7. TESTITEORIAA NORMAALIJAKAUMALLE

[vrt. Arjas-Sirén, jaksot 4.1-4.3]

Esim. (jatkoa bussiesimerkkiin, s. 23)

Valmistaja on väittänyt, että keskikulutus on (korkeintaan) 34. Kokeessa (s. 23) havaitsimme (otoskoko $n = 20$), että keskikulutus oli $\bar{x} = 36$ ja $s = 3.5$.

Miten suhtautua valmistajan väitteeseen "$\mu = 34$" ja saatuun koetulokseen?

a) Tekikö sattuma vain meille kiusaa?

b) Vai onko väite "$\mu = 34$" syytä hylätä ja päätellä, että $\mu > 34$?

Idea: Jos $\bar{x}$ on "paljon suurempi" kuin 34, päädymme (b):hen. Mitä on "paljon suurempi"? Onko $\bar{x} = 36$ sellainen?

Olet. hetkeksi, että $\mu = 34$ (so. valmistajan väite pätee). Tark. sat. muuttujaa

$$T = \frac{\bar{X} - 34}{S/\sqrt{20}}, \quad \text{joka} \sim t_{19} \quad (\text{s. 22 mukaan})$$

Nyt käsillä olevassa tapauksessa tämän havaittu arvo on

$$t = \frac{\bar{x} - 34}{s/\sqrt{20}} = \frac{36-34}{3.5/\sqrt{20}} \approx 2.56$$

t_{19}-jakauman taulukon mukaan (ks. esim. 3. harjoitukset) näin iso tai vielä isompi T:n arvo saadaan (edelleen olettaen että $\mu = 34$) tn:llä

$$P(T \geq 2.56) < 0.01$$

[taulukon mukaan $P(T \geq 2.539) \approx 0.01$]

Siis alle 1 %!

Johtopäätös: Vaihtoehto (a) ei ole kovin uskottava vaan päädymme (b):hen. Virheellisen päätelmän riski on alle 1 %.

Yo. menettely on esimerkki tilastollisesta testaamisesta, tarkemmin ns. t-testistä. Esittelemme sen lyhyesti.

t-testi normaalijakauman odotusarvolle

Mallina edelleen $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ $\perp\!\!\!\perp$
(sekä μ että σ^2 tuntemattomia parametreja)

On formuloitu hypoteesit

$$(*) \quad \begin{cases} H_0: \ \mu = \mu_0 \quad (\text{tai } \mu \le \mu_0) & \text{nollahypoteesi} \\ H_1: \ \mu > \mu_0 & \text{vastahypoteesi} \end{cases}$$

Havaitun aineiston $\underset{\sim}{x} = (x_1, \dots, x_n)$ perusteella on testattava, voidaanko H_0:aa pitää totena vai onko $\underset{\sim}{x}$ ristiriidassa sen kanssa ja tukee paremminkin H_1:tä.

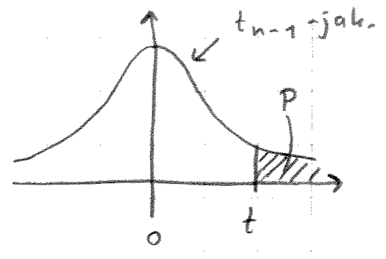
Muodostetaan t-testisuure

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \quad \text{jossa} \quad \begin{cases} \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \\ s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2} \end{cases}$$

Tämän suuret (positiiviset) arvot ovat kriittisiä H_0:lle ja tukevat H_1:tä !

Jos H_0 pätee eli $\mu = \mu_0$, vastaava sm T noudattaa t_{n-1}-jakaumaa (ks. s. 22). Tällöin voidaan laskea tn

$$p = P_{H_0}(T \ge t)$$



t_{n-1}-jakauman "oikeana häntätodennäköisyytenä"
(merkintä P_{H_0} korostaa että kyseinen tn lasketaan olettaen että H_0 pätee)

p on nimeltään testin p-arvo.

