Ilkka Mellin

Tilastolliset menetelmät

Osa 3: Tilastolliset testit

Tilastollinen testaus

Tilastolliset testit

>> Tilastollinen testaus

Tilastolliset hypoteesit

Tilastolliset testit ja testisuureet

Virheet testauksessa

Merkitsevyystaso ja testin hylkäysalue

Testin *p*-arvo

Testin suorittaminen

Esimerkki: Normaalijakauman parametreja koskevat testit

Tilastolliset testit ja mitta-asteikot

Tilastollisten hypoteesien testaaminen 1/5

- 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*.

Tilastollisten hypoteesien testaaminen 2/5

- Olkoon X tutkimuksen kohteena olevan perusjoukon jonkin ominaisuuden vaihtelua perusjoukossa kuvaava *satunnaismuuttuja*.
- Olkoon satunnaismuuttujan X *todennäköisyysjakauman pistetodennäköisyys*- 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 .

Tilastollisten hypoteesien testaaminen 3/5

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

$$\theta = \theta_0$$

voidaan *testata tilastollisesti*?

- **Tilastollisessa testauksessa** *hypoteesi*

$$\theta = \theta_0$$

asetetaan koetteelle havaintojen todennäköisyys-jakaumasta $f(x\,;\,\theta)$ sisältämää informaatiota vastaan.

Tilastollisten hypoteesien testaaminen 4/5

- Oletamme jatkossa, että *havainnot*

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

  muodostavat (**yksinkertaisen**) **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* piste-todennäköisyys- tai tiheysfunktio $f(x\,;\,\theta)$:

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

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

Tilastollisten hypoteesien testaaminen 5/5

- 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 $\theta = \theta_0$ *yhteensopivuutta*.
- *Hyvä yhteensopivuus* merkitsee sitä, että havainnot ovat *sopusoinnussa* oletuksen $\theta = \theta_0$ kanssa.
- *Huono yhteensopivuus* merkitsee sitä, että havainnot ja oletus $\theta = \theta_0$ ovat *ristiriidassa* keskenään.

Tilastolliset testit

Tilastollinen testaus

>> Tilastolliset hypoteesit

Tilastolliset testit ja testisuureet

Virheet testauksessa

Merkitsevyystaso ja testin hylkäysalue

Testin *p*-arvo

Testin suorittaminen

Esimerkki: Normaalijakauman parametreja koskevat testit

Tilastolliset testit ja mitta-asteikot

Testausasetelmaa koskevat hypoteesit

- Kun todennäköisyysjakauman parametreja koskevia väitteitä tai oletuksia testataan tilastollisesti, **testaus-asetelmasta** on tehtävä asetelman kiinnittämiseksi 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 1/2

- *Yleiset testausasetelmaa koskevat oletukset* muodostavat testin **yleisen hypoteesin** H.
- *Yleinen hypoteesi* H sisältää oletukset
  - *perusjoukosta*
  - *käytetystä otantamenetelmästä*
  - *perusjoukon jakaumasta*
- **Yleisen hypoteesin H oletuksista *pidetään kiinni koko testauksen ajan*, mikä merkitsee sitä, että testi tehdään aina ehdollisesti yleisen hypoteesin H oletusten suhteen**.

Yleinen hypoteesi 2/2

- Huomaa, että 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*.

Tilastolliset hypoteesit
Nollahypoteesin muoto yksinkertaisissa testausasetelmissa

- Olkoon

  $$f(x\ ;\ \theta)$$

  tutkimuksen kohteena olevaa perusjoukon ominaisuutta kuvaavan *todennäköisyysjakauman pistetodennäköisyys*- 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.
- Huomautuksia:
  - Jos nollahypoteesi on muotoa ”*on sama*” tai ”*ei ole eroa*”, vaihtoehtoinen hypoteesi on tavallisesti muotoa ”*ei ole sama*” tai ”*on eroa*”.
  - Tilastollista testiä tehtäessä toivotaan usein, että *nollahypoteesi voidaan hylätä* ja *vaihtoehtoinen hypoteesi hyväksyä*.
  - Vaihtoehtoisen hypoteesin *hyväksyminen* merkitsee yleensä *informaation lisääntymistä*.

Vaihtoehtoisen hypoteesin muoto yksinkertaisissa testausasetelmissa 1/2

- Jos nollahypoteesi on yksinkertaista muotoa

  $$H_0 : \theta = \theta_0$$

  vaihtoehtoinen hypoteesi voidaan muotoilla kolmella eri tavalla.
- Huomautus:

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

Vaihtoehtoisen hypoteesin muoto yksinkertaisissa testausasetelmissa 2/2

- Jos *vaihtoehtoinen hypoteesi* on muotoa

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

  tai muotoa

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

  vaihtoehtoista hypoteesia kutsutaan **yksisuuntaiseksi**.
- Jos *vaihtoehtoinen hypoteesi* on muotoa

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

  vaihtoehtoista hypoteesia kutsutaan **kaksisuuntaiseksi**.

Tilastolliset testit

Tilastollinen testaus

Tilastolliset hypoteesit

>> Tilastolliset testit ja testisuureet

Virheet testauksessa

Merkitsevyystaso ja testin hylkäysalue

Testin *p*-arvo

Testin suorittaminen

Esimerkki: Normaalijakauman parametreja koskevat testit

Tilastolliset testit ja mitta-asteikot

Tilastollinen testi päätössääntönä

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

Testisuure

- Tilastollinen testi perustetaan **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*.
- Siten yhteensopivuuden mittaaminen vaatii *testisuureen jakauman* tuntemista.

Testisuure ja testi päätössääntönä

- 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 normaaliarvo

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

Tilastolliset testit

Virheet testauksessa

Hylkäysvirhe ja sen todennäköisyys 1/2

- 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*.

Virheet testauksessa

Hylkäysvirhe ja sen todennäköisyys 2/2

- Tilastollisessa tutkimuksessa noudatetaan tieteen yleistä *varovaisuusperiaatetta*:

  Hypoteeseja ei pidä hylätä ilman riittäviä syitä.
- Siksi nollahypoteesin H_0 *virheellisen hylkäyksen todennäköisyys halutaan tehdä* tilastollisessa testauksessa *mahdollisimman pieneksi*.
- Jotta hylkäysvirheen todennäköisyys saataisiin mahdollisimman pieneksi, *havainnoilta on vaadittava vahvoja todisteita nollahypoteesia* H_0 *vastaan* ennen sen hylkäämistä.

Hyväksymisvirhe ja sen todennäköisyys

- 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 *komplementti-todennäköisyyttä*

$$\Pr(\mathrm{H}_0 \text{ hylätään} \mid \mathrm{H}_0 \text{ ei ole tosi}) = 1 - \beta$$

  kutsutaan testin **voimakkuudeksi**.
- Hyvä testi on *voimakas*, koska voimakkaalla testillä on pieni hyväksymisvirheen todennäköisyys β .
- Testin voimakkuus (1 – β) riippuu tavallisesti mm. testattavan parametrin *todellisesta arvosta*.
- Testin voimakkuutta testattavan parametrin arvojen funktiona kutsutaan **voimakkuusfunktioksi**; esimerkki: ks. kappaletta **Yhden otoksen *t*-testi** luvussa **Testit suhde-asteikollisille muuttujille.**

Ensimmäisen ja toisen 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 *hyväksymisvirhettä* eli sitä, että nollahypoteesi H_0 *hyväksytään* silloin, kun se *ei ole tosi*, kutsutaan **toisen lajin virheeksi**.

