

Statisztika II - hipotézisvizsgálat

Sztochasztika

Horváth Illés

2024/11/28

- (1) Hipotézisvizsgálat általában
- (2) A próba szerkezete
- (3) Próbák a várható értékre (u -próba, t -próba)
- (4) Nemparaméteres próbák (χ^2 -próbák)

Egy zacskó cukor névleges tömege 1000 g. Azt szeretnénk ellenőrizni, hogy a cukorgyár nem csal-e azzal, hogy kevesebb cukrot tölt a zacskókba.

A csomagolási technológia miatt az egy zacskóba kerülő cukor mennyisége véletlenszerű, ezért az, hogy egy-két zacskóban kevesebb a cukor tömege, mint 1000 g, még nem feltétlen jelent tendeciózus csalást.

Általában az ilyen helyzetet úgy vizsgáljuk, hogy veszünk egy mintát, és az alapján hozunk egy döntést: elfogadjuk vagy elutasítjuk az adott információt. Ez a *hipotézisvizsgálat*.

Példa. Megmérjük 5 zacskóban a cukor tömegét, és a következőket kapjuk: 986, 992, 1003, 976, 968 g. Ez alapján elfogadjuk-e, hogy a zacskókba kerülő cukor várható értéke 1000 g?

Az általános keret a következő. Adott egy kezdeti hipotézis (információ), amit tesztelni szeretnénk. Ez a *nullhipotézis*, amit H_0 -al jelölünk. Egy bináris döntést akarunk hozni arról, hogy H_0 elfogadható-e vagy nem. A cukros példában

- H_0 : az egy zacskóba kerülő cukor tömegének várható értéke 1000 g.

Az *ellenhipotézis* a másik döntési lehetőség, amit H_1 -gyel jelölünk. Gyakran ez a nullhipotézis komplementere, de ez nem szükségszerű, elég az, hogy H_0 és H_1 nem teljesülhet egyszerre. A cukros példában egy természetes választás H_1 -re

- H_1 : az egy zacskóba kerülő cukor tömegének várható értéke nem egyenlő 1000 grammal.

Egy másik lehetséges választás H_1 -re

- H_1 : az egy zacskóba kerülő cukor tömegének várható értéke kevesebb, mint 1000 g.

Ezután veszünk egy mintát, és a minta alapján végrehajtunk egy próbát. A próba általában a mintán, H_0 -on és H_1 -en alapuló számítás, ami egy döntést eredményez: vagy

- elfogadjuk H_0 -t, vagy
- elvetjük H_0 -t és H_1 -et fogadjuk el.

A próba tulajdonképpen felosztja a lehetséges minták terét két tartományra: az elfogadási tartományban elfogadjuk H_0 -t, a másik tartományban elvetjük H_0 -t H_1 javára.

A lehetőségek:

	H_0 igaz	H_0 hamis
elfogadjuk H_0 -t	✓	másodfajú hiba
elvetjük H_0 -t	elsőfajú hiba	✓

Ideális esetben egy próba 0 elsőfajú és 0 másodfajú hibát eredményezne, de ez nem lehetséges a minta véletlensége miatt.

Két triviális próba következik.

Ha H_0 -t mindig elfogadjuk, a mintától függetlenül, akkor sosem vétünk elsőfajú hibát.

Ha ellenben H_0 -t mindig elvetjük, sosem vétünk másodfajú hibát.

A tipikus próba a kettő között van: elfogadjuk H_0 -t bizonyos mintákra, és elvetjük H_0 -t más mintákra.

Van egy egyensúly: ha az elfogadási tartomány nagyobb, akkor az elsőfajú hiba valószínűsége kisebb és a másodfajú hiba valószínűsége nagyobb, míg ha az elfogadási tartomány kisebb, akkor fordítva.

Egy próbánál a döntést fel lehet úgy is fogni, hogy azt vizsgáljuk, hogy a minta tipikusnak tekinthető-e H_0 szerint. Ha igen, elfogadjuk H_0 -t, ha nem, elvetjük H_0 -t.

Az elfogadási tartomány méretét a *szignifikanciaszint* (más néven konfidenciaszint) révén adjuk meg. Egy próba szignifikanciaszintje egy 0 és 100% közötti érték; például 95%-os szignifikanciaszint azt jelenti, hogy H_0 -t elfogadjuk a minta alapján, ha beleesik az összes minta közül a H_0 szerint legtipikusabb 95%-ba.

A szignifikanciaszint tehát a H_0 -ba vetett bizalom szintje; az értékét mi döntjük el előre. Semmiképpen sem jellemzi azt, hogy mennyire jó a próba.

Annak eldöntése, hogy egy minta tipikusnak tekinthető-e, általában egy statisztika kiszámításával történik, amiről már könnyebb eldönteni, hogy tipikus-e (H_0 teljesülése esetén).

A legtöbb próba szerkezete a következő:

- kiszámítunk egy *statisztikát* a mintából,
- kiszámítunk egy *percentilist* (más néven kvantilist) a szignifikanciaszint és a statisztika H_0 szerinti elméleti eloszlása alapján,
- majd a döntést a statisztika és percentilis összehasonlítása alapján hozzuk meg.

A szignifikanciaszint az elsőfajú hibát kontrollálja; ha a szignifikanciaszint $1 - \varepsilon$, akkor az elsőfajú hiba valószínűsége ε . A másodfajú hibát nem kontrolláljuk.

Statisztikai programok általában a következő módon működnek: az adott szignifikanciaszinthez tartozó percentilis kiszámítása helyett azt adják meg, mi a legkisebb szignifikanciaszint, amin H_0 -t még elfogadjuk az adott minta alapján. Ez az érték az ún. *p-értéke* a mintának. Ezután a *p-érték* közvetlenül összehasonlítható a kívánt szignifikanciaszinttel.

A következő próba az *egymintás, kétoldali u -próba* (angolul z -test), amivel a minta várható értékét lehet tesztelni egy rögzített érték ellenében. A CHT-n alapul.

Legyen adott egy $X_1, \dots, X_n$ fae minta $\mathbb{D}(X_1) = \sigma$ ismert szórással és $\mathbb{E}(X_1) = m$ ismeretlen várható értékkel. Legyen

- $H_0: m = \mu$, ahol μ egy adott érték;
- $H_1: m \neq \mu$.

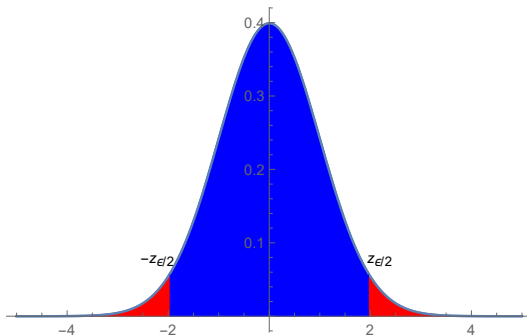
Ahhoz, hogy H_0 -t teszteljük H_1 ellen $1 - \varepsilon$ szignifikanciaszinten, a következőket végezzük el:

- kiszámítjuk az $u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n}$ statisztikát a mintából;
- kiszámítjuk az $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$ percentilist a standard normális eloszlás táblázatából és a szignifikanciaszintből, és
- ha $u \in [-u_{\varepsilon/2}, u_{\varepsilon/2}]$ teljesül, elfogadjuk H_0 -t; ha nem teljesül, elvetjük H_0 -t.

Próbák a várható értékre – egymintás, kétoldali u -próba

Ha u elég közel van 0-hoz, akkor az u és $u_{\varepsilon}/2$ eltérést elfogadjuk a véletlen mintavételből származónak, és elfogadjuk H_0 -t.

Ha u túl távol van 0-tól, akkor nem fogadjuk el, hogy az eltérés a véletlen mintavételből származik, és elvetjük H_0 -t.



A cukros példában a minta

$$X_1 = 986, \quad X_2 = 992, \quad X_3 = 1003, \quad X_4 = 976, \quad X_5 = 968.$$

Tegyük fel, hogy $\sigma = 20$ ismert a csomagolási technológia révén.

- $H_0: m = 1000$;
- $H_1: m \neq 1000$.

Teszteljük H_0 -t H_1 ellenében 95% szignifikanciaszinten.

Először kiszámítjuk a statisztikát:

$$u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{985 - 1000}{20} \sqrt{5} = -1.677.$$

Majd kiszámítjuk a percentilist:

$$u_{\varepsilon/2} = \Phi^{-1}(0.975) = 1.96.$$

Végül ellenőrizzük, hogy $u \in [-u_{\varepsilon/2}, u_{\varepsilon/2}]$ teljesül-e.

$$-1.677 \in [-1.96, 1.96]$$

teljesül, így elfogadjuk H_0 -t 95%-os szignifikanciaszinten.

Próbák a várható értékre – egymintás, egyoldali u -próba

Legyen adott egy $X_1, \dots, X_n$ fae minta, ahol $\mathbb{D}(X_1) = \sigma$ ismert és $\mathbb{E}(X_1) = m$ ismeretlen. Legyen

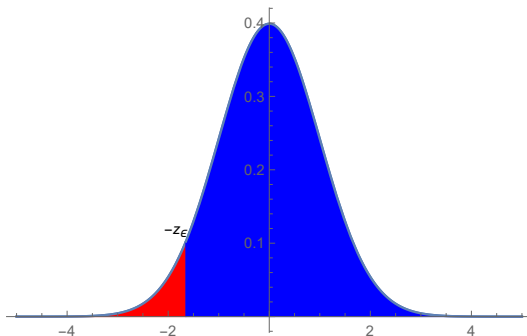
- H_0 : $m = \mu$, ahol μ adott konstans;
- H_1 : $m < \mu$.

Ahhoz, hogy H_0 -t teszteljük H_1 ellenében $1 - \varepsilon$ szignifikanciaszinten, a következőket csináljuk:

- kiszámítjuk az $u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n}$ statisztikát a mintából;
- kiszámítjuk a $u_\varepsilon = \Phi^{-1}(1 - \varepsilon)$ percentilist a standard normális eloszlás táblázatából és a szignifikanciaszintből, és
- ha $u \in [-u_\varepsilon, \infty)$ teljesül, elfogadjuk H_0 -t; ha nem, akkor elvetjük H_0 -t H_1 javára.

Próbák a várható értékre – egymintás, egyoldali u -próba

Ha u túl távol van 0-tól **balra**, akkor nem fogadjuk el, hogy az eltérés a véletlen mintavételből származik, és így elvetjük H_0 -t H_1 javára. (Az ábrán a teljes ε valószínűséget a bal oldalra csoportosítjuk.)



Próbák a várható értékre – egymintás, egyoldali u -próba

Legyen adott egy $X_1, \dots, X_n$ fae minta, ahol $\mathbb{D}(X_1) = \sigma$ ismert és $\mathbb{E}(X_1) = m$ ismeretlen. Legyen

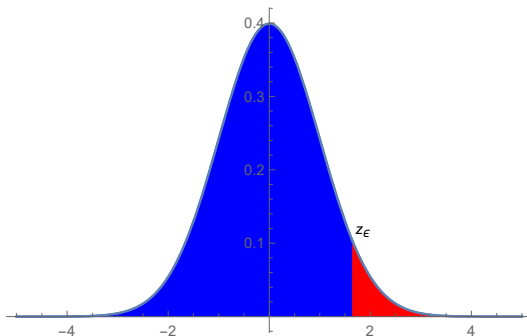
- H_0 : $m = \mu$, ahol μ adott konstans;
- H_1 : $m > \mu$.

Ahhoz, hogy H_0 -t teszteljük H_1 ellenében $1 - \varepsilon$ szignifikanciaszinten, a következőket csináljuk:

- kiszámítjuk az $u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n}$ statisztikát a mintából;
- kiszámítjuk a $u_\varepsilon = \Phi^{-1}(1 - \varepsilon)$ percentilist a standard normális eloszlás táblázatából és a szignifikanciaszintből, és
- ha $u \in (-\infty, u_\varepsilon]$ teljesül, elfogadjuk H_0 -t; ha nem, akkor elvetjük H_0 -t H_1 javára.

Próbák a várható értékre – egymintás, egyoldali u -próba

Ha u túl távol van 0-tól **jobbra**, akkor nem fogadjuk el, hogy az eltérés a véletlen mintavételből származik, és így elvetjük H_0 -t H_1 javára. (Az ábrán a teljes ε valószínűséget a jobb oldalra csoportosítjuk.)



Az u -próbák a CHT-n alapulnak, ami garantálja, hogy az $u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n}$ statisztika közel $N(0,1)$ eloszlású, ha H_0 teljesül.

Általában a legtöbb próba esetében van egy határeloszlás-tétel a háttérben. Lehet másik tétel, mint a CHT, és a határeloszlás is lehet $N(0,1)$ -től eltérő. Ilyenkor a percentilis értékét egy másik eloszláshoz tartozó táblázatból vesszük.

Mindazonáltal a legtöbb próbának a szerkezete ugyanaz:

- kiszámítjuk a statisztikát a minta alapján;
- kiszámítjuk a percentilist a releváns eloszlás táblázata és a szignifikanciaszint alapján, és
- összehasonlítjuk a statisztikát és a percentilist, ami alapján elfogadjuk vagy elvetjük H_0 -t.

Mi erre a szerkezetre fogunk koncentrálni, nem a háttérben meghúzódó határeloszlás-tételekre.

Próbák a várható értékre – kétmintás, kétoldali u -próba

A kétmintás, kétoldali u -próba esetében 2 különálló minta adott:

- $X_1, \dots, X_{n_1}$ fae, ismert σ_1 szórással és ismeretlen m_1 várható értékkel;
- $Y_1, \dots, Y_{n_2}$ fae (és az X -ektől is független), ismert σ_2 szórással és ismeretlen m_2 várható értékkel.

A próba célja a két minta várható értékének összevetése:

- $H_0: m_1 = m_2$;
- $H_1: m_1 \neq m_2$;

A próba a következő:

- a mintákból kiszámítjuk a statisztikát:

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- $N(0,1)$ táblázatából és az $1 - \varepsilon$ szignifikanciaszintből kiszámítjuk az $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$ percentilist;
- ha $u \in [-u_{\varepsilon/2}, u_{\varepsilon/2}]$ teljesül, elfogadjuk H_0 -t, egyébként elvetjük H_0 -t.

Próbák a várható értékre – kétmintás, egyoldali u -próba

A kétmintás, egyoldali u -próba esetében 2 különálló minta adott:

- $X_1, \dots, X_n$ fae, ismert σ_1 szórással és ismeretlen m_1 várható értékkel;
- $Y_1, \dots, Y_m$ fae (és az X -ektől is független), ismert σ_2 szórással és ismeretlen m_2 várható értékkel.

A próba célja a két minta várható értékének összevetése:

- $H_0: m_1 = m_2$;
- $H_1: m_1 > m_2$;

A próba a következő:

- a mintákból kiszámítjuk a statisztikát:

$$u = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- $N(0,1)$ táblázatából és az $1 - \varepsilon$ szignifikanciaszintből kiszámítjuk az $u_\varepsilon = \Phi^{-1}(1 - \varepsilon)$ percentilist;
- ha $u \in (-\infty, u_\varepsilon]$ teljesül, elfogadjuk H_0 -t, egyébként elvetjük H_0 -t.

Próbák a várható értékre – t -próba

Az u -próbában mindig feltettük, hogy a σ szórás ismert. De mi a helyzet, ha σ nem ismert?

Esetleg használhatjuk σ helyett a minta s_n empirikus szórását vagy az s_n^* korrigált empirikus szórást (az utóbbit fogjuk). Ez lesz a t -próba.

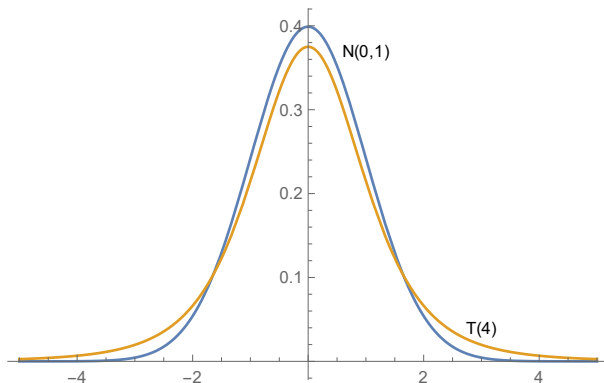
A fő különbség az u -próbához képest, hogy amíg n kicsi, addig $t = \frac{\bar{x} - \mu}{s_n^*} \sqrt{n}$ eloszlása $N(0,1)$ helyett a *Student-féle t -eloszlás* vagy egyszerűen csak t -eloszlás. Ennek megfelelően a percentilis is a t -eloszlás táblázatából jön majd.

Minden n -re van egy külön t -eloszlás, ezért a *szabadsági fokot* is könyvelnünk kell: ha a minta mérete n , akkor a percentilist az $n - 1$ szabadsági fokú t -eloszlás táblázatából vesszük.

Amint $n \rightarrow \infty$, a t -eloszlás tart az $N(0,1)$ eloszláshoz, így ha n nagy (pl. $n > 30$), az u -próba és a t -próba gyakorlatilag megegyezik.

Próbák a várható értékre – t -próba

Az $N(0,1)$ eloszlás és a 4 szabadsági fokú t -eloszlás sűrűségfüggvényeinek összehasonlítása:



Legyen adott egy $X_1, \dots, X_n$ fae minta ismeretlen $\mathbb{D}(X_1)$ szórással és ismeretlen $\mathbb{E}(X_1) = m$ várható értékkel. Legyen

- $H_0: m = \mu$, ahol μ egy adott érték;
- $H_1: m \neq \mu$.

Ahhoz, hogy H_0 -t teszteljük H_1 ellen $1 - \varepsilon$ szignifikanciaszinten, a következőket végezzük el:

- kiszámítjuk a $t = \frac{\bar{x} - \mu}{s_n^*} \sqrt{n}$ statisztikát a mintából;
- kiszámítjuk a $t_{\varepsilon/2}$ percentilist az $n - 1$ szabadsági fokú t -eloszlás táblázatából és a szignifikanciaszintből, és
- ha $t \in [-t_{\varepsilon/2}, t_{\varepsilon/2}]$ teljesül, elfogadjuk H_0 -t; ha nem teljesül, elvetjük H_0 -t.

Próbák a várható értékre – egymintás, egyoldali t -próba

Legyen adott egy $X_1, \dots, X_n$ fae minta ismeretlen $\mathbb{D}(X_1)$ szórással és ismeretlen $\mathbb{E}(X_1) = m$ várható értékkel. Legyen

- $H_0: m = \mu$, ahol μ egy adott érték;
- $H_1: m > \mu$.

Ahhoz, hogy H_0 -t teszteljük H_1 ellen $1 - \varepsilon$ szignifikanciaszinten, a következőket végezzük el:

- kiszámítjuk a $t = \frac{\bar{x} - \mu}{s_n^*} \sqrt{n}$ statisztikát a mintából;
- kiszámítjuk a t_ε percentilist az $n - 1$ szabadsági fokú t -eloszlás táblázatából és a szignifikanciaszintből, és
- ha $t \in (-\infty, t_\varepsilon]$ teljesül, elfogadjuk H_0 -t; ha nem teljesül, elvetjük H_0 -t.

A kétmintás t -próbákhoz a hipotézisvizsgálatos táblázatra hivatkozunk.

A lila bab fogyasztása gyógyítja a rákot!!!4!

Tanmese következik. Egy újságíró megkeres egy tudóst, hogy szerinte a bab fogyasztása gyógyítja a rákot.

A tudós csinál egy tesztet, amiben azt teszteli, hogy a babfogyasztás hatása a rák gyógyulására pozitívnak tekinthető-e (azaz H_0 az, hogy a hatás 0, H_1 az, hogy pozitív). 95%-os konfidenciaszinten azt kapja, hogy a hatás nulla.

Erre az újságíró azt mondja, hogy de ő úgy tudja, hogy a bab színe is számít. Erre a tudós csinál tesztet, amiben 20 különböző színű babot (fehér, barna, lila stb.) külön-külön tesztel.

19 színnél az jött ki, hogy 95%-os konfidenciaszinten a hatás nulla, de egy színnél, a lila babnál az jött ki, hogy 95%-os konfidenciaszinten a lila bab hatása pozitív a rák gyógyulására.

Tanulságok. 1. Ne teszteljünk túl! 2. Nagy horderejű, pl. gyógyszeripari tesztek magas konfidenciaszinten végezzünk (pl. 99,99%).

Az egymintás, kétoldali u - és t -próbákhoz lehet számolni konfidenciaintervallumot is. A konfidenciaintervallum egy, a minta alapján számítható intervallum (azaz nem pontbecslés, hanem intervallumbecslés), az u -próba esetén

$$\left[\bar{x} - u_{\varepsilon/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + u_{\varepsilon/2} \frac{\sigma}{\sqrt{n}} \right],$$

míg a t -próba esetén

$$\left[\bar{x} - t_{\varepsilon/2} \frac{s_n^*}{\sqrt{n}}, \bar{x} + t_{\varepsilon/2} \frac{s_n^*}{\sqrt{n}} \right].$$

A naiv interpretáció az, hogy a konfidenciaintervallum $1 - \varepsilon$ valószínűséggel tartalmazza a valódi várható értéket.

Ez igaz is meg nem is.

Formálisan például az u -próba, azaz ismert σ szórás esetén teljesül, hogy akármilyen is a valódi m várható érték,

$$\mathbb{P}_m \left(m \in \left[\bar{x} - u_{\varepsilon/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + u_{\varepsilon/2} \frac{\sigma}{\sqrt{n}} \right] \right) = 1 - \varepsilon.$$

Azonban a véletlen nem az m paraméter értékében van, hanem az $\bar{x}$ mintaátlagban.

Tehát amíg a mintát nem ismerjük, addig a fenti valószínűség teljesül, de amint a mintát megismertük, onnantól már nincs véletlen a kérdésben; $m \in \left[\bar{x} - u_{\varepsilon/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + u_{\varepsilon/2} \frac{\sigma}{\sqrt{n}} \right]$ vagy teljesül, vagy nem.

Egy másik lehetséges interpretáció, hogy a konfidenciaintervallum azokat az m értékeket tartalmazza, amelyeket elfogadnánk a H_0 nullhipotézisben az adott minta esetén.

2. feladat

Egy gyárban a cementet 25 kg névleges súlyú zsákokba csomagolják. A csomagolási technológiából eredően az egy zsákba kerülő cement súlyának szórása 0.5 kg, a várható értéke azonban ismeretlen, jelölje m . Megvizsgálunk 25 zsákot, és azt tapasztaljuk, hogy a bennük lévő cement mennyisége átlagosan 24.82 kg.

- (a) Elfogadjuk-e 95%-os szinten az $m = 25$ hipotézist az $m \neq 25$ hipotézis ellenében?
- (b) Elfogadjuk-e 90%-os szinten az $m = 25$ hipotézist az $m \neq 25$ hipotézis ellenében?
- (c) Tegyük fel, hogy a szórás csak 0.3 kg. Elfogadjuk-e ekkor az 95%-os szinten az $m = 25$ hipotézist az $m \neq 25$ hipotézis ellenében?

2. feladat

Megoldás.

(a) Milyen próbát alkalmazzunk?

- σ ismert, tehát u -próba;
- egy minta várható értékét teszteljük egy rögzített érték ellen, tehát egymintás u -próba;
- H_1 -ben $\mu \neq 25$ szerepel, tehát kétoldali egymintás u -próba.

A nullhipotézis és az ellenhipotézis

- $H_0: \mu = 25$,
- $H_1: \mu \neq 25$.

Kiszámítjuk a statisztikát:

$$u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{24.82 - 25}{0.5} \sqrt{25} = -1.8.$$

Kiszámítjuk a percentilist:

$$u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2) = \Phi^{-1}(0.975) = 1.96.$$

2. feladat

(a) Az összehasonlítás:

$$u = -1.8 \in [-u_{\varepsilon/2}, u_{\varepsilon/2}] = [-1.96, 1.96]$$

teljesül, tehát elfogadjuk H_0 -t 95%-os szignifikanciaszinten.

(b) A különbség az, hogy a 90%-os szignifikanciaszinthez tartozó percentilis

$$u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2) = \Phi^{-1}(0.95) = 1.65,$$

és

$$u = -1.8 \in [-u_{\varepsilon/2}, u_{\varepsilon/2}] = [-1.65, 1.65]$$

már nem teljesül, tehát 90%-os szignifikanciaszinten H_0 -t elvetjük.

Szokás azt is mondani, hogy az elméleti várható értéktől való eltérés 90%-os szinten szignifikáns (de 95%-os szinten már nem). Általában minél magasabb szignifikanciaszinten vetjük el H_0 -t egy mintára, annál erősebben utasítjuk el H_0 -t.

(c) Ha $\sigma = 0.3$, akkor a statisztika ezúttal

$$u = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{24.82 - 25}{0.3} \sqrt{25} = -3,$$

és

$$u = -3 \in [-u_{\varepsilon/2}, u_{\varepsilon/2}] = [-1.96, 1.96]$$

nem teljesül, tehát H_0 -t 95% szignifikanciaszinten elvetjük.

4. feladat

Egy sóoldat koncentrációját mérjük. 5 mérés eredménye a következő (g/l): 7.7, 8.1, 7.7, 7.5, 7.0. Korábban azt az információt kaptuk, hogy az oldat koncentrációja 7.2 g/l. Elfogadjuk-e ezt 95%-os szinten azon hipotézis ellenében, hogy a koncentráció nem 7.2 g/l? Mi lenne a helyzet a következő mintával: 7.5, 7.4, 7.3, 7.4, 7.5?

Megoldás. σ ismeretlen, ezért t -próbát végzünk; H_1 szerint $c \neq 7.2$, így egymintás, kétoldali t -próbára van szükség. A koncentrációt jelölje c .

- $H_0: c = 7.2$;
- $H_1: c \neq 7.2$.

4. feladat

A minta átlaga

$$\bar{x} = \frac{7.7 + 8.1 + 7.7 + 7.5 + 7.0}{5} = 7.6,$$

és korrigált empirikus szórása

$$(s_n^*)^2 = \frac{1}{5-1} ((7.7-7.6)^2 + (8.1-7.6)^2 + (7.7-7.6)^2 + (7.7-7.5)^2 + (7.0-7.6)^2) = 0.16,$$

tehát

$$s_n^* = 0.4,$$

és a statisztika

$$t = \frac{\bar{x} - \mu}{s_n^*} \sqrt{n} = \frac{7.6 - 7.2}{0.4} \sqrt{5} = 2.236.$$

A percentilis a 95%-os kétoldali kvantilise az $n - 1 = 4$ szabadsági fokú t -eloszlásnak:

$$t_{\varepsilon/2} = 2.776.$$

Az összehasonlítás:

$$t = 2.236 \in [-t_{\varepsilon/2}, t_{\varepsilon/2}] = [-2.776, 2.776]$$

teljesül, tehát elfogadjuk H_0 -t 95%-os szignifikanciaszinten.

4. feladat

A második minta 7.5, 7.4, 7.3, 7.4, 7.5, erre

$$\bar{x} = 7.42, \quad s_n^* = 0.0837,$$

és a statisztika

$$t = \frac{\bar{x} - \mu}{s_n^*} \sqrt{n} = \frac{7.42 - 7.2}{0.0837} \sqrt{5} = 5.880,$$

az összehasonlítás pedig

$$t = 5.880 \in [-t_{\varepsilon/2}, t_{\varepsilon/2}] = [-2.776, 2.776]$$

nem teljesül, tehát ez alapján a minta alapján elvetjük H_0 -t 95%-os szignifikanciaszinten.

6. feladat

Egy cég motiválni akarja az alkalmazottait a termelékenység növelése érdekében. Kétféle módszert tesztelnek: az A módszer a munkahelyi körülmények javítása, a B módszer a fizetésemelés. A termelékenység változását mind a 6 alkalmazottra megmérték mindkét módszer esetén, és a következő eredményeket kapták:

alkalmazott	1	2	3	4	5	6
körülmények javítása	1.2	1.0	0.8	0.6	0.9	0.9
fizetésemelés	-0.2	0.3	3.6	1.4	-0.1	1.6

- 1 Teszteljük 95%-os szinten, hogy a körülmények javítása növeli-e a termelékenységet. (Mi a nullhipotézis?)
- 2 Teszteljük 95%-os szinten, hogy a fizetésemelés növeli-e a termelékenységet.
- 3 Teszteljük 95%-os szinten azt a hipotézist, hogy a fizetésemelés jobban növeli a termelékenységet, mint a körülmények javítása.

Megoldás.

(a) Egymintás, egyoldali t -próbát végzünk a következő mintára:

alkalmazott	1	2	3	4	5	6
körülmények javítása	1.2	1.0	0.8	0.6	0.9	0.9

Az m várható érték ismeretlen; $m = 0$ -t teszteljük $m > 0$ ellenében. Mindig H_0 tartalmazza az egyenlőséget és H_1 az egyenlőtlenséget:

- $H_0: m = 0$;
- $H_1: m > 0$.

$$\bar{x} = 0.9, \quad s_n^* = 0.2,$$

tehát a statisztika

$$t = \frac{\bar{x} - \mu}{s_n^*} \sqrt{n} = \frac{0.9 - 0}{0.2} \sqrt{6} = 10.06.$$

- (a) A percentilis az $n - 1 = 5$ szabadsági fokú t -eloszlás 95%-os egyoldali kvantilise:

$$t_{\varepsilon} = 2.015.$$

Az összehasonlítás

$$t = 10.06 \in (-\infty, t_{\varepsilon}] = (-\infty, 2.015)$$

nem teljesül, tehát elvetjük H_0 -t H_1 ellenében 95%-os szignifikanciaszinten, és arra következtetünk, hogy a munkahelyi körülmények javítása növeli a termelékenységet.

- (b) Az előzőhöz hasonlóan egymintás, egyoldali t -próbát végzünk a következő mintára:

alkalmazott	1	2	3	4	5	6
fizetésemelés	-0.2	0.3	3.6	1.4	-0.1	1.6

Ismét

- $H_0: m = 0$;
- $H_1: m > 0$.

Ezúttal

$$\bar{x} = 1.1, \quad s_n^* = 1.439,$$

tehát a statisztika

$$t = \frac{\bar{x} - \mu}{s_n^*} \sqrt{n} = \frac{1.1 - 0}{1.439} \sqrt{6} = 1.872.$$

(b) Az összehasonlítás

$$t = 1.872 \in (-\infty, t_{\varepsilon}] = (-\infty, 2.015)$$

most teljesül, tehát elfogadjuk H_0 -t 95%-os szignifikanciaszinten, azaz a fizetésemelés nem javítja a termelékenységet szignifikánsan.

Annak ellenére jutottunk erre a következtetésre, hogy $\bar{x}$ értéke magasabb a második mintára (0.9 az első és 1.1 a második mintára); ez lényegében azon múlik, hogy s_n^* értéke jóval magasabb a második mintára (0.2 illetve 1.439). Magasabb empirikus szórás mellett még a kicsit magasabb átlag is tekinthető a mintavételből származó véletlen ingadozás eredményének.

Másképpen mondva a t -próba az $\bar{x} - \mu$ eltérést a minta empirikus szórásához képest teszteli.

6. feladat

(c) Ezúttal a minta

alkalmazott	1	2	3	4	5	6
körülmények javítása	1.2	1.0	0.8	0.6	0.9	0.9
fizetésemelés	-0.2	0.3	3.6	1.4	-0.1	1.6

Milyen próbát végezzünk?

Természetes választás lenne a kétmintás t -próba, de itt lehet jobbat is: egymintás t -próba a két minta különbségére.

Ennek oka, hogy a két minta nem két teljesen független forrásból származik, mivel ugyanazokon az alkalmazottakon teszteltünk.

- (c) Ez behoz egy extra véletlent az alkalmazottak miatt, mi viszont erre nem vagyunk kíváncsiak, csak a két módszer összehasonlítására. Az egymintás t -próba a két minta különbségére kiejti az alkalmazottakból származó extra véletlent (amennyiben azt additív jellegűnek tételezzük fel).

Kétmintás próba használata akkor lenne indokolt, ha például a két módszert két különböző csoporton tesztelnék.

A különbség minta

alkalmazott	1	2	3	4	5	6
$A - B$	1.4	0.7	-2.8	-0.8	1.0	-0.7

- $H_0: m = 0$;
- $H_1: m < 0$.

(c)

$$\bar{x} = -0.2, \quad s_n^* = 1.538,$$

és a statisztika

$$t = \frac{\bar{x} - \mu}{s_n^*} \sqrt{n} = \frac{-0.2 - 0}{1.538} \sqrt{6} = -0.318.$$

A percentilis az $n - 1 = 5$ szabadsági fokú t -eloszlás 95%-os egyoldali kvantilise:

$$t_\epsilon = 2.015.$$

Az összehasonlítás

$$t = -0.318 \in [-\infty, t_\epsilon) = (-\infty, 2.015]$$

teljesül, tehát elfogadjuk H_0 -t 95%-os szignifikanciaszinten, és arra következtetünk, hogy a fizetésemelés nem javítja jobban a termelékenységet, mint a munkahelyi körülmények javítása.