

3. lecke

A hipotézisvizsgálat alapjai, várható értékek vizsgálatára irányuló próbák

Az előző leckében megismerkedhettek a különböző becslési formákkal. **Akkor készítünk becsléseket, amikor minta alapján próbáljuk becsülni a sokaság egy jellemzőjét, vagyis nem tudunk semmit a sokasági jellemzőről.** Ebben a leckében megismerkedhetnek a hipotézisvizsgálat alapjaival.

A gyakorlatban sokszor előfordul, hogy egy sokaság valamely paraméterére vonatkozóan van egy feltételezett érték, vagy valamilyen jellemzőjéről (például eloszlásáról) van egy előzetes elképzelésünk és csak azt szeretnénk eldönteni, hogy ez megfelel-e a valóságnak. Ha a sokaság teljes körű megfigyelésére nincs módunk, akkor a mintavétel módszeréhez folyamodhatunk. Ilyenkor egy véletlen minta alapján azt fogjuk megvizsgálni, hogy a mintánk támogatja-e a hipotézisünket, vagy ellentmond neki. Így bizonyos megbízhatósággal állíthatjuk majd, hogy hipotézisünk teljesül vagy sem.

A felállított hipotézisek helyességének véletlen mintákra alapozott vizsgálatát hipotézisvizsgálatnak nevezzük. Az ennek során alkalmazott eljárások a statisztikai próbák vagy tesztek.

Hipotézisvizsgálattal számos dolog tesztelhető, erre néhány példa:

- Tekinthető-e a férfiak várható élettartama 70 évnek, vagy szignifikánsan eltér ettől?
- Szignifikáns különbségek vannak-e régiók számítógépes ellátottságában?
- Szignifikánsan különbözik-e két párt szavazóinak aránya?
- Szignifikáns kapcsolat van-e munkanélküliségi ráta és az egy főre jutó GDP között?
- Szignifikáns kapcsolat van-e a nem és a beosztás között?
- Egyenletesen érkeznek-e az ügyfelek egy ügyfélszolgálatra?

A hipotézisvizsgálatot több probléma vizsgára is alkalmazhatjuk.

- Amennyiben arra vagyunk kíváncsiak, hogy **szignifikáns különbség van-e statisztikai jellemzők között**, úgy **különbözőségvizsgálatot** hajthatunk végre.
- Amennyiben pedig arra vagyunk kíváncsiak, hogy **szignifikáns kapcsolat van-e változók között**, arra a **kapcsolatvizsgálat vagy összefüggésvizsgálat** eszközei fognak választ adni.

A gyakorlatban sokszor előfordul, hogy egy sokaság valamely paraméterére vonatkozóan van egy előzetes feltételezