

(5) Määrätään valitun **testisuureen arvo** havainnoista.

(6) Tehdään **päätös** nollahypoteesin hylkäämisestä.

- Jos testisuureen arvo *joutuu hylkäysalueelle*, *hylätään* nollahypoteesi H_0 ja *hyväksytään* vaihtoehtoinen hypoteesi H_1 .
- Jos testisuureen arvo *ei joudu hylkäysalueelle*, *jätetään* nollahypoteesi H_0 *voimaan*.

Testin suorittaminen *p*-arvon valintaan perustuvissa testausasetelmissa

Jos testi perustetaan testisuureen arvoa vastaaviin *p-arvoihin*, testin suorittamisessa on seuraavat vaiheet:

(1) Asetetaan testin **hypoteesit**:

- *Yleinen hypoteesi* H
- Testauksen kohteena oleva *nollahypoteesi* H_0
- *Vaihtoehtoinen hypoteesi* H_1

(2) Valitaan testiä varten **testisuure**.

- Testisuureen tehtävänä on mitata havaintojen ja nollahypoteesin H_0 *yhteensopivuutta*.

(3) Poimitaan **otos** niin, että *yleisen hypoteesin* H oletukset *pätevät*.

- Jos havaintojen sisältämät todisteet nollahypoteesia H_0 vastaan ovat testisuureella mitattuna *kyllin vahvoja*, nollahypoteesi H_0 *hylätään* ja vaihtoehtoinen hypoteesi H_1 *hyväksytään*.

(4) Määrätään valitun **testisuureen arvo** havainnoista.

(5) Määrätään testisuureen havaittua arvoa vastaava ***p*-arvo**.

(6) Tehdään **päätös** nollahypoteesin hylkäämisestä.

- Jos testin *p*-arvo *on kyllin pieni*, *hylätään* nollahypoteesi H_0 ja *hyväksytään* vaihtoehtoinen hypoteesi H_1 .
- Jos testin *p*-arvo *ei ole kyllin pieni*, *jätetään* nollahypoteesi H_0 *voimaan*.

8.8. Normaalijakauman parametreja koskevat testit: Esimerkki

Esitämme tässä kappaleessa testiteorian havainnollistuksena *testit normaalijakauman* **odotusarvolle** ja **varianssille** yksinkertaisen *laadunvalvontaa* koskevan esimerkin tapauksessa. Normaalijakauman parametreja koskevia testejä käsitellään perusteellisesti kappaleessa **Testejä suhdeasteikollisille muuttujille**.

Testausasetelma

Kone tekee ruuveja, joiden **tavoitepituutena** on 10 cm. Ruuvien *pituus vaihtelee kuitenkin satunnaisesti*, mutta voidaan olettaa, että pituuden vaihtelua voidaan kuvata *normaalijakaumalla*.

Ruuveja tehdään koneella erinä. Valmistuserää pidetään myyntikelpoisena, jos erän ruuvien pituudet *eivät vaihtele liian paljon* ja ruuvit ovat *keskimäärin* oikean mittaisia:

> **Ruuvien pituuksien varianssi ei saa ylittää *tilastollisesti merkitsevästi* arvoa 0.01 cm^2 ja ruuvien keskipituus ei saa poiketa *tilastollisesti merkitsevästi* pituuden tavoitearvostaan 10 cm**.

Ruuvien pituutta *valvotaan* seuraavalla tavalla:

(i) Jokaisesta valmistuserästä poimitaan joukko ruuveja tutkittavaksi käyttäen *satunnaisotantaa*.

(ii) Otokseen poimittujen ruuvien *pituudet mitataan*.

(iii) Otokseen poimittujen ruuvien **pituuksien otosvarianssia verrataan arvoon 0.01 cm^2** ja **pituuksien aritmeettista keskiarvoa verrataan ruuvien tavoitepituuteen 10 cm**.

(v) Jos otokseen poimittujen ruuvien pituuksien varianssi on *liian suuri* tai pituuksien aritmeettinen keskiarvo poikkeaa pituuden tavoitearvosta *liian paljon*, niin koko *valmistuserä hylätään*.

Seuraavassa katsotaan, miten ruuvien pituuden valvonnassa käytetään hyväksi **tilastollista testausta**.

Havainnot

Oletetaan, että erään valmistuserän ruuvien joukosta on poimittu *satunnaisotos*, jonka koko on

$$n = 30$$

ja otokseen poimittujen ruuvien pituudet on mitattu.

Otosta kuvaavat seuraavat *tunnusluvut*: Pituuksien *aritmeettinen keskiarvo* on

$$\bar{X} = 10.09 \text{ cm}$$

ja pituuksien *otoskeskihajonta* on

$$s = 0.1038 \text{ cm}$$

Huomautus:

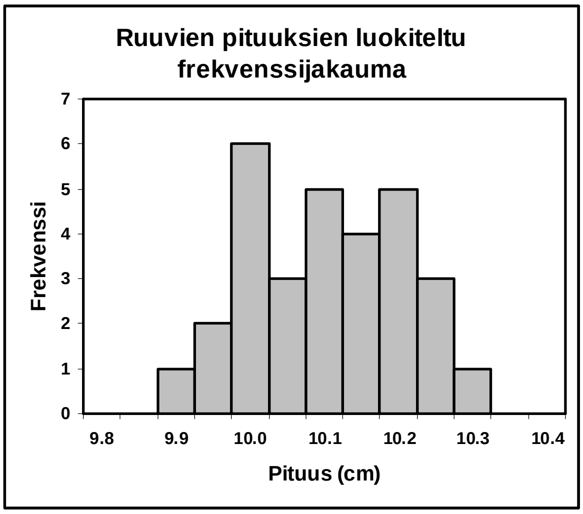
- Samaa aineistoa on tarkasteltu myös luvuissa **Tilastollisten aineistojen kuvaaminen** ja **Väliestimointi**.

Taulukko oikealla esittää otokseen poimittujen ruuvien pituuksien *luokiteltua frekvenssijakaumaa*.

Kuva oikealla esittää otokseen poimittujen ruuvien pituuksien luokiteltua frekvenssijakaumaa vastaavaa *histogrammia*.

Luokkavälit määräävät kuvion suorakaiteiden kannat ja suorakaiteiden korkeudet on valittu niin, että suorakaiteiden pinta-alat suhtautuvat toisiinsa kuten vastaavat luokkafrekvenssit.

Luokkavälit	Luokkafrekvenssit
(9.85,9.90]	1
(9.90,9.95]	2
(9.95,10.00]	6
(10.00,10.05]	3
(10.05,10.10]	5
(10.10,10.15]	4
(10.15,10.20]	5
(10.20,10.25]	3
(10.25,10.30]	1



Huomautus:

- Jos otantaa toistetaan, kaikki otosta koskevat tiedot (sekä havaintoarvot että niistä määrätyt otossuureet kuten aritmeettiset keskiarvot ja otoskeskihajonnat sekä havaintoarvojen jakaumaa kuvaavat graafiset esitykset kuten hisrogrammit) vaihtelevat satunnaisesti otoksesta toiseen.

Ongelma: Onko otosinformaatio *sopusoinnussa* ruuvien pituuden varianssille ja odotusarvolle asetettujen tavoitearvojen kanssa?

Ratkaisu: Konstruoidaan otosinformaation ja ruuvien pituuden varianssin ja odotusarvon tavoitearvoille yhteensopivuuden tutkimista varten tarkoitukseen sopivat **tilastolliset testit**.

Testausasetelmaa koskevat hypoteesit

Määritellään satunnaismuuttuja X:

$$X = \text{ruuvin pituus}$$

Yleinen hypoteesi H :

Otokseen poimittujen ruuvien pituudet *eivät riipu toisistaan* ja ruuvien pituudet *vaihtelevat satunnaisesti noudattaen normaalijakaumaa*:

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

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

Pidämme testauksen aikana kiinni yleisestä hypoteesista H.

Nollahypoteesi $\mathbf{H_{10}}$:

Ruuvien pituuksien varianssi *on korkeintaan* 0.01 cm^2 :

$$\mathrm{H}_{10} : \sigma^2 \le \sigma_0^2 = 0.01 \text{ cm}^2$$

Vaihtoehtoinen hypoteesi $\mathbf{H_{11}}$:

Ruuvien pituuksien varianssi *on suurempi kuin* 0.01 cm^2 :

$$\mathrm{H}_{11} : \sigma^2 > \sigma_0^2 = 0.01 \text{ cm}^2$$

Ruuvien pituuksien *varianssia* koskevaa nollahypoteesia H_{10} voidaan testata ns. **χ^2-testillä**; ks. alla.

Nollahypoteesi $\mathbf{H_{20}}$:

Ruuvien pituuksien odotusarvo *yhtyy tavoitearvoonsa* 10 cm:

$$\mathrm{H}_{20} : \mu = \mu_0 = 10 \text{ cm}$$

Vaihtoehtoinen hypoteesi $\mathbf{H_{21}}$:

Ruuvien pituuksien odotusarvo *poikkeaa pituuden tavoitearvosta* 10 cm:

$$\mathrm{H}_{21} : \mu \ne \mu_0 = 10 \text{ cm}$$

Ruuvien pituuksien *odotusarvoa* koskevaa nollahypoteesia H_{20} voidaan testata ns. ***t*-testillä**: ks. alla.

χ^2-testi varianssille

Yleinen hypoteesi H :

Otokseen poimittujen ruuvien pituudet *eivät riipu toisistaan* ja ruuvien pituudet *vaihtelevat satunnaisesti noudattaen normaalijakaumaa*:

$$X_1, X_2, \dots , X_n \perp$$

$$X_i \sim \mathrm{N}(\mu, \sigma^2) \text{ , } i = 1, 2, \dots , n$$

Nollahypoteesi H$_{10}$:

Ruuvien pituuksien varianssi *on korkeintaan* 0.01 cm^2 :

$$\mathrm{H}_{10} : \sigma^2 \le \sigma_0^2 = 0.01 \text{ cm}^2$$

Vaihtoehtoinen hypoteesi H$_{11}$:

Ruuvien pituuksien varianssi *on suurempi kuin* 0.01 cm^2 :

$$\mathrm{H}_{11} : \sigma^2 > \sigma_0^2 = 0.01 \text{ cm}^2$$

Käytetään **testisuureena** χ^2*-testisuuretta*

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

jossa

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

on havaintojen (harhaton) *otosvarianssi* ja σ_0^2 on *nollahypoteesin* H$_{10}$ *kiinnittämä parametrin* σ^2 *arvo*.

Voidaan osoittaa, että *testisuure* χ^2 noudattaa **χ^2-jakaumaa** vapaustein (n – 1), *jos yleinen hypoteesi* H ja *ehto*

$$\sigma^2 = \sigma_0^2$$

pätevät (ks. lukua **Otokset ja otosjakaumat**):

$$\chi^2 \quad \chi^2(n-1)$$

Esimerkin tapauksessa otokseen poimittujen ruuvien pituuksien *aritmeettinen keskiarvo* oli

$$\bar{X} = 10.09 \text{ cm}$$

otokseen poimittujen ruuvien pituuksien *otoskeskihajonta* oli

$$s = 0.1038 \text{ cm}$$

ja *nollahypoteesin* H$_{10}$ *kiinnittämä parametrin* σ^2 *arvo* oli

$$\sigma_0^2 = 0.01 \text{ cm}^2$$

Siten χ^2*-testisuureen arvoksi* saadaan

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(30-1)\times 0.1038^2}{0.01} = 31.246$$

Voidaan osoittaa, että yllä määritellyn χ^2-*testisuureen* **normaaliarvo** eli *testisuureen odotusarvo ehdon*

$$\sigma^2 = \sigma_0^2$$

pätiessä on

$$\mathrm{E}(\chi^2 \mid \sigma^2 = \sigma_0^2) = n-1$$

χ^2-testisuureen normaaliarvoonsa (n – 1) verrattuna *suuret* ja *pienet* arvot viittaavat siihen, että *nollahypoteesi* H_{10} *ei päde*.

Valitaan **merkitsevyystasoksi**

$$\alpha = 0.05$$

Koska *vaihtoehtoinen hypoteesi*

$$\mathrm{H}_{11}: \ \sigma^2 > \sigma_0^2 = 0.01 \text{ cm}^2$$

on *yksisuuntainen*, *hylkäysalueen määräämistä varten valitaan* **kriittinen arvo** tai **raja** χ_α^2 siten, että

$$\Pr(\chi^2 \geq \chi_\alpha^2) = \alpha = 0.05$$

jossa satunnaismuuttuja χ^2 noudattaa χ^2-*jakaumaa* vapausastein (n – 1) = 29. Kriittinen arvo χ_α^2 toteuttaa ehdon

$$\Pr(\chi^2 \leq \chi_\alpha^2) = 1-\alpha = 0.95$$

χ^2-jakauman *taulukoista* nähdään, että

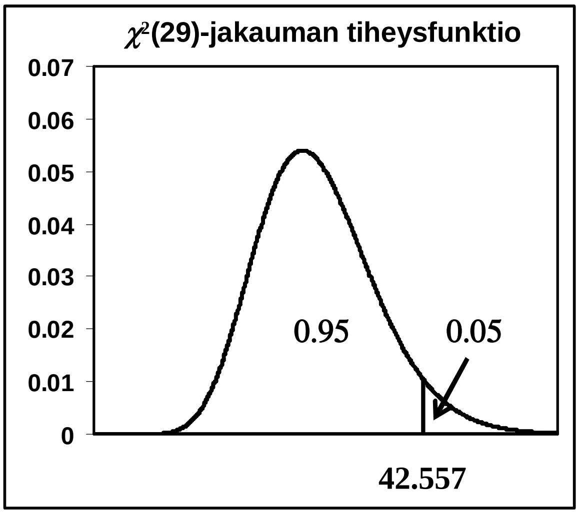
$$\Pr(\chi^2 \geq 42.557) = 0.05$$

kun vapausasteiden lukumäärä (n – 1) = 29.

Siten haluttu *kriittinen arvo* on:

$$\chi_\alpha^2 = 42.557$$

Kuvio oikealla havainnollistaa kriittisen arvon määräämistä.



Valitaan χ^2-testin **hylkäysalueeksi**

$$(\chi_\alpha^2, +\infty)$$

Jos χ^2-testisuureen arvo joutuu hylkäysalueelle, *nollahypoteesi*

$$\mathrm{H}_{10}: \ \sigma^2 \leq \sigma_0^2$$

hylätään merkitsevyystasolla α = 0.05. Todennäköisyys, että χ^2-testisuureen arvo joutuu ehdon

$$\sigma^2 = \sigma_0^2$$

pätiessä hylkäysalueelle on $\alpha = 0.05$.

χ^2-testin **hyväksymisalue** on muotoa

$$[0, \chi_\alpha^2]$$

Jos χ^2-testisuureen arvo joutuu hyväksymisalueelle, *nollahypoteesi*

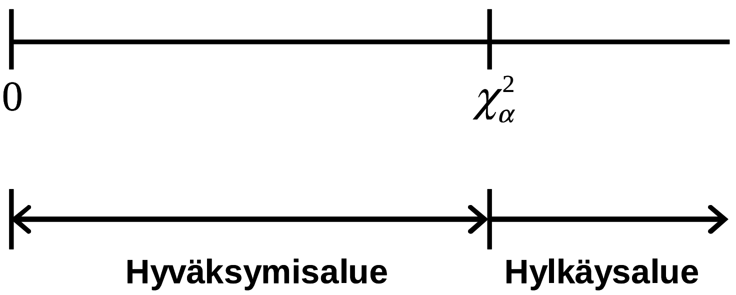
$$\mathrm{H}_{10}: \sigma^2 \le \sigma_0^2$$

jätetään voimaan merkitsevyystasolla α.

χ^3-testin **hylkäys-** ja **hyväksymisalueita** voidaan kuvata *yksisuuntaisen vaihtoehtoisen hypoteesin*

$$\mathrm{H}_{11}: \sigma^2 > \sigma_0^2$$

tapauksessa alla olevalla kuviolla:



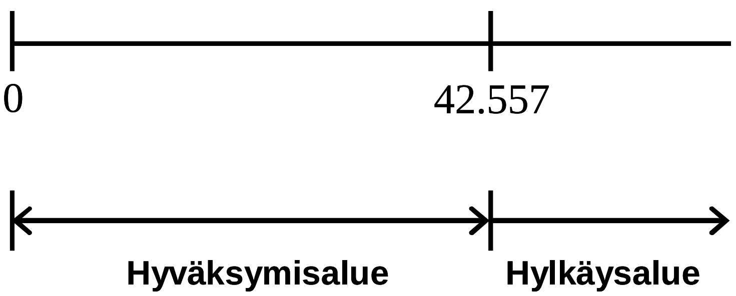
Kriittinen arvo χ_α^2 on määrätty siten, että

$$\Pr(\chi^2 \ge \chi_\alpha^2) = \alpha$$

jolloin

$$\Pr(\chi^2 \le \chi_\alpha^2) = 1 - \alpha$$

Esimerkin tapauksessa χ^2-testin **hylkäys-** ja **hyväksymisalueet** saavat seuraavan muodon:



Kriittinen arvo 42.557 on siis määrätty siten, että

$$\Pr(\chi^2 \ge 42.557) = 0.05$$

jolloin

$$\Pr(\chi^2 \le 42.557) = 0.95$$

Esimerkin tapauksessa otoksesta määrätty χ^2-testisuureen arvo *on pienempi kuin kriittinen arvo* χ^2_α:

$$\chi^2 = 31.246 < 42.557 = \chi^2_\alpha$$

Siten *testisuureen arvo on joutunut hyväksymisalueelle* ja **voimme jättää nollahypoteesin**

$$\mathrm{H}_{10} : \sigma^2 \le \sigma_0^2 = 0.01 \text{ cm}^2$$

voimaan *merkitsevyystasolla* $\alpha = 0.05$.

Johtopäätös:

Ruuvien pituuden varianssi *ei ole* tilastollisesti merkitsevästi arvoa 0.01 cm^2 suurempi.

Oletetaan, että ruuveja tekevä kone toimii niin, että ruuvien *pituuden varianssi on jatkuvasti hyväksyttävän suuruista*. Tällöin siis nollahypoteesi

$$\mathrm{H}_{10} : \sigma^2 \le \sigma_0^2 = 0.01 \text{ cm}^2$$

pätee koko ajan. Oletetaan nyt, että poimimme koneen tekemien ruuvien joukosta toistuvasti uusia otoksia ja testaamme jokaisen otoksen perusteella nollahypoteesia H_{10} käyttämällä merkitsevyystasona lukua

$$\alpha = 0.05$$

Tällöin joudumme hylkäämään nollahypoteesin H_{10} keskimäärin 5 kertaa 100:sta, **vaikka nollahypoteesi H_{10} pätee koko ajan**.

Esimerkin tapauksessa otoksesta määrättyä χ^2-testisuureen arvoa 31.270 vastaava ***p*-arvo** on χ^2-jakauman taulukoiden mukaan

$$p = \Pr(\chi^2 > 31.270) > 0.1$$

Tämä merkitsee sitä, että χ^2-testisuure saa normaaliarvoonsa $(n - 1) = 29$ nähden arvoa 31.270 poikkeuksellisempia arvoja todennäköisyydellä, joka on suurempi kuin 0.1, *jos nollahypoteesi*

$$\mathrm{H}_{10} : \sigma^2 \le \sigma_0^2 = 0.01 \text{ cm}^2$$

pätee. Siten *emme voi hylätä nollahypoteesia* H_{01} *millään tavanomaisella merkitsevyystasolla*. Saatu tulos on sopusoinnussa merkistevyystasoa käyttävän tekniikan avulla saadun tuloksen kanssa.

t-testi odotusarvolle

Yleinen hypoteesi H :

Otokseen poimittujen ruuvien pituudet *eivät riipu toisistaan* ja ruuvien pituudet *vaihtelevat satunnaisesti noudattaen normaalijakaumaa*:

$$X_1, X_2, \dots , X_n \perp$$

$$X_i \sim \mathrm{N}(\mu, \sigma^2) \text{ , } i = 1, 2, \dots , n$$

Nollahypoteesi H_{20} :

Ruuvien pituuksien odotusarvo *yhtyy tavoitearvoonsa* 10 cm:

$$\mathrm{H}_{20} : \mu = \mu_0 = 10 \text{ cm}$$

Vaihtoehtoinen hypoteesi H_{21} :

Ruuvien pituuksien odotusarvo *poikkeaa pituuden tavoitearvosta* 10 cm:

$$\mathrm{H}_{21}:\ \mu \neq \mu_0 = 10 \text{ cm}$$

Käytetään **testisuureena** *t-testisuuretta*

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

jossa

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

on havaintojen *aritmeettinen keskiarvo*,

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

on havaintojen (harhaton) *otosvarianssi* ja μ_0 *nollahypoteesin* H_{20} *kiinnittämä odotusarvo-parametrin* μ *arvo*.

Voidaan osoittaa, että testisuure t noudattaa ***t*-jakaumaa** vapaustein ($n - 1$), *jos yleinen hypoteesi* H ja *nollahypoteesi*

$$\mathrm{H}_{20}:\ \mu = \mu_0$$

pätevät (ks. lukua **Otokset ja otosjakaumat**):

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

t-testisuure mittaa havaintojen aritmeettisen keskiarvon $\bar{X}$ ja nollahypoteesin H_{20} kiinnittämän odotusarvoparametrin μ arvon μ_0 *tilastollista etäisyyttä*, jossa mittayksikkönä on aritmeettisen keskiarvon $\bar{X}$ keskivirheen $\sigma/\sqrt{n}$ estimaattori $s/\sqrt{n}$.

Esimerkin tapauksessa otokseen poimittujen ruuvien pituuksien *aritmeettinen keskiarvo* oli

$$\bar{X} = 10.09 \text{ cm}$$

otokseen poimittujen ruuvien pituuksien *otoskeskihajonta* oli

$$s = 0.1038 \text{ cm}$$

ja *nollahypoteesin* H_{10} *kiinnittämä parametrin* μ *arvo* oli

$$\mu_0 = 10 \text{ cm}$$

Siten *testisuureen t arvoksi* saadaan

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{10.09 - 10}{0.1038/\sqrt{30}} = 4.749$$

Voidaan osoittaa, että edellä määritellyn *t-testisuureen* **normaaliarvo** eli *testisuureen odotusarvo nollahypoteesin*

$$\mathrm{H}_{20}:\ \mu = \mu_0$$

pätiessä on

$$\mathrm{E}(t \mid \mathrm{H}_{20}) = 0$$

t-testisuureen itseisarvoltaan *suuret* arvot viittaavat siihen, että *nollahypoteesi* H_{20} *ei päde*.

Huomautus:

- Testisuureen t jakauma on *symmetrinen* origon suhteen.

Valitaan **merkitsevyystasoksi**

$$\alpha = 0.05$$

Koska *vaihtoehtoinen hypoteesi*

$$\mathrm{H}_{21}: \ \mu \neq \mu_0 = 10 \text{ cm}$$

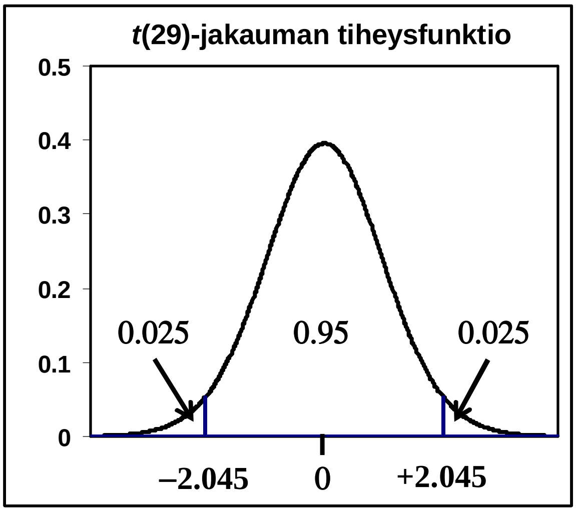
on *kaksisuuntainen*, *hylkäysalueen määräämistä varten valitaan* **kriittiset arvot** $-t_{\alpha/2}$ ja $+t_{\alpha/2}$ siten, että

$$\Pr(t \leq -t_{\alpha/2}) = \Pr(t \geq +t_{\alpha/2}) = 0.025$$

jossa satunnaismuuttuja t noudattaa t-*jakaumaa* vapausastein $(n - 1) = 29$. Kriittiset arvot $-t_{\alpha/2}$ ja $+t_{\alpha/2}$ toteuttavat ehdon

$$\Pr(-t_{\alpha/2} \leq t \leq +t_{\alpha/2}) = 1 - \alpha = 0.95$$

Huomaa, että merkitsevyystasoon α liittyvät kriittiset arvot ovat tässä (kaksisuuntaisen vaihtoehtoisen hypoteesin) tapauksessa täsmälleen samat kuin *luottamustasoon* $(1 - \alpha)$ liittyvät *luottamuskertoimet*; ks. lukua **Väliestimointi**.



t-jakauman *taulukoista* nähdään, että

$$\Pr(t \geq +2.045) = 0.025$$

$$\Pr(t \leq -2.045) = 0.025$$

kun vapausasteiden lukumäärä $(n - 1) = 29$.

Siten *kriittiset arvot* ovat:

$$+t_{\alpha/2} = +2.045$$

$$-t_{\alpha/2} = -2.045$$

Kuvio oikealla havainnollistaa kriittisten rajojen määräämistä.

t-testin **hylkäysalue** on muotoa

$$(-\infty, -t_{\alpha/2}) \cup (+t_{\alpha/2}, +\infty)$$

Jos t-testisuureen arvo joutuu hylkäysalueelle, *nollahypoteesi*

$$\mathrm{H}_{20}: \ \mu = \mu_0$$

hylätään merkitsevyystasolla α. Todennäköisyys, että t-testisuureen arvo joutuu nollahypoteesin H_{20} pätiessä hylkäysalueelle on α.

t-testin **hyväksymisalue** on muotoa

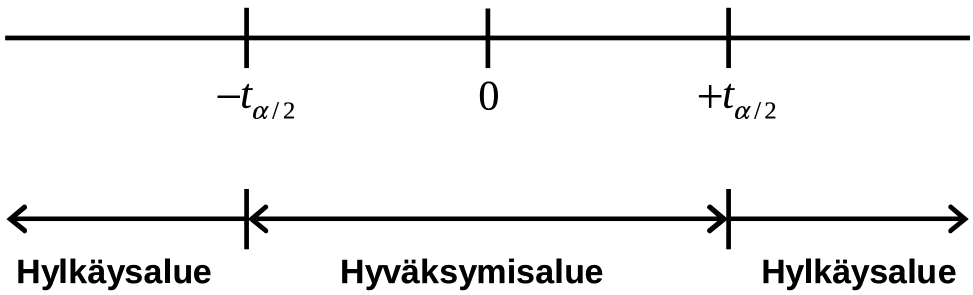
$$[-t_{\alpha/2}, +t_{\alpha/2}]$$

Jos t-testisuureen arvo joutuu hyväksymisalueelle, *nollahypoteesi*

$$\mathrm{H}_{20}: \ \mu = \mu_0$$

jätetään voimaan merkitsevyystasolla α.

t-testin **hylkäys-** ja **hyväksymisalueita** voidaan kuvata *kaksisuuntaisen vaihtoehtoisen hypoteesin tapauksessa* alla olevalla kaaviolla:



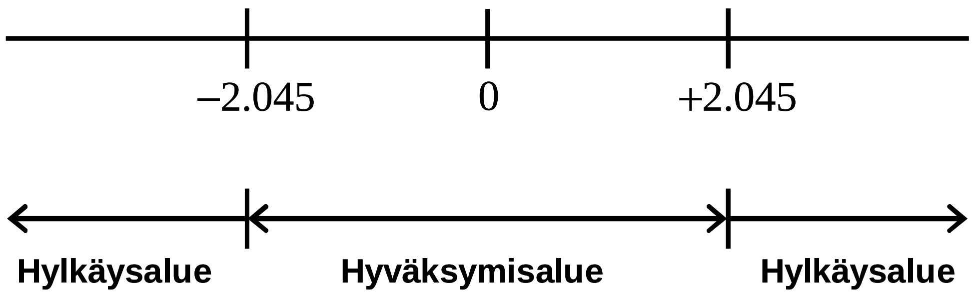
Kriittiset arvot $-t_{\alpha/2}$ ja $+t_{\alpha/2}$ on määrätty siten, että

$$\Pr(t \le -t_{\alpha/2}) = \Pr(t \ge +t_{\alpha/2}) = \alpha/2$$

jolloin

$$\Pr(-t_{\alpha/2} \le t \le +t_{\alpha/2}) = 1-\alpha$$

Esimerkin tapauksessa t-testin **hylkäys-** ja **hyväksymisalueet** saavat seuraavan muodon:



Kriittiset arvot -2.045 ja $+2.045$ on siis määrätty siten, että

$$\Pr(t \le -2.045) = \Pr(t \ge +2.045) = 0.025$$

jolloin

$$\Pr(-2.045 \le t \le +2.045) = 0.95$$

Esimerkin tapauksessa otoksesta määrätty t-testisuureen arvo *on suurempi kuin kriittinen arvo* $+t_{\alpha/2}$:

$$t = 4.749 > 2.045 = +t_{0.025}$$

Koska testisuureen arvo on joutunut hylkäysalueelle, **voimme hylätä nollahypoteesin**

$$\mathrm{H}_{20}: \ \mu = \mu_0 = 10 \text{ cm}$$

ja **hyväksyä vaihtoehtoisen hypoteesin**

$$\mathrm{H}_{21}: \ \mu \neq \mu_0 \ = 10 \text{ cm}$$

merkitsevyystasolla $\alpha = 0.05$.