Maailman tila ja testin tulos

- *Maailman tilat* ja *testin tulokset* voidaan luokitella seuraavaksi nelikentäksi:

		Maailman tila	
		Nollahypoteesi *pätee*	Nollahypoteesi *ei päde*
Testin tulos	Nollahypoteesi *jää voimaan*	**Oikea johtopäätös**	*Hyväksymis-virhe*
	Nollahypoteesi *hylätään*	*Hylkäysvirhe*	**Oikea johtopäätös**

Testin hylkäys- ja hyväksymisalueet ja testi päätössääntönä

- 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:

  Jako hylkäys- ja hyväksymisalueisiin *ei saa riippua havainnoista*.

Tilastolliset testit

Tilastollinen testaus

Tilastolliset hypoteesit

Tilastolliset testit ja testisuureet

Virheet testauksessa

>> Merkitsevyystaso ja testin hylkäysalue

Testin *p*-arvo

Testin suorittaminen

Esimerkki: Normaalijakauman parametreja koskevat testit

Tilastolliset testit ja mitta-asteikot

Merkitsevyystaso

- Testin **merkitsevyystaso** α on todennäköisyys sille, 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* (ennen havaintojen keräämistä); ks. kohtaa Esimerkki: Normaalijakauman parametrien testaaminen.

Merkitsevyystason frekvenssitulkinta

- Oletetaan, että *nollahypoteesi* H_0 *pätee* testausasetelmassa.
- Valitaan testin *merkitsevyystasoksi* α.
- *Toistetaan* otantaa ja sovelletaan jokaiseen otokseen *samaa testiä*.
- **Tällöin joudumme virheellisesti hylkäämään nollahypoteesin** H_0 **keskimäärin**

  α %:ssa

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

Merkitsevyystason valinta: Tavanomaiset merkitsevyystasot 1/2

- Koska testeissä halutaan ensisijaisesti suojautua *hylkäysvirhettä* vastaan, testin merkitsevyystasoksi α on tapana valita pieniä lukuja.
- Ns. **tavanomaiset merkitsevyystasot** ovat

$$\alpha = 0.05$$

$$\alpha = 0.01$$

$$\alpha = 0.001$$

- Testin merkitsevyystasoa α valittaessa on aina syytä ottaa huomioon *väärän päätöksen seuraukset*.

Merkitsevyystason valinta: Tavanomaiset merkitsevyystasot 2/2

- Jos nollahypoteesi H_0 voidaan hylätä merkitsevyystasolla $\alpha = 0.05$, sanotaan:

  Testisuureen arvo (tai testin tulos) on **melkein merkitsevä**.

- Jos nollahypoteesi H_0 voidaan hylätä merkitsevyystasolla $\alpha = 0.01$, sanotaan:

  Testisuureen arvo (tai testin tulos) on **merkitsevä**.

- Jos nollahypoteesi H_0 voidaan hylätä merkitsevyystasolla $\alpha = 0.001$, sanotaan:

  Testisuureen arvo (tai testin tulos) on **erittäin merkitsevä**.

Testin hylkäysalueen määrääminen yksinkertaisissa testausasetelmissa 1/6

- 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.

Testin hylkäysalueen määrääminen yksinkertaisissa testausasetelmissa 2/6

- Oletetaan, että *testisuureena* on (jatkuva) satunnais-muuttuja 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.