Tulkinta: Jos p on hyvin pieni, nyt havaitunlainen — tai vielä suurempi eli "kriittisempi" — t-testisuureen arvo on H_0:n pätressä hyvin epätodennäköinen. Tällöin H_0 voidaan hylätä ja H_1 hyväksytään

Yleensä verrataan p:tä ennaltavalittuun merkitsevyystasoon α (tyypillisesti 0.05 tai 0.01). Jos $p \leq \alpha$, sanotaan, että H_0 hylätään (ja H_1 hyväksytään) merkitsevyystasolla α. Muuten H_0 jää voimaan.

(Yleensä H_0 edustaa "neutraalia" tai "oletusarvoista" asiaintilaa, joten sen hylkäämiseen vaaditaan kyllin "vahva näyttö": virheellisen hylkäyspäätöksen riski α on siksi aiheellista valita varsin pieneksi.)

Esim. Sivun 24 esimerkissä oli $H_0: \mu = 34$ ja $H_1: \mu > 34$. Havaittiin t-testisuureen arvo $t \approx 2.56$ ja p-arvo $p = P_{H_0}(T \geq 2.56)$ oli alle 0.01, joten H_0 voitiin hylätä ja H_1 hyväksyä vaikkapa merkitsevyystasolla $\alpha = 0.01$.

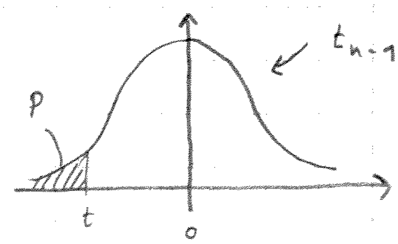
Hypoteesien vaihtoehtoisia muotoja:

1. Jos (*):n sijasta hypoteesit ovat muotoa

$$\begin{cases} H_0: \mu = \mu_0 \quad (\text{tai } \mu \geq \mu_0) \\ H_1: \mu < \mu_0 \end{cases}$$

H_0:lle kriittisiä ja H_1:tä tukevia ovat t:n pienet (voimakkaasti negatiiviset) arvot. Tällöin p-arvo on "vasen häntätodennäköisyys"

$$p = P_{H_0}(T \leq t)$$



Muuten menetellään kuten edellä.

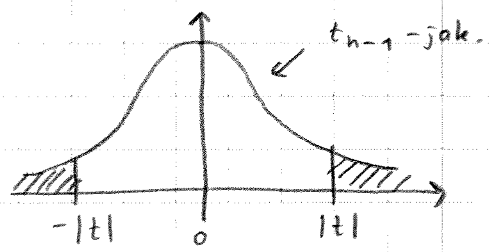
2. Ns. <u>kaksisuuntaisessa testausasetelmassa</u> hypoteesit ovat muotoa

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

H_0:lle kriittisiä ja H_1:tä tukevia ovat nyt <u>sekä</u> t:n pienet (negatiiviset) <u>että</u> suuret (positiiviset) arvot. p-arvo lasketaan "molemmilta hänniltä":

$$p = P_{H_0}(|T| \geq |t|)$$
$$= 2\, P_{H_0}(T \geq |t|)$$

↑ symmetrian nojalla!



<u>Huom.</u> p-arvon tulkinnasta:

* p-arvon määritelmän todennäköisyyskäsite viittaa aineistoon, ei parametriin μ.

* p-arvo <u>ei</u> ole "tn sille, että H_0 pätee" <u>vaan</u> se on tn saada testisuureelle sellainen arvo (tai vielä poikkeavampi arvo) kuin nyt aineistosta on saatu, <u>olettaen että H_0 pätee</u>.

* Siispä: hyvin pieni p-arvo <u>ei</u> tarkoita, että H_0 on "hyvin epätodennäköinen" <u>vaan</u>, että on saatu sellainen aineisto, joka H_0:n pätiessä on hyvin epätodennäköinen!

<u>Huom.</u> Mikäli σ^2 on <u>tunnettu</u> luku, voidaan t-testisuureen sijasta käyttää testisuureena

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

jolloin p-arvot lasketaan $N(0,1)$-jakaumasta t_{n-1}-jakauman sijaan.
[Vrt. tilanne luottamusvälejä muodostettaessa jaksossa 6.]