Johtopäätös:

> **Ruuvien keskimääräinen pituus *poikkeaa* tilastollisesti merkitsevästi tavoitearvostaan 10 cm. Saattaa olla syytä pysäyttää ruuveja valmistava kone tarkistusta varten**.

Oletetaan, että ruuveja tekevä kone toimii niin, että ruuvien *pituuden odotusarvo on jatkuvasti hyväksyttävän suuruinen*. Tällöin siis nollahypoteesi

$$\mathrm{H}_{20}: \ \mu = \mu_0 \ = 10 \text{ cm}$$

pätee koko ajan. Oletetaan nyt, että poimimme koneen tekemien ruuvien joukosta toistuvasti uusia otoksia ja testaamme jokaisen otoksen perusteella nollahypoteesia H_{20} käyttämällä merkitsevyys-tasona lukua

$$\alpha = 0.05$$

Tällöin joudumme hylkäämään nollahypoteesin $\mathbf{H_{20}}$ keskimäärin 5 kertaa 100:sta, **vaikka nollahypoteesi $\mathbf{H_{20}}$ pätee koko ajan**.

Esimerkin tapauksessa otoksesta määrätty t-testisuureen arvoa 4.749 vastaava ***p*-arvo** on t-jakauman taulukoiden mukaan

$$p = 2\times\Pr(t > |4.749|) < 2\times0.0005 = 0.001$$

Tämä merkitsee sitä, että t-testisuure saa normaaliarvoonsa 0 nähden arvoa 4.749 poikkeuksellisempia arvoja todennäköisyydellä, joka on pienempi kuin 0.001, *jos nollahypoteesi*

$$\mathrm{H}_{20}: \ \mu = \mu_0 \ = 10 \text{ cm}$$

pätee. Siten *voimme hylätä nollahypoteesin* H_{20} *kaikilla tavanomaisilla merkitsevyystasoilla*. Saatu tulos on sopusoinnussa merkistevyystasoa käyttävän tekniikan avulla saadun tuloksen kanssa.

8.9. Tilastolliset testit ja havaintojen mitta-asteikoilliset ominaisuudet

Havaintojen *mitta-asteikolliset ominaisuudet* ohjaavat testin valintaa; lisätietoja mitta-asteikoista: ks. lukua **Tilastollisten aineistojen kerääminen ja mittaaminen**. Tilastolliset testit voidaan ryhmitellä havaintojen mitta-asteikollisten ominaisuuksien suhteen seuraavalla tavalla:

- **Testit suhde- tai välimatka-asteikollisille muuttujille**
- **Testit järjestysasteikollisille muuttujille**
- **Testit laatueroasteikollisille muuttujille**

Seuraavissa kolmessa luvussa tarkastellaan suurta joukkoa erilaisia testejä. Tarkasteltavat testit ovat seuraavat:

Testejä suhde- tai välimatka-asteikollisille muuttujille:

- **Yhden otoksen *t*-testi normaalijakauman odotusarvolle**
- **Kahden riippumattoman otoksen *t*-testi normaalijakaumien odotusarvoille: Yleinen tapaus**

- **Kahden riippumattoman otoksen *t*-testi normaalijakaumien odotusarvoille: Yhtä suurten varianssien tapaus**
- ***t*-testi parivertailuille**
- **Yhden otoksen testi normaalijakauman varianssille**
- **Kahden riippumattoman otoksen testi normaalijakauman variansseille**

Ks. lukua **Testejä suhdeasteikollisille muuttujille**.

Testejä järjestysasteikollisille muuttujille:

- **Merkkitesti**
- **Wilcoxonin rankitesti**
- **Mannin ja Whitneyn testi** eli **Wilcoxonin rankisummatesti**

Ks. lukua **Testejä järjestysasteikollisille muuttujille**.

Testejä laatueroasteikollisille muuttujille:

- **Testi suhteelliselle osuudelle**
- **Suhteellisten osuuksien vertailutesti**

Ks. lukua **Testejä laatueroasteikollisille muuttujille**.

9. Testejä suhdeasteikollisille muuttujille

Tarkastelemme tässä luvussa seuraavia **testejä normaalijakauman parametreille**:

- **Yhden otoksen *t*-testi odotusarvolle**
- **Kahden riippumattoman otoksen *t*-testi *A* odotusarvoille: Yleinen tapaus**
- **Kahden riippumattoman otoksen *t*-testi *B* odotusarvoille: Yhtä suurten varianssien tapaus**
- ***t*-testi parivertailuille**
- **Yhden otoksen χ^2-testi varianssille**
- **Kahden riippumattoman otoksen *F*-testi variansseille**

Testejä normaalijakauman odotusarvolle *voidaan soveltaa myös monissa sellaisissa tilanteissa, joissa havainnot eivät noudata normaalijakaumaa*. Siten testejä normaalijakauman odotusarvolle voidaan pitää **yleisinä testeinä** *suhde-* (tai *välimatka-*) *asteikollisten satunnaismuuttujien* **odotusarvoparametrille**.

Avainsanat:

Aritmeettinen keskiarvo, Asymptoottien testi, *F*-jakauma, *F*-testi, Hylkäysalue, Hylkäysvirhe, Hyväksymisalue, Hyväksymisvirhe, χ^2-jakauma, χ^2-testi, Kahden otoksen testi, Kriittinen arvo, Merkitsevyystaso, Nollahypoteesi, Normaaliarvo, Normaalijakauma, Odotusarvo, Otos, Otoskeskihajonta, Otosvarianssi, *p*-arvo, Parametri, Parivertailu, Riippumattomat otokset, Sovitettu pari, Standardipoikkeama, Suhdeasteikko, *t*-jakauma, *t*-testi, Testisuure, Todennäköisyysjakauma, Varianssi, Vertailutesti, Voimakkuus, Välimatka-asteikko, Yhden otoksen testi, Yleinen hypoteesi

9.1. Suhdeasteikollisten muuttujien testit

Normaalijakauma on tilastotieteen tärkein jakauma. Oletetaan, että satunnaismuuttuja X noudattaa normaalijakaumaa **parametrein** μ ja σ^2:

$$X \quad \mathrm{N}(\mu, \sigma^2)$$

Tällöin parametri μ on normaalijakauman *odotusarvo* ja paramaetri σ^2 on normaalijakauman **varianssi**:

$$\mu = \mathrm{E}(X)$$

$$\sigma^2 = \mathrm{Var}(X)$$

Parametrit μ ja σ^2 määräävät täysin normaalijakauman.

Normaalijakauman parametreja koskevat testit voidaan jakaa kahteen ryhmään:

- **Yhden otoksen testit**
- **Kahden otoksen testit**

Yhden otoksen testeissä testataan *yksinkertaisia nollahypoteeseja*, jotka koskevat normaalijakauman *odotusarvo*- tai *varianssiparametria*. Kahden otoksen testit ovat *vertailutestejä*, joissa verrataan kahden normaalijakauman *odotusarvo*- tai *varianssiparametreja* toisiinsa.

Tässä luvussa tarkastellaan seuraavia **testejä normaalijakauman parametreille**:

- **Yhden otoksen *t*-testi odotusarvolle**
- **Kahden riippumattoman otoksen *t*-testi *A* odotusarvoille: Yleinen tapaus**
- **Kahden riippumattoman otoksen *t*-testi *B* odotusarvoille: Yhtä suurten varianssien tapaus**
- ***t*-testi parivertailuille**
- **Yhden otoksen χ^2-testi varianssille**
- **Kahden riippumattoman otoksen *F*-testi variansseille**

Testejä normaalijakauman odotusarvolle *voidaan soveltaa myös monissa sellaisissa tilanteissa, joissa havainnot eivät noudata normaalijakaumaa*. Tämä perustuu seuraaviin seikkoihin:

(i) Testit odotusarvolle perustuvat havaintojen *aritmeettisiin keskiarvoihin*.

(ii) *Keskeisen raja-arvolauseen mukaan* myös *ei-normaalisten havaintojen aritmeettiset keskiarvot noudattavat* – tietyin ehdoin – *suurissa otoksissa approksimatiivisesti normaalijakaumaa*.

Tämä merkitsee sitä, että testejä normaalijakauman odotusarvolle voidaan pitää **yleisinä testeinä** *suhde*- (tai *välimatka*-) *asteikollisten satunnaismuuttujien* **odotusarvoparametrille**; mitta-asteikot: ks. lukua **Tilastollisten aineistojen kerääminen ja mittaaminen**.

Sen sijaan **testit normaalijakauman varianssille** *eivät yleensä ole käyttökelpoisia ei-normaalisille havainnoille* ja tilanne ei välttämättä parane suurillakaan havaintojen lukumäärillä.

9.2. Yhden otoksen *t*-testi odotusarvolle

Testausasetelma yhden otoksen *t*-testissä

Olkoon

$$X_i\ ,\ i = 1, 2, \ldots , n$$

satunnaisotos normaalijakaumasta $N(\mu,\sigma^2)$. Tällöin satunnaismuuttujat X_i , $i = 1, 2, \ldots , n$ ovat *riippumattomia* ja noudattavat *samaa normaalijakaumaa* $N(\mu,\sigma^2)$:

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

Asetetaan normaalijakauman $N(\mu,\sigma^2)$ *odotusarvo*- eli *paikkaparametrille* μ *nollahypoteesi*

$$H_0 : \mu = \mu_0$$

Ongelma: Ovatko havainnot *sopusoinnussa* hypoteesin H_0 kanssa?

Ratkaisu: Yhden otoksen *t*-testi.

Hypoteesit yhden otoksen *t*-testissä

Yleinen hypoteesi H :

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

Nollahypoteesi:

$$H_0 : \mu = \mu_0$$

Vaihtoehtoiset hypoteesit:

$$\left.\begin{array}{l} H_1 : \mu > \mu_0 \\ H_1 : \mu < \mu_0 \end{array}\right\} \text{1-suuntaiset vaihtoehtoiset hypoteesit}$$

$$H_1 : \mu \neq \mu_0 \qquad \text{2-suuntainen vaihtoehtoinen hypoteesi}$$

Parametrien estimointi yhden otoksen *t*-testissä

Olkoot

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

ja

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

tavanomaiset *harhattomat estimaattorit* normaalijakauman parametreille μ ja σ^2. Tunnusluku $\bar{X}$ on havaintojen X_i , $i = 1, 2, \ldots , n$ *aritmeettinen keskiarvo* ja s^2 on havaintojen X_i , $i = 1, 2, \ldots , n$ (harhaton) *otosvarianssi*.

Testisuure ja sen jakauma yhden otoksen *t*-testissä

Määritellään ***t*-testisuure**

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

Jos *nollahypoteesi*

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

pätee, niin *testisuure t noudattaa t-jakaumaa* vapausastein $(n - 1)$:

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

Perustelu:

Oletetaan, että testin yleinen hypoteesi H ja nollahypoteesi H_0 pätevät:

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

Koska tällöin (ks. lukua **Otokset ja otosjakaumat**)

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \mathrm{N}\left(\mu_0, \frac{\sigma^2}{n}\right)$$

niin

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad \mathrm{N}(0,1)$$

Koska standardipoikkeama σ on *tuntematon*, satunnaismuuttuja Z on testisuureena *epä-operationaalinen*.

Jos standardipoikkeama σ korvataan satunnaismuuttujan Z lausekkeessa vastaavalla *otos-suureella*

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

niin saadaan *t-testisuure*

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

joka *nollahypoteesin* H_0 *pätiessä noudattaa t-jakaumaa* vapausastein $(n - 1)$:

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

Todistus: ks. lukua **Otokset ja otosjakaumat**.

■

Testisuure t mittaa havaintoarvojen aritmeettisen keskiarvon $\bar{X}$ ja nollahypoteesin H_0 kiinnittämän odotusarvoparametrin μ arvon μ_0 *tilastollista etäisyyttä*. *Mittayksikkönä* on erotuksen

$$\bar{X} - \mu_0$$

standardipoikkeaman

$$\sigma / \sqrt{n}$$

estimaattori, jota määrättäessä on oletettu, että nollahypoteesi H_0 pätee.

Testisuureen *t normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$E(t) = 0$$

Siten itseisarvoltaan *suuret* testisuureen *t* arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*.

Hylkäysalueen määrääminen yhden otoksen *t*-testissä

Valitaan testin **merkitsevyystasoksi** α.

(i) Jos vaihtoehtoinen hypoteesi on muotoa

$$H_1 : \mu > \mu_0$$

niin kriittinen arvo $+t_\alpha$ saadaan ehdosta

$$\Pr(t \geq +t_\alpha) = \alpha$$

jossa $t \quad t(n-1)$. Testin hylkäysalue on tällöin muotoa

$$(+t_\alpha, +\infty)$$

(ii) Jos vaihtoehtoinen hypoteesi on muotoa

$$H_1 : \mu < \mu_0$$

niin kriittinen arvo $-t_\alpha$ saadaan ehdosta

$$\Pr(t \leq -t_\alpha) = \alpha$$

jossa $t \quad t(n-1)$. Testin hylkäysalue on tällöin muotoa

$$(-\infty, -t_\alpha)$$

(iii) Jos vaihtoehtoinen hypoteesi on muotoa

$$H_1 : \mu \neq \mu_0$$

niin kriittiset arvot $-t_{\alpha/2}$ ja $+t_{\alpha/2}$ saadaan ehdoista

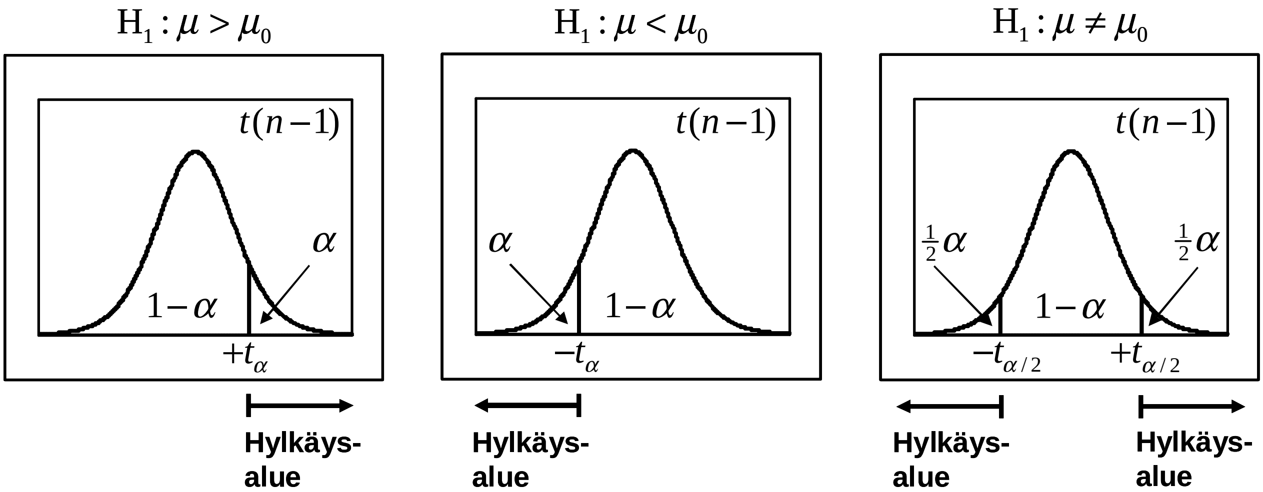
$$\Pr(t \leq -t_{\alpha/2}) = \alpha/2$$
$$\Pr(t \geq +t_{\alpha/2}) = \alpha/2$$

jossa $t \quad t(n-1)$. Testin hylkäysalue on tällöin muotoa

$$(-\infty, -t_{\alpha/2}) \cup (+t_{\alpha/2}, +\infty)$$

Nollahypoteesi hylätään, jos testisuureen arvo osuu hylkäysalueelle.

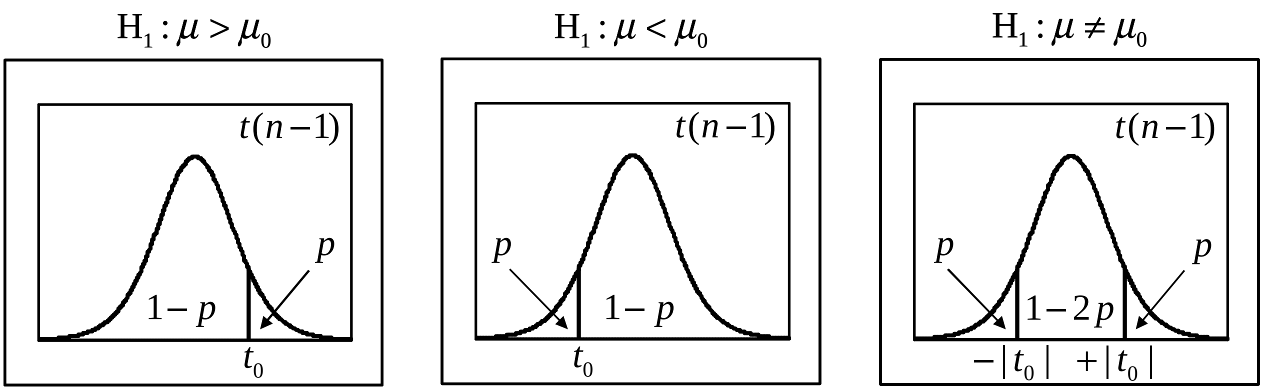
Alla olevat kuviot havainnollistavat testin hylkäysalueen määräämistä:



p-arvon määrääminen yhden otoksen *t*-testissä

Olkoon *t-testisuureen* havaittu arvo t_0 .

Alla olevat kuviot havainnollistavat testin *p-arvon* määräämistä:



Testin *p*-arvo = *p* Testin *p*-arvo = *p* Testin *p*-arvo = 2*p*

Nollahypoteesi *hylätään, jos testin p-arvo on kyllin pieni*.

Normaalisuusoletuksen merkitys yhden otoksen *t*-testissä

Yhden otoksen *t*-testin yleisessä hypoteesissa oletetaan, että havainnot ovat *normaalijakautuneita*. *t*-testi *ei* kuitenkaan *ole herkkä poikkeamille normaalisuudesta*, jos havaintojen lukumäärä *n* on ”*kyllin suuri*”.

Testiä on melko turvallista käyttää, kun havaintojen lukumäärä

$$n > 15$$

ellei havaintojen jakauma *ole kovin vino* ja havaintojen joukossa *ole poikkeavia havaintoja*. Jos havaintojen lukumäärä

$$n > 40$$

testiä voidaan melko turvallisesti käyttää jopa selvästi vinoille havaintojen jakaumille.

Yhden otoksen *t*-testin hyväksymisvirheen todennäköisyys ja voimakkuus

Tarkastellaan t-testin **hyväksymisvirheen todennäköisyyttä** ja **voimakkuutta** tilanteessa, jossa normaalijakauman *varianssi* σ^2 *on tunnettu*.

Huomautus:

- Jos normaalijakauman varianssia σ^2 *tunneta*, vaatii t-testin voimakkuuden määrääminen ns. *epäkeskisen t-jakauman* määrittelemistä. Sivuutamme tämän yleisen tapauksen käsittelyn tässä esityksessä.

Olkoon *nollahypoteesi* muotoa

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

ja *vaihtoehtoinen hypoteesi* muotoa

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

Koska normaalijakauman varianssi σ^2 on oletettu tunnetuksi, nollahypoteesia voidaan testata *t-testisuureella*

$$t = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

Testitisuure t *noudattaa nollahypoteesin* H_0 *pätiessä standardoitua normaalijakaumaa* (ks. lukua **Otokset ja otosjakaumat**):

$$t \sim \mathrm{N}(0, 1)$$

Koska *vaihtoehtoinen hypoteesi* on tässä muotoa

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

voimme valita testin *päätössäännöksi* seuraavan säännön: *Hylkää nollahypoteesi*

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

jos

$$t = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} < -z_\alpha$$

Kriittinen arvo $-z_\alpha$ saadaan ehdosta

$$\Pr(Z \le -z_\alpha) = \alpha$$

jossa satunnaismuuttuja Z noudattaa standardoitua normaalijakaumaa:

$$Z \sim \mathrm{N}(0, 1)$$

Testin *päätössääntö* voidaan kirjoittaa myös seuraavaan muotoon: *Hylkää nollahypoteesi*

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

jos

$$\bar{X} < \mu_0 - z_\alpha \sigma / \sqrt{n} = \bar{X}_c$$

jossa kriittinen arvo $-z_\alpha$ on määrätty kuten edellä.

Testin **hyväksymisvirheen todennäköisyys** β on ehdollinen todennäköisyys

$$\begin{aligned}\beta &= \Pr(\mathrm{H}_0 \text{ jätetään voimaan} \,|\, \mathrm{H}_0 \text{ ei ole tosi}) \\ &= \Pr(\bar{X} \geq \bar{X}_c \,|\, \mu = \mu^*) \\ &= \Pr\left(\frac{\bar{X} - \mu^*}{\sigma / \sqrt{n}} \geq \frac{\bar{X}_c - \mu^*}{\sigma / \sqrt{n}} \right) \\ &= \Pr\left(t^* \geq \frac{\bar{X}_c - \mu^*}{\sigma / \sqrt{n}} \right)\end{aligned}$$

jossa

$$t^* = \frac{\bar{X} - \mu^*}{\sigma / \sqrt{n}}$$

ja

$$\mu^* < \mu_0$$

Testin **voimakkuus** $1 - \beta$ on ehdollinen todennäköisyys

$$\begin{aligned}1 - \beta &= \Pr(\mathrm{H}_0 \text{ hylätään} \,|\, \mathrm{H}_0 \text{ ei ole tosi}) \\ &= \Pr(\bar{X} < \bar{X}_c \,|\, \mu = \mu^*) \\ &= \Pr\left(\frac{\bar{X} - \mu^*}{\sigma / \sqrt{n}} < \frac{\bar{X}_c - \mu^*}{\sigma / \sqrt{n}} \right) \\ &= \Pr\left(t^* < \frac{\bar{X}_c - \mu^*}{\sigma / \sqrt{n}} \right)\end{aligned}$$

jossa

$$t^* = \frac{\bar{X} - \mu^*}{\sigma / \sqrt{n}}$$

kuten edellä ja

$$\mu^* < \mu_0$$

Havainnollistus:

Kuvio alla havainnollistaa t-testin *hyväksymisvirheen todennäköisyyttä* β ja *voimakkuutta* $1 - \beta$.

Yleinen hypoteesi H :

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

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

Nollahypoteesi H_0 :

$$\mu = \mu_0$$

Vaihtoehtoinen hypoteesi H_1 :

$\mu < \mu_0$

Merkitsevyystasoa α vastaava *kriittinen arvo* $-z_\alpha$ määrätään siten, että

$$\Pr(Z \le -z_\alpha) = \alpha$$

jossa

$$Z \sim \mathrm{N}(0, 1)$$

Testin *päätössääntö* voidaan ilmaista seuraavassa muodossa:

Hylkää nollahypoteesi H_0, jos

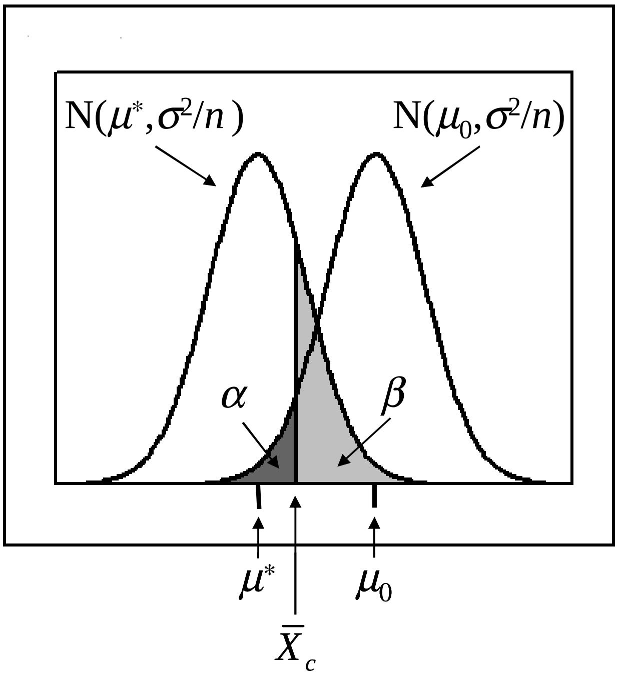
$$\bar{X} < \bar{X}_c = \mu_0 - z_\alpha \sigma / \sqrt{n}$$

Hyväksymisvirheen todennäköisyys β :

$$\beta = \Pr(\bar{X} \ge \bar{X}_c \mid \mu = \mu^*)$$

Voimakkuus $1 - \beta$:

$$1 - \beta = \Pr(\bar{X} < \bar{X}_c \mid \mu = \mu^*)$$



9.3. Kahden riippumattoman otoksen *t*-testi *A* odotusarvoille: Yleinen tapaus

Testausasetelma kahden riippumattoman otoksen *t*-testissä *A*

Olkoon

$$X_{i1}\ ,\ i = 1, 2, \ldots, n_1$$

satunnaisotos normaalijakaumasta $\mathrm{N}(\mu_1, \sigma_1^2)$. Tällöin satunnaismuuttujat X_{i1} , $i = 1, 2, \ldots, n_1$ ovat *riippumattomia* ja noudattavat *samaa normaalijakaumaa* $\mathrm{N}(\mu_1, \sigma_1^2)$:

$$X_{11}, X_{21}, \ldots, X_{n_1 1} \perp$$

$$X_{i1} \sim \mathrm{N}(\mu_1, \sigma_1^2)\ ,\ i = 1, 2, \ldots, n_1$$

Olkoon

$$X_{j2}\ ,\ j = 1, 2, \ldots, n_2$$

satunnaisotos normaalijakaumasta $\mathrm{N}(\mu_2, \sigma_2^2)$. Tällöin satunnaismuuttujat X_{j2} , $j = 1, 2, \ldots, n_2$ ovat *riippumattomia* ja noudattavat *samaa normaalijakaumaa* $\mathrm{N}(\mu_2, \sigma_2^2)$:

$$X_{12}, X_{22}, \ldots, X_{n_2 2} \perp$$

$$X_{j2} \sim \mathrm{N}(\mu_2, \sigma_2^2)\ ,\ j = 1, 2, \ldots, n_2$$

Oletetaan lisäksi, että otokset

$$X_{i1}\ ,\ i = 1, 2, \ldots, n_1$$

ja

$$X_{j2}\ ,\ j = 1, 2, \ldots, n_2$$

ovat *riippumattomia*. Otosten riippumattomuudella tarkoitetaan sitä, että se millainen havainto on saatu jakaumasta 1 ei ole vaikuttanut siihen millainen havainto on saatu jakaumasta 2, ja kääntäen, se millainen havainto on saatu jakaumasta 2 ei ole vaikuttanut siihen millainen havainto on saatu jakaumasta 1.

Huomautus:

- Jakaumien varianssit *saavat* tässä esitettävässä *t*-testissä *A erota toisistaan*. Jos varianssit voidaan olettaa *yhtä suuriksi*, kannattaa käyttää seuraavassa kappaleessa esitettävää *t*-testiä *B*.

Asetetaan normaalijakaumien $\mathrm{N}(\mu_1,\sigma_1^2)$ ja $\mathrm{N}(\mu_2,\sigma_2^2)$ *odotusarvo*- eli *paikkaparametreille* μ_1 ja μ_2 *nollahypoteesi*

$$\mathrm{H}_0 : \mu_1 = \mu_2 = \mu$$

Ongelma: Ovatko havainnot *sopusoinnussa* hypoteesin H_0 kanssa?

Ratkaisu: **Kahden riippumattoman otoksen *t*-testi *A*.**

Hypoteesit kahden riippumattoman otoksen *t*-testissä *A*

Yleinen hypoteesi H :

(i) $X_{i1} \sim \mathrm{N}(\mu_1,\sigma_1^2)\,,\ i = 1, 2, \ldots, n_1$

(ii) $X_{j2} \sim \mathrm{N}(\mu_2,\sigma_2^2)\,,\ j = 1, 2, \ldots, n_2$

(iii) Havainnot X_{i1} ja X_{j2} ovat riippumattomia kaikille i ja j

Nollahypoteesi:

$$\mathrm{H}_0 : \mu_1 = \mu_2 = \mu$$

Vaihtoehtoiset hypoteesit:

$$\left.\begin{array}{l}\mathrm{H}_1 : \mu_1 > \mu_2 \\ \mathrm{H}_1 : \mu_1 < \mu_2\end{array}\right\} \text{ 1-suuntaiset vaihtoehtoiset hypoteesit}$$

$$\mathrm{H}_1 : \mu_1 \neq \mu_2 \qquad \text{2-suuntainen vaihtoehtoinen hypoteesi}$$

Parametrien estimointi kahden riippumattoman otoksen *t*-testissä *A*

Olkoot

$$\bar{X}_k = \frac{1}{n_k}\sum_{i=1}^{n_k} X_{ik}\,,\ k = 1, 2$$

ja

$$s_k^2 = \frac{1}{n_k - 1}\sum_{i=1}^{n_k} (X_{ik} - \bar{X}_k)^2\,,\ k = 1, 2$$

tavanomaiset *harhattomat estimaattorit* normaalijakauman parametreille μ_k ja σ_k^2. Tunnusluku $\bar{X}_k$ on havaintojen X_{ik} , i = 1, 2, … , n_k , k = 1, 2 *aritmeettinen keskiarvo* ja s_k^2 on havaintojen X_{ik} , i = 1, 2, … , n_k , k = 1, 2 *otosvarianssi*.