Testin hylkäysalueen määrääminen yksinkertaisissa testausasetelmissa 3/6

- Jos vaihtoehtoisena hypoteesina on *yksisuuntainen vaihtoehto*

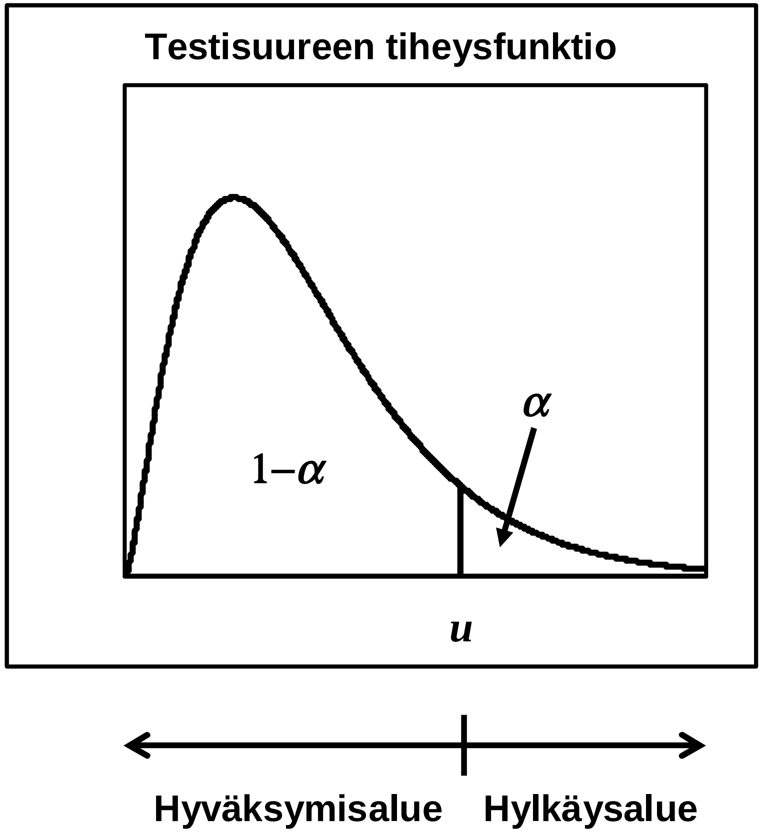
  $$\mathrm{H}_1 : \theta > \theta_0$$

  hylkäysalue on muotoa

  $$(u, b)$$

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

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



Merkitsevyystaso ja testin hylkäysalue

Testin hylkäysalueen määrääminen yksinkertaisissa testausasetelmissa 4/6

- Jos vaihtoehtoisena hypoteesina on *yksi-suuntainen vaihtoehto*

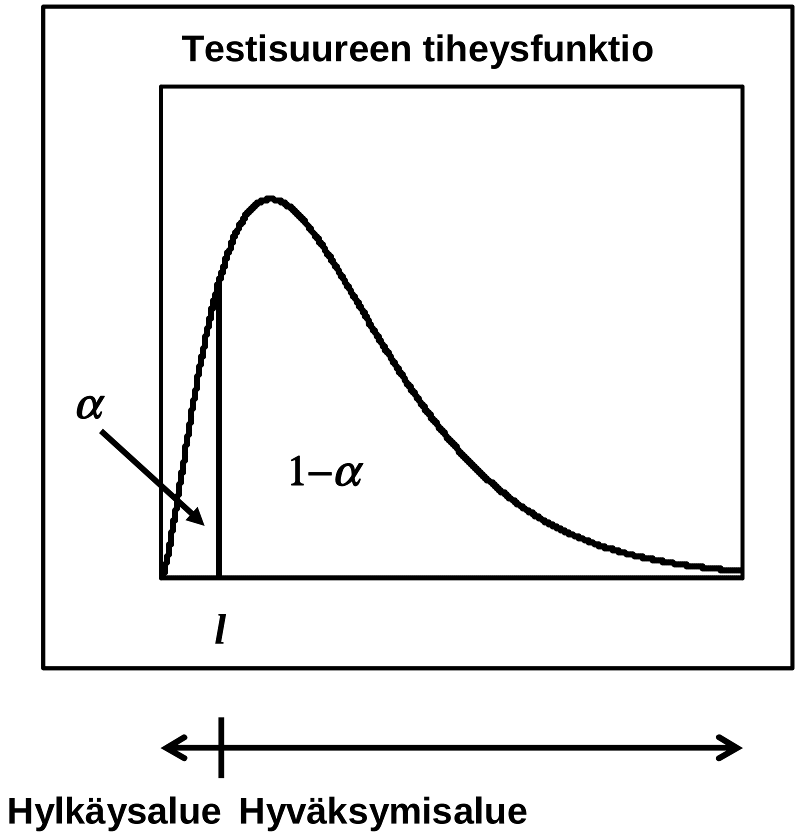
  $$\mathrm{H}_1 : \theta < \theta_0$$

  hylkäysalue on muotoa

  $$(a, l)$$

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

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



Testin hylkäysalueen määrääminen yksinkertaisissa testausasetelmissa 5/6

- Jos vaihtoehtoisena hypoteesina on *kaksi-suuntainen vaihtoehto*

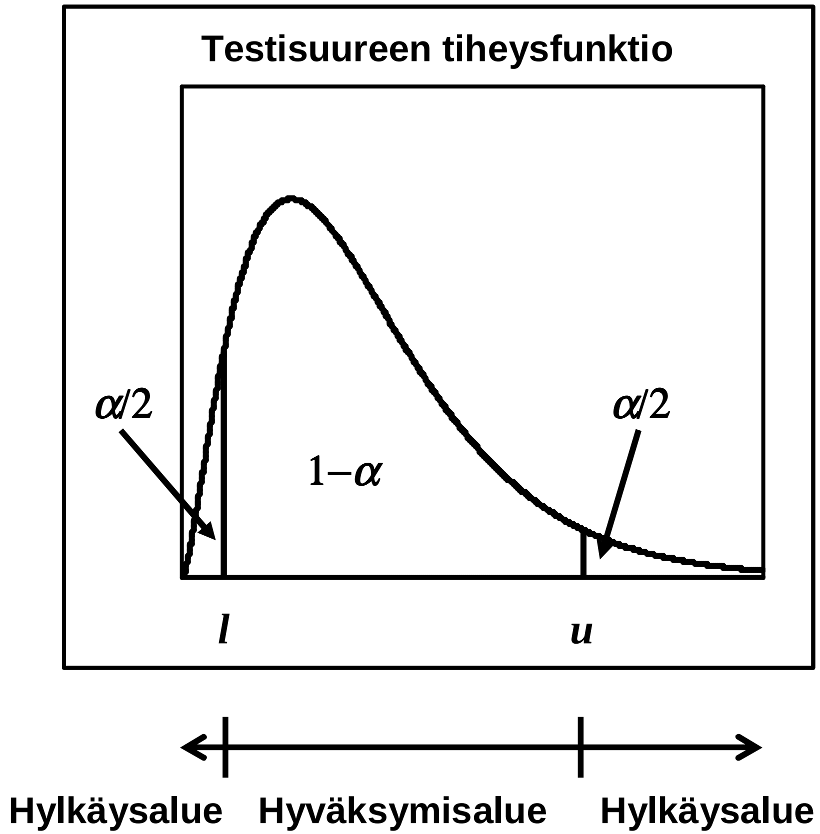
  $$\mathrm{H}_1 : \theta \neq \theta_0$$

  hylkäysalue on muotoa

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

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

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



Testin hylkäysalueen määrääminen yksinkertaisissa testausasetelmissa 6/6

- Oletetaan, että testisuureen Z jakauma on *symmetrinen* origon suhteen.
- Tällöin kalvoilla 3/6-5/6 esitetyille *kriittisille rajoille* pätee:

$$l = -u$$

- Huomautus:

  Kalvojen 3/6-5/6 todennäköisyydet 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*.

Tilastolliset testit

Tilastollinen testaus
Tilastolliset hypoteesit
Tilastolliset testit ja testisuureet
Virheet testauksessa
Merkitsevyystaso ja testin hylkäysalue
>> Testin *p*-arvo
Testin suorittaminen
Esimerkki: Normaalijakauman parametreja koskevat testit
Tilastolliset testit ja mitta-asteikot

p-arvo ja merkitsevyystasot

- Nollahypoteesin hylkääminen voidaan perustaa etukäteen valitun merkitsevyystason ja sitä vastaavan hylkäysalueen määräämisen sijasta testin *p-arvoon*.
- **Testin *p*-arvo on *pienin merkitsevyystaso*, jolla nolla-hypoteesi H_0 voidaan *hylätä*.**
- *Tilastolliset ohjelmistot* tulostavat nykyään lähes aina sovellettavan testin *p*-arvon ja siksi *p*-arvojen käyttö *on lähes kokonaan syrjäyttänyt* etukäteen valittujen kiinteiden merkitsevyystasojen käytön.

***p*-arvo 1/2**

- 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*.

p-arvo 2/2

- 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*.

Testin *p*-arvo

p-arvon frekvenssitulkinta

- Oletetaan, että testausasetelma on sellainen, että yleisen hypoteesin H lisäksi myös *nollahypoteesi* H_0 *pätee*.
- *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 valitessa on syytä ottaa huomioon *väärän päätöksen seuraukset*.

p-arvon määrääminen yksinkertaisissa testausasetelmissa 1/6

- Testin *p-arvo* riippuu yksinkertaisissa testausasetelmissa *vaihtoehtoisen hypoteesin muodosta*.
- Olkoon parametria θ koskeva nollahypoteesi yksinkertaista muotoa

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

- Oletetaan yksinkertaisuuden vuoksi, että testisuureen jakauma on *symmetrinen* origon suhteen.

p-arvon määrääminen yksinkertaisissa testausasetelmissa 2/6

- 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 $b = +\infty$.

(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.

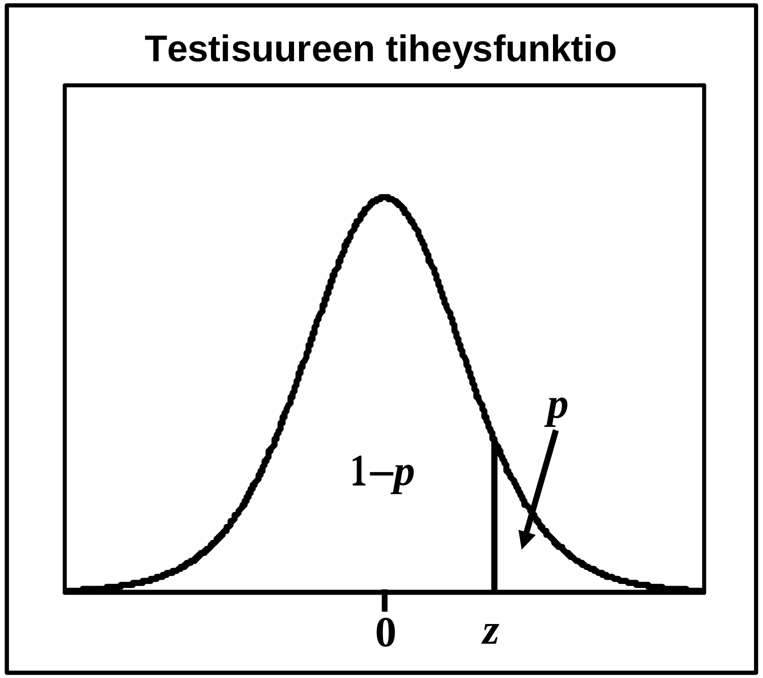
p-arvon määrääminen yksinkertaisissa testausasetelmissa 3/6

- Olkoon testisuureen *Z* *havainnoista määrätty arvo z*.
- Jos vaihtoehtoisena hypoteesina on *yksisuuntainen vaihtoehto*

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

niin testin *p-arvo* on

$$p = \Pr(Z \geq z \mid \mathrm{H}_0)$$



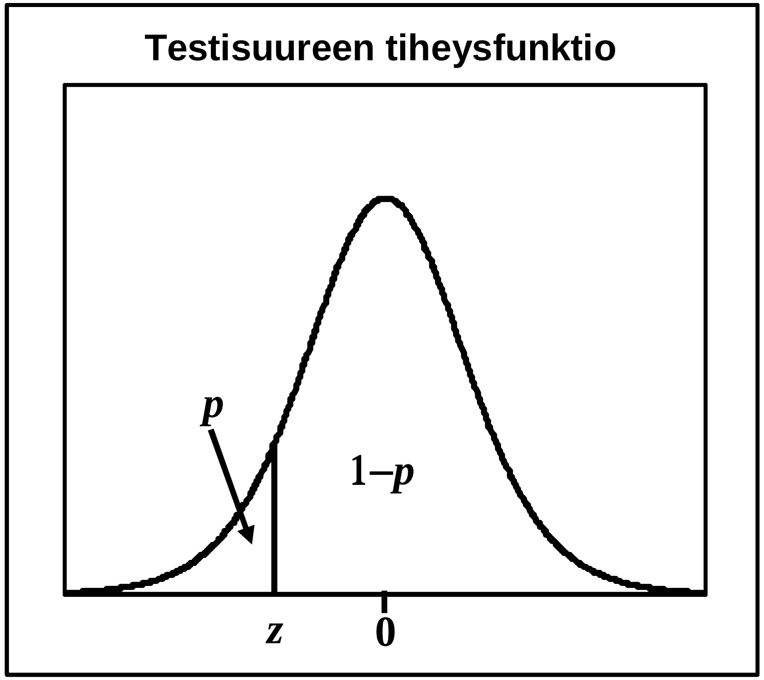
***p*-arvon määrääminen yksinkertaisissa testausasetelmissa 4/6**

- Olkoon testisuureen Z *havainnoista määrätty arvo z*.
- Jos vaihtoehtoisena hypoteesina on *yksisuuntainen vaihtoehto*

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

niin testin *p-arvo* on

$$p = \Pr(Z \leq z \mid \mathrm{H}_0)$$



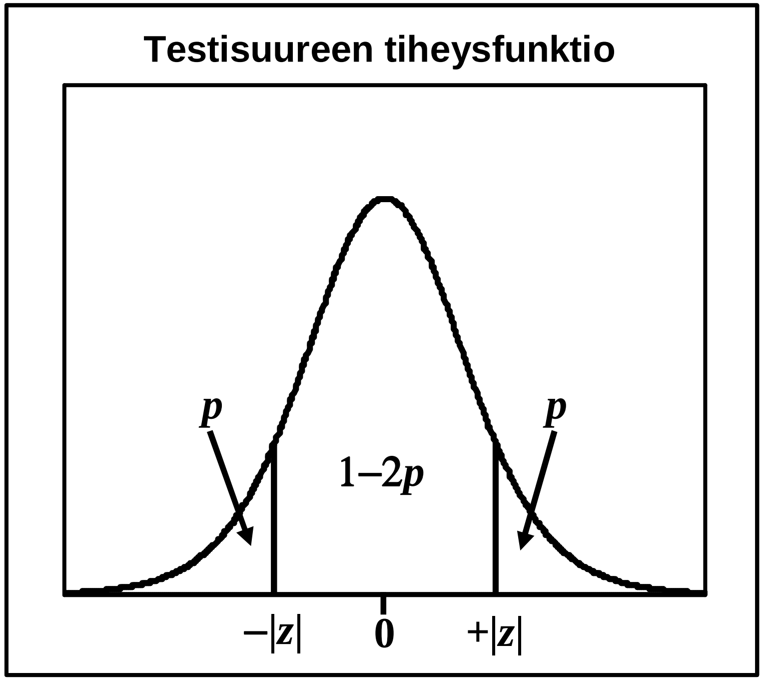
p-arvon määrääminen yksinkertaisissa testausasetelmissa 5/6

- Olkoon testisuureen Z *havainnoista määrätty arvo z*.
- Jos vaihtoehtoisena hypoteesina on *kaksisuuntainen vaihtoehto*

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

niin testin *p-arvo* on

$$2p = 2\times\Pr(Z \geq |z| \mid \mathrm{H}_0)$$



p-arvon määrääminen yksinkertaisissa testausasetelmissa 6/6

- Huomautus:

  Kalvojen 3/6-5/6 todennäköisyydet 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*.

Tilastolliset testit

Tilastollinen testaus
Tilastolliset hypoteesit
Tilastolliset testit ja testisuureet
Virheet testauksessa
Merkitsevyystaso ja testin hylkäysalue
Testin *p*-arvo
>> Testin suorittaminen
Esimerkki: Normaalijakauman parametreja koskevat testit
Tilastolliset testit ja mitta-asteikot

Testin suorittaminen merkitsevyystason valintaan perustuvassa testausmenettelyssä 1/3

- 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*.

Testin suorittaminen merkitsevyystason valintaan perustuvassa testausmenettelyssä 2/3

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

(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.

Testin suorittaminen merkitsevyystason valintaan perustuvassa testausmenettelyssä 3/3

(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 määräämiseen perustuvassa testausmenettelyssä 1/3

- 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*.

Testin suorittaminen *p*-arvon määräämiseen perustuvassa testausmenettelyssä 2/3

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

– Jos havaintojen sisältämät todisteet nolla-hypoteesia 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**.

Testin suorittaminen *p*-arvon määräämiseen perustuvassa testausmenettelyssä 3/3

(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*.

Tilastolliset testit

Tilastollinen testaus

Tilastolliset hypoteesit

Tilastolliset testit ja testisuureet

Virheet testauksessa

Merkitsevyystaso ja testin hylkäysalue

Testin *p*-arvo

Testin suorittaminen

>> Esimerkki: Normaalijakauman parametreja koskevat testit

Tilastolliset testit ja mitta-asteikot

Esimerkki: Normaalijakauman parametreja koskevat testit

Testausasetelma 1/6

- Kone tekee ruuveja, joiden **tavoitepituus** on 10 cm.
- Oletetaan, että ruuvien *pituus vaihtelee satunnaisesti* noudattaen normaalijakaumaa.
- 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 tavoitearvosta 10 cm.

Esimerkki: Normaalijakauman parametreja koskevat testit

Testausasetelma 2/6

- Ruuvien pituutta *valvotaan* seuraavalla tavalla:
  - (i) Jokaisesta valmistuserästä ruuveja poimitaan *yksinkertainen satunnaisotos*.
  - (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 *valmistuserä hylätään*.
- Seuraavassa näytetään, miten ruuvien pituuden valvonnassa käytetään hyväksi **tilastollista testausta**.

Esimerkki: Normaalijakauman parametreja koskevat testit

Testausasetelma 3/6

- Oletetaan, että valmistuserän ruuvien joukosta on poimittu **yksinkertainen satunnaisotos**, jonka *koko* n = 30 ja otokseen poimittujen ruuvien pituudet on mitattu.
- Taulukko oikealla esittää otokseen poimittujen ruuvien pituuksien *luokiteltua frekvenssi-jakaumaa*.
- Huomautus:

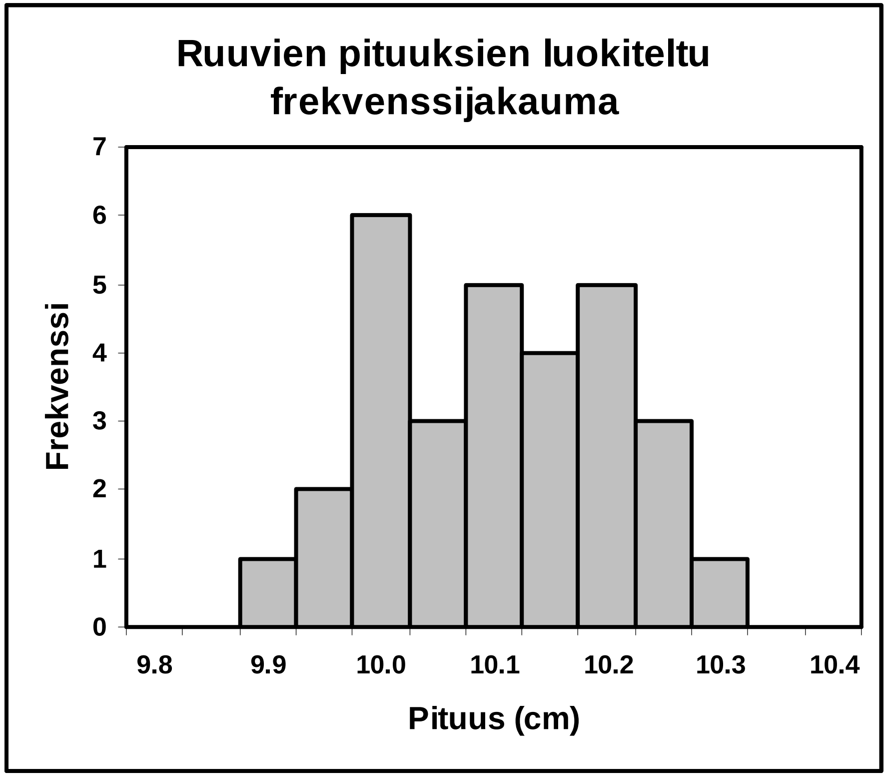
  Aineistoa on käsitelty myös luvuissa **Tilastollisten aineistojen kuvaaminen** ja **Väliestimointi**.

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

Esimerkki: Normaalijakauman parametreja koskevat testit

Testausasetelma 4/6

- Kuva oikealla esittää otokseen poimittujen ruuvien pituuksien luokiteltua frekvenssijakaumaa vastaavaa *histogrammia*.
- *Luokkavälit* määräävät histogrammin suorakaiteiden *kannat*.
- Suorakaiteiden *korkeudet* on valittu niin, että suorakaiteiden *pinta-alat* suhtautuvat toisiinsa kuten vastaavat *luokka-frekvenssit*.



Testausasetelma 5/6

- Yhteenveto otostiedoista:

  Pituuksien *aritmeettinen keskiarvo*:

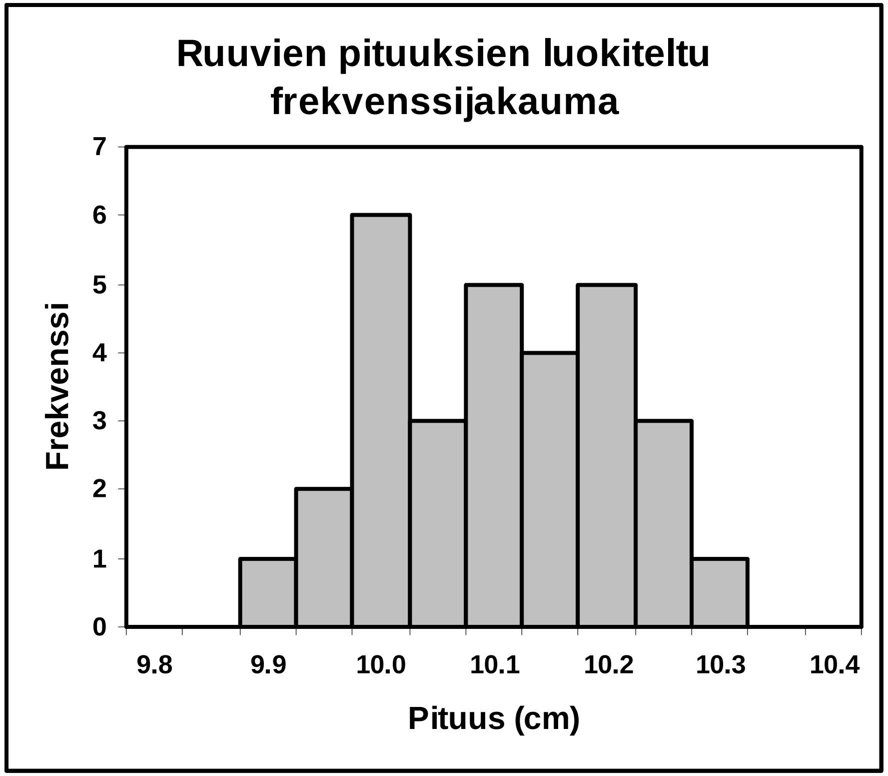
  $$\bar{X} = 10.09 \text{ cm}$$

  Pituuksien *otoskeskihajonta*:

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

- Huomautus:

  Jos otantaa toistetaan, kaikki otoksia koskevat tiedot (sekä havaintoarvot että havaintoarvoista lasketut otostunnusluvut) *vaihtelevat satunnaisesti* otoksesta toiseen.

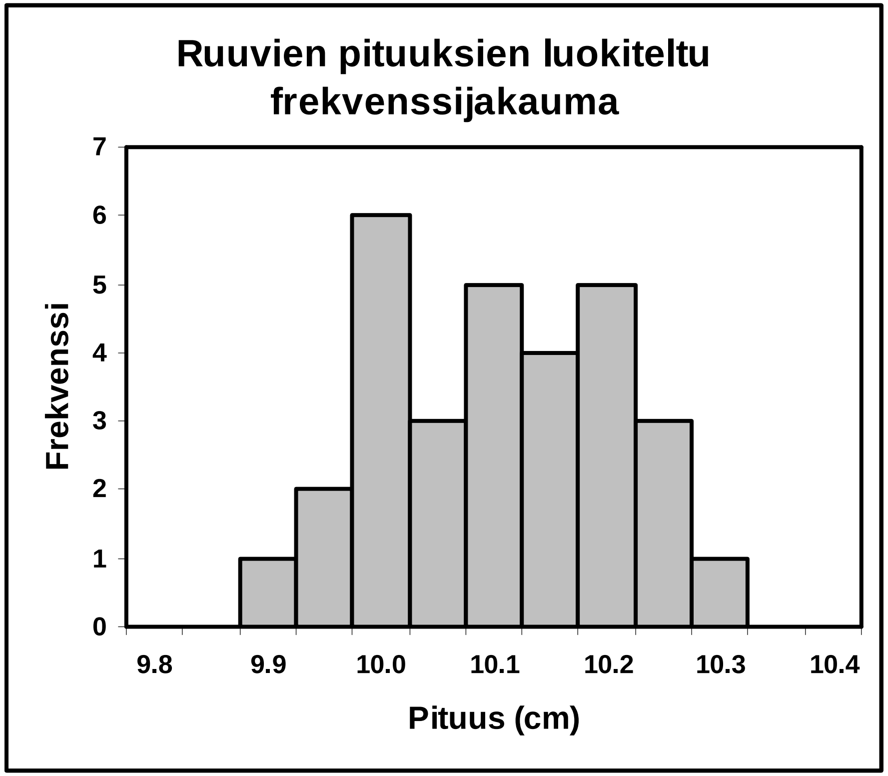


Esimerkki: Normaalijakauman parametreja koskevat testit

Testausasetelma 6/6

- Kysymys:

  Onko otosinformaatio sopusoinnussa ruuvien pituuden varianssille ja odotusarvolle asetettujen tavoitearvojen kanssa?
- Vastataan kysymykseen konstruoimalla tarkoitukseen sopivat **tilastolliset testit**.



Testausasetelmaa koskevat hypoteesit 1/2

- Määritellään satunnaismuuttuja X:

  X = ruuvin pituus

- **Yleinen hypoteesi H** :

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

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

  Pidämme koko testauksen ajan kiinni yleisestä hypoteesista H.

Esimerkki: Normaalijakauman parametreja koskevat testit

Testausasetelmaa koskevat hypoteesit 2/2

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

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

  $$H_{10} : \sigma^2 = \sigma_0^2 \leq 0.01 \text{ cm}^2$$

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

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

  $$H_{11} : \sigma^2 = \sigma_0^2 > 0.01 \text{ cm}^2$$

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

  Ruuvien pituuksien odotusarvo *yhtyy pituuden tavoitearvoon* 10 cm:

  $$H_{20} : \mu = \mu_0 = 10 \text{ cm}$$

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

  Ruuvien pituuksien odotusarvo *poikkeaa pituuden tavoitearvosta* 10 cm:

  $$H_{21} : \mu \neq \mu_0 = 10 \text{ cm}$$

Esimerkki: Normaalijakauman parametreja koskevat testit

Testit

- Ruuvien pituuksien *varianssia* koskevaa nollahypoteesia

$$H_{10} : \sigma^2 = \sigma_0^2 \leq 0.01 \text{ cm}^2$$

  voidaan testata ns. **χ^2-testillä**.

- Ruuvien pituuksien *odotusarvoa* koskevaa nollahypoteesia

$$H_{20} : \mu = \mu_0 = 10 \text{ cm}$$

  voidaan testata ns. ***t*-testillä**.

χ^2-testi varianssille: Hypoteesit

- **Yleinen hypoteesi H** :

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

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

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

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

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

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

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

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

Esimerkki: χ^2-testi varianssille
Testisuure ja sen jakauma 1/3

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

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

jossa

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

havaintojen (harhaton) *otosvarianssi* ja

$$\sigma_0^2$$

nollahypoteesin H_{10} *kiinnittämä parametrin* σ^2 *arvo*.

Esimerkki: χ^2-testi varianssille
Testisuure ja sen jakauma 2/3

- Voidaan osoittaa, että χ^2-*testisuure*

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

  noudattaa **χ^2-jakaumaa** vapaustein $(n - 1)$, *jos yleinen hypoteesi* H ja *oletus*

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

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

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

Esimerkki: χ^2-testi varianssille
Testisuure ja sen jakauma 3/3

- Esimerkin tapauksessa otokseen poimittujen ruuvien pituuksien *aritmeettinen keskiarvo* on

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

  otokseen poimittujen ruuvien pituuksien *otoskeskihajonta* on

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

  ja *nollahypoteesin* H_{10} *kiinnittämä parametrin* σ^2 *arvo* on

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

- Siten χ^2*-testisuureen arvoksi* saadaan

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

Esimerkki: χ^2-testi varianssille

Testisuureen normaaliarvo

- Voidaan osoittaa, että χ^2-*testisuureen*

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

normaaliarvo eli χ^2-testisuureen *odotusarvo oletuksen*

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

pätiessä on

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

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

Esimerkki: χ^2-testi varianssille
Testin hylkäysalueen määrääminen 1/5

- 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** χ^2_α siten, että

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

  jossa satunnaismuuttuja χ^2 noudattaa χ^2-*jakaumaa* vapausastein $n - 1 = 29$.

- Kriittinen raja χ^2_α toteuttaa ehdon

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

Esimerkki: χ^2-testi varianssille

Testin hylkäysalueen määrääminen 2/5

- χ^2-jakauman *taulukoista* nähdään, että

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

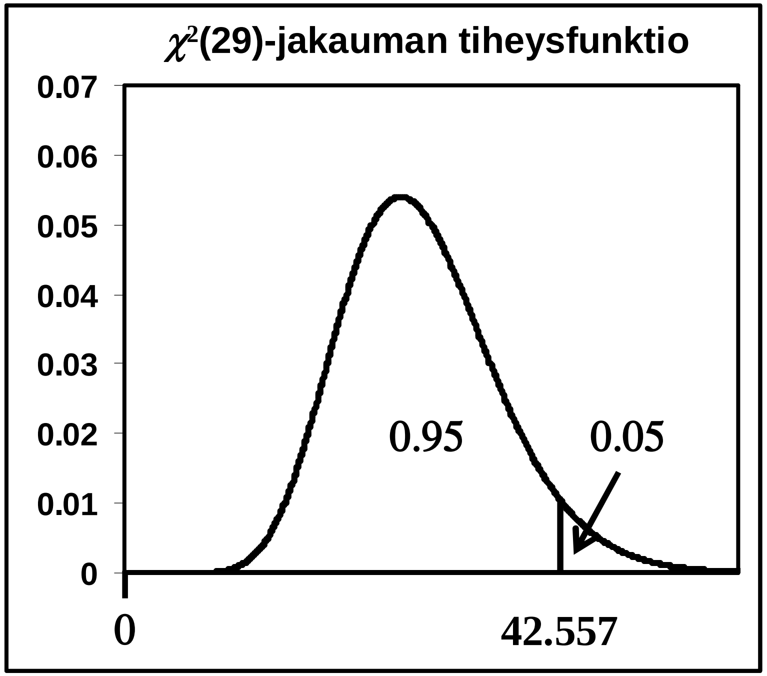
  kun vapausasteiden lukumäärä

  $$n - 1 = 29$$

- Siten *kriittinen raja* on:

  $$\chi^2_{0.05} = 42.557$$

- Kuvio oikealla havainnollistaa kriittisen rajan määräämistä.



Esimerkki: χ^2-testi varianssille

Testin hylkäysalueen määrääminen 3/5

- Valitaan χ^2-testin **hylkäysalueeksi**

$$(\chi^2_\alpha, +\infty)$$

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

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

hylätään merkitsevyystasolla α.

- Todennäköisyys, että χ^2-testisuureen arvo joutuu ehdon $\sigma^2 = \sigma_0^2$ pätiessä hylkäysalueelle on α.
- χ^2-testin **hyväksymisalue** on muotoa

$$[0, \chi^2_\alpha]$$

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

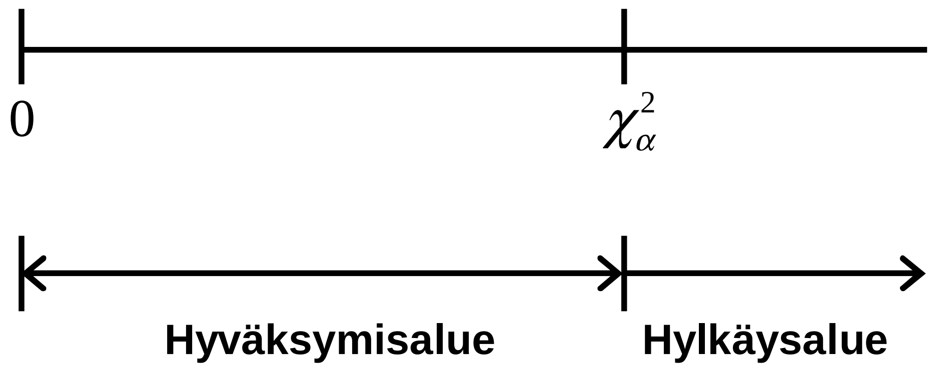
$$\mathrm{H}_{10} : \sigma^2 = \sigma_0^2 \le 0.01 \text{ cm}^2$$

jätetään voimaan merkitsevyystasolla α.

Esimerkki: χ^2-testi varianssille

Testin hylkäysalueen määrääminen 4/5

- χ^2-testin **hylkäys-** ja **hyväksymisalueita** voidaan kuvata *yksisuuntaisen vaihtoehtoisen hypoteesin tapauksessa* alla olevalla kuviolla:



- Kriittinen raja χ^2_α määrätään niin, että

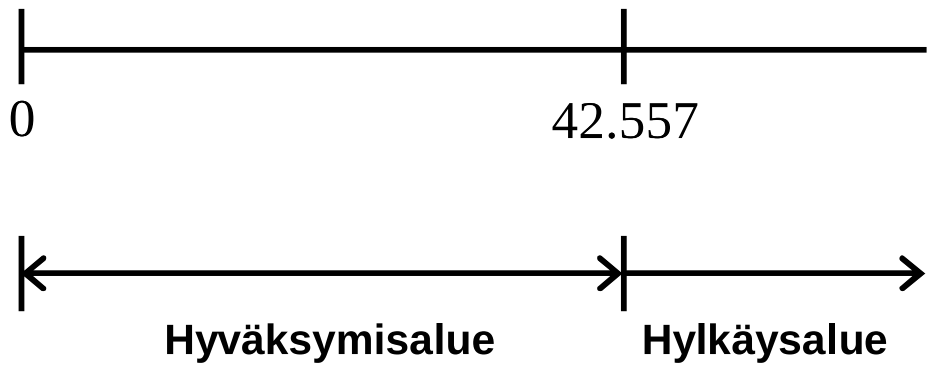
$$\Pr(\chi^2 \geq \chi^2_\alpha) = \alpha$$

jolloin

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

Testin hylkäysalueen määrääminen 5/5

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



- Kriittinen raja 42.557 on määrätty niin, että

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

jolloin

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

Esimerkki: χ^2-testi varianssille

Testin tulos

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

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

- *Koska testisuureen arvo on joutunut hyväksymisalueelle*, **voimme jättää nollahypoteesin**

  $$H_{10} : \sigma^2 = \sigma_0^2 \leq 0.01 \text{ cm}^2$$

  voimaan *merkitsevyystasolla*

  $$\alpha = 0.05$$

- Johtopäätös:

  Ruuvien pituuden varianssi ei ole tilastollisesti merkitsevästi arvoa 0.01 cm² suurempi.

Esimerkki: χ^2-testi varianssille

Merkitsevyystason frekvenssitulkinta

- Oletetaan, että kone tekee jatkuvasti ruuveja, joiden *pituuden varianssi on hyväksyttävän suuruista*.
- Tällöin siis nollahypoteesi

$$H_{10} : \sigma^2 = \sigma_0^2 \leq 0.01 \text{ cm}^2$$

pätee koko ajan.

- Oletetaan, että poimimme koneen tekemien ruuvien joukosta toistuvasti uusia, samankokoisia otoksia ja testaamme nollahypoteesia H_{10} jokaisen otoksen perusteella käyttämällä merkitsevyystasona lukua

$$\alpha = 0.05$$

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

Esimerkki: χ^2-testi varianssille
Testin *p*-arvo

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

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

  mikä merkitsee sitä, että χ^2-testisuure saa normaaliarvoonsa nähden arvoa 31.270 poikkeuksellisempia arvoja todennäköisyydellä, joka on suurempi kuin 0.1, *jos nollahypoteesi*

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

  pätee.
- Siten *emme voi hylätä nollahypoteesia* H_{01} *millään tavanomaisella merkitsevyystasolla*.

t-testi odotusarvolle: Hypoteesit

- **Yleinen hypoteesi H** :

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

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

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

  Ruuvien pituuksien odotusarvo *yhtyy pituuden tavoitearvoon* 10 cm:

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

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

  Ruuvien pituuksien odotusarvo *poikkeaa pituuden tavoitearvosta* 10 cm:

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

Esimerkki: *t*-testi odotusarvolle
Testisuure ja sen jakauma 1/3

- Käytetään **testisuureena** *Studentin 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*

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

havaintojen (harhaton) *otosvarianssi* ja

$$\mu_0$$

nollahypoteesin H_{20} *kiinnittämä parametrin* μ *arvo*.

Esimerkki: *t*-testi odotusarvolle
Testisuure ja sen jakauma 2/3

- Voidaan osoittaa, että *t*-testisuure

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

  noudattaa Studentin ***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 parametrin μ arvon μ_0 *tilastollista etäisyyttä*, jossa mittayksikkönä on aritmeettisen keskiarvon $\bar{X}$ keskivirheen $\sigma/\sqrt{n}$ estimaattori $s/\sqrt{n}$.

Esimerkki: *t*-testi odotusarvolle
Testisuure ja sen jakauma 3/3

- Esimerkin tapauksessa otokseen poimittujen ruuvien pituuksien *aritmeettinen keskiarvo* on

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

  otokseen poimittujen ruuvien pituuksien *otoskeskihajonta* on

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

  ja *nollahypoteesin* H_{20} *kiinnittämä parametrin* μ *arvo* on

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

- Siten *t-testisuureen arvoksi* saadaan

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

Esimerkki: *t*-testi odotusarvolle

Testisuureen normaaliarvo

- Voidaan osoittaa, että *t*-testisuureen

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

normaaliarvo eli *t*-testisuureen *odotusarvo nollahypoteesin*

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

pätiessä on

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

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

  t-testisuureen jakauma on *symmetrinen* origon suhteen.

Esimerkki: *t*-testi odotusarvolle

Testin hylkäysalueen määrääminen 1/5

- Valitaan **merkitsevyystasoksi**

$$\alpha = 0.05$$

- Koska *vaihtoehtoinen hypoteesi* $H_{21} : \mu \neq \mu_0 = 10$ 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 *Studentin t-jakaumaa* vapausastein $n - 1 = 29$.

- Kriittiset rajat $-t_{\alpha/2}$ 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**.

Esimerkki: *t*-testi odotusarvolle

Testin hylkäysalueen määrääminen 2/5

- *t*-jakauman *taulukoista* nähdään, että

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

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

  kun vapausasteiden lukumäärä

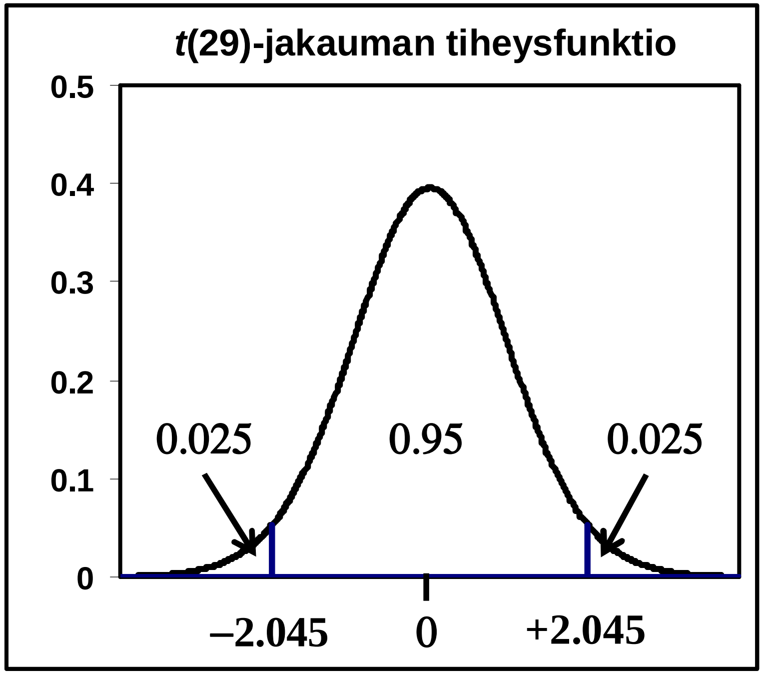
  $$n - 1 = 29$$

- Siten *kriittiset rajat* ovat:

  $$+t_{0.025} = +2.045$$

  $$-t_{0.025} = -2.045$$

- Kuvio oikealla havainnollistaa kriittisten rajojen määräämistä.



Esimerkki: *t*-testi odotusarvolle

Testin hylkäysalueen määrääminen 3/5

- Valitaan *t*-testin **hylkäysalueeksi**

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

- 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*

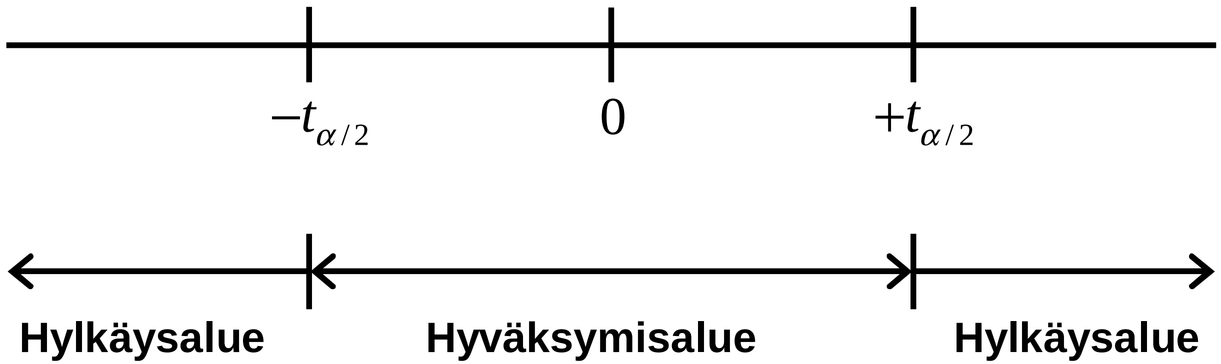
$$\mathrm{H}_{20} : \mu = \mu_0$$

jätetään voimaan merkitsevyystasolla α.

Esimerkki: *t*-testi odotusarvolle

Testin hylkäysalueen määrääminen 4/5

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



- Kriittiset rajat $-t_{\alpha/2}$ ja $+t_{\alpha/2}$ määrätään niin, että

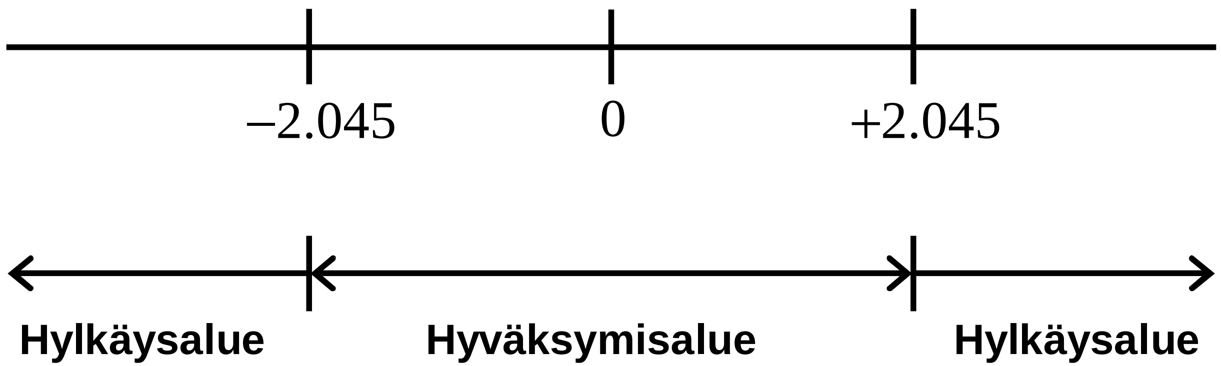
$$\Pr(t \le -t_{\alpha/2}) = \Pr(t \ge +t_{\alpha/2}) = \alpha/2$$

  jolloin

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

Testin hylkäysalueen määrääminen 5/5

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



- Kriittiset rajat –2.045 ja +2.045 on määrätty niin, että

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

jolloin

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

Esimerkki: *t*-testi odotusarvolle

Testin tulos

- 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**

  $$H_0 : \mu = \mu_0 = 10 \text{ cm}$$

  ja **hyväksyä vaihtoehtoisen hypoteesin**

  $$H_0 : \mu \neq \mu_0 = 10 \text{ cm}$$

  merkitsevyystasolla

  $$\alpha = 0.05$$

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

Esimerkki: *t*-testi odotusarvolle
Merkitsevyystason frekvenssitulkinta

- Oletetaan, että kone tekee jatkuvasti *keskimäärin oikean mittaisia ruuveja*.
- Tällöin siis nollahypoteesi

$$H_{20} : \mu = \mu_0 = 10 \text{ cm}$$

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

$$\alpha = 0.05$$

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

Esimerkki: *t*-testi odotusarvolle

Testin *p*-arvo

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

  $$p = 2\times\Pr(t > |4.749|) < 2\times 0.0005 = 0.001$$

  mikä merkitsee sitä, että *t*-testisuure saa normaaliarvoonsa 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*.

Tilastolliset testit

Tilastollinen testaus
Tilastolliset hypoteesit
Tilastolliset testit ja testisuureet
Virheet testauksessa
Merkitsevyystaso ja testin hylkäysalue
Testin *p*-arvo
Testin suorittaminen
Esimerkki: Normaalijakauman parametreja koskevat testit
>> Tilastolliset testit ja mitta-asteikot

Mitta-asteikot ja tilastolliset testit

- Havaintojen *mitta-asteikolliset ominaisuudet ohjaavat testin valintaa*.

  Mitta-asteikot: ks. lukua **Tilastollisten aineistojen kerääminen ja mittaaminen**.

- Tilastolliset testit voidaan ryhmitellä havaintojen mitta-asteikollisten ominaisuuksien suhteen seuraavalla tavalla:
  - **Suhde-** (ja välimatka-) **asteikollisten muuttujien testit**
  - **Järjestysasteikollisten muuttujien testit**
  - **Laatueroasteikollisten muuttujien testit**

Testejä suhdeasteikollisille muuttujille

- **Suhde-** (ja välimatka-) **asteikollisten muuttujien** testejä:
  - **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 varianseille eli varianssien vertailutesti**
- Ks. lukua **Testejä suhdeasteikollisille muuttujille.**

Testejä järjestysasteikollisille muuttujille

- **Järjestysasteikollisten muuttujien** testejä:
  - **Merkkitesti**
  - **Wilcoxonin rankitesti**
  - **Mannin ja Whitneyn testi** eli **Wilcoxonin rankisummatesti**
- Ks. lukua **Testejä järjestysasteikollisille muuttujille.**

Testejä laatueroasteikollisille muuttujille

- **Laatueroasteikollisten muuttujien** testejä:
  - **Testi suhteelliselle osuudelle**
  - **Suhteellisten osuuksien vertailutesti**
- Ks. lukua **Testejä laatueroasteikollisille muuttujille.**