Testisuure ja sen jakauma kahden riippumattoman otoksen *t*-testissä *A*

Määritellään ***t*-testisuure *A*** kaavalla

$$t_A = \frac{\bar{X}_1 - \bar{X}_1}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

Jos *nollahypoteesi*

$$\mathrm{H}_0 : \mu_1 = \mu_2 = \mu$$

pätee, niin *testisuure* t_A *noudattaa suurissa otoksissa approksimatiivisesti standardoitua normaalijakaumaa* N(0,1):

$$t_A \quad {}_a \; \mathrm{N}(0,1)$$

Perustelu:

Oletetaan, että testin yleinen hypoteesi H ja nollahypoteesi H_0 pätevät:

$$X_{11}, X_{21}, \ldots, X_{n_1 1}, X_{12}, X_{22}, \ldots, X_{n_2 2} \perp$$

$$X_{i1} \quad \mathrm{N}(\mu, \sigma_1^2)\,,\; i = 1, 2, \ldots, n_1$$

$$X_{j2} \quad \mathrm{N}(\mu, \sigma_2^2)\,,\; j = 1, 2, \ldots, n_2$$

Tällöin (ks. lukua **Otokset ja otosjakaumat**)

$$\bar{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{i1} \quad \mathrm{N}\left(\mu, \frac{\sigma_1^2}{n_1}\right)$$

$$\bar{X}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} X_{j2} \quad \mathrm{N}\left(\mu, \frac{\sigma_2^2}{n_2}\right)$$

Koska $\bar{X}_1 \perp \bar{X}_2$, niin

$$\bar{X}_1 - \bar{X}_2 \quad \mathrm{N}\left(0, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

Edellä esitetystä seuraa, että

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \quad \mathrm{N}(0,1)$$

Koska varianssit σ_1^2 ja σ_2^2 ovat *tuntemattomia*, satunnaismuuttuja Z on testisuureena *epäoperationaalinen*. Jos varianssit σ_1^2 ja σ_2^2 korvataan satunnaismuuttujan Z lausekkeessa niiden *harhattomilla estimaattoreilla*

$$s_k^2 = \frac{1}{n_k - 1}\sum_{i=1}^{n_k} (X_{ik} - \bar{X}_k)^2\,,\; k = 1, 2$$

niin saadaan *t-testisuure*

$$t_A = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

joka *nollahypoteesin* H_0 *pätiessä noudattaa suurissa otoksissa approksimatiivisesti standardoitua normaalijakaumaa* N(0,1):

$$t_A \sim_a \mathrm{N}(0,1)$$

Todistus sivuutetaan.

■

Testisuure t_A mittaa otoksien 1 ja 2 aritmeettisten keskiarvojen *tilastollista etäisyyttä*. *Mitta-yksikkönä* on erotuksen

$$\bar{X}_1 - \bar{X}_2$$

standardipoikkeaman

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

estimaattori, jota määrättäessä on oletettu, että nollahypoteesi H_0 pätee.

Testisuureen t_A *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$\mathrm{E}(t_A) = 0$$

Siten itseisarvoltaan *suuret* testisuureen t_A arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*.

Pienissä otoksissa testisuureen t_A jakaumalle saadaan parempi approksimaatio käyttämällä approksimoivana jakaumana *t-jakaumaa* vapausastein (ns. *Satterthwaiten approksimaatio*)

$$\nu = \frac{\left[\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right]^2}{\dfrac{1}{n_1 - 1}\left[\dfrac{s_1^2}{n_1}\right]^2 + \dfrac{1}{n_2 - 1}\left[\dfrac{s_2^2}{n_2}\right]^2}$$

Jos ν *ei ole kokonaisluku*, ν:n arvo pyöristetään *alaspäin* lähimpään kokonaislukuun.

Approksimatiivisena jakaumana käytetään joskus myös *t-jakaumaa* vapausastein

$$\nu = \min\{n_1, n_2\}$$

Tämä approksimaatio *ei ole yhtä hyvä kuin* Satterthwaiten approksimaatio.

Hylkäysalueen määrääminen kahden riippumattoman otoksen *t*-testissä *A*

Valitaan testin merkitsevyystasoksi α.

Käsittelemme tässä kriittisten rajojen määräämistä vain, kun testisuureen t_A jakaumaa approksimoidaan *normaalijakaumalla*. Jos testisuureen t_A approksimoidaan t-jakaumalla, niin kriittisten rajojen määrääminen tapahtuu täsmälleen samalla tavalla kuin yhden otoksen t-testin tapauksessa.

(i) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : \mu_1 > \mu_2$$

niin kriittinen arvo $+z_\alpha$ saadaan ehdosta

$$\Pr(Z \geq +z_\alpha) = \alpha$$

jossa Z $\mathrm{N}(0,1)$. Testin hylkäysalue on tällöin muotoa

$$(+z_\alpha, +\infty)$$

(ii) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : \mu_1 < \mu_2$$

niin kriittinen arvo $-z_\alpha$ saadaan ehdosta

$$\Pr(Z \leq -z_\alpha) = \alpha$$

jossa Z $\mathrm{N}(0,1)$. Testin hylkäysalue on tällöin muotoa

$$(-\infty, -z_\alpha)$$

(iii) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : \mu_1 \neq \mu_2$$

niin kriittiset arvot $-z_{\alpha/2}$ ja $+z_{\alpha/2}$ saadaan ehdoista

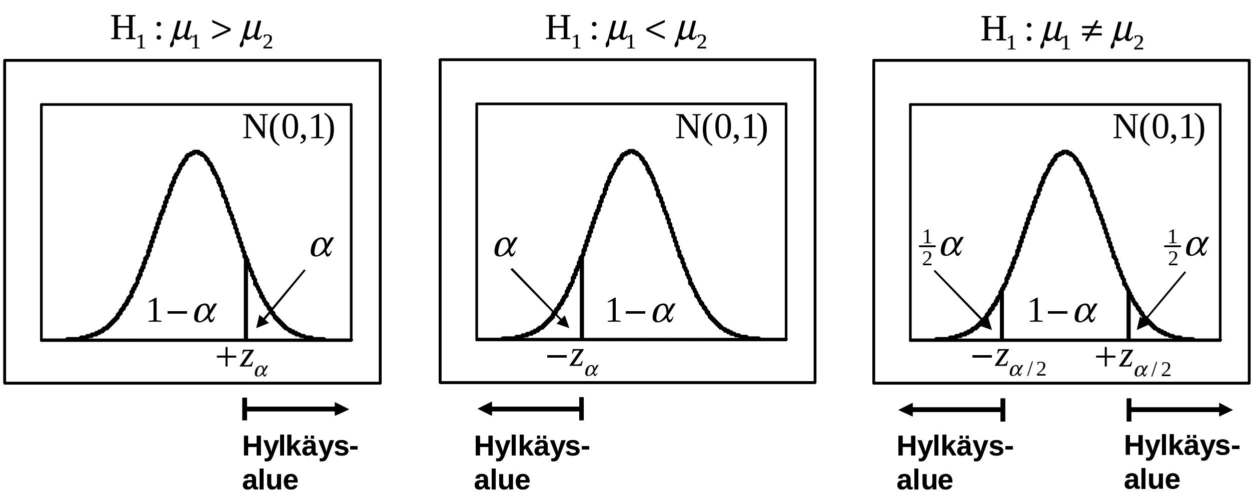
$$\Pr(Z \leq -z_{\alpha/2}) = \alpha/2$$
$$\Pr(Z \geq +z_{\alpha/2}) = \alpha/2$$

jossa Z $\mathrm{N}(0,1)$. Testin hylkäysalue on tällöin muotoa

$$(-\infty, -z_{\alpha/2}) \cup (+z_{\alpha/2}, +\infty)$$

Nollahypoteesi hylätään, jos testisuureen arvo osuu hylkäysalueelle.

Alla olevat kuviot havainnollistavat testin hylkäysalueen valintaa:



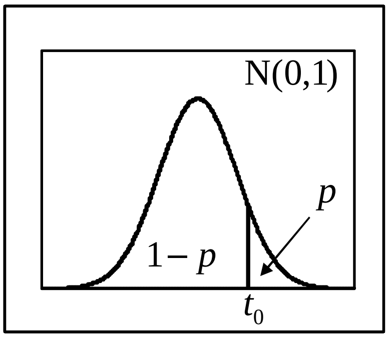
p*-arvon määrääminen kahden riippumattoman otoksen *t*-testissä *A

Olkoon *t-testisuureen* t_A havaittu arvo t_0 .

Käsittelemme tässä testin p-arvon määräämistä vain, kun testisuureen t_A jakaumaa approksimoidaan *normaalijakaumalla*. Jos testisuureen t_A jakaumaa approksimoidaan t-jakaumalla, niin p-arvon määrääminen tapahtuu täsmälleen samalla tavalla kuin yhden otoksen t-testin tapauksessa.

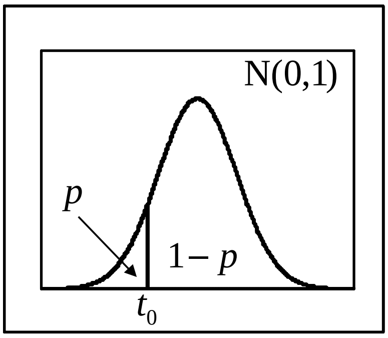
Alla olevat kuviot havainnollistavat testin *p-arvon* määräämistä:

$\mathrm{H}_1 : \mu_1 > \mu_2$



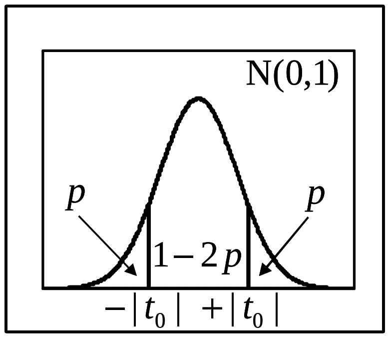
Testin p-arvo = p

$\mathrm{H}_1 : \mu_1 < \mu_2$



Testin p-arvo = p

$\mathrm{H}_1 : \mu_1 \neq \mu_2$



Testin p-arvo = $2p$

Nollahypoteesi *hylätään, jos testin p-arvo on kyllin pieni*.

Normaalisuusoletuksen merkitys kahden riippumattoman otoksen *t*-testissä *A*

Kahden otoksen t-testin yleisen hypoteesin mukaan havainnot ovat molemmissa otoksissa *normaalijakautuneita*. Testi *ei* kuitenkaan *ole herkkä poikkeamille normaalisuudesta*, jos molempien otosten otoskoot ovat ”*kyllin suuria*”.

Testiä on melko turvallista käyttää, kun

$$n_1 > 15 \text{ ja } n_2 > 15$$

ja n_1 ja n_2 eivät eroa toisistaan kovin paljon, *elleivät* havaintojen jakaumat *ole kovin vinoja* ja *ellei* havaintojen joukossa *ole poikkeavia havaintoja*. Jos

$$n_1 > 40 \text{ ja } n_2 > 40$$

testiä voidaan melko turvallisesti käyttää jopa selvästi vinoille havaintojen jakaumille.

9.4. Kahden riippumattoman otoksen *t*-testi *B* odotusarvoille: Yhtä suurten varianssien tapaus

Testausasetelma kahden riippumattoman otoksen *t*-testissä *B*

Oletetaan, että edellisessä kappaleessa esitettyjen oletusten lisäksi hypoteesi

$$\sigma_1^2 = \sigma_2^2 = \sigma^2$$

pätee. Tällöin testausastelmaa voidaan kuvata seuraavalla tavalla:

Olkoon

$$X_{i1}\,,\ i = 1, 2, \ldots, n_1$$

satunnaisotos normaalijakaumasta $\mathrm{N}(\mu_1, \sigma^2)$. Tällöin satunnaismuuttujat X_{i1} , $i = 1, 2, \ldots, n_1$ ovat *riippumattomia* ja noudattavat *samaa normaalijakaumaa* $\mathrm{N}(\mu_1, \sigma^2)$:

$$X_{11}, X_{21}, \ldots, X_{n_1 1} \perp$$
$$X_{i1} \quad \mathrm{N}(\mu_1, \sigma^2),\ i = 1, 2, \ldots, n_1$$

Olkoon

$$X_{j2},\ j = 1, 2, \ldots, n_2$$

satunnaisotos normaalijakaumasta $\mathrm{N}(\mu_2, \sigma^2)$. Tällöin satunnaismuuttujat X_{j2}, $j = 1, 2, \ldots, n_2$ ovat *riippumattomia* ja noudattavat *samaa normaalijakaumaa* $\mathrm{N}(\mu_2, \sigma^2)$:

$$X_{12}, X_{22}, \ldots, X_{n_2 2} \perp$$
$$X_{j2} \quad \mathrm{N}(\mu_2, \sigma^2),\ j = 1, 2, \ldots, n_2$$

Oletetaan lisäksi, että otokset

$$X_{i1},\ i = 1, 2, \ldots, n_1$$

ja

$$X_{j2},\ j = 1, 2, \ldots, n_2$$

ovat *riippumattomia*. Otosten riippumattomuudella tarkoitetaan sitä, että se millainen havainto on saatu jakaumasta 1 ei ole vaikuttanut siihen millainen havainto on saatu jakaumasta 2, ja kääntäen, se millainen havainto on saatu jakaumasta 2 ei ole vaikuttanut siihen millainen havainto on saatu jakaumasta 1.

Huomautus:

- Jakaumien varianssit on oletettu tässä esitettävässä t-testissä B *yhtä suuriksi*. Jos varianssit saavat *erota toisistaan*, on käytettävä edellisessä kappaleessa esitettyä t-testiä A.

Asetetaan normaalijakaumien $\mathrm{N}(\mu_1, \sigma^2)$ ja $\mathrm{N}(\mu_2, \sigma^2)$ *odotusarvo*- eli *paikkaparametreille* μ_1 ja μ_2 *nollahypoteesi*

$$\mathrm{H}_0 : \mu_1 = \mu_2 = \mu$$

Ongelma: Ovatko havainnot *sopusoinnussa* hypoteesin H_0 kanssa?

Ratkaisu: **Kahden riippumattoman otoksen *t*-testi *B*.**

Hypoteesit kahden riippumattoman otoksen *t*-testissä *B*

Yleinen hypoteesi H :

(i) $X_{i1} \sim \mathrm{N}(\mu_1, \sigma^2),\ i = 1, 2, \ldots, n_1$

(ii) $X_{j2} \sim \mathrm{N}(\mu_2, \sigma^2),\ j = 1, 2, \ldots, n_2$

(iii) Havainnot X_{i1} ja X_{j2} ovat riippumattomia kaikille i ja j

Nollahypoteesi:

$$\mathrm{H}_0 : \mu_1 = \mu_2 = \mu$$

Vaihtoehtoiset hypoteesit:

$$\left.\begin{array}{l} \mathrm{H}_1 : \mu_1 > \mu_2 \\ \mathrm{H}_1 : \mu_1 < \mu_2 \end{array}\right\} \text{ 1-suuntaiset vaihtoehtoiset hypoteesit}$$

$$\mathrm{H}_1 : \mu_1 \neq \mu_2 \qquad \text{2-suuntainen vaihtoehtoinen hypoteesi}$$

Parametrien estimointi kahden riippumattoman otoksen *t*-testissä *B*

Olkoot

$$\bar{X}_k = \frac{1}{n_k}\sum_{i=1}^{n_k} X_{ik} \ , \ k = 1, 2$$

ja

$$s_k^2 = \frac{1}{n_k - 1}\sum_{i=1}^{n_k} (X_{ik} - \bar{X}_k)^2 \ , \ k = 1, 2$$

tavanomaiset *harhattomat estimaattorit* normaalijakauman parametreille μ_k ja σ_k^2. Tunnusluku $\bar{X}_k$ on havaintojen X_{ik} , $i = 1, 2, \ldots , n_k$, $k = 1, 2$ *aritmeettinen keskiarvo* ja s_k^2 on havaintojen X_{ik} , $i = 1, 2, \ldots , n_k$, $k = 1, 2$ *otosvarianssi*.

Testisuure ja sen jakauma kahden riippumattoman otoksen *t*-testissä *B*

Määritellään ***t*-testisuure *B*** kaavalla

$$t_B = \frac{\bar{X}_1 - \bar{X}_1}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

jossa

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

on ns. *yhdistetty* (engl. *pooled*) *varianssi*. Jos *nollahypoteesi*

$$\mathrm{H}_0 : \mu_1 = \mu_2 = \mu$$

pätee, niin *testisuure* t_B *noudattaa t-jakaumaa* vapausastein $(n_1 + n_2 - 2)$:

$$t_B \quad t(n_1 + n_2 - 2)$$

Perustelu:

Oletetaan, että testin yleinen hypoteesi H ja nollahypoteesi H_0 pätevät:

$$X_{11}, X_{21}, \ldots , X_{n_1 1}, X_{12}, X_{22}, \ldots , X_{n_2 2} \perp$$

$$X_{i1} \quad \mathrm{N}(\mu, \sigma^2) \ , \ i = 1, 2, \ldots , n_1$$

$$X_{j2} \quad \mathrm{N}(\mu, \sigma^2) \ , \ j = 1, 2, \ldots , n_2$$

Tällöin (ks. lukua **Otokset ja otosjakaumat**)

$$\bar{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{i1} \quad \mathrm{N}\left(\mu, \frac{\sigma^2}{n_1}\right)$$

$$\bar{X}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} X_{j2} \qquad \mathrm{N}\left(\mu, \frac{\sigma^2}{n_2}\right)$$

Koska $\bar{X}_1 \perp \bar{X}_2$, niin

$$\bar{X}_1 - \bar{X}_2 \qquad \mathrm{N}\left(0, \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$$

Edellä esitetystä seuraa, että

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad \mathrm{N}(0,1)$$

Koska standardipoikkeama σ on tuntematon, satunnaismuuttuja Z on testisuureena *epä-operationaalinen*.

Määritellään otosvarianssit

$$s_k^2 = \frac{1}{n_k - 1}\sum_{i=1}^{n_k} (X_{ik} - \bar{X}_k)^2 \ , \ k = 1, 2$$

Jos satunnaismuuttujan Z lausekkeessa standardipoikkeama σ korvataan otossuureella

$$s_P = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

niin saadaan t-testisuure

$$t_B = \frac{\bar{X}_1 - \bar{X}_2}{s_P\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

joka nollahypoteesin H_0 pätiessä noudattaa t-jakaumaa vapausastein $(n_1 + n_2 - 2)$:

$$t_B \sim t(n_1 + n_2 - 2)$$

Todistus sivuutetaan; ks. kuitenkin vastaavaa todistusta yhden otoksen t-testin tapauksessa kappeleessa **Yhden otoksen *t*-testi**.

■

Testisuure t_B mittaa otoksien 1 ja 2 aritmeettisten keskiarvojen *tilastollista etäisyyttä*. *Mitta-yksikkönä* on erotuksen

$$\bar{X}_1 - \bar{X}_2$$

standardipoikkeaman

$$\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

estimaattori, jota määrättäessä on oletettu, että nollahypoteesi H_0 pätee.

Testisuureen t_B *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$E(t_B) = 0$$

Siten itseisarvoltaan *suuret* testisuureen t_B arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*.

Hylkäysalueen määrääminen kahden riippumattoman otoksen *t*-testissä *B*

Kriittisten rajojen määrääminen tapahtuu täsmälleen samalla tavalla kuin yhden otoksen *t*-testin tapauksessa.

p-arvon määrääminen kahden riippumattoman otoksen *t*-testissä *B*

Testin *p*-arvon määrääminen tapahtuu täsmälleen samalla tavalla kuin yhden otoksen *t*-testin tapauksessa.

Normaalisuusoletuksen merkitys kahden riippumattoman otoksen *t*-testissä *B*

Kahden otoksen *t*-testin yleisen hypoteesin mukaan havainnot ovat molemmissa otoksissa *normaali-jakautuneita*. Testi *ei* kuitenkaan *ole herkkä poikkeamille normaalisuudesta*, jos molempien otosten otoskoot ovat ”*kyllin suuria*”.

Testiä on melko turvallista käyttää, kun

$$n_1 > 15 \text{ ja } n_2 > 15$$

ja n_1 ja n_2 eivät eroa toisistaan kovin paljon, *elleivät* havaintojen jakaumat *ole kovin vinoja* ja *ellei* havaintojen joukossa *ole poikkeavia havaintoja*. Jos

$$n_1 > 40 \text{ ja } n_2 > 40$$

testiä voidaan melko turvallisesti käyttää jopa selvästi vinoille havaintojen jakaumille.

9.5. *t*-testi parivertailuille

Parivertailuasetelma

Parivertailuasetelma syntyy tilastollisessa tutkimuksessa esimerkiksi seuraavissa tilanteissa:

(i) Tavoitteena on *verrata kahta mittaria* mittaamalla molemmilla mittareilla *samoja kohteita samoissa olosuhteissa*.

(ii) Tavoitteena on *tutkia* jonkin *käsittelyn vaikutusta* mittaamalla *samoja kohteita ennen käsittelyä ja käsittelyn jälkeen*.

(iii) Päämääränä on *vertailla kahta perusjoukkoa* mittaamalla *saman muuttujan arvoja perusjoukkojen alkioiden sovitetuissa pareissa*.

Kaikissa näissä tilanteissa mittaukset muodostavat *mittauspareja*.

Saattaisi olla houkuttelevaa soveltaa testauasetelmaan *kahden riippumattoman otoksen t-testejä* muodostamalla otokset niin, että toinen otos muodostetaan ensimmäisistä mittauksista pareissa ja toinen otos jälkimmäisistä mittauksista pareissa. **Näin *ei saa* kuitenkaan tehdä, koska parivertailutilanteissa samaan pariin liittyvät mittaukset riippuvat yleensä toisistaan ja *siksi kahden riippumattoman otoksen t-testin soveltaminen saattaa johtaa täysin virheelliseen johtopäätökseen*.**

Testausasetelma *t*-testissä parivertailuille

Oletetaan, että havainnot muodostuvat muuttujaa X koskevista mittaustuloksien *pareista*

$$(X_{i1}, X_{i2}),\ i = 1, 2, \ldots, n$$

jotka ovat toisistaan *riippumattomia*. Tavoitteena on *verrata mittauksia toisiinsa*: Antavatko mittaukset *keskimäärin saman tuloksen*?

Kuten edellä todettiin, **riippumattomien otosten *t*-testiä ei saa tällaisessa tilanteessa käyttää**. Testausongelma saadaan ratkaistuksi siirtymällä tarkastelemaan mittaustuloksien X_{i1} ja X_{i2} *erotuksia*

$$D_i = X_{i1} - X_{i2},\ i = 1, 2, \ldots, n$$

Mittaukset 1 ja 2 antavat *keskimäärin saman tuloksen*, jos erotukset D_i saavat *keskimäärin arvon nolla*. Testausongelman ratkaisuna on tavanomainen *yhden otoksen t-testi mittaustuloksien* X_{i1} ja X_{i2} *erotuksien* D_i *odotusarvolle*.

Hypoteesit *t*-testissä parivertailuille

Yleinen hypoteesi H :

$$D_1, D_2, \ldots, D_n \perp$$
$$D_i \sim \mathrm{N}(\mu_D, \sigma_D^2),\ i = 1, 2, \ldots, n$$

Nollahypoteesi:

$$\mathrm{H}_0 : \mu_D = 0$$

Vaihtoehtoiset hypoteesit:

$$\left.\begin{array}{l}\mathrm{H}_1 : \mu_D > 0 \\ \mathrm{H}_1 : \mu_D < 0\end{array}\right\}\ \text{1-suuntaiset vaihtoehtoiset hypoteesit}$$

$$\mathrm{H}_1 : \mu_D \neq 0 \quad \text{2-suuntainen vaihtoehtoinen hypoteesi}$$

Parametrien estimointi *t*-testissä parivertailuille

Olkoot

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$$

ja

$$s_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2$$

tavanomaiset *harhattomat estimaattorit* normaalijakauman parametreille μ_D ja σ_D^2. Tunnusluku $\bar{D}$ on erotusten $D_i = X_{i1} - X_{i2},\ i = 1, 2, \ldots, n$ *aritmeettinen keskiarvo* ja s_D^2 on erotusten $D_i = X_{i1} - X_{i2},\ i = 1, 2, \ldots, n$ *otosvarianssi*.

Testisuure ja sen jakauma *t*-testissä parivertailuille

Määritellään ***t*-testisuure**

$$t = \frac{\bar{D}}{s_D / \sqrt{n}}$$

Jos *nollahypoteesi*

$$\mathrm{H}_0 : \mu_D = 0$$

pätee, niin *testisuure t noudattaa t-jakaumaa* vapausastein ($n - 1$):

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

Perustelu:

Oletetaan, että testin yleinen hypoteesi H ja nollahypoteesi H_0 pätevät:

$$D_1, D_2, \ldots, D_n \perp$$

$$D_i \sim \mathrm{N}(0, \sigma_D^2)\,,\; i = 1, 2, \ldots, n$$

Tällöin (ks. lukua **Otokset ja otosjakaumat**)

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i \sim \mathrm{N}\left(0, \frac{\sigma_D^2}{n}\right)$$

ja siten

$$Z = \frac{\bar{D}}{\sigma_D / \sqrt{n}} \sim \mathrm{N}(0,1)$$

Koska standardipoikkeama σ_D on *tuntematon*, satunnaismuuttuja Z on testisuureena *epä-operationaalinen*. Jos standardipoikkeama σ_D korvataan satunnaismuuttujan Z lausekkeessa vastaavalla *otossuureella*

$$s_D = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2}$$

niin saadaan *t-testisuure*

$$t = \frac{\bar{D}}{s_D / \sqrt{n}}$$

joka *nollahypoteesin* H_0 *pätiessä noudattaa t-jakaumaa* vapausastein ($n - 1$):

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

Todistus: ks. kappaletta **Yhden otoksen *t*-testi**.

■

Testisuure t mittaa havaintoarvojen erotuksien aritmeettisen keskiarvon *tilastollista etäisyyttä* nollasta. *Mittayksikkönä* on erotuksien D_i aritmeettisen keskiarvon $\bar{D}$ standardipoikkeaman

$$\sigma_D / \sqrt{n}$$

estimaattori, jota määrättäessä on oletettu, että nollahypoteesi H_0 pätee.

Testisuureen t *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$\mathrm{E}(t) = 0$$

Siten itseisarvoltaan *suuret* testisuureen t arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*.

Hylkäysalueen määrääminen *t*-testissä parivertailuille

Kriittisten rajojen määrääminen tapahtuu täsmälleen samalla tavalla kuin yhden otoksen *t*-testin tapauksessa.

p-arvon määrääminen *t*-testissä parivertailuille

Testin *p*-arvon määrääminen tapahtuu täsmälleen samalla tavalla kuin yhden otoksen *t*-testin tapauksessa.

Normaalisuusoletuksen merkitys *t*-testissä parivertailuille

Parivertailuasetelman *t*-testin yleisessä hypoteesissa oletetaan, että havaintoarvojen erotukset ovat *normaali-jakautuneita*. Testi *ei* kuitenkaan *ole herkkä poikkeamille normaalisuudesta*, jos havaintojen lukumäärä n on ”*kyllin suuri*”.

Testiä on melko turvallista käyttää, kun

$$n > 15$$

ellei erotusten jakauma *ole kovin vino* ja erotuksien joukossa *ole poikkeavia erotuksia*. Jos havaintojen lukumäärä

$$n > 40$$

testiä voidaan melko turvallisesti käyttää jopa selvästi vinoille erotuksien jakaumille.

9.6. Yhden otoksen χ^2-testi varianssille

Testausasetelma yhden otoksen testissä varianssille

Olkoon

$$X_i\,,\ i = 1, 2, \dots , n$$

satunnaisotos normaalijakaumasta $\mathrm{N}(\mu,\sigma^2)$. Tällöin satunnaismuuttujat X_i , $i = 1, 2, \dots , n$ ovat *riippumattomia* ja noudattavat *samaa normaalijakaumaa* $\mathrm{N}(\mu,\sigma^2)$:

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

Asetetaan normaalijakauman $\mathrm{N}(\mu,\sigma^2)$ *varianssiparametrille* σ^2 *nollahypoteesi*

$$\mathrm{H}_0 : \sigma^2 = \sigma_0^2$$

Ongelma: Ovatko havainnot *sopusoinnussa* hypoteesin H_0 kanssa?

Ratkaisu: **Yhden otoksen χ^2-testi varianssille**.

Hypoteesit yhden otoksen χ^2-testissä varianssille

Yleinen hypoteesi H :

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

Nollahypoteesi:

$$\mathrm{H}_0 : \sigma^2 = \sigma_0^2$$

Vaihtoehtoiset hypoteesit:

$$\left.\begin{aligned} &\mathrm{H}_1 : \sigma^2 > \sigma_0^2 \\ &\mathrm{H}_1 : \sigma^2 < \sigma_0^2 \end{aligned}\right\} \text{1-suuntaiset vaihtoehtoiset hypoteesit}$$

$$\mathrm{H}_1 : \sigma^2 \neq \sigma_0^2 \quad \text{2-suuntainen vaihtoehtoinen hypoteesi}$$

Parametrien estimointi yhden χ^2-testissä varianssille

Olkoot

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

ja

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

tavanomaiset *harhattomat estimaattorit* normaalijakauman parametreille μ ja σ^2. Tunnusluku $\bar{X}$ on havaintojen X_i , i = 1, 2, … , n *aritmeettinen keskiarvo* ja s^2 on havaintojen X_i , i = 1, 2, … , n *otosvarianssi*.

Testisuure ja sen jakauma yhden otoksen χ^2-testissä varianssille

Määritellään **χ^2-testisuure**

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Jos *nollahypoteesi*

$$\mathrm{H}_0 : \sigma^2 = \sigma_0^2$$

pätee, niin *testisuure* χ^2 *noudattaa* χ^2*-jakaumaa* vapausastein (n – 1):

$$\chi^2 \sim \chi^2(n-1)$$

Perustelu:

Oletetaan, että testin yleinen hypoteesi H ja nollahypoteesi H_0 pätevät:

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim \mathrm{N}(\mu, \sigma_0^2),\ i = 1, 2, \ldots, n$$

Tällöin (ks. lukua **Otokset ja otosjakaumat**)

$$Y = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma_0} \right)^2 \quad \chi^2(n)$$

Koska odotusarvo μ on *tuntematon*, satunnaismuuttuja Y on testisuureena *epä-operationaalinen*. Jos satunnaismuuttujan Y lausekkeessa odotusarvo μ korvataan vastaavalla *otossuureella*

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

niin saadaan χ^2*-testisuure*

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma_0} \right)^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

jossa

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

Jos *nollahypoteesi* H_0 *pätee*, testisuure χ^2 *noudattaa* χ^2*-jakaumaa* vapausastein $(n - 1)$:

$$\chi^2 \sim \chi^2(n - 1)$$

Todistus: ks. lukua **Otokset ja otosjakaumat**.

■

Testisuureen χ^2 *normaaliarvo* = $(n - 1)$, koska *nollahypoteesin* H_0 *pätiessä*

$$E(\chi^2) = n - 1$$

Siten sekä *pienet* että *suuret* testisuureen χ^2 arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*.

Hylkäysalueen määrääminen yhden otoksen χ^2-testissä varianssille

Valitaan testin merkitsevyystasoksi α.

(i) Jos vaihtoehtoinen hypoteesi on muotoa

$$H_1 : \sigma^2 > \sigma_0^2$$

niin kriittinen arvo χ_α^2 saadaan ehdosta

$$\Pr(\chi^2 \geq \chi_\alpha^2) = \alpha$$

jossa $\chi^2 \quad \chi^2(n-1)$. Testin hylkäysalue on tällöin muotoa

$$(\chi_\alpha^2, +\infty)$$

(ii) Jos vaihtoehtoinen hypoteesi on muotoa

$$H_1 : \sigma^2 < \sigma_0^2$$

niin kriittinen arvo $\chi_{1-\alpha}^2$ saadaan ehdosta

$$\Pr(\chi^2 \leq \chi_{1-\alpha}^2) = \alpha$$

jossa $\chi^2 \quad \chi^2(n-1)$. Testin hylkäysalue on tällöin muotoa

$$(0, \chi^2_{1-\alpha})$$

(iii) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : \sigma^2 \neq \sigma_0^2$$

niin kriittiset arvot $\chi^2_{1-\alpha/2}$ ja $\chi^2_{\alpha/2}$ saadaan ehdoista

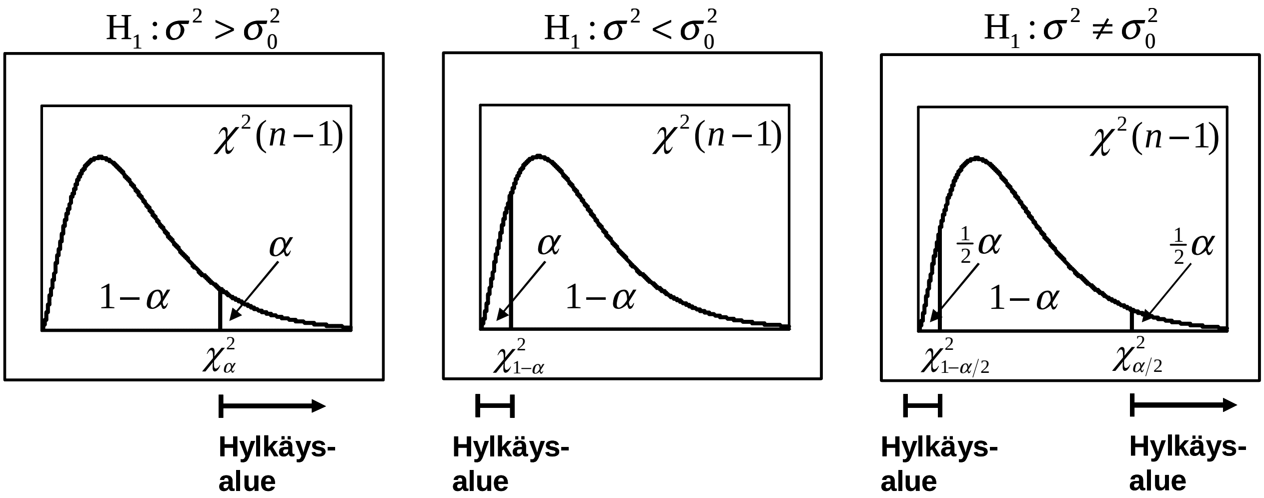
$$\Pr(\chi^2 \leq \chi^2_{1-\alpha/2}) = \alpha/2$$
$$\Pr(\chi^2 \geq \chi^2_{\alpha/2}) = \alpha/2$$

jossa $\chi^2 \quad \chi^2(n-1)$. Testin hylkäysalue on tällöin muotoa

$$(0, \chi^2_{1-\alpha/2}) \cup (\chi^2_{\alpha/2}, +\infty)$$

Nollahypoteesi hylätään, jos testisuureen arvo osuu hylkäysalueelle.

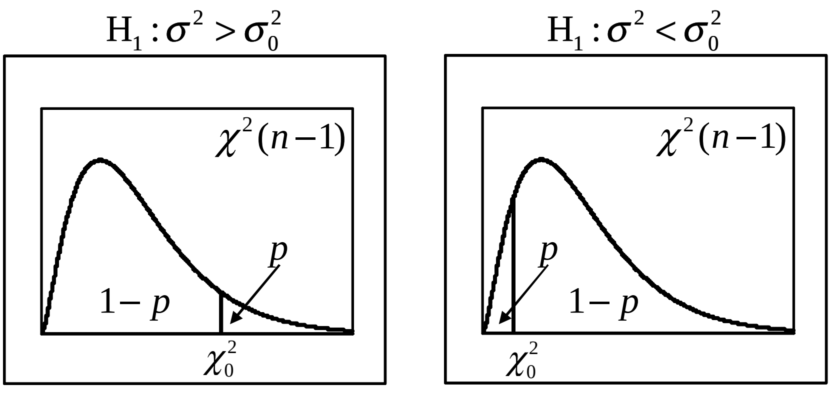
Alla olevat kuviot havainnollistavat testin hylkäysalueen määräämistä:



p-arvon määrääminen yhden otoksen χ^2-testissä varianssille

Olkoon χ^2-*testisuureen* havaittu arvo χ_0^2 .

Alla olevat kuviot havainnollistavat testin *p-arvon* määräämistä, kun vaihtoehtoinen hypoteesi on yksisuuntainen:



Testin *p*-arvo = *p* Testin *p*-arvo = *p*

Kaksisuuntaisen vaihtoehtoisen hypoteesin

$$\mathrm{H}_1 : \sigma^2 \neq \sigma_0^2$$

tapauksessa testin *p*-arvo on

$$p = 2 \times \min\left\{\Pr(\chi^2 \geq \chi_0^2), \Pr(\chi^2 \leq \chi_0^2)\right\}$$

jossa

$$\chi^2 \quad \chi^2(n-1)$$

Nollahypoteesi *hylätään, jos testin p-arvo on kyllin pieni*.

Normaalisuusoletuksen merkitys yhden otoksen χ^2-testissä varianssille

Tässä esitetyn varianssitestin yleisessä hypoteesissa oletetaan, että havainnot ovat *normaalijakautuneita*. Testi *on aika herkkä poikkeamille normaalisuudesta* ja siten *testi ei toimi kovinkaan hyvin*, jos havaintojen jakauma *on vino* tai havaintojen joukossa *on poikkeavia havaintoja*. Suuretkaan havaintojen lukumäärät eivät yleensä paranna tilannetta.

9.7. Kahden riippumattoman otoksen *F*- testi varianseille

Testausasetelma kahden riippumattoman otoksen *F*-testissä varianseille

Olkoon

$$X_{i1}\,,\ i = 1, 2, \ldots, n_1$$

satunnaisotos normaalijakaumasta $\mathrm{N}(\mu_1, \sigma_1^2)$. Tällöin satunnaismuuttujat X_{i1} , $i = 1, 2, \ldots, n_1$ ovat *riippumattomia* ja noudattavat *samaa normaalijakaumaa* $\mathrm{N}(\mu_1, \sigma_1^2)$:

$$X_{11}, X_{21}, \ldots, X_{n_1 1} \perp$$

$$X_{i1} \quad \mathrm{N}(\mu_1, \sigma_1^2)\,,\ i = 1, 2, \ldots, n_1$$

Olkoon

$$X_{j2}\,,\ j = 1, 2, \ldots, n_2$$

satunnaisotos normaalijakaumasta $\mathrm{N}(\mu_2, \sigma_2^2)$. Tällöin satunnaismuuttujat X_{j2} , $j = 1, 2, \ldots, n_2$ ovat *riippumattomia* ja noudattavat *samaa normaalijakaumaa* $\mathrm{N}(\mu_2, \sigma_2^2)$:

$$X_{12}, X_{22}, \ldots, X_{n_2 2} \perp$$

$$X_{j2} \quad \mathrm{N}(\mu_2, \sigma_2^2)\,,\ j = 1, 2, \ldots, n_2$$

Oletetaan lisäksi, että otokset

$$X_{i1}\,,\ i = 1, 2, \ldots, n_1$$

ja

$$X_{j2}\,,\ j = 1, 2, \ldots, n_2$$

ovat *riippumattomia*. Otosten riippumattomuudella tarkoitetaan sitä, että se millainen havainto on saatu jakaumasta 1 ei ole vaikuttanut siihen millainen havainto on saatu jakaumasta 2, ja kääntäen, se millainen havainto on saatu jakaumasta 2 ei ole vaikuttanut siihen millainen havainto on saatu jakaumasta 1.

Asetetaan normaalijakaumien $\mathrm{N}(\mu_1,\sigma_1^2)$ ja $\mathrm{N}(\mu_2,\sigma_2^2)$ *varianssiparametreille* σ_1^2 ja σ_2^2 *nollahypoteesi*

$$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$$

Ongelma: Ovatko havainnot *sopusoinnussa* hypoteesin H_0 kanssa?

Ratkaisu: **Kahden riippumattoman otoksen *F*-testi variansseille**.

Hypoteesit kahden riippumattoman otoksen *F*-testissä variansseille

Yleinen hypoteesi H :

(i) $X_{i1} \sim \mathrm{N}(\mu_1,\sigma_1^2)$, $i = 1, 2, \ldots, n_1$

(ii) $X_{j2} \sim \mathrm{N}(\mu_2,\sigma_2^2)$, $j = 1, 2, \ldots, n_2$

(iii) Havainnot X_{i1} ja X_{j2} ovat riippumattomia kaikille i ja j

Nollahypoteesi:

$$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$$

Vaihtoehtoiset hypoteesit:

$$\left.\begin{array}{l} \mathrm{H}_1 : \sigma_1^2 > \sigma_2^2 \\ \mathrm{H}_1 : \sigma_1^2 < \sigma_2^2 \end{array}\right\} \text{1-suuntaiset vaihtoehtoiset hypoteesit}$$

$$\mathrm{H}_1 : \sigma_1^2 \neq \sigma_2^2 \quad \text{2-suuntainen vaihtoehtoinen hypoteesi}$$

Parametrien estimointi kahden riippumattoman otoksen *F*-testissä variansseille

Olkoot

$$\bar{X}_k = \frac{1}{n_k}\sum_{i=1}^{n_k} X_{ik} \ , \ k = 1, 2$$

ja

$$s_k^2 = \frac{1}{n_k - 1}\sum_{i=1}^{n_k} (X_{ik} - \bar{X}_k)^2 \ , \ k = 1, 2$$

tavanomaiset *harhattomat estimaattorit* normaalijakauman parametreille μ_k ja σ_k^2. Tunnusluku $\bar{X}_k$ on havaintojen X_{ik} , i = 1, 2, … , n_k , k = 1, 2 *aritmeettinen keskiarvo* ja s_k^2 on havaintojen X_{ik} , i = 1, 2, … , n_k , k = 1, 2 *otosvarianssi*.

Testisuure ja sen jakauma kahden riippumattoman otoksen *F*-testissä variansseille

Määritellään ***F*-testisuure**

$$F = \frac{s_1^2}{s_2^2}$$

Jos *nollahypoteesi*

$$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$$

pätee, niin *testisuure F noudattaa F-jakaumaa* vapausastein ($n_1 - 1$) ja ($n_2 - 1$) :

$$F \quad F(n_1 - 1, n_2 - 1)$$

Perustelu:

Oletetaan, että testin yleinen hypoteesi H ja nollahypoteesi H_0 pätevät:

$$X_{11}, X_{21}, \mathrm{K}, X_{n_1 1}, X_{12}, X_{22}, \mathrm{K}, X_{n_2 2} \perp$$

$$X_{i1} \quad \mathrm{N}(\mu_1, \sigma^2)\ ,\ i = 1, 2, \mathrm{K}, n_1$$

$$X_{j2} \quad \mathrm{N}(\mu_2, \sigma^2)\ ,\ j = 1, 2, \mathrm{K}, n_2$$

Tällöin (ks. lukua **Otokset ja otosjakaumat**)

$$Y_1 = \sum_{i=1}^{n_1} \left(\frac{X_{i1} - \mu_1}{\sigma} \right)^2 \quad \chi^2(n_1)$$

$$Y_2 = \sum_{j=1}^{n_2} \left(\frac{X_{j2} - \mu_2}{\sigma} \right)^2 \quad \chi^2(n_2)$$

Koska $Y_1 \perp Y_2$, niin

$$Y = \frac{Y_1 / n_1}{Y_2 / n_2} \quad F(n_1, n_2)$$

Koska odotusarvot μ_1 ja μ_2 ovat *tuntemattomia*, satunnaismuuttuja Y on testisuureena *epä-operationaalinen*.

Korvataan satunnaismuuttujien Y_1 ja Y_2, lausekkeissa odotusarvot μ_1 ja μ_2 vastaavilla *otossuureilla*

$$\bar{X}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} X_{ik}\ ,\ k = 1, 2$$

Saamme satunnaismuuttujat

$$V_1 = \frac{(n_1 - 1)s_1^2}{\sigma^2} = \sum_{i=1}^{n_1} \left(\frac{X_{i1} - \bar{X}_1}{\sigma} \right)^2$$

$$V_2 = \frac{(n_2 - 1)s_2^2}{\sigma^2} = \sum_{j=1}^{n_2} \left(\frac{X_{j2} - \bar{X}_2}{\sigma} \right)^2$$

jossa

$$s_k^2 = \frac{1}{n_k - 1} \sum_{i=1}^{n_k} (X_{ik} - \bar{X}_k)^2\ ,\ k = 1, 2$$

Jos *nollahypoteesi* H_0 *pätee* (ks. lukua **Otokset ja otosjakaumat**)

$$V_1 \quad \chi^2(n_1-1)$$
$$V_2 \quad \chi^2(n_2-1)$$

Lisäksi satunnaismuuttujat V_1 ja V_2 ovat *riippumattomia*:

$$V_1 \perp V_2$$

Määritellään *F-testisuure*

$$F = \frac{V_1/(n_1-1)}{V_2/(n_2-1)} = \frac{s_1^2}{s_2^2}$$

Jos *nollahypoteesi* H_0 *pätee,* testisuure F *noudattaa suoraan F-jakauman määritelmän mukaan F-jakaumaa* vapausastein $(n_1 - 1)$ ja $(n_2 - 1)$:

$$F \quad F(n_1-1, n_2-1)$$

■

Suurille n_2 testisuureen on F *normaaliarvo* ≈ 1, koska *nollahypoteesin* H_0 *pätiessä*

$$\mathrm{E}(F) = \frac{n_2-1}{n_2-3}$$

Siten sekä *pienet* että *suuret* testisuureen F arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*.

Hylkäysalueen määrääminen kahden riippumattoman otoksen *F*-testissä varianssseille

Valitaan testin merkitsevyystasoksi α.

(i) Jos vaihtoehtoinen hypoteesi on muotoa

$$H_1 : \sigma_1^2 > \sigma_2^2$$

niin kriittinen arvo F_α saadaan ehdosta

$$\Pr(F \ge F_\alpha) = \alpha$$

jossa $F \quad F(n_1-1, n_2-1)$. Testin hylkäysalue on tällöin muotoa

$$(F_\alpha, +\infty)$$

(ii) Jos vaihtoehtoinen hypoteesi on muotoa

$$H_1 : \sigma_1^2 < \sigma_2^2$$

niin kriittinen arvo $F_{1-\alpha}$ saadaan ehdosta

$$\Pr(F \le F_{1-\alpha}) = \alpha$$

jossa $F \quad F(n_1-1, n_2-1)$. Testin hylkäysalue on tällöin muotoa

$$(0, F_{1-\alpha})$$

(iii) Jos vaihtoehtoinen hypoteesi on muotoa

$$H_1 : \mu_1 \neq \mu_2$$

niin kriittiset arvot $F_{1-\alpha/2}$ ja $F_{\alpha/2}$ saadaan ehdoista

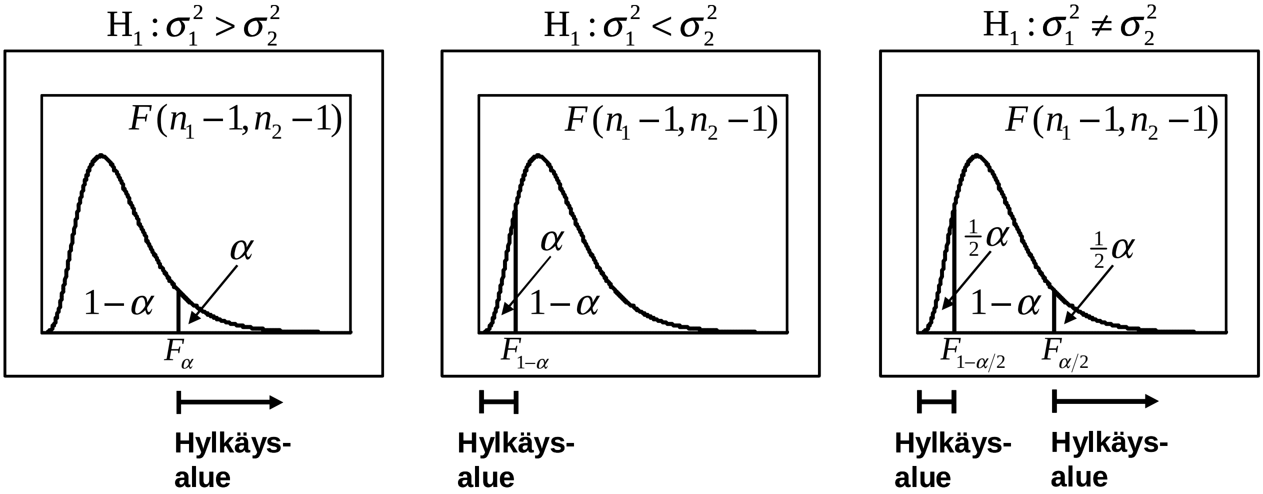
$$\Pr(F \leq F_{1-\alpha/2}) = \alpha/2$$
$$\Pr(F \geq F_{\alpha/2}) = \alpha/2$$

jossa F $F(n_1-1, n_2-1)$. Testin hylkäysalue on tällöin muotoa

$$(0, F_{1-\alpha/2}) \cup (F_{\alpha/2}, +\infty)$$

Nollahypoteesi hylätään, jos testisuureen arvo osuu hylkäysalueelle.

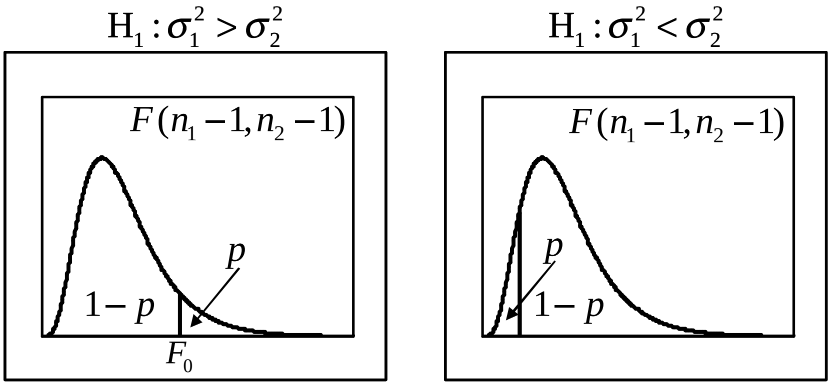
Alla olevat kuviot havainnollistavat testin hylkäysalueen valintaa:



p-arvon määrääminen kahden riippumattoman otoksen *F*-testissä variansseille

Olkoon *F-testisuureen* F havaittu arvo F_0 .

Alla olevat kuviot havainnollistavat testin *p-arvon* määräämistä, kun vaihtoehtoinen hypoteesi on yksisuuntainen:



Testin *p*-arvo = *p*   Testin *p*-arvo = *p*

Kaksisuuntaisen vaihtoehtoisen hypoteesin

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

tapauksessa testin *p*-arvo on

$$p = 2 \times \min\left\{\Pr(F \geq F_0), \Pr(F \leq F_0)\right\}$$

jossa

$$F \quad F(n_1 - 1, n_2 - 1)$$

Nollahypoteesi *hylätään, jos testin p-arvo on kyllin pieni*.

Normaalisuusoletuksen merkitys kahden riippumattoman otoksen *F*-testissä varianсseille

Tässä esitetyn varianssien vertailutestin yleisessä hypoteesissa oletetaan, että havainnot ovat molemmissa otoksissa *normaalijakautuneita*. Testi *on aika herkkä poikkeamille normaalisuudesta* ja siten *testi ei toimi kovinkaan hyvin*, jos havaintojen jakauma *on vino* tai havaintojen joukossa *on poikkeavia havaintoja*. Suuretkaan havaintojen lukumäärät eivät yleensä paranna tilannetta.

10. Testejä järjestysasteikollisille muuttujille

Tarkastelemme tässä luvussa seuraavia **testejä järjestysasteikollisille muuttujille**:

- **Merkkitesti**
- **Wilcoxonin rankitesti**
- **Mannin ja Whitneyn** testi eli **Wilcoxonin rankisummatesti**

Testeissä testataan tarkemmin määrittelemättömän todennäköisyysjakauman **sijaintiparametria** (**mediaania**) koskevia hypoteeseja. Vaikka testit on muotoiltu *järjestyasteikollisille muuttujille*, niitä voidaan käyttää myös *suhde*- tai *välimatka-asteikollisille muuttujille*.

Avainsanat:

Asymptoottinen testi, Eksakti testi, Hylkäysalue, Järjestysasteikko, Kahden otoksen testi, Mannin ja Whitneyn testi, Mediaani, Merkitsevyystaso, Merkkitesti, Nollahypoteesi, Normaaliarvo, Normaalijakauma, Odotusarvo, Otos, Otosvarianssi, *p*-arvo, Parametri, Parivertailu, Riippumattomat otokset, Sovitettu pari, Standardipoikkeama, Standardointi, Suhdeasteikko, *t*-testi, Testisuure, Todennäköisyys-jakauma, Varianssi, Vertailutesti, Välimatka-asteikko, Wilcoxonin rankisummatesti, Wilcoxonin rankitesti, Yhden otoksen testi, Yleinen hypoteesi

10.1. Järjestysasteikollisten muuttujien testit

Tässä luvussa tarkastellaan seuraavia **testejä järjestysasteikollisille muuttujille**:

- **Merkkitesti**
- **Wilcoxonin rankitesti**
- **Mannin ja Whitneyn** testi eli **Wilcoxonin rankisummatesti**

Kaikissa käsiteltävissä testeissä testataan tarkemmin määrittelemättömän todennäköisyysjakauman **sijaintiparametria** (**mediaania**) koskevia hypoteeseja. Vaikka testit on muotoiltu *järjestyasteikollisille muuttujille*, niitä voidaan hyvin käyttää – ja on monissa tilanteissa järkevää käyttää – myös *suhde-* tai *välimatka-asteikollisille muuttujille*; mitta-asteikot: ks. lukua **Tilastollisten aineistojen kerääminen ja mittaaminen**.

Testit ovat luonteeltaan *jakaumista riippumattomia*, millä tarkoitetaan sitä, että testien yleiset hypoteesit eivät määrittele tarkasti perusjoukon jakaumaa. Testejä voidaan kutsua myös *ei-parametrisisiksi*, millä taas tarkoitetaan sitä, että testattavat hypoteesit eivät kohdistu minkään todennäköisyysjakauman parametreihin.

Merkkitesti ja *Wilcoxonin rankitesti* ovat **yhden otoksen testejä**, mutta niitä voidaan soveltaa myös **parivertailuasetelmissa**. *Mannin ja Whitneyn testi* eli *Wilcoxonin rankisummatesti* on **kahden otoksen testi**.

10.2. Merkkitesti

Olkoon

$$X_1, X_2, \dots , X_n$$

satunnaisotos jakaumasta, jonka pistetodennäköisyys- tai tiheysfunktio $f(x)$ on *symmetrinen*. Asetetaan jakauman *mediaanille Me nollahypoteesi*

$$\mathrm{H}_0 : Me = Me_0$$

Ovatko havainnot *sopusoinnussa* nollahypoteesin H_0 kanssa? Tämän testausongelman eräänä ratkaisuna on **merkkitesti**, joka on eräs mahdollinen ei-parametrinen vastine *yhden otoksen t-testille normaalijakauman odotusarvolle*; ks. lukua **Testit suhdeasteikollisille muuttujille**.

Testisuureet S^- ja S^+

Määritellään *erotukset*

$$D_i = X_i - Me_0 \text{ , } i = 1, 2, \dots , n'$$

ja olkoon

$$n \le n'$$

niiden erotusten D_i lukumäärä, jotka ovat $\neq 0$. On ilmeistä, että *nollahypoteesin* H_0 *pätiessä positiivisten ja negatiivisten erotusten on jakauduttava suunnilleen tasan*.

Määritellään **testisuureet** S^- ja S^+ :

$$S^- = \textit{negatiivisten} \text{ erotusten } D_i = X_i - Me_0 \text{ lukumäärä}$$

$$S^+ = \textit{positiivisten} \text{ erotusten } D_i = X_i - Me_0 \text{ lukumäärä}$$

Testisuureiden S^- ja S^+ ominaisuudet

(i) $S^- + S^+ = n$

(ii) Jos *nollahypoteesi* H_0 *pätee*, testisuureet S^- ja S^+ noudattavat *binomijakaumaa* Bin(n,q) parametrein n ja q = 1/2:

$$S^- \sim \mathrm{Bin}(n,1/2)$$

$$S^+ \sim \mathrm{Bin}(n,1/2)$$

(iii) Jos *nollahypoteesi* H_0 *pätee*, niin

$$\mathrm{E}(S^-) = \mathrm{E}(S^+) = nq = \tfrac{1}{2}n$$

(iv) Jos *nollahypoteesi* H_0 *pätee*,

$$\mathrm{D}^2(S^-) = \mathrm{D}^2(S^+) = nq(1-q) = \tfrac{1}{4}n$$

Eksakti merkkitesti

Testisuureiden S^- ja S^+ jakaumat on taulukoitu ja monet tietokoneohjelmat laskevat testin p-arvoja. Merkkitestin *p-arvot* määrätään seuraavilla kaavoilla, joissa s^- ja s^+ ovat testisuureiden S^- ja S^+ havaitut arvot:

(i) *Vaihtoehtoinen hypoteesi*:

$$\mathrm{H}_1 : Me > Me_0$$

Testin *p-arvo*:

$$p = \Pr(S^+ > s^+)$$

(ii) *Vaihtoehtoinen hypoteesi*:

$$\mathrm{H}_1 : Me < Me_0$$

Testin *p-arvo*:

$$p = \Pr(S^- < s^-)$$

(iii) *Vaihtoehtoinen hypoteesi*:

$$\mathrm{H}_1 : Me \neq Me_0$$

Testin *p-arvo*:

$$p = 2\times\min\{\Pr(S^+ > s^+)\ ,\ \Pr(S^- < s^-)\}$$

Standardoitu merkkitestisuure ja sen asymptoottinen jakauma

Jos *nollahypoteesi* H_0 *pätee*, niin

$$\mathrm{E}(S^-) = \mathrm{E}(S^+) = \tfrac{1}{2}n$$

$$\mathrm{D}^2(S^-) = \mathrm{D}^2(S^+) = \tfrac{1}{4}n$$

Määritellään *standardoitu satunnaismuuttuja*

$$Z = \frac{S^* - \mathrm{E}(S^*)}{\mathrm{D}(S^*)}$$

jossa

$$S^* = S^- \text{ tai } S^+$$

Jos *nollahypoteesi* H_0 *pätee*, niin testisuure Z noudattaa suurissa otoksissa approksimatiivisesti standardoitua normaalijakaumaa N(0,1):

$$Z \sim_a \mathrm{N}(0,1)$$

Approksimaatio on tavallisesti *riittävän hyvä*, jos $n > 20$. *Pienissä otoksissa* nojataan testisuureen S^* (= S^- tai S^+) *tarkkaan jakaumaan*.

Testisuureen Z *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$\mathrm{E}(Z) = 0$$

Siten itseisarvoltaan *suuret* testisuureen Z arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*. *Nollahypoteesi* H_0 *hylätään*, jos testin p-arvo on *kyllin pieni*. Hylkäysalueen valinta ja p-arvon määrääminen: ks. lukua **Tilastollinen testaus**.

Kommentteja

Merkkitesti voidaan tulkita *ei-parametriseksi vastineeksi normaalijakauman odotusarvoa koskevalle yhden otoksen t-testille*. Merkkitestissä *ei tehdä* – toisin kuin yhden otoksen t-testissä – mitään *oletuksia perusjoukon jakauman tyypistä*. Merkkitestin testisuureen arvo *ei riipu havaintoarvoista*, vaan ainoastaan niiden keskinäisestä *järjestyksestä*.

Merkkitesti ja parivertailuasetelmat

Merkkitestiä voidaan soveltaa *parivertailuasetelmiin*, joissa havainnot muodostuvat toisistaan *riippumattomista mittauspareista*

$$(X_i\,,\,Y_i)\,,\; i = 1, 2, \ldots, n'$$

Oletetaan, että X- ja Y-mittausten jakaumat ovat muuten samat, mutta niiden *mediaaneilla* (*sijaintiparametreilla*) *saattaa olla eri arvot.*

Määritellään havaintojen X_i ja Y_i *erotukset*

$$D_i = X_i - Y_i\,,\; i = 1, 2, \ldots, n'$$

ja olkoon n niiden erotusten D_i lukumäärä, jotka ovat $\neq 0$ ja oletetaan, että erotusten D_i , $i = 1, 2, \ldots, n$ jakauma on *symmetrinen*.

Määritellään **testisuureet** S^- ja S^+ erotuksille D_i kuten edellä. Olkoon

$$Me_D$$

erotusten D_i , $i = 1, 2, \ldots, n$ *mediaani*. Tällöin *nollahypoteesin*

$$\mathrm{H}_0 : Me_D = 0$$

testaamiseen voidaan soveltaa **merkkitestiä**.

10.3. Wilcoxonin rankitesti

Olkoon

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

satunnaisotos jakaumasta, jonka pistetodennäköisyys- tai tiheysfunktio $f(x)$ on *symmetrinen*. Asetetaan jakauman *mediaanille Me nollahypoteesi*

$$\mathrm{H}_0 : Me = Me_0$$

Ovatko havainnot *sopusoinnussa* nollahypoteesin H_0 kanssa? Tämän testausongelman eräänä ratkaisuna on **Wilcoxonin rankitesti**, joka on eräs mahdollinen ei-parametrinen vastine *yhden otoksen t-testille normaalijakauman odotusarvolle*; ks. lukua **Testit suhdeasteikollisille muuttujille**.

Testisuureet W^- ja W^+

Olkoon

$$|D_i| = |X_i - Me_0|\ ,\ i = 1, 2, \ldots , n'$$

ja olkoon

$$n \leq n'$$

niiden erotusten D_i lukumäärä, jotka ovat $\neq 0$.

Olkoot

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

itseisarvot $|D_i|$ järjestettyinä *suuruusjärjestykseen* pienimmästä suurimpaan ja olkoon

$$R(Z_i) = \text{itseisarvon } Z_i \text{ järjestysnumero eli } \textit{ranki},\ i = 1, 2, \ldots , n$$

Määritellään **testisuure**

$$W^- = \sum_{D_i<0} R(Z_i)$$

on niiden *rankien summa*, joita vastaavat erotukset

$$D_i = X_i - Me_0 < 0$$

Määritellään **testisuure**

$$W^+ = \sum_{D_i>0} R(Z_i)$$

on niiden *rankien summa*, joita vastaavat erotukset

$$D_i = X_i - Me_0 > 0$$

Testisuureiden W^- ja W^+ ominaisuudet

(i) $$W^- + W^+ = \tfrac{1}{2}n(n+1)$$

(ii) Jos *nollahypoteesi* H_0 *pätee*,

$$\mathrm{E}(W^-) = \mathrm{E}(W^+) = \tfrac{1}{4}n(n+1)$$

(iii) Jos *nollahypoteesi* H_0 *pätee*,

$$\mathrm{D}^2(W^-) = \mathrm{D}^2(W^+) = \tfrac{1}{24}n(n+1)(2n+1)$$

Eksakti Wilcoxonin rankitesti

Testisuureiden W^- ja W^+ jakaumat on taulukoitu ja monet tietokoneohjelmat laskevat testin p-arvoja. Wilcoxonin rankitestin *p-arvot* määrätään seuraavilla kaavoilla, joissa w^- ja w^+ ovat testisuureen W^- ja W^+ havaitut arvot:

(i) *Vaihtoehtoinen hypoteesi*:

$$\mathrm{H}_1 : Me > Me_0$$

Testin *p-arvo*:

$$p = \Pr(W^+ > w^+)$$

(ii) *Vaihtoehtoinen hypoteesi*:

$$\mathrm{H}_1 : Me < Me_0$$

Testin *p-arvo*:

$$p = \Pr(W^- < w^-)$$

(iii) *Vaihtoehtoinen hypoteesi*:

$$\mathrm{H}_1 : Me \neq Me_0$$

Testin *p-arvo*:

$$p = 2\times\min\{\Pr(W^+ > w^+)\ ,\ \Pr(W^- < w^-)\}$$

Standardoitu *W*-testisuure ja sen asymptoottinen jakauma

Jos *nollahypoteesi* H_0 *pätee*,

$$\mathrm{E}(W^-) = \mathrm{E}(W^+) = \tfrac{1}{4}n(n+1)$$
$$\mathrm{D}^2(W^-) = \mathrm{D}^2(W^+) = \tfrac{1}{24}n(n+1)(2n+1)$$

Määritellään **testisuure**

$$Z = \frac{W^* - \mathrm{E}(W^*)}{\mathrm{D}(W^*)}$$

jossa

$$W^* = W^- \text{ tai } W^+$$

Jos *nollahypoteesi* H_0 *pätee*, niin testisuure Z noudattaa suurissa otoksissa *approksimatiivisesti standardoitua normaalijakaumaa* N(0,1):

$$Z \sim_a \mathrm{N}(0,1)$$

Approksimaatio on tavallisesti *riittävän hyvä*, jos $n > 20$. *Pienissä otoksissa* nojataan testisuureen W^* (= W^- tai W^+) *tarkkaan jakaumaan*.

Testisuureen z *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$\mathrm{E}(Z) = 0$$

Siten itseisarvoltaan *suuret* testisuureen Z arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*. *Nollahypoteesi* H_0 *hylätään*, jos testin p-arvo on *kyllin pieni*. Hylkäysalueen valinta ja p-arvon määrääminen: ks. lukua **Tilastollinen testaus**.

Kommentteja

Wilcoxonin rankitesti voidaan tulkita *ei-parametriseksi vastineeksi normaalijakauman odotusarvoa koskevalle yhden otoksen t-testille*. Wilcoxonin rankitestissä *ei tehdä* – toisin kuin yhden otoksen *t*-testissä – *mitään oletuksia perusjoukon jakauman tyypistä*. Wilcoxonin rankitestin testisuureen arvo ei riipu *havaintoarvoista*, vaan ainoastaan niiden keskinäisestä *järjestyksestä*. Wilcoxonin rankitesti käyttää kuitenkin *enemmän informaatiota havaintojen järjestyksestä* kuin *merkkitesti*. Siksi Wilcoxonin rankitesti on *voimakkaampi* kuin merkkitesti.

Wilcoxonin rankitesti ja parvertailuasetelmat

Wilcoxonin rankitestiä voidaan soveltaa *parivertailuasetelmiin*, joissa havainnot muodostuvat *toisistaan riippumattomista mittauspareista*

$$(X_i\,,\,Y_i)\,,\; i = 1, 2, \ldots , n'$$

Oletetaan, että *X*- ja *Y*-mittausten jakaumat ovat muuten samat, mutta niiden *mediaaneilla* (*sijaintiparametreilla*) *saattaa olla eri arvot.*

Määritellään havaintojen X_i ja Y_i *erotukset*

$$D_i = X_i - Y_i\,,\; i = 1, 2, \ldots , n'$$

ja olkoon *n* niiden erotusten D_i lukumäärä, jotka ovat $\neq 0$ ja oletetaan, että erotusten D_i , $i = 1, 2, \ldots , n$ jakauma on *symmetrinen*.

Määritellään **testisuureet** W^- ja W^+ erotuksille D_i kuten edellä. Olkoon

$$Me_D$$

erotusten D_i , $i = 1, 2, \ldots , n$ *mediaani*. Tällöin *nollahypoteesin*

$$\mathrm{H}_0 : Me_D = 0$$

testaamiseen voidaan soveltaa **Wilcoxonin rankitestiä**.

10.4. Mannin ja Whitneyn testi

Olkoot

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

riippumattomia havaintoja satunnaismuuttujan *X* jakaumasta perusjoukossa S_1 (otos 1).

Olkoot

$$Y_1, Y_2, \ldots , Y_m$$

ovat *riippumattomia havaintoja* satunnaismuuttujan *Y* jakaumasta perusjoukossa S_2 (otos 2).

Oletetaan lisäksi, että otokset 1 ja 2 ovat *riippumattomia*. Otosten riippumattomuudella tarkoitetaan sitä, että se millainen havainto on saatu jakaumasta 1 ei ole vaikuttanut siihen millainen havainto on saatu jakaumasta 2, ja kääntäen, se millainen havainto on saatu jakaumasta 2 ei ole vaikuttanut siihen millainen havainto on saatu jakaumasta 1.

Tehdään oletus, että satunnaismuuttujat *X* ja *Y* noudattavat muuten *samaa jakaumaa*, mutta niiden *mediaanit* (*sijaintiparametrit*) saattavat *erota toisistaan* ja asetetaan *nollahypoteesi*, että satunnais-muuttujilla *X* ja *Y* on sama *mediaani* (*sijaintiparametri*).

Ovatko havainnot *sopusoinnussa* nollahypoteesin H_0 kanssa? Tämän testausongelman eräänä ratkaisuna on **Mannin ja Whitneyn testi**, joka on eräs mahdollinen ei-parametrinen vastine *kahden riippumattoman otoksen t-testiä normaalijakaumien odotusarvoille*; ks. lukua **Testit suhdeasteikollisille muuttujille**.

Hypoteesit

Yleinen hypoteesi H :

(1) Havainnot

$$X_i \quad F_X \,, i = 1, 2, \ldots, n$$

(2) Havainnot

$$Y_j \quad F_Y \,, j = 1, 2, \ldots, m$$

(3) Jakaumat F_X ja F_Y ovat muuten samat, mutta niiden *mediaanit* (*sijaintiparametrit*) saattavat *erota toisistaan*.

(4) Havainnot X_i ja Y_j ovat *riippumattomia* kaikille i ja j.

Nollahypoteesi H_0 :

$$H_0 : F_X = F_Y$$

Vaihtoehtoinen hypoteesi H_1 :

$$H_1 : F_X \neq F_Y$$

Mannin ja Whitneyn testin idea

Yhdistetään X- ja *Y*-havainnot *yhdeksi otokseksi* ja *järjestetään* yhdistetyn otoksen havainnot *suuruus-järjestykseen* pienimmästä suurimpaan ja tarkastellaan miten *X*- ja *Y*-havainnot *seuraavat* yhdistetyssä otoksessa *toisiaan*.

Jos kaikki *X*-havainnot (*Y*-havainnot) *edeltävät* kaikkia *Y*-havaintoja (*X*-havaintoja), *ei ole uskottavaa*, *että nollahypoteesi* H_0 *pätee*. Jos satunnaismuuttujat X ja Y noudattavat samaa jakaumaa, on ilmeistä, että *X*- ja *Y*-havaintojen *on sekoituttava* sopivasti toisiinsa. Mannin ja Whitneyn testisuure *mittaa* tätä sekoittumista.

Testisuure U_1 – muoto 1

Määritellään satunnaismuuttujat

$$D_{ij}^{(1)} = \begin{cases} 1 \,, \text{ jos } X_i < Y_j \\ 0 \,, \text{ jos } X_i > Y_j \end{cases}$$

$$i = 1, 2, \ldots, n \,;\; j = 1, 2, \ldots, m$$

ja **testisuure**

$$U_1 = \sum_{i=1}^{n} \sum_{j=1}^{m} D_{ij}^{(1)}$$

Testisuure U_1 – muoto 2

Määritellään satunnaismuuttujat

$R(X_i)$ = havainnon X_i järjestysnumero eli *ranki* yhdistetyssä otoksessa

$i = 1, 2, \ldots , n$

ja **testisuure**

$$U_1 = nm + \tfrac{1}{2}n(n+1) - \sum_{i=1}^{n} R(X_i)$$

Testisuureen U_1 ominaisuudet

Testisuureen U_1 muodot 1 ja 2 ovat *ekvivalentteja*. Testisuureen U_1 arvo ei riipu X- ja Y-havaintoarvojen *suuruudesta*, vaan ainoastaan niiden keskinäisestä *järjestyksestä*. Aina pätee

$$0 \le U_1 \le nm$$

ja erityisesti

$U_1 = 0$, jos $X_i > Y_j$ kaikille i ja j

$U_1 = nm$, jos $X_i < Y_j$ kaikille i ja j

Testisuure U_2 – muoto 1

Määritellään satunnaismuuttujat

$$D_{ji}^{(2)} = \begin{cases} 1 \text{ , jos } Y_j < X_i \\ 0 \text{ , jos } Y_j > X_i \end{cases}$$

$$j = 1, 2, \ldots , m \,;\, i = 1, 2, \ldots , n$$

ja **testisuure**

$$U_2 = \sum_{j=1}^{m} \sum_{i=1}^{n} D_{ji}^{(2)}$$

Testisuure U_2 – muoto 2

Määritellään satunnaismuuttujat

$R(Y_j)$ = havainnon Y_j järjestysnumero eli *ranki* yhdistetyssä otoksessa

$j = 1, 2, \ldots , m$

ja **testisuure**

$$U_2 = nm + \tfrac{1}{2}m(m+1) - \sum_{j=1}^{m} R(Y_j)$$

Testisuureen U_2 ominaisuudet

Testisuureen U_2 muodot 1 ja 2 ovat *ekvivalentteja*. Testisuureen U_2 arvo ei riipu X- ja Y-havaintoarvojen *suuruudesta*, vaan ainoastaan niiden keskinäisestä *järjestyksestä*. Aina pätee

$$0 \le U_2 \le nm$$

ja erityisesti

$U_2 = 0$, jos $Y_j > X_i$ kaikille i ja j

$U_2 = nm$, jos $Y_j < X_i$ kaikille i ja j

Testisuureiden U_1 ja U_2 ominaisuudet

(i) $U_1 + U_2 = nm$

(ii) Jos *nollahypoteesi* H_0 *pätee*, niin

$$\mathrm{E}(U_1) = \mathrm{E}(U_2) = \tfrac{1}{2}nm$$

(iii) Jos *nollahypoteesi* H_0 *pätee*, niin

$$\mathrm{D}^2(U_1) = \mathrm{D}^2(U_2) = \tfrac{1}{12}nm(n+m+1)$$

Standardoitu U_1-testisuure ja sen asymptoottinen jakauma

Jos *nollahypoteesi* H_0 *pätee*, niin *standardoitu satunnaismuuttuja*

$$Z_1 = \frac{U_1 - \mathrm{E}(U_1)}{\mathrm{D}(U_1)}$$

noudattaa suurissa otoksissa *approksimatiivisesti standardoitua normaalijakaumaa* N(0,1):

$$Z_1 \sim_a \mathrm{N}(0,1)$$

Approksimaatio on tavallisesti *riittävän hyvä*, jos $n > 10$ ja $m > 10$. *Pienissä otoksissa* nojataan testisuureen U_1 *tarkkaan jakaumaan*.

Testisuureen Z_1 *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$\mathrm{E}(Z_1) = 0$$

Siten itseisarvoltaan *suuret* testisuureen Z_1 arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*. *Nollahypoteesi* H_0 *hylätään*, jos testin p-arvo on *kyllin pieni*. Hylkäysalueen valinta ja p-arvon määrääminen: ks. lukua **Tilastollinen testaus**.

Standardoitu U_2-testisuure ja sen asymptoottinen jakauma

Jos *nollahypoteesi* H_0 *pätee*, niin *standardoitu satunnaismuuttuja*

$$Z_2 = \frac{U_2 - \mathrm{E}(U_2)}{\mathrm{D}(U_2)}$$

noudattaa suurissa otoksissa *approksimatiivisesti standardoitua normaalijakaumaa* N(0,1):

$$Z_2 \sim_a \mathrm{N}(0,1)$$

Approksimaatio on tavallisesti *riittävän hyvä*, jos $n > 10$ ja $m > 10$. *Pienissä otoksissa* nojataan testisuureen U_2 *tarkkaan jakaumaan*.

Testisuureen Z_2 *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$\mathrm{E}(Z_2) = 0$$

Siten itseisarvoltaan *suuret* testisuureen z_2 arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*. *Nollahypoteesi* H_0 *hylätään*, jos testin p-arvo on *kyllin pieni*. Hylkäysalueen valinta ja p-arvon määrääminen: ks. lukua **Tilastollinen testaus**.

Kommentteja

Mannin ja Whitneyn testi voidaan tulkita *ei-parametriseksi vastineeksi kahden riippumattoman otoksen t-testille normaalijakaumien odotusarvoille*. Mannin ja Whitneyn testissä *ei tehdä* – toisin kuin kahden riippumattoman otoksen *t*-testissä – *mitään oletuksia perusjoukkojen jakaumasta*. Mannin ja Whitneyn testisuureiden arvo ei riipu muuttujien X ja Y *arvoista*, vaan ainoastaan niiden keskinäisestä *järjestyksestä*.

Jos havainnot *ovat normaalijakautuneita*, Mannin ja Whitneyn testi *ei ole yhtä voimakas* kuin kahden riippumattoman otoksen *t*-testi. Jos havainnot *eivät ole normaalijakautuneita*, Mannin ja Whitneyn testi *saattaa olla paljon voimakkaampi* kuin kahden riippumattoman otoksen *t*-testi.

Mannin ja Whitneyn testi *on varteenotettava vaihtoehto kahden riippumattoman otoksen t-testille*, jos *otoskoot eivät ole kovin isoja* ja *perusjoukot eivät ole normaalijakautuneita*.

10.5. Wilcoxonin rankisummatesti

Mannin ja Whitneyn testi ja Wilcoxonin rankisummatesti

Wilcoxonin rankisummatesti perustuu Mannin ja Whitneyn testisuureiden U_1 ja U_2 muodoissa 2 esiintyviin havaintojen *rankisummiin* eli *järjestyslukujen summiin*. Wilcoxonin rankisummatesti on *ekvivalentti* Mannin ja Whitneyn testin kanssa.

Testisuure T_1

Määritellään satunnaismuuttujat

$R(X_i)$ = havainnon X_i järjestysnumero eli *ranki* yhdistetyssä otoksessa
$i = 1, 2, \dots , n$

ja **testisuure**

$$T_1 = \sum_{i=1}^{n} R(X_i)$$

Testisuure T_2

Määritellään satunnaismuuttujat

$R(Y_j)$ = havainnon Y_j järjestysnumero eli *ranki* yhdistetyssä otoksessa
$j = 1, 2, \dots , m$

ja **testisuure**

$$T_2 = \sum_{j=1}^{m} R(Y_j)$$

Testisuureiden T_1 ja T_2 ominaisuudet

(i) $$T_1 + T_2 = \tfrac{1}{2}(n+m)(n+m+1)$$

(ii) Jos *nollahypoteesi* H_0 *pätee,*

$$\mathrm{E}(T_1) = \tfrac{1}{2}n(n+m+1)$$
$$\mathrm{E}(T_2) = \tfrac{1}{2}m(n+m+1)$$

(iii) Jos *nollahypoteesi* H_0 *pätee*,

$$D^2(T_1) = D^2(T_2) = \tfrac{1}{12} nm(n+m+1)$$

Standardoitu T_1-testisuure ja sen asymptoottinen jakauma

Jos *nollahypoteesi* H_0 *pätee*, niin standardoitu satunnaismuuttuja

$$Z_1 = \frac{T_1 - E(T_1)}{D(T_1)}$$

noudattaa suurissa otoksissa *approksimatiivisesti standardoitua normaalijakaumaa* N(0,1):

$$Z_1 \sim_a N(0,1)$$

Approksimaatio on tavallisesti *riittävän hyvä*, jos $n > 10$ ja $m > 10$. *Pienissä otoksissa* nojataan testisuureen T_1 *tarkkaan jakaumaan*.

Testisuureen Z_1 *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$E(Z_1) = 0$$

Siten itseisarvoltaan *suuret* testisuureen Z_1 arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*. *Nollahypoteesi* H_0 *hylätään*, jos testin *p*-arvo on *kyllin pieni*. Hylkäysalueen valinta ja *p*-arvon määrääminen: ks. lukua **Tilastollinen testaus**.

Standardoitu T_2-testisuure ja sen asymptoottinen jakauma

Jos *nollahypoteesi* H_0 *pätee*, niin standardoitu satunnais-muuttuja

$$Z_2 = \frac{T_2 - E(T_2)}{D(T_2)}$$

noudattaa suurissa otoksissa *approksimatiivisesti standardoitua normaalijakaumaa* N(0,1):

$$Z_2 \sim_a N(0,1)$$

Approksimaatio on tavallisesti *riittävän hyvä*, jos $n > 10$ ja $m > 10$. *Pienissä otoksissa* nojataan testisuureen T_2 *tarkkaan jakaumaan*.

Testisuureen Z_2 *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$E(Z_2) = 0$$

Siten itseisarvoltaan *suuret* testisuureen Z_2 arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*. *Nollahypoteesi* H_0 *hylätään*, jos testin *p*-arvo on *kyllin pieni*. Hylkäysalueen valinta ja *p*-arvon määrääminen: ks. lukua **Tilastollinen testaus**.

11. Testejä laatueroasteikollisille muuttujille

Tarkastelemme tässä luvussa seuraavia **testejä laatueroasteikollisille muuttujille**:

- **Testi suhteelliselle osuudelle**
- **Suhteellisten osuuksien vertailutesti**

Molemmissa käsiteltävissä testeissä testataan **suhteellista osuutta** tai **Bernoulli-jakauman odotusarvoa** koskevia hypoteeseja. Vaikka testit on muotoiltu *laatueroasteikollisille muuttujille*, niitä voidaan hyvin käyttää myös *suhde*-, *välimatka*- ja *järjestysasteikollisille muuttujille*.

Avainsanat:

Asymptoottinen testi, Bernoulli-jakauma, Hylkäysalue, Järjestyasteikko, Kahden otoksen testi, Kriittinen arvo, Laatueroasteikko, Merkitsevyystaso, Nollahypoteesi, Normaaliarvo, Normaalijakauma, Odotusarvo, Otos, Otoskeskihajonta, Otosvarianssi, *p*-arvo, Parametri, Riippumattomat otokset, Standardi-poikkeama, Suhdeasteikko, *t*-testi, Testisuure, Todennäköisyysjakauma, Varianssi, Vertailutesti, Välimatka-asteikko, Yhden otoksen testi, Yleinen hypoteesi

11.1. Laatueroasteikollisten muuttujien testit

Tässä luvussa tarkastelemme seuraavia **testejä laatueroasteikollisille muuttujille**:

- **Testi suhteelliselle osuudelle**
- **Suhteellisten osuuksien vertailutesti**

Molemmissa käsiteltävissä testeissä testataan **suhteellista osuutta** tai **Bernoulli-jakauman odotusarvoa** koskevia hypoteeseja. Vaikka testit on muotoiltu *laatueroasteikollisille muuttujille*, niitä voidaan hyvin käyttää – ja on monissa tilanteissa järkevää käyttää – myös *suhde*-, *välimatka*- ja *järjestysasteikollisille muuttujille*; mitta-asteikot: ks. lukua **Tilastollisten aineistojen kerääminen ja mittaaminen**.

Testi suhteelliselle osuudelle on **yhden otoksen testi**. *Suhteellisten osuuksien vertailutesti* on **kahden otoksen testi**.

11.2. Testi suhteelliselle osuudelle

Testausasetelma testissä suhteelliselle osuudelle

Olkoon A perusjoukon S *tapahtuma* ja olkoon

$$\Pr(A) = p$$

$$\Pr(A^c) = 1 - p = q$$

Määritellään satunnaismuuttuja

$$X = \begin{cases} 1, & \text{jos tapahtuma } A \text{ sattuu} \\ 0, & \text{jos tapahtuma } A \text{ ei satu} \end{cases}$$

Tällöin satunnaismuuttuja X noudattaa *Bernoulli-jakaumaa* parametrinaan p:

$$X \sim \mathrm{Ber}(p)$$

ja

$$\mathrm{E}(X) = p$$

$$\mathrm{Var}(X) = \mathrm{D}^2(X) = pq$$

Oletetaan, että tapahtuma A on muotoa

A = ”Perusjoukon S alkiolla on ominaisuus P”

Tällöin

$$p = \Pr(A)$$

on todennäköisyys poimia perusjoukosta S satunnaisesti alkio, jolla on ominaisuus P. Jos perusjoukko S on *äärellinen*, niin todennäköisyys p kuvaa niiden perusjoukon S alkioiden *suhteellista osuutta*, joilla on ominaisuus P.

Olkoon $X_1, X_2, \ldots, X_n$ *satunnaisotos* perusjoukosta S, joka noudattaa *Bernoulli-jakaumaa* Bernoulli(p). Tällöin

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \quad \text{Bernoulli}(p),\ i = 1, 2, \ldots, n$$

Asetetaan Bernoulli-jakauman *odotusarvoparametrille p nollahypoteesi*

$$\mathrm{H}_0 : p = p_0$$

Ongelma: Ovatko havainnot *sopusoinnussa* hypoteesin H_0 kanssa?

Ratkaisu: **Testi suhteelliselle osuudelle**.

Hypoteesit testissä suhteelliselle osuudelle

Yleinen hypoteesi H :

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \quad \text{Bernoulli}(p),\ i = 1, 2, \ldots, n$$

Nollahypoteesi:

$$\mathrm{H}_0 : p = p_0$$

Vaihtoehtoiset hypoteesit:

$$\left.\begin{array}{l} \mathrm{H}_1 : p > p_0 \\ \mathrm{H}_1 : p < p_0 \end{array}\right\} \quad \text{1-suuntaiset vaihtoehtoiset hypoteesit}$$

$$\mathrm{H}_1 : p \neq p_0 \qquad \text{2-suuntainen vaihtoehtoinen hypoteesi}$$

Parametrien estimointi testissä suhteelliselle osuudelle

Olkoon

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

tavanomainen *harhaton estimaattori* Bernoulli-jakauman parametrille *p*.

Huomaa, että

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

on tapahtuman *A frekvenssi* siinä *n*-kertaisessa riippumattomien Bernoulli-kokeiden sarjassa, jota yksinkertaisen satunnaisotoksen poiminta Bernoulli-jakaumasta Bernoulli(*p*) merkitsee.

Siten

$$\hat{p} = \frac{f}{n}$$

on tapahtuman *A suhteellinen frekvenssi* ja

$$f = \sum_{i=1}^{n} X_i \quad \text{Bin}(n, p)$$

Testisuure ja sen jakauma testissä suhteelliselle osuudelle

Määritellään *Z-testisuure*

$$Z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$$

Jos *nollahypoteesi*

$$\mathrm{H}_0 : p = p_0$$

pätee, niin *testisuure Z noudattaa suurissa otoksissa approksimatiivisesti standardoitua normaalijakaumaa* N(0,1):

$$Z \sim_a \mathrm{N}(0,1)$$

Perustelu:

Oletetaan, että testin yleinen hypoteesi H ja nollahypoteesi H_0 pätevät:

$$X_1, X_2, \ldots, X_n \perp$$
$$X_i \sim \mathrm{Bernoulli}(p_0),\ i = 1, 2, \ldots, n$$

Tällöin pätee (ks. monisteen **Todennäköisyyslaskenta** lukua **Stokastiikan konvergenssi-käsitteet ja raja-arvolauseet** sekä lukuja **Otokset ja otosjakaumat** ja **Väliestimointi**):

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{f}{n} \sim_a \mathrm{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$$

jolloin

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim_a \mathrm{N}(0,1)$$

■

Approksimaatio on tavallisesti *riittävän hyvä*, jos

$$n\hat{p} \ge 10$$
$$n(1-\hat{p}) \ge 10$$

Testisuure Z mittaa parametrin p estimaatin $\hat{p}$ ja nollahypoteesin kiinnittämän parametrin p arvon p_0 *tilastollista etäisyyttä*. *Mittayksikkönä* on erotuksen

$$\hat{p} - p_0$$

standardipoikkeaman

$$\sqrt{\frac{p(1-p)}{n}}$$

estimaattori, joka on määrätty olettaen, että nollahypoteesi H_0 pätee

Testisuureen Z *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$\mathrm{E}(Z) = 0$$

Siten itseisarvoltaan *suuret* testisuureen Z arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*.

Hylkäysalueen määrääminen testissä suhteelliselle osuudelle

Valitaan testin merkitsevyystasoksi α.

(i) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : p > p_0$$

niin kriittinen arvo $+z_\alpha$ saadaan ehdosta

$$\Pr(Z \geq +z_\alpha) = \alpha$$

jossa Z $\mathrm{N}(0,1)$. Testin hylkäysalue on tällöin muotoa

$$(+z_\alpha, +\infty)$$

(ii) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : p < p_0$$

niin kriittinen arvo $-z_\alpha$ saadaan ehdosta

$$\Pr(Z \leq -z_\alpha) = \alpha$$

jossa Z $\mathrm{N}(0,1)$. Testin hylkäysalue on tällöin muotoa

$$(-\infty, -z_\alpha)$$

(iii) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : p \neq p_0$$

niin kriittiset arvot $-z_{\alpha/2}$ ja $+z_{\alpha/2}$ saadaan ehdoista

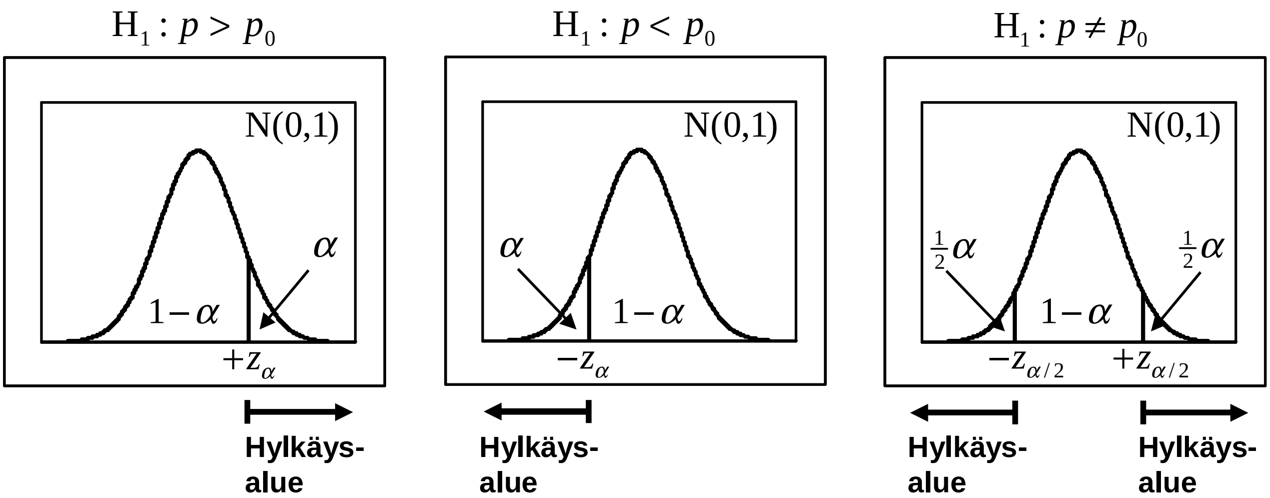
$$\Pr(Z \leq -z_{\alpha/2}) = \alpha/2$$
$$\Pr(Z \geq +z_{\alpha/2}) = \alpha/2$$

jossa Z $\mathrm{N}(0,1)$. Testin hylkäysalue on tällöin muotoa

$$(-\infty, -z_{\alpha/2}) \cup (+z_{\alpha/2}, +\infty)$$

Nollahypoteesi hylätään, jos testisuureen Z arvo osuu hylkäysalueelle.

Alla olevat kuviot havainnollistavat testin hylkäysalueen valintaa:

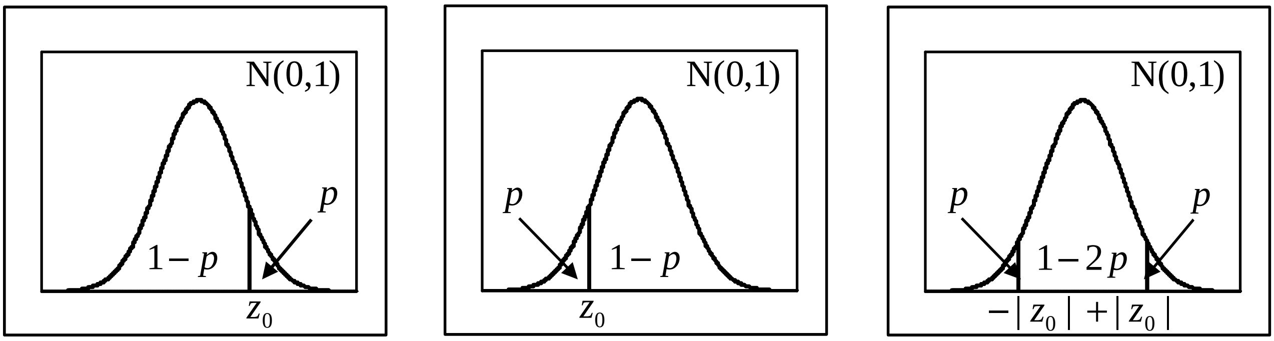


p-arvon määrääminen testissä suhteelliselle osuudelle

Olkoon *Z-testisuureen Z* havaittu arvo z_0 .

Alla olevat kuviot havainnollistavat testin *p-arvon* määräämistä:

$\mathrm{H}_1 : p > p_0$ $\mathrm{H}_1 : p < p_0$ $\mathrm{H}_1 : p \neq p_0$



p-arvo = p p-arvo = p p-arvo = $2p$

Nollahypoteesi *hylätään, jos testin p-arvo on kyllin pieni*.

11.3. Suhteellisten osuuksien vertailutesti

Testausasetelma suhteellisten osuuksien vertailutestissä

Olkoon A perusjoukon S_k , $k = 1, 2$ *tapahtuma* ja olkoot

$$\Pr(A) = p_k$$

$$\Pr(A^c) = 1 - p_k = q_k$$

Määritellään satunnaismuuttujat X_k , $k = 1, 2$:

$$X_k = \begin{cases} 1, \text{ jos } A \text{ tapahtuu perusjoukossa } S_k \\ 0, \text{ jos } A \text{ ei tapahdu perusjoukossa } S_k \end{cases}$$

Tällöin

$$X_k \sim \mathrm{Bernoulli}(p_k) \text{ , } k = 1, 2$$

ja

$$\mathrm{E}(X_k) = p_k$$

$$\mathrm{Var}(X_k) = \mathrm{D}^2(X_k) = p_k q_k$$

Oletetaan, että tapahtuma A on muotoa

A = ”Perusjoukon alkiolla on ominaisuus P”

Tällöin

$$p_k = \Pr(A)$$

on todennäköisyys poimia perusjoukosta S_k , $k = 1, 2$ satunnaisesti alkio, jolla on ominaisuus P. Jos perusjoukko S_k , $k = 1, 2$ on *äärellinen*, niin todennäköisyys p_k kuvaa niiden perusjoukon S_k alkioiden *suhteellista osuutta*, joilla on ominaisuus P.

Olkoon

$$X_{11}, X_{21}, \ldots, X_{n_1 1}$$

satunnaisotos perusjoukosta S_1, joka noudattaa *Bernoulli-jakaumaa* Bernoulli(p_1). Tällöin

$$X_{11}, X_{21}, \ldots, X_{n_1 1} \perp$$
$$X_{i1} \quad \text{Bernoulli}(p_1)\,,\ i = 1, 2, \ldots, n_i$$

Olkoon

$$X_{12}, X_{22}, \ldots, X_{n_2 2}$$

satunnaisotos perusjoukosta S_2, joka noudattaa *Bernoulli-jakaumaa* Bernoulli(p_2). Tällöin

$$X_{12}, X_{22}, \ldots, X_{n_2 2} \perp$$
$$X_{i2} \quad \text{Bernoulli}(p_2)\,,\ i = 1, 2, \ldots, n_2$$

Olkoot otokset lisäksi toisistaan *riippumattomia*.

Asetetaan Bernoulli-jakaumien *parametreille* p_1 ja p_2 *nollahypoteesi*

$$\mathrm{H}_0 : p_1 = p_2 = p$$

Ongelma: Ovatko havainnot *sopusoinnussa* hypoteesin H_0 kanssa?

Ratkaisu: **Suhteellisten osuuksien vertailutesti**.

Hypoteesit suhteellisten osuuksien vertailutestissä

Yleinen hypoteesi H :

(i)
$$X_{11}, X_{21}, \ldots, X_{n_1 1} \perp$$
$$X_{i1} \quad \text{Bernoulli}(p_1)\,,\ i = 1, 2, \ldots, n_i$$

(ii)
$$X_{12}, X_{22}, \ldots, X_{n_2 2} \perp$$
$$X_{i2} \quad \text{Bernoulli}(p_2)\,,\ i = 1, 2, \ldots, n_2$$

(iii) $X_{11}, X_{21}, \ldots, X_{n_1 1}, X_{12}, X_{22}, \ldots, X_{n_2 2} \perp$

Nollahypoteesi:

$$\mathrm{H}_0 : p_1 = p_2 = p$$

Vaihtoehtoiset hypoteesit:

$$\left.\begin{array}{l} \mathrm{H}_1 : p_1 > p_2 \\ \mathrm{H}_1 : p_1 < p_2 \end{array}\right\} \text{ 1-suuntaiset vaihtoehtoiset hypoteesit}$$
$$\mathrm{H}_1 : p_1 \neq p_2 \qquad \text{2-suuntainen vaihtoehtoinen hypoteesi}$$

Parametrien estimointi suhteellisten osuuksien vertailutestissä

Olkoon

$$\hat{p}_k = \frac{1}{n_k}\sum_{i=1}^{n_k} X_{ik} \; , \; k = 1, 2$$

tavanomainen *harhaton estimaattori* Bernoulli-jakauman parametrille p_k , k =1, 2.

Huomaa, että

$$\sum_{i=1}^{n_k} X_{ik} = f_k \; , \; k = 1, 2$$

on tapahtuman *A frekvenssi* siinä *n*-kertaisessa riippumattomien Bernoulli-kokeiden sarjassa, jota yksinkertaisen satunnaisotoksen poiminta Bernoulli-jakaumasta Bernoulli(p_k) , k = 1, 2 merkitsee. Siten

$$\hat{p}_k = \frac{f_k}{n_k} \; , \; k = 1, 2$$

on tapahtuman *A suhteellinen frekvenssi* otoksessa k = 1, 2 ja

$$f_k = \sum_{i=1}^{n_k} X_{ik} \quad \mathrm{Bin}(n_k, p_k)$$

Jos *nollahypoteesi*

$$\mathrm{H}_0 : p_1 = p_2 = p$$

pätee, voidaan otokset *yhdistää* ja parametrin *p harhaton estimaattori* on tapahtuman *A* suhteellinen frekvenssi yhdistetyssä otoksessa:

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{f_1 + f_2}{n_1 + n_2}$$

Jos *nollahypoteesi* H_0 *pätee*, niin

$$\mathrm{Var}(\hat{p}_1 - \hat{p}_2) = \frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2} = p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

Testisuure suhteellisten osuuksien vertailutestissä

Määritellään *Z-testisuure*

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Jos *nollahypoteesi*

$$\mathrm{H}_0 : p_1 = p_2 = p$$

pätee, niin *testisuure Z noudattaa suurissa otoksissa approksimatiivisesti standardoitua normaalijakaumaa*:

$$Z \sim_a \mathrm{N}(0,1)$$

Perustelu:

Oletetaan, että testin yleinen hypoteesi H ja nollahypoteesi H_0 pätevät:

$$X_{11}, X_{21}, \ldots, X_{n_1 1}, X_{12}, X_{22}, \ldots, X_{n_2 2} \perp$$
$$X_{i1} \sim \text{Bernoulli}(p)\,,\ i = 1, 2, \ldots, n_1$$
$$X_{j2} \sim \text{Bernoulli}(p)\,,\ j = 1, 2, \ldots, n_2$$

Tällöin pätee (ks. monisteen **Todennäköisyyslaskenta** lukua **Stokastiikan konvergenssikäsitteet ja raja-arvolauseet** sekä lukuja **Otokset ja otosjakaumat** ja **Väliestimointi**):

$$\hat{p}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{i1} = \frac{f_1}{n_1} \sim_a \mathrm{N}\left(p, \frac{p(1-p)}{n_1}\right)$$
$$\hat{p}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} X_{j2} = \frac{f_2}{n_2} \sim_a \mathrm{N}\left(p, \frac{p(1-p)}{n_2}\right)$$

Koska

$$\hat{p}_1 \perp \hat{p}_2$$

niin

$$Y = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p(1-p)\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} \sim_a \mathrm{N}(0,1)$$

Koska todennäköisyys p on *tuntematon*, satunnaismuuttuja Y on testisuureena *epäoperationaalinen*.

Jos satunnaismuuttujan Y lausekkeessa todennäköisyys p korvataan *otossuureella*

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{f_1 + f_2}{n_1 + n_2}$$

saadaan *testisuure*

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

joka *nollahypoteesin* H_0 *pätiessä noudattaa suurissa otoksissa standardoitua normaalijakaumaa* N(0,1):

$$Z \sim_a \mathrm{N}(0, 1)$$

Todistus sivuutetaan.

■

Approksimaatio on tavallisesti *riittävän hyvä*, jos

$$n_1\hat{p}_1 \geq 5$$
$$n_1(1-\hat{p}_1) \geq 5$$

ja

$$n_2\hat{p}_2 \geq 5$$
$$n_2(1-\hat{p}_2) \geq 5$$

Testisuure Z mittaa tapahtuman A otoksista 1 ja 2 määrättyjen suhteellisten frekvenssien *etäisyyttä*. *Mittayksikkönä* on erotuksen

$$\hat{p}_1 - \hat{p}_2$$

standardipoikkeaman

$$\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

estimaattori, joka on määrätty olettaen, että nollahypoteesi H_0 pätee

Testisuureen Z *normaaliarvo* = 0, koska *nollahypoteesin* H_0 *pätiessä*

$$\mathrm{E}(Z) = 0$$

Siten itseisarvoltaan *suuret* testisuureen z arvot viittaavat siihen, että *nollahypoteesi* H_0 *ei päde*.

Hylkäysalueen määrääminen suhteellisten osuuksien vertailutestissä

Valitaan testin merkitsevyystasoksi α.

(i) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : p_1 > p_2$$

niin kriittinen arvo $+z_\alpha$ saadaan ehdosta

$$\Pr(Z \geq +z_\alpha) = \alpha$$

jossa $Z \quad \mathrm{N}(0,1)$. Testin hylkäysalue on tällöin muotoa

$$(+z_\alpha, +\infty)$$

(ii) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : p_1 < p_2$$

niin kriittinen arvo $-z_\alpha$ saadaan ehdosta

$$\Pr(Z \leq -z_\alpha) = \alpha$$

jossa $Z \quad \mathrm{N}(0,1)$. Testin hylkäysalue on tällöin muotoa

$$(-\infty, -z_\alpha)$$

(iii) Jos vaihtoehtoinen hypoteesi on muotoa

$$\mathrm{H}_1 : p_1 \neq p_2$$

niin kriittiset arvot $-z_{\alpha/2}$ ja $+z_{\alpha/2}$ saadaan ehdoista

$$\Pr(Z \leq -z_{\alpha/2}) = \alpha/2$$
$$\Pr(Z \geq +z_{\alpha/2}) = \alpha/2$$

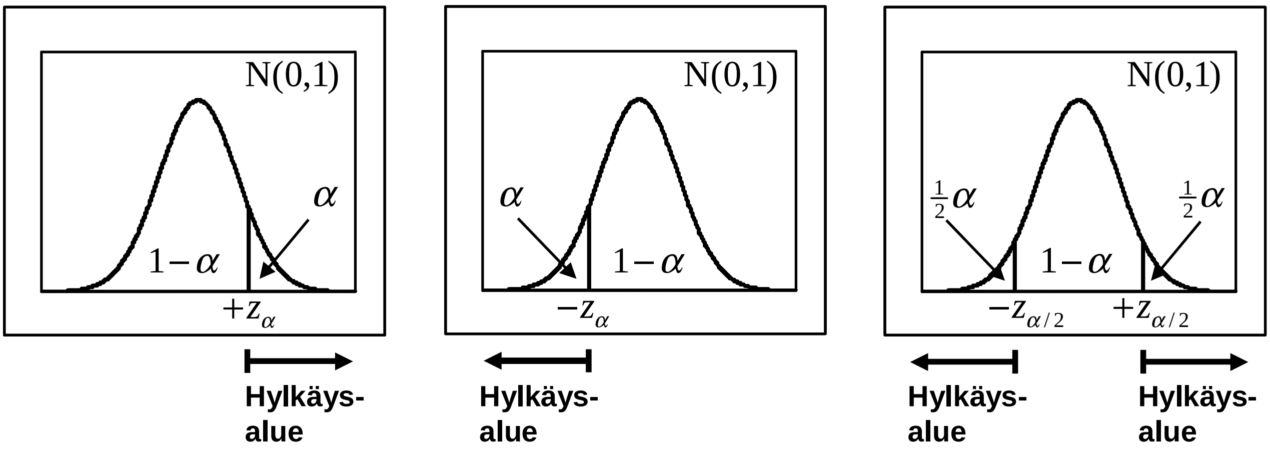
jossa Z $\mathrm{N}(0,1)$. Testin hylkäysalue on tällöin muotoa

$$(-\infty, -z_{\alpha/2}) \cup (+z_{\alpha/2}, +\infty)$$

Nollahypoteesi hylätään, jos testisuureen Z arvo osuu hylkäysalueelle.

Alla olevat kuviot havainnollistavat testin hylkäysalueen valintaa:

$\mathrm{H}_1 : p_1 > p_2$ $\mathrm{H}_1 : p_1 < p_2$ $\mathrm{H}_1 : p_1 \neq p_2$

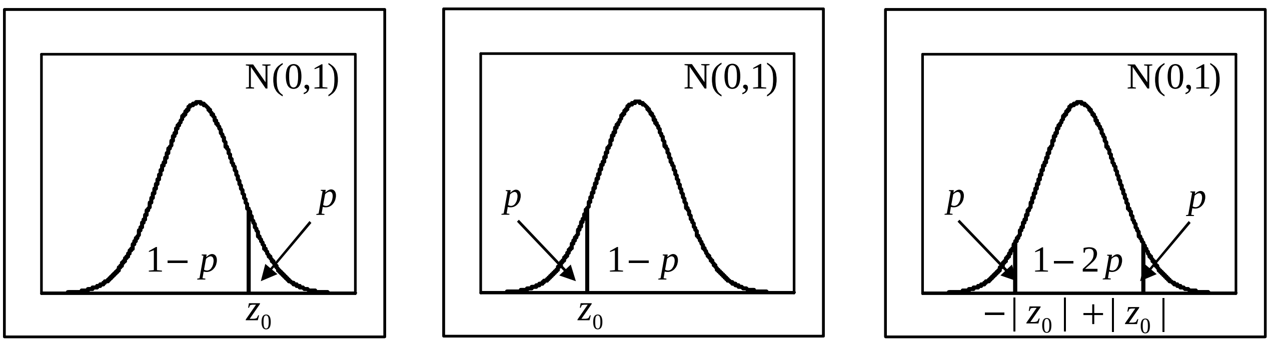


p-arvon määrääminen suhteellisten osuuksien vertailutestissä

Olkoon *Z-testisuureen Z* havaittu arvo z_0 .

Alla olevat kuviot havainnollistavat testin *p-arvon* määräämistä:

$\mathrm{H}_1 : p_1 > p_2$ $\mathrm{H}_1 : p_1 < p_2$ $\mathrm{H}_1 : p_1 \neq p_2$



p-arvo = p p-arvo = p p-arvo = $2p$

Nollahypoteesi *hylätään, jos testin p-arvo on kyllin pieni*.

12. Yhteensopivuuden, homogeenisuuden ja riippumattomuuden testaaminen

Tarkastelemme tässä luvussa seuraavia testejä:

- **χ^2-testi yhteensopivuudelle**
- **χ^2-testi homogeenisuudelle**
- **χ^2-testi riippumattomuudelle**
- **Bowmanin ja Shentonin sekä Wilkin ja Shapiron testit normaalisuudelle**

χ^2-yhteensopivuustestillä pyritään selvittämään *ovatko havainnot sopusoinnussa havaintojen jakaumasta tehdyn* **jakaumaoletuksen** *kanssa*.

Läheistä sukua χ^2-testille yhteensopivuudelle ovat testit *homogeenisuudelle* ja *riippumattomuudelle*. **χ^2-testissä homogeenisuudelle** perusjoukko on jaettu kahteen tai usempaan ryhmään ja testattavana hypoteesina on se, että tarkasteltava muuttuja noudattaa ryhmissä *samaa jakaumaa*. **χ^2-testissä riippumattomuudelle** perusjoukon alkiot voidaan luokitella ristiin *kahden tekijän suhteen* ja testattavana hypoteesina on se, että tekijät ovat *riippumattomia*.

Koska *normaalijakaumalla* on niin keskeinen asema tilastotieteessä, havaintojen **normaalisuudelle** on kehitetty useita erilaisia testejä. Tässä luvussa tarakstellaan **Bowmanin ja Shentonin** ja **Wilkin ja Shapiron testejä**.

Avainsanat:

Bowmanin ja Shentonin testi, Ei-parametrinen testi, Frekvenssitaulukko, Havaittu frekvenssi, Homogeenisuuden testaaminen, Huipukkuus, Hylkäysalue, Htväksymisalue, Jakaumaoletuksen testaaminen, Jakaumista riippumaton testi, χ^2-etäisyys, χ^2-jakauma, χ^2-testi, χ^2-testi homogeenisuudelle, χ^2-testi riippumattomuudelle, χ^2-testi yhteensopivuudelle, Kriittinen arvo, Merkitsevyystaso, Normaalisuuden testaaminen, Odotettu frekvenssi, p-arvo, Rankit Plot -kuvio, Solu, Solufrekvenssi, Suhteellinen frekvenssi, Vinous, Wilkin ja Shapiron testi, Yhteensopivuuden testaaminen

12.1. Jakaumaoletuksien testaus

Tässä luvussa tarkastellaan seuraavia sekä *diskreeteille* että *jatkuville* muuttujille tarkoitettuja testejä **yhteensopivuudelle**, **homogeenisuudelle** ja **riippumattomuudelle**:

- **χ^2-testi yhteensopivuudelle**
- **χ^2-testi homogeenisuudelle**
- **χ^2-testi riippumattomuudelle**
- **Bowmanin ja Shentonin sekä Wilkin ja Shapiron testit normaalisuudelle**

Tilastollisen testauksen perusmuodossa testataan todennäköisyysjakaumien **parametreja** koskevia nollahypoteeseja; ks. lukuja **Tilastollinen testaus** ja **Testejä suhdeasteikollisille muuttujille**. Tällöin testin *yleinen hypoteesi kiinnittää havaintojen jakauman*.

Kysymys: *Voidaanko yleisen hypoteesin jakaumaoletusta testata tilastollisesti*?

Vastaus: *Kyllä*!

Jakaumaoletuksia koskevia tilastollisia testejä kutsutaan tavallisesti **yhteensopivuustesteiksi**. *Yhteensopivuustesteillä* pyritään selvittämään *ovatko havainnot sopusoinnussa tehdyn jakauma-oletuksen kanssa*. *Yleisenä yhteensopivuustestinä* käytetään **χ^2-testiä**.

Läheistä sukua χ^2-testille yhteensopivuudelle ovat testit *homogeenisuudelle* ja *riippumattomuudelle*.

χ^2-testissä homogeenisuudelle testausasetelma on seuraava:

(i) Perusjoukko voidaan jakaa *kahteen* tai *useampaan ryhmään*.

(ii) Testattavana hypoteesina on se, että tarkasteltava muuttuja noudattaa ryhmissä *samaa jakaumaa*.

χ^2-testissä riippumattomuudelle testausasetelma on seuraava:

(i) Perusjoukon alkiot voidaan luokitella ristiin *kahden tekijän suhteen*.

(ii) Testattavana hypoteesina on se, että tekijät ovat *riippumattomia*.

χ^2-testit homogeenisuudelle ja *riippumattomuudelle* ovat erilaisista lähtökohdistaan huolimatta läheistä sukua toisilleen – esimerkiksi niihin liittyvät laskutoimitukset ovat täysin samat.

Koska *normaalijakaumalla* on niin keskeinen asema tilastotieteessä, havaintojen **normaalisuudelle** on kehitetty useita erilaisia testejä. Tässä luvussa tarkastellaan kahta normaalisuustestiä:

(i) **Bowmanin ja Shentonin testi** perustuu havaintojen vinouden ja huipukkuuden mittoihin.

(ii) **Wilkin ja Shapiron testi** perustuu ns. *rankit plot -kuvioon*.

12.2. Yhteensopivuuden testaaminen

Testausasetelma χ^2-yhteensopivuustestissä

Tarkastellaan tutkimusasetelmaa, jossa perusjoukon S alkioita kuvataan *faktorilla* eli *tekijällä* A, joka saa olla *laatuero-*, *järjestys-*, *välimatka-* tai *suhdeasteikollinen muuttuja*.

Oletetaan, että perusjoukosta S on poimittu *yksinkertainen satunnaisotos*. Haluamme *testata jakaumaoletusta*, jonka mukaan tekijä *A noudattaa perusjoukossa S jotakin määrättyä todennäköisyysjakaumaa*.

Testattava oletus voidaan muotoilla seuraavilla tavoilla:

(i) Voidaanko havaintojen jakaumaa *kuvata* oletuksen määrittelemällä todennäköisyysjakaumalla?

(ii) Voiko otos olla oletuksen määrittelemän todennäköisyysjakauman *generoima* eli *tuottama*?

Yhteensopivuustestissä tutkitaan ovatko otos ja tehty jakaumaoletus *yhteensopivia*.

χ^2*-yhteensopivuustestissä* havaintojen ja havaintojen jakaumasta tehdyn oletuksen *yhteensopivuutta mitataan* seuraavalla tavalla:

(1) Valitaan havainnoille sopiva *luokitus*.

(2) Määrätään *havaintojen luokkafrekvenssit*.

(3) Määrätään tehdyn jakaumaoletuksen mukaiset *odotetut luokkafrekvenssit*.

(4) *Verrataan* havaittuja ja odotettuja luokkafrekvenssejä toisiinsa χ^2*-testisuureella*.

Hypoteesit χ^2-yhteensopivuustestissä

Yleinen hypoteesi H :

Havainnot $X_1, X_2, \dots, X_n$ on saatu poimimalla satunnaisotos perusjoukosta S.

Nollahypoteesi H_0 :

Havainnot $X_1, X_2, \dots, X_n$ noudattavat todennäköisyysjakaumaa $f(x;\theta)$, jonka parametrit eivät välttämättä ole tunnettuja.

Vaihtoehtoinen hypoteesi H_1 :

Havainnot $X_1, X_2, \dots, X_n$ eivät noudata nollahypoteesin H_0 määrittelemää todennäköisyysjakaumaa.

Havaitut luokkafrekvenssit

Luokitellaan havainnot $X_1, X_2, \dots, X_n$ *toisensa poissulkeviin luokkiin*, joiden lukumäärä on m ja olkoon

$$O_k \,,\; k = 1, 2, \dots, m$$

niiden havaintojen *frekvenssi* eli *lukumäärä* jotka kuuluvat luokkaan k. Frekvenssi O_k on luokkaan k kuuluvien havaintojen **havaittu frekvenssi**.

Oletetaan, että havainnot $X_1, X_2, \dots, X_n$ ovat *diskreetin* satunnaismuuttujan X havaittuja arvoja ja, että satunnaismuuttujan X *mahdolliset arvot* ovat

$$y_1, y_2, \dots, y_m$$

Tällöin havainto X_i *luokitellaan* luokkaan k, jos

$$X_i = y_k \,,\; i = 1, 2, \dots, n \,,\; k = 1, 2, \dots, m$$

Luokkaan k kuuluvien havaintojen X_i *havaittu frekvenssi* O_k on niiden havaintojen *lukumäärä*, jotka saavat arvon y_k .

Oletetaan, että havainnot $X_1, X_2, \ldots , X_n$ ovat *jatkuvan* satunnaismuuttujan X havaittuja arvoja ja, että

$$X \in (a,b)$$

Jaetaan väli (a,b) pisteillä

$$a = a_0 < a_1 < a_2 < \cdots < a_{m-1} < a_m = b$$

pistevieraisiin *osaväleihin*

$$(a_{k-1},a_k] \text{ , } k = 1, 2, \ldots , m$$

Tällöin havainto X_i *luokitellaan* luokkaan k, jos

$$X_i \in (a_{k-1},a_k] \text{ , } i = 1, 2, \ldots , n \text{ , } k = 1, 2, \ldots , m$$

Luokkaan k kuuluvien havaintojen X_i *havaittu frekvenssi* O_k on niiden havaintojen *lukumäärä,* jotka kuuluvat väliin k.

Havaitut luokkafrekvenssit O_k voidaan esittää *frekvenssitaulukkona* seuraavassa muodossa:

Luokka	1	2	…	m	**Summa**
Havaittu frekvenssi	O_1	O_2	…	O_m	n

Frekvenssiä O_k kutsutaan *havaituksi solufrekvenssiksi frekvenssitaulukon solussa k*. Havaitut solufrekvenssit O_k toteuttavat yhtälön

$$\sum_{k=1}^{m} O_k = n$$

Odotetut luokkafrekvenssit

Oletetaan, että havainnot $X_1, X_2, \ldots , X_n$ ovat satunnaismuuttujan X havaittuja arvoja ja, että havainnot on *luokiteltu* toisensa poissulkeviin luokkiin, joiden lukumäärä on m.

Oletetaan, että *nollahypoteesi* H_0 *määrää täydellisesti satunnaismuuttujan X jakauman* ja olkoon P_k *todennäköisyys* sille, että satunnaismuuttuja X saa arvon luokasta k, kun *nollahypoteesi* H_0 *pätee*. Tällöin luokkaan k kuuluvien havaintojen **odotettu frekvenssi** E_k on

$$E_k = nP_k \text{ , } k = 1, 2, \ldots , m$$

Oletetaan, että *nollahypoteesi* H_0 *määrää satunnaismuuttujan X jakauman tyypin, mutta jakauman parametrit ovat tuntemattomia*. Jos jakauman parametreja ei tunneta, jakauman parametreja on ensin *estimoida* havainnoista. Olkoon P_k *estimoitu todennäköisyys* sille, että satunnaismuuttuja X saa arvon luokasta k, kun *nollahypoteesi* H_0 *pätee*. Tällöin luokkaan k kuuluvien havaintojen **odotettu frekvenssi** E_k on

$$E_k = nP_k \text{ , } k = 1, 2, \ldots , m$$

(i) Oletetaan, että X on *diskreetti* satunnaismuuttuja, jonka mahdolliset arvot ovat

$$y_1, y_2, \ldots , y_m$$

Tällöin

$$P_k = \Pr(X = y_k)\ ,\ k = 1, 2, \ldots , m$$

jossa todennäköisyys $\Pr(X = y_k)$ määrätään olettaen, että *nollahypoteesi* H_0 *pätee*. Todennäköisyydet $\Pr(X = y_k)$ voidaan määrätä satunnaismuuttujan X *kertymäfunktion* tai *pistetodennäköisyysfunktion* avulla.

(ii) Oletetaan, että X on *jatkuva* satunnaismuuttuja, joka saa arvoja väliltä (a,b) ja, että väli (a,b) on jaettu pisteillä

$$a = a_0 < a_1 < a_2 < \cdots < a_{m-1} < a_m = b$$

pistevieraisiin osaväleihin

$$(a_{k-1}, a_k]\ ,\ k = 1, 2, \ldots , m$$

Tällöin

$$P_k = \Pr(a_{k-1} < X \le a_k)\ ,\ k = 1, 2, \ldots , m$$

jossa todennäköisyys $\Pr(a_{k-1} < X \le a_k)$ määrätään olettaen, että *nollahypoteesi* H_0 *pätee*. Todennäköisyydet $\Pr(a_{k-1} < X \le a_k)$ voidaan määrätä satunnaismuuttujan X *kertymäfunktion* tai *tiheysfunktion* avulla.

Odotetut frekvenssit E_k voidaan esittää *frekvenssitaulukkona* seuraavassa muodossa:

Luokka	1	2	…	m	**Summa**
Odotettu frekvenssi	E_1	E_2	…	E_m	n

Frekvenssiä E_k kutsutaan *odotetuksi solufrekvenssiksi frekvenssitaulukon solussa k*. Odotetut solufrekvenssit E_k toteuttavat yhtälön

$$\sum_{k=1}^{m} E_k = n$$

Testisuure ja sen jakauma χ^2-yhteensopivuustestissä

Testi nollahypoteesille H_0 *perustuu* havaittujen frekvenssien O_k ja odotettujen frekvenssien E_k *vertaamiseen*. *Jos* havaittujen frekvenssien O_k ja odotettujen frekvenssien E_k *jakaumat muistuttavat toisiaan*, havainnot *ovat sopusoinnussa* nollahypoteesin H_0 kanssa.

Määritellään **χ^2-testisuure**

$$\chi^2 = \sum_{k=1}^{m} \frac{(O_k - E_k)^2}{E_k}$$

jossa

O_k = *havaittu frekvenssi* luokassa k

E_k = *odotettu frekvenssi* luokassa k

m = luokkien lukumäärä

Testisuure χ^2 *mittaa* havaittujen ja odotettujen frekvenssien *jakaumien yhteensopivuutta* tai *etäisyyttä* ja siksi sitä kutsutaan usein χ^2*-etäisyydeksi*.

χ^2-testisuure voidaan kirjoittaa myös muotoon

$$\chi^2 = n\sum_{k=1}^{m}\frac{(\hat{p}_k - P_k)^2}{P_k}$$

jossa

$\hat{p}_k$ = O_k/n = *havaittu suhteellinen frekvenssi* luokassa k

P_k = *todennäköisyys*, että havainto kuuluu luokkaan k,
kun nollahypoteesi H_0 *pätee*

m = luokkien lukumäärä

Jos nollahypoteesi H_0 *pätee*, testisuure χ^2 noudattaa *suurissa otoksissa approksimatiivisesti* χ^2*-jakaumaa vapausastein* $f = (m - 1 - p)$:

$$\chi^2 \quad {}_{a}\,\chi^2(f)$$

jossa

f = $m - 1 - p$

m = luokkien lukumäärä

p = odotettujen frekvenssien E_k määräämiseksi estimoitujen parametrien lukumäärä

Approksimaatio on tavallisesti *riittävän hyvä*, jos *odotetut frekvenssit* E_k toteuttavat ehdot

$$E_k > 5\,,\ k = 1, 2, \ldots, m$$

Ehdot saadaan toteutumaan valitsemalla havainnoille *sopiva luokitus*.

Testisuureen χ^2 *normaaliarvo* eli *odotusarvo nollahypoteesin* H_0 *pätiessä* on

$$\mathrm{E}(\chi^2) = f$$

jossa

$$f = m - 1 - p$$

Normaaliarvoaan *merkitsevästi suuremmat* χ^2-testisuureen arvot viittaavat siihen, että nollahypoteesi H_0 *ei päde*. Normaaliarvoaan *merkitsevästi pienemmät* χ^2-testisuureen arvot viittaavat siihen, että nollahypoteesi H_0 *pätee liian hyvin*: Havainnot saattavat olla *väärennettyjä*.

χ^2*-yhteensopivuustesti* on **jakaumista riippumaton**, **ei-parametrinen testi**:

(i) Testin yleinen hypoteesi *ei kiinnitä havaintojen jakaumaa*, joten se soveltuu *kaikille* todennäköisyysjakaumille.

(ii) Testissä *ei testata todennäköisyysjakauman parametreja koskevaa hypoteesia*, vaan oletusta havaintojen jakaumasta.

χ^2-yhteensopivuustesti: Sovellus

Luvuissa **Tilastollisten aineistojen kuvaaminen**, **Väliestimointi** ja **Tilastollinen testaus** on käsitelty seuraavaa esimerkkiä :

Kone tekee *ruuveja*, joiden *tavoitepituutena* on 10 cm. *Ruuvien pituus saa vaihdella satunnaisesti* jonkin verran, *kunhan valmistettujen ruuvien keskimääräinen pituus on mahdollisimman lähellä tavoitearvoaan*. Ruuvien laadunvalvonnassa niiden *keskimääräistä pituutta* tutkitaan siten, että jokaisesta valmistetusta ruuvierästä poimitaan *satunnaisotos* ja otokseen poimittujen ruuvien pituudet mitataan.

Olemme aikaisemmin soveltaneet näin kerättyyn esimerkkiaineistoon seuraavia tilastollisia menetelmiä:

(i) Luvussa **Tilastollisten aineistojen kuvaaminen** näytettiin, miten ruuvien *pituuden jakaumaa otoksessa voidaan kuvata luokitellulla frekvenssijakaumalla* ja sitä vastaavalla *histogrammilla*.

(ii) Luvussa **Väliestimointi** näytettiin, miten koneen tekemien ruuvien *keskipituudelle voidaan konstruoida* otostietojen perusteella *luottamusväli*.

(iii) Luvussa **Tilastollinen testaus** näytettiin, miten voidaan *testata ovatko otoksessa saadut tiedot* ruuvien *keskipituudesta sopusoinnussa* ruuvien *tavoitepituuden kanssa*.

Sekä ruuvien keskipituuden *luottamusväli* (ks. lukua **Väliestimointi**) että *testi* ruuvien keskipituuden tavoitearvolle (ks. lukua **Tilastollinen testaus**) perustuivat oletukseen, jonka mukaan *ruuvien pituus vaihtelee normaalijakauman mukaan*.

Tätä **jakaumaoletusta** *voidaan testata* **χ^2-yhteensopivuustestillä** seuraavassa esitettävällä tavalla.

Yleinen hypoteesi H :

Ruuvit on poimittu *satunnaisotannalla* koneen tekemien ruuvien joukosta.

Nollahypoteesi H_0 :

Koneen tekemien ruuvien pituudet *noudattavat normaalijakaumaa*.

Vaihtoehtoinen hypoteesi H_1 :

Koneen tekemien ruuvien pituudet *eivät noudata normaalijakaumaa*.

Koneen valmistamien ruuvien joukosta poimittiin siis *satunnaisotos*, jonka *koko*

$$n = 30$$

ja otokseen poimittujen ruuvien pituudet mitattiin. Ruuvien pituuksien *aritmeettinen keskiarvo* otoksessa oli

$$\bar{X} = 10.09 \text{ cm}$$

ja *otoskeskihajonta* oli

$$s = 0.1038 \text{ cm}$$

Taulukko oikealla esittää pituuksien *luokiteltua frekvenssijakaumaa*.

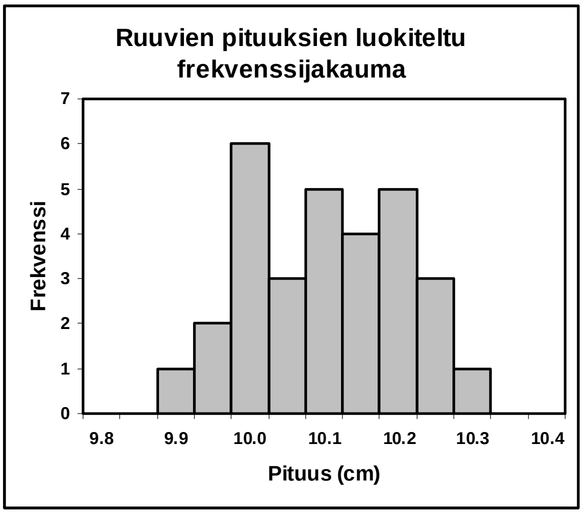
Luokkavälit	Luokkafrekvenssit
(9.85,9.90]	1
(9.90,9.95]	2
(9.95,10.00]	6
(10.00,10.05]	3
(10.05,10.10]	5
(10.10,10.15]	4
(10.15,10.20]	5
(10.20,10.25]	3
(10.25,10.30]	1

Kuva oikealla esittää otokseen poimittujen ruuvien pituuksien luokiteltua frekvenssijakaumaa vastaavaa *histogrammia*.

Voisiko tällainen pituuksien jakauma syntyä *normaalijakautuneesta perusjoukosta* poimitusta satunnaisotoksesta, ts. *ovatko havainnot sopusoinnussa* niistä tehdyn *normaalisuusoletuksen kanssa*?

Tähän kysymykseen antaa vastauksen **χ^2-yhteensopivuustesti**.

Tarkastelemme alla testin tekemistä esimerkin havaintoaineiston tapauksessa.



χ^2-yhteensopivuustestin vaatimat laskutoimitukset voidaan järjestää seuraavan taulukon muotoon (taulukko on luotu *Microsoft Excel* -ohjelmalla):

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Luokka	Luokan yläraja	Havaittu luokka-frekvenssi	Standardoitu luokan yläraja	Kertymä-funktion arvo	Kertymä-funktion arvojen erotus	Odotettu luokka-frekvenssi	Khi^2-arvo
k	a_k	O_k	z_k	$F(z_k)$	$F(z_k)-F(z_{k-1})$	E_k	χ_k^2
1	9.90	1	-1.79444	0.03637	0.03637	1.09115	0.00762
2	9.95	2	-1.31292	0.09460	0.05823	1.74698	0.03665
3	10.00	6	-0.83141	0.20287	0.10827	3.24799	2.33177
4	10.05	3	-0.34990	0.36321	0.16034	4.81010	0.68116
5	10.10	5	0.13161	0.55235	0.18915	5.67443	0.08016
6	10.15	4	0.61313	0.73010	0.17775	5.33245	0.33295
7	10.20	5	1.09464	0.86316	0.13306	3.99177	0.25466
8	10.25	3	1.57615	0.94250	0.07934	2.38026	0.16136
9	10.30	1	2.05766	0.98019	0.05750	1.72487	0.30462
Summa	*	30	*	*	1	30	4.190937

Taulukon *sarakkeiden* selitykset:

(1) *Luokka*: $k = 1, 2, \dots , m = 9$

(2) Luokan k *yläraja*: a_k

(3) *Havaittu luokkafrekvenssi* luokassa k : O_k

$$\sum_{k=1}^{m} O_k = n = 30$$

(4) Luokan k yläraja a_k *standardoituna*:

$$z_k = \frac{a_k - \bar{X}}{s} = \frac{a_k - 10.09}{0.1038}$$

jossa

$$\bar{X} = 10.09 \text{ cm}$$

on ruuvien pituuksien aritmeettinen keskiarvo otoksessa ja

$$s = 0.1038 \text{ cm}$$

on pituuksien keskihajonta otoksessa.

(5) Standardoidun normaalijakauman N(0,1) *kertymäfunktion* $F(\cdot)$ *arvo* pisteessä z_k : $F(z_k)$

(6) Kertymäfunktion arvojen erotus

$$F(z_k) - F(z_{k-1}) = \Pr(z_{k-1} < z \le z_k) = P_k$$

on *todennäköisyys*, että havainto kuuluu luokkaan k, *jos nollahypoteesi* H_0 *ruuvien pituuden normaalijakautuneisuudesta pätee*. Todennäköisyydet P_k toteuttavat ehdon

$$\sum_{k=1}^{m} P_k = 1$$

koska valitun luokituksen ulkopuolelle jääneet normaalijakauman häntäalueiden todennäköisyysmassat on yhdistetty reunaluokkiin.

(7) *Odotettu luokkafrekvenssi* luokassa k : $E_k = nP_k$

$$\sum_{k=1}^{m} E_k = n = 30$$

(8) Luokan k χ^2*-arvo*:

$$\chi_k^2 = \frac{(O_k - E_k)^2}{E_k}$$

jossa

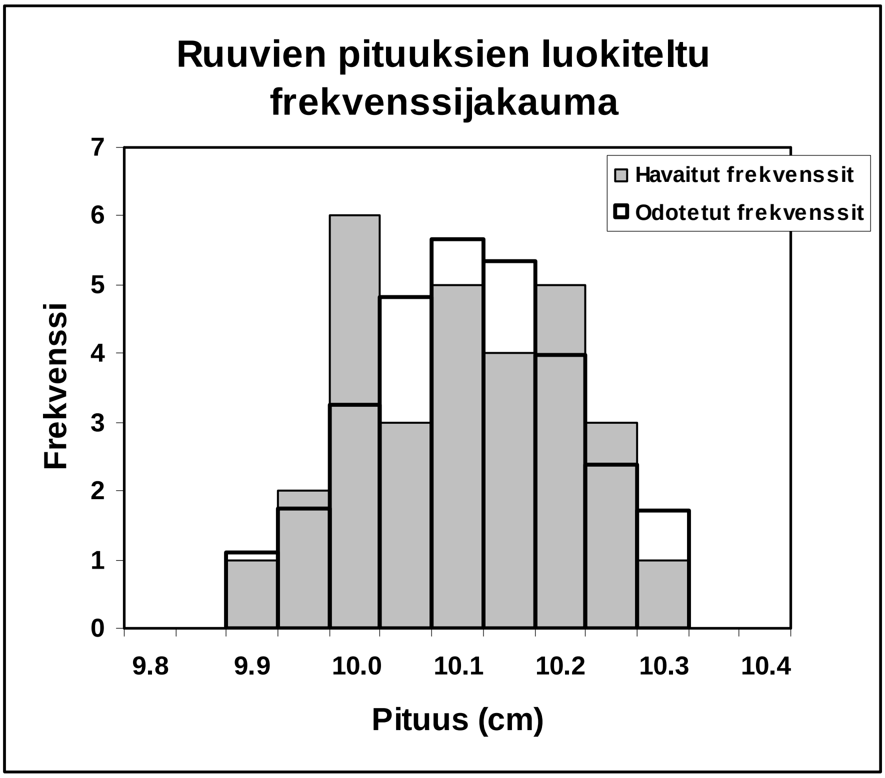
O_k = havaittu frekvenssi luokassa k

E_k = odotettu frekvenssi luokassa k

χ^2-yhteensopivuustestin *testisuureen arvo* saadaan sarakkeen (8) lukujen sarakesummana:

$$\chi^2 = \sum_{k=1}^{m} \chi_k^2 = \sum_{k=1}^{m} \frac{(O_k - E_k)^2}{E_k} = 4.20$$

χ^2-yhteensopivuustesti *vertaa havaittuja frekvenssejä* O_k ja *nollahypoteesin mukaan odotettuja frekvenssejä* E_k *toisiinsa*. *Geometrisesti* vertailu merkitsee havaittuja luokkafrekvenssejä O_k vastaavien suorakaiteiden *pintaalojen vertaamista* odotettuja frekvenssejä E_k vastaavien suorakaiteiden *pinta-aloihin*; ks. kuvaa alla.



Nollahypoteesin

H_0 : Ruuvien pituudet noudattavat *normaalijakaumaa*

pätiessä testisuure χ^2 *noudattaa* χ^2**-jakaumaa** vapausastein $(m - 1 - p)$:

$$\chi^2 \sim \chi^2(m-1-p)$$

jossa

m = luokkien lukumäärä

p = odotettujen frekvenssien E_k määräämiseksi estimoitujen parametrien lukumäärä

Esimerkissä

$$m - 1 - p = 9 - 1 - 2 = 6$$

koska.

Valitaan **merkitsevyystasoksi**

$$\alpha = 0.05$$

Merkitsevyystasoa $\alpha = 0.05$ vastaava **kriittinen arvo** on

$$\chi^2_{0.05} = 12.592$$

koska χ^2-jakauman taulukoiden mukaan

$$\Pr(\chi^2 \geq 12.592) = 0.05$$

jossa

$$\chi^2 \quad \chi^2(6)$$

Merkitsevyystasoa $\alpha = 0.05$ vastaava **hylkäysalue** on siten muotoa

$$(12.592, +\infty)$$

Koska χ^2-yhteensopivuustestin testisuureen arvo

$$\chi^2 = 4.20 < 12.592$$

kuuluu siis testin **hyväksymisalueeseen** ja siten *nollahypoteesi jää voimaan* merkitsevyystasolla α = 0.05:

Havainnot ovat sopusoinnussa normaalisuusoletuksen kanssa.

Huomautuksia:

- Microsoft Excel -ohjelman mukaan χ^2-yhteensopivuustestin testisuureen arvoa 4.20 vastaava ***p*-arvo** on 0.65. Siten normaalisuusoletuksen hylkäämiseen ei ole myöskään testin *p*-arvon mukaan mitään perusteita.
- Tarkkaan ottaen *luokkia olisi pitänyt yhdistää* niin, että ehdot

  $$E_k > 5 , k = 1, 2, \ldots , m$$

  olisivat toteutuneet. Tällä ei kuitenkaan pitäisi *tässä* olla vaikutusta testin tulokseen.

12.3. Homogeenisuuden testaaminen

Testausasetelma χ^2-homogeenisuustestissä

Tarkastellaan tutkimusasetelmaa, jossa perusjoukon *S* alkioita kuvataan *yhdellä faktorilla* eli *tekijällä A*, joka saa olla *laatuero*-, *järjestys*-, *välimatka*- tai *suhdeasteikollinen muuttuja*.

Jaetaan perusjoukko *S* jakaa *kahteen tai useampaan ryhmään*, poimitaan ryhmistä *toisistaan riippumattomat satunnaisotokset* ja tarkastellaan *tekijän A vaihtelua otoksissa*. Tehdään oletus, että *tekijä A noudattaa kaikissa ryhmissä samaa*, tarkemmin määrittelemätöntä *todennäköisyys-jakaumaa*.

Haluamme *testata* tehtyä *jakaumaoletusta*:

(i) Voidaanko eri otoksista (ryhmistä) saatuja havaintoarvojen jakaumia *kuvata samalla todennäköisyysjakaumalla*?

(ii) Voivatko otokset olla *saman todennäköisyysjakauman generoimia* eli *tuottamia*?

Jos tehty jakaumaoletus *pätee* ja *tekijä A noudattaa siis kaikissa ryhmissä samaa jakaumaa*, perusjoukko on **homogeeninen** ja *perusjoukkoa ei tarvitse jakaa tekijää A koskevissa tarkasteluissa erillisiksi ryhmiksi*. Jos tehty jakaumaoletus *ei päde* ja *tekijä A noudattaa siis eri ryhmissä eri jakaumia*, perusjoukko on **heterogeeninen** ja *ryhmiä on syytä tarkastella tekijää A koskevissa tarkasteluissa erillisinä*. Tällaisten jakaumaoletuksen testaamiseen tarkoitettuja testejä kutsutaan **homogeenisuustesteiksi**.

χ^2-homogeenisuustestin suorittaminen

χ^2-homogeenisuustestissä havaintojen ja niiden jakaumasta eri ryhmissä tehdyn homogeenisuus-oletuksen *yhteensopivuutta mitataan* seuraavalla tavalla:

(1) Valitaan havainnoille *yhteinen luokitus*.

(2) Määrätään *havaintojen luokkafrekvenssit* jokaisesta otoksesta.

(3) Määrätään tehdyn homogeenisuusoletuksen mukaiset *odotetut luokkafrekvenssit*.

(4) *Verrataan* havaittuja ja odotettuja luokkafrekvenssejä toisiinsa χ^2*-testisuureella*.

Hypoteesit χ^2-homogeenisuustestissä

Yleinen hypoteesi H :

Perusjoukosta voidaan jakaa r ryhmään, joista on poimittu (toisistaan riippumattomat) *satunnaisotokset*.

Nollahypoteesi H_0 :

Otokset $i = 1, 2, \dots , r$ on poimittu *samasta todennäköisyysjakaumasta*.

Vaihtoehtoinen hypoteesi H_1 :

Otokset $i = 1, 2, \dots , r$ on poimittu *eri todennäköisyysjakaumista*.

Havaitut frekvenssit

Oletetaan, että tutkimuksen kohteena oleva perusjoukko S on jaettu r *ryhmään* ja poimitaan ryhmistä toisistaan riippumattomat *satunnaisotokset*, joiden koot ovat

$$n_i \,,\; i = 1, 2, \dots , r.$$

Luokitellaan havainnot jokaisessa otoksessa *samaa luokitusta käyttäen* toisensa poissulkeviin luokkiin, joiden lukumäärä olkoon c ja määrätään ryhmän i luokkaan j kuuluvien havaintojen **havaittu frekvenssi** eli *lukumäärä*

$$O_{ij} \,,\; i = 1, 2, \dots , r \,,\; j = 1, 2, \dots , c.$$

Muodostetaan *havaituista frekvensseistä* O_{ij} $(r \times c)$*-frekvenssitaulukko* $[O_{ij}]$:

Luokat

Ryhmät		**1**	**2**	**…**	**c**	**Summa**
	1	O_{11}	O_{12}	…	O_{1c}	n_1
	2	O_{21}	O_{22}	…	O_{2c}	n_2
	…	…	…	…	…	…
	r	O_{r1}	O_{r2}	…	O_{rc}	n_r
	Summa	C_1	C_2	…	C_c	n

Taulukossa:

r = ryhmien lukumäärä

c = luokkien lukumäärä

O_{ij} = *havaittu frekvenssi* ryhmän i luokassa j, $i = 1, 2, \ldots , r$, $j = 1, 2, \ldots , c$

n_i = *otoskoko* ryhmässä i

C_j = *havaittu frekvenssi* yhdistetyn havaintoaineiston luokassa j

n = havaintojen kokonaislukumäärä

Frekvenssiä O_{ij} kutsutaan tavallisesti **havaituksi solufrekvenssiksi** frekvenssitaulun solussa (i, j).

Havaittujen frekvenssien O_{ij} taulukossa pätee:

(i) *Rivisummat* yhtyvät ryhmäkohtaisiin otoskokoihin:

$$\sum_{j=1}^{c} O_{ij} = n_i \ , \ i = 1, 2, \ldots , r$$

(ii) *Sarakesummat* yhtyvät yhdistetyn havaintoaineiston luokkafrekvensseihin:

$$\sum_{i=1}^{r} O_{ij} = C_j \ , \ j = 1, 2, \ldots , c$$

(iii) *Havaintojen kokonaislukumäärä*:

$$\sum_{i=1}^{r}\sum_{j=1}^{c} O_{ij} = \sum_{i=1}^{r} n_i = \sum_{j=1}^{c} C_j = n$$

Nollahypoteesin tulkinta χ^2-homogeenisuustestissä

Jos nollahypoteesi H_0 pätee, havaintojen pitää jakautua (satunnaisvaihtelua lukuun ottamatta) jokaisessa ryhmässä $i = 1, 2, \ldots , r$ *samalla tavalla* luokkiin $j = 1, 2, \ldots , c$.

Jos nollahypoteesi H_0 pätee, havaintojen jakautuminen luokkiin $j = 1, 2, \ldots , c$ *ei saa riippua* siitä, mihin ryhmään $i = 1, 2, \ldots , r$ ne kuuluvat.

Jos nollahypoteesi H_0 pätee, todennäköisyys, että havainto kuuluu luokkaan $j = 1, 2, \ldots , c$ *ei saa riippua* siitä, mihin ryhmään $i = 1, 2, \ldots , r$ se kuuluu.

Odotettujen frekvenssien määrääminen

Olkoon x on tarkastelun kohteena olevan perusjoukon S alkio.

Määritellään seuraavat todennäköisyydet:

$$p_{ij} = \Pr(x \text{ kuuluu ryhmään } i \textit{ ja } \text{luokkaan } j)$$

$$p_{j|i} = \Pr(x \text{ kuuluu luokkaan } j \mid x \text{ kuuluu ryhmään } i)$$

$$p_{i} = \Pr(x \text{ kuuluu ryhmään } i)$$

$$p_{j} = \Pr(x \text{ kuuluu luokkaan } j)$$

$$i = 1, 2, \ldots , r \ , \ j = 1, 2, \ldots , c$$

Todennäköisyyslaskennan *yleisen tulosäännön* mukaan

$$p_{ij} = p_{j|i}\, p_i \ , \ i = 1, 2, \ldots , r \ , \ j = 1, 2, \ldots , c$$

Jos nollahypoteesi H_0 *pätee*, todennäköisyys, että perusjoukon S alkio x kuuluu luokkaan j ei saa riippua siitä, mihin ryhmään i alkio x kuuluu. Siten nollahypoteesi H_0 *perusjoukon homogeenisuudesta* voidaan ilmaista muodossa

$$\mathrm{H}_0 : p_{j|i} = p_{\bullet j} \,,\; i = 1, 2, \ldots, r \,,\; j = 1, 2, \ldots, c$$

tai muodossa

$$\mathrm{H}_0 : p_{ij} = p_{i\bullet}\, p_{\bullet j} \,,\; i = 1, 2, \ldots, r \,,\; j = 1, 2, \ldots, c$$

Todennäköisyydet p_{ij}, $p_{i\bullet}$, $p_{\bullet j}$ voidaan estimoida havaituista frekvensseistä O_{ij} kaavoilla

$$\hat{p}_{ij} = \frac{O_{ij}}{n} \qquad \hat{p}_{i\bullet} = \frac{n_i}{n} \qquad \hat{p}_{\bullet j} = \frac{C_j}{n}$$

Jos nollahypoteesi H_0 *pätee, solutodennäköisyydet* p_{ij} *voidaan estimoida kaavoilla*

$$P_{ij} = \frac{n_i}{n} \times \frac{C_j}{n} = \hat{p}_{i\bullet}\, \hat{p}_{\bullet j}$$

Määrätään *nollahypoteesin* H_0 *pätiessä* **odotetut solufrekvenssit** E_{ij} yhtälöillä

$$E_{ij} = nP_{ij} = \frac{C_j}{n} \times n_i = \frac{n_i C_j}{n} \,,\; i = 1, 2, \ldots, r \,,\; j = 1, 2, \ldots, c$$

Muodostetaan *odotetuista frekvensseistä* E_{ij} $(r \times c)$-*frekvenssitaulukko* $[E_{ij}]$:

Luokat

Ryhmät	1	2	…	c	Summa
1	E_{11}	E_{12}	…	E_{1c}	n_1
2	E_{21}	E_{22}	…	E_{2c}	n_2
…	…	…	…	…	…
r	E_{r1}	E_{r2}	…	E_{rc}	n_r
Summa	C_1	C_2	…	C_c	n

Taulukossa

r = ryhmien lukumäärä

c = luokkien lukumäärä

E_{ij} = *odotettu frekvenssi* ryhmän i luokassa j, $i = 1, 2, \ldots, r$, $j = 1, 2, \ldots, c$

n_i = *otoskoko* ryhmässä i

C_j = *havaittu frekvenssi* yhdistetyn havaintoaineiston luokassa j

n = havaintojen kokonaislukumäärä

Frekvenssiä E_{ij} kutsutaan tavallisesti **odotetuksi solufrekvenssiksi** frekvenssitaulun solussa (i, j).

Odotettujen frekvenssien E_{ij} taulukossa pätee:

(i) *Rivisummat* yhtyvät ryhmäkohtaisiin otoskokoihin:

$$\sum_{j=1}^{c} E_{ij} = n_i \;,\; i = 1, 2, \ldots, r$$

(ii) *Sarakesummat* yhtyvät yhdistetyn havaintoaineiston luokkafrekvensseihin:

$$\sum_{i=1}^{r} E_{ij} = C_j \;,\; j = 1, 2, \ldots, c$$

(iii) *Havaintojen kokonaislukumäärä*:

$$\sum_{i=1}^{r}\sum_{j=1}^{c} E_{ij} = \sum_{i=1}^{r} n_i = \sum_{j=1}^{c} C_j = n$$

Testisuure ja sen jakauma χ^2-homogeenisuustestissä

Määritellään χ^2**-testisuure**

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

jossa

O_{ij} = *havaittu frekvenssi* solussa (i, j)

E_{ij} = *odotettu frekvenssi* solussa (i, j)

r = ryhmien lukumäärä

c = luokkien lukumäärä

Testisuure χ^2 *mittaa* havaittujen ja odotettujen frekvenssien *jakaumien yhteensopivuutta* tai *etäisyyttä* ja siksi sitä kutsutaan usein χ^2*-etäisyydeksi*.

Homogeenisuustestin χ^2-testisuure voidaan kirjoittaa myös muotoon

$$\chi^2 = n\sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(\hat{p}_{ij} - P_{ij})^2}{P_{ij}}$$

jossa

$$\hat{p}_{ij} = \frac{O_{ij}}{n}$$

$$P_{ij} = \frac{E_{ij}}{n} = \frac{n_i}{n} \times \frac{C_j}{n} = \hat{p}_i \, \hat{p}_j$$

$$\hat{p}_i = \frac{n_i}{n}$$

$$\hat{p}_j = \frac{C_j}{n}$$

$$i = 1, 2, \ldots, r \;,\; j = 1, 2, \ldots, c$$

Jos nollahypoteesi H_0 *pätee*, testisuure χ^2 noudattaa *suurissa otoksissa approksimatiivisesti* χ^2*-jakaumaa vapausastein* $f = (r - 1)(c - 1)$:

$$\chi^2 \ {}_a\ \chi^2(f)$$

jossa

r = *ryhmien* lukumäärä

c = *luokkien* lukumäärä

Approksimaatio on tavallisesti *riittävän hyvä*, jos *odotetut frekvenssit* E_{ij} ja *keskimääräiset odotetut frekvenssit* C_j/r toteuttavat ehdot

$$E_{ij} > 1 \,,\; i = 1, 2, \ldots, r \,,\; j = 1, 2, \ldots, c$$

$$C_j / r > 5 \,,\; j = 1, 2, \ldots, c$$

Ehdot saadaan toteutumaan valitsemalla havainnoille *sopiva luokitus*.

Testisuureen χ^2 *normaaliarvo* eli *odotusarvo nollahypoteesin* H_0 *pätiessä* on

$$\mathrm{E}(\chi^2) = f$$

jossa

$$f = (r - 1)(c - 1)$$

Normaaliarvoaan *merkitsevästi suuremmat* χ^2-testisuureen arvot viittaavat siihen, että nollahypoteesi H_0 *ei päde*.

χ^2-*homogeenisuustesti* on **jakaumista riippumaton**, **ei-parametrinen testi**:

(i) Testin yleinen hypoteesi *ei kiinnitä havaintojen jakaumaa*.

(ii) Testissä *ei testata todennäköisyysjakauman parametreja koskevaa hypoteesia*, vaan oletusta havaintojen jakaumasta.

12.4. Riippumattomuuden testaaminen

Testausasetelma χ^2-riippumattomuustestissä

Tarkastellaan tutkimusasetelmaa, jossa perusjoukon S alkioita kuvataan *kahdella faktorilla* eli *tekijällä* A ja B, jotka saavat olla *laatuero-*, *järjestys-*, *välimatka-* tai *suhdeasteikollisia muuttujia*.

Poimitaan perusjoukosta S *satunnaisotos* ja tarkastellaan tekijöiden A ja B *vaihtelua* otoksessa. Tehdään oletus, että *tekijät* A ja B *ovat riippumattomia*. Haluamme *testata* tehtyä *riippumattomuus-oletusta*: Ovatko havainnot *sopusoinnussa* tehdyn riippumattomuusoletuksen kanssa?

Jos tehty oletus *pätee* ja *tekijät* A ja B *ovat siis* **riippumattomia**, *tekijöitä* A ja B *voidaan tarkastella erillisinä*. Jos tehty oletus *ei päde* ja *tekijät* A ja B *eivät siis ole riippumattomia*, *tekijät* A ja B *ovat* **assosioituneita**.

Riippumattomuusoletuksen testaamiseen tarkoitettuja testejä kutsutaan **riippumattomuustesteiksi**.

χ^2-riippumattomuustestin suorittaminen

χ^2-*riippumattomuustestissä* havaintojen ja tehdyn riippumattomuusoletuksen *yhteensopivuutta mitataan* seuraavalla tavalla:

(1) Valitaan havainnoille sopivat *luokitukset* tekijöiden A ja B suhteen.

(2) Luokitellaan havainnot tekijöiden A ja B suhteen *ristiin* ja määrätään *havaitut luokkafrekvenssit*.

(3) Määrätään tehdyn riippumattomuusoletuksen mukaiset *odotetut luokkafrekvenssit*.

(4) *Verrataan* havaittuja ja odotettuja luokkafrekvenssejä toisiinsa χ^2-*testisuureella*.

Hypoteesit χ^2-riippumattomuustestissä

Yleinen hypoteesi H :

Perusjoukosta on poimittu *satunnaisotos* ja havaintoyksiköt on *luokiteltu ristiin* kahden tekijän A ja B suhteen.

Nollahypoteesi H_0 :

Tekijät A ja B *ovat* riippumattomia.

Vaihtoehtoinen hypoteesi H_1 :

Tekijät A ja B *eivät ole* riippumattomia.

Havaitut frekvenssit

Poimitaan tutkimuksen kohteena olevasta perusjoukosta S *satunnaisotos*, jonka koko on n.

Luokitellaan havaintoyksiköt *tekijän A suhteen* toisensa poissulkeviin luokkiin, joiden lukumäärä on r. Määrätään tekijän A luokkaan i kuuluvien havaintojen *havaittu frekvenssi* eli lukumäärä R_i , kun $i = 1, 2, \dots , r$.

Luokitellaan havaintoyksiköt *tekijän B suhteen* toisensa poissulkeviin luokkiin, joiden lukumäärä on c. Määrätään tekijän B luokkaan j kuuluvien havaintojen *havaittu frekvenssi* eli lukumäärä C_j , kun $j = 1, 2, \dots , c$.

Luokitellaan havaintoyksiköt *tekijöiden A ja B suhteen ristiin* toisensa poissulkeviin luokkiin, joiden lukumäärä on $r \times c$. Määrätään tekijän A luokkaan i *ja* tekijän B luokkaan j kuuluvien havaintojen **havaittu frekvenssi** eli *lukumäärä* O_{ij} , kun $i = 1, 2, \dots , r$ ja $j = 1, 2, \dots , c$.

Muodostetaan *havaituista frekvensseistä* O_{ij} $(r \times c)$-*frekvenssitaulukko* $[O_{ij}]$:

***B*-luokat**

A-luokat	1	2	…	*c*	Summa
1	O_{11}	O_{12}	…	O_{1c}	R_1
2	O_{21}	O_{22}	…	O_{2c}	R_2
…	…	…	…	…	…
r	O_{r1}	O_{r2}	…	O_{rc}	R_r
Summa	C_1	C_2	…	C_c	n

Taulukossa:

r = A-luokkien lukumäärä

c = B-luokkien lukumäärä

O_{ij} = *havaittu frekvenssi* luokassa, jonka määrää A-luokka i *ja* B-luokka j, $i = 1, 2, \ldots, r$, $j = 1, 2, \ldots, c$

R_i = *havaittu frekvenssi* A-luokassa i

C_j = *havaittu frekvenssi* B-luokassa j

n = havaintojen kokonaislukumäärä

Frekvenssiä O_{ij} kutsutaan tavallisesti **havaituksi solufrekvenssiksi** frekvenssitaulun solussa (i, j).

Havaittujen frekvenssien O_{ij} taulukossa pätee:

(i) *Rivisummat* yhtyvät havaittuihin frekvensseihin A-luokituksessa:

$$\sum_{j=1}^{c} O_{ij} = R_i \,,\; i = 1, 2, \ldots, r$$

(ii) *Sarakesummat* yhtyvät havaittuihin frekvensseihin B-luokituksessa:

$$\sum_{i=1}^{r} O_{ij} = C_j \,,\; j = 1, 2, \ldots, c$$

(iii) *Havaintojen kokonaislukumäärä*:

$$\sum_{i=1}^{r}\sum_{j=1}^{c} O_{ij} = \sum_{i=1}^{r} R_i = \sum_{j=1}^{c} C_j = n$$

Nollahypoteesin tulkinta χ^2-riippumattomuustestissä

Jos nollahypoteesi H_0 *pätee*, *havaintojen jakautuminen* A-luokkiin *ei saa riippua* siitä, mihin B-luokkaan havainnot kuuluvat, toisin sanoen, *jos nollahypoteesi* H_0 *pätee*, *todennäköisyys*, että havainto kuuluu A-luokkaan $i = 1, 2, \ldots, r$ *ei saa riippua* siitä, mihin B-luokkaan $j = 1, 2, \ldots, c$ se kuuluu.

Jos nollahypoteesi H_0 *pätee*, *havaintojen jakautuminen* B-luokkiin *ei saa riippua* siitä, mihin A-luokkaan havainnot kuuluvat, toisin sanoen, *jos nollahypoteesi* H_0 *pätee*, *todennäköisyys*, että havainto kuuluu B-luokkaan $j = 1, 2, \ldots, c$ *ei saa riippua* siitä, mihin A-luokkaan $i = 1, 2, \ldots, r$ se kuuluu.

Odotettujen frekvenssien määrääminen

Olkoon x on tarkastelun kohteena olevan perusjoukon S alkio.

Määritellään seuraavat todennäköisyydet:

$$p_{ij} = \Pr(x \text{ kuuluu } A\text{-luokkaan } i \text{ ja } B\text{-luokkaan } j)$$

$$p_{i\cdot} = \Pr(x \text{ kuuluu } A\text{-luokkaan } i)$$

$$p_{\cdot j} = \Pr(x \text{ kuuluu } B\text{-luokkaan } j)$$

$$i = 1, 2, \ldots, r \,,\; j = 1, 2, \ldots, c$$

Jos nollahypoteesi H_0 *pätee*, *tapahtumat*

$$\{x \in S \mid x \text{ kuuluu } A\text{-luokkaa } i\}$$

$$\{x \in S \mid x \text{ kuuluu } B\text{-luokkaa } j\}$$

ovat riippumattomia kaikille $i = 1, 2, \ldots , r$, $j = 1, 2, \ldots , c$.

Siten nollahypoteesi H_0 *tekijöiden A* ja *B riippumattomuudesta* voidaan ilmaista muodossa

$$H_0 : p_{ij} = p_{i\cdot}\, p_{\cdot j}\, , \; i = 1, 2, \ldots, r\, , \; j = 1, 2, \ldots, c$$

Todennäköisyydet p_{ij} , $p_{i\cdot}$, $p_{\cdot j}$ voidaan estimoida havaituista frekvensseistä O_{ij} kaavoilla

$$\hat{p}_{ij} = \frac{O_{ij}}{n} \qquad \hat{p}_{i\cdot} = \frac{R_i}{n} \qquad \hat{p}_{\cdot j} = \frac{C_j}{n}$$

Jos nollahypoteesi H_0 *pätee, solutodennäköisyydet* p_{ij} *voidaan estimoida kaavoilla*

$$P_{ij} = \frac{R_i}{n} \times \frac{C_j}{n} = \hat{p}_{i\cdot}\, \hat{p}_{\cdot j}$$

Määrätään *nollahypoteesin* H_0 *pätiessä odotetut solufrekvenssit* E_{ij} yhtälöillä

$$E_{ij} = nP_{ij} = \frac{R_i C_j}{n}\, , \; i = 1, 2, \ldots, r\, , \; j = 1, 2, \ldots, c$$

Muodostetaan *odotetuista frekvensseistä* E_{ij} ($r \times c$)-*frekvenssitaulukko* $[E_{ij}]$:

***B*-luokat**

A-luokat	**1**	**2**	**…**	***c***	**Summa**
1	E_{11}	E_{12}	…	E_{1c}	R_1
2	E_{21}	E_{22}	…	E_{2c}	R_2
…	…	…	…	…	…
r	E_{r1}	E_{r2}	…	E_{rc}	R_r
Summa	C_1	C_2	…	C_c	n

Taulukossa

r = *A*-luokkien lukumäärä

c = *B*-luokkien lukumäärä

E_{ij} = *odotettu frekvenssi* luokassa, jonka määrää *A*-luokka *i ja B*-luokka *j*, $i = 1, 2, \ldots , r$, $j = 1, 2, \ldots , c$

R_i = *havaittu frekvenssi A*-luokassa *i*

C_j = *havaittu frekvenssi B*-luokassa *j*

n = havaintojen kokonaislukumäärä

Frekvenssiä E_{ij} kutsutaan tavallisesti **odotetuksi solufrekvenssiksi** frekvenssitaulun solussa (i, j).

Odotettujen frekvenssien E_{ij} taulukossa pätee:

(i) *Rivisummat* yhtyvät havaittuihin frekvensseihin *A*-luokituksessa:

$$\sum_{j=1}^{c} E_{ij} = R_i \ , \ i = 1, 2, \ldots, r$$

(ii) *Sarakesummat* yhtyvät havaittuihin frekvensseihin *B*-luokituksessa:

$$\sum_{i=1}^{r} E_{ij} = C_j \ , \ j = 1, 2, \ldots, c$$

(iii) *Havaintojen kokonaislukumäärä*:

$$\sum_{i=1}^{r}\sum_{j=1}^{c} E_{ij} = \sum_{i=1}^{r} R_i = \sum_{j=1}^{c} C_j = n$$

Testisuure ja sen jakauma χ^2-riippumattomuustestissä

Määritellään χ^2-**testisuure**

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

jossa

O_{ij} = *havaittu frekvenssi* solussa (i, j)

E_{ij} = *odotettu frekvenssi* solussa (i, j)

r = *A*-luokkien lukumäärä

c = *B*-luokkien lukumäärä

Testisuure χ^2 *mittaa* havaittujen ja odotettujen frekvenssien *jakaumien yhteensopivuutta* tai *etäisyyttä* ja siksi sitä kutsutaan usein χ^2-*etäisyydeksi*.

χ^2-testisuure voidaan kirjoittaa myös muotoon

$$\chi^2 = n\sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(\hat{p}_{ij} - P_{ij})^2}{P_{ij}}$$

jossa

$$\hat{p}_{ij} = \frac{O_{ij}}{n}$$

$$P_{ij} = \frac{E_{ij}}{n} = \frac{R_i}{n} \times \frac{C_j}{n} = \hat{p}_{i\bullet}\, \hat{p}_{\bullet j}$$

$$\hat{p}_{i\bullet} = \frac{R_i}{n}$$

$$\hat{p}_{\bullet j} = \frac{C_j}{n}$$

$$i = 1, 2, \ldots, r \ , \ j = 1, 2, \ldots, c$$

Jos nollahypoteesi H_0 *pätee*, testisuure χ^2 noudattaa *suurissa otoksissa approksimatiivisesti* χ^2-*jakaumaa vapausastein* $f = (r-1)(c-1)$:

$$\chi^2 \underset{a}{\sim} \chi^2(f)$$

jossa

r = A-luokkien lukumäärä

c = B-luokkien lukumäärä

Approksimaatio on tavallisesti *riittävän hyvä*, jos *odotetut frekvenssit* E_{ij} ja *keskimääräiset odotetut frekvenssit* R_i/c ja C_j/r toteuttavat ehdot

$$E_{ij} > 1\,,\ i = 1, 2, \ldots, r\,,\ j = 1, 2, \ldots, c$$

$$R_i / c > 5\,,\ i = 1, 2, \ldots, r$$

$$C_j / r > 5\,,\ j = 1, 2, \ldots, c$$

Ehdot saadaan toteutumaan valitsemalla havainnoille *sopiva luokitus*.

Testisuureen χ^2 *normaaliarvo* eli *odotusarvo nollahypoteesin* H_0 *pätiessä* on

$$\mathrm{E}(\chi^2) = f$$

jossa

$$f = (r - 1)(c - 1)$$

Normaaliarvoaan *merkitsevästi suuremmat* χ^2-testisuureen arvot viittaavat siihen, että nollahypoteesi H_0 *ei päde*.

χ^2*-riippumattomuustesti* on **jakaumista riippumaton**, **ei-parametrinen testi**:

(i) Testin yleinen hypoteesi *ei kiinnitä havaintojen jakaumaa*.

(ii) Testissä *ei testata todennäköisyysjakauman parametreja koskevaa hypoteesia*, vaan oletusta havaintojen jakaumasta.

12.5. χ^2-homogeenisuustesti ja χ^2-riippumattomuustesti

χ^2*-homogeenisuustesti* ja χ^2*-riippumattomuustesti* tehdään teknisesti *täsmälleen samalla tavalla* ja niinpä frekvenssitaulukosta *ei voi sellaisenaan nähdä* kummasta testausasetelmasta on kyse:

- *Odotetut frekvenssit* määrätään *samalla kaavalla*.
- *Testisuureet* lasketaan *samalla kaavalla*.
- *Testisuureet* noudattavat *nollahypoteesin pätiessä* approksimatiivisesti *samaa jakaumaa*.

Testien testausasetelmat ovat kuitenkin *täysin erilaiset*.

Homogeenisuustestin testausasetelma:

(i) Perusjoukko koostuu *r ryhmästä* ja testissä tarkastellaan perusjoukon alkioiden *jakautumista luokkiin eri ryhmissä, kun alkiot on luokittelu yhden tekijän A suhteen käyttäen kaikissa ryhmissä samaa luokitusta*.

(ii) Havaintoaineisto muodostuu *toisistaan riippumattomista ryhmäkohtaisista satunnaisotoksista*.

(iii) Sekä ryhmäkohtaiset otoskoot n_i että havaintojen kokonaislukumäärä n ovat *kiinteitä* eli *ei-satunnaisia* (eli valittuja) lukuja, kun taas *sattuma määrää miten havainnot jakautuvat luokkiin* ryhmien sisällä.

Riippumattomuustestin testausasetelma:

(i) Testissä tarkastellaan *kahden tekijän A ja B assosiaatiota* eli *riippuvuutta*, kun havainnot on luokiteltu *ristiin* tekijöiden A ja B suhteen.

(ii) Havaintoaineisto muodostuu *yhdestä satunnaisotoksesta*.

(iii) Vain havaintojen kokonaislukumäärä n on *kiinteä* eli *ei-satunnainen* (eli valittu) luku, kun taas *sattuma määrää miten havainnot jakautuvat luokkiin tekijöiden A ja B ristiluokituksen suhteen*.

12.6. Normaalisuuden testaaminen

Normaalijakaumalla on *keskeinen asema tilastotieteessä*. Esimerkiksi tavanomaisen *t*-testin yleisessä hypoteesissa oletetaan, että havainnot noudattavat normaalijakaumaa. Siksi tilastotieteessä on kehitetty useita erilaisia menetelmiä havaintojen normaalisuuden tutkimiseen.

Normaalisuutta voidaan testata edellä esitetyllä **χ^2-testillä**, joka on yleinen yhteensopivuustesti. Seuraavassa tarkastellaan kuitenkin seuraavia, erityisesti normaalisuuden tutkimiseen tarkoitettuja menetelmiä:

- **Bowmanin ja Shentonin testi**
- **Rankit Plot -kuvio** ja **Wilkin ja Shapiron testi**

Bowmanin ja Shentonin testi normaalisuudelle

Olkoot γ_1 ja γ_2 tavanomaiset keskusmomentteihin perustuvat tunnusluvut todennäköisyysjakaumien *vinoudelle* ja *huipukkuudelle*. *Normaalijakaumalle*

$$\gamma_1 = \gamma_2 = 0$$

Bowmanin ja Shentonin testissä havaintojen normaalisuuden testaaminen perustuu testisuureeseen, joka on *vastaavien otossuureiden funktio*. Testisuure saa *suuria arvoja*, jos havaintojen vinous ja/tai huipukkuus *poikkeavat* paljon normaalijakautuneen satunnaismuuttujan vinoudesta ja/tai huipukkuudesta.

Todennäköisyysjakauman vinous ja huipukkuus

Satunnaismuuttujan X ***k*. origomomentti** on

$$\alpha_k = \mathrm{E}(X^k)\,,\ k = 1, 2, 3, \ldots$$

Satunnaismuuttujan X ***k*. keskusmomentti** on

$$\mu_k = \mathrm{E}\left[(X-\alpha_1)^k\right],\ k = 1, 2, 3, \ldots$$

Erityisesti

$$\alpha_1 = \mathrm{E}(X) = \mu_X$$

on satunnaismuuttujan X *odotusarvo* ja

$$\mu_2 = \mathrm{E}\left[(X-\mu_X)^2\right] = \sigma_X^2$$

on satunnaismuuttujan X *varianssi*.

Todennäköisyysjakauman **vinouden** *mittana* käytetään tunnuslukua

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}$$

ja *todennäköisyysjakauman* **huipukkuuden** *mittana*. käytetään tunnuslukua

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

Normaalijakaumalle

$$\gamma_1 = \gamma_2 = 0$$

Havaintojen jakauman vinous ja huipukkuus

Olkoot

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

välimatka- tai *suhdeasteikollisen satunnaismuuttujan* X havaittuja arvoja.

Havaintojen $X_1, X_2, \ldots, X_n$ ***k*. origomomentti** on

$$a_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \ , \ k = 1, 2, \ldots$$

Havaintojen $X_1, X_2, \ldots, X_n$ ***k*. keskusmomentti** on

$$m_k = \frac{1}{n}\sum_{i=1}^{n} (X_i - a_1)^k \ , \ k = 1, 2, \ldots$$

Erityisesti

$$a_1 = \overline{X}$$

on havaintojen $X_1, X_2, \ldots, X_n$ *aritmeettinen keskiarvo* ja

$$m_2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})^2 = \hat{\sigma}_X^2$$

on havaintojen $X_1, X_2, \ldots, X_n$ *otosvarianssi*.

Havaintojen jakauman **vinouden** *mittana* käytetään tunnuslukua

$$c_1 = \frac{m_3}{m_2^{3/2}}$$

ja *havaintojen jakauman* **huipukkuuden** *mittana* käytetään tunnuslukua

$$c_2 = \frac{m_4}{m_2^2} - 3$$

Hypoteesit

Yleinen hypoteesi H :

Havainnot X_1 , X_2 , … , X_n on poimittu *yksinkertaisella satunnaisotannalla* perusjoukosta S.

Nollahypoteesi H_0 :

Havainnot X_1 , X_2 , … , X_n *noudattavat normaalijakaumaa*.

Vaihtoehtoinen hypoteesi H_1 :

Havainnot X_1 , X_2 , … , X_n *eivät noudata normaali-jakaumaa*.

Bowmanin ja Shentonin testi

Määritellään **χ^2-testisuure**

$$\chi^2 = \frac{n}{6}c_1^2 + \frac{n}{24}c_2^2$$

Jos nollahypoteesi H_0 *pätee*, testisuure χ^2 noudattaa suurissa otoksissa *approksimatiivisesti* χ^2-*jakaumaa* 2:lla *vapausasteella*:

$$\chi^2 \;{}_{a}\; \chi^2(2)$$

Testisuureen χ^2 *normaaliarvo* eli *odotusarvo nollahypoteesin* H_0 *pätiessä* on

$$\mathrm{E}(\chi^2) = 2$$

Normaaliarvoaan *merkitsevästi suuremmat* χ^2-testisuureen arvot viittaavat siihen, että nollahypoteesi H_0 *ei päde*. Normaaliarvoaan *merkitsevästi pienemmät* χ^2-testisuureen arvot viittaavat siihen, että nollahypoteesi H_0 *pätee liian hyvin*, jolloin havainnot saattavat olla *väärennettyjä*.

Rankit Plot -kuvio sekä Wilkin ja Shapiron testi normaalisuudelle

Tietokonegrafiikka mahdollistaa havaintoaineiston normaalisuuden tutkimisen *graafisin keinoin*. Normaalisuuden tutkimiseen tarkoitetut graafiset menetelmät perustuvat kuvioihin, joiden ideana on se, että *normaalijakautuneet havainnot asettuvat* (*satunnaisvaihtelua lukuun ottamatta*) *suoralle viivalle* ja *havaintojen epänormaalisuus tulee kuviossa esille poikkeamina tästä suorasta*. Tällainen kuvio voidaan muodostaa usealla eri periaatteella; tässä tarkastellaan ns. **Rankit Plot -kuviota**.

Olkoot

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

havainnot X_1 , X_2 , … , X_n *suuruusjärjestyksessä* pienimmästä suurimpaan. Olkoon

$$\mathrm{E}(Y_i) , i = 1, 2, \ldots , n$$

i. havainnon Y_i *odotusarvo*, jossa Y_i on suuruusjärjestyksessä *i*. havainto standardoidusta normaalijakaumasta N(0,1) poimitusta satunnaisotoksesta.

Piirretään *pistediagrammi*

$$(\mathrm{E}(Y_i), Z_i) , i = 1, 2, \ldots , n$$

Jos havainnot X_1 , X_2 , … , X_n ovat *normaalijakautuneen satunnaismuuttujan X havaittuja arvoja, pisteet*

$$(\mathrm{E}(Y_i), Z_i) , i = 1, 2, \ldots , n$$

asettuvat (*satunnaisvaihtelua lukuun ottamatta*) *suoralle viivalle*. *Poikkeamat* suorasta viittaavat epänormaalisuuteen.

Kuviosta voidaan tunnistaa:

- Havaintoarvojen jakauman *vinous*
- Havaintoarvojen jakauman *huipukkuus*
- *Poikkeavat havainnot* (*engl*. outliers)

Wilkin ja Shapiron testisuure on *Rankit Plot -kuvion pisteistä*

$$(\mathrm{E}(Y_i), Z_i)\ ,\ i = 1, 2, \ldots , n$$

lasketun *otoskorrelaatiokertoimen neliö*. *Pienet* testisuureen arvot viittaavat siihen, että *normaalisuusoletus ei päde*. *Suuret* testisuureen arvot ovat sopusoinnussa *normaalisuusoletuksen* kanssa. Testisuureen jakauma on *epästandardi*, mutta monet tilasto-ohjelmistot osaavat laskea Wilkin ja Shapiron testisuureen arvoa vastaavia *p-arvoja*.

Rankit Plot -kuvio ja Wilkin ja Shapiron testi: Sovellus

Kappaleessa **Yhteensopivuuden testaaminen** tarkasteltiin seuraavaa esimerkkiä:

(i) Kone tekee *ruuveja*, joiden *tavoitepituutena* on 10 cm.

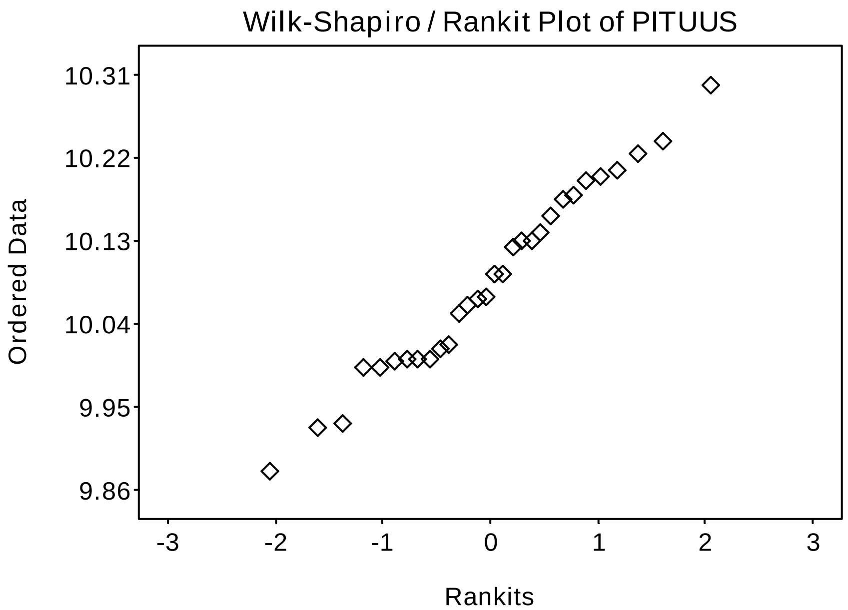
(ii) Ruuvien *pituus vaihtelee satunnaisesti* jonkin verran.

(iii) Ruuvien pituuksien oletetaan kuitenkin noudattavan *normaalijakaumaa*.

Tällöin todettiin, että *havainnot ovat yleisen* χ^2*-yhteensopivuustestin perusteella sopusoinnussa normaalisuusoletuksen kanssa*.

Tarkastellaan normaalisuusoletusta vielä **Rankit Plot -kuvion** ja **Shapiron ja Wilkin testin** valossa.

Alla on otokseen poimittujen ruuvien pituuksista piirretty *Rankit Plot -kuvio*.



Havaintoja vas… *normaalisuusoletusta ei ole mitään syytä asettaa kyseenalaiseksi*. Rankit Plot -kuvioon liittyvä *Wilkin ja Shapiron testisuureen arvo* on esimerkkiaineiston tapauksessa

$$0.9779$$

ja testisuureen arvoa vastaava *p-arvo* on

$$0.7674$$

Siten *normaalisuusoletusta ei ole syytä asettaa kyseenalaiseksi myöskään Wilkin ja Shapiron testin perusteella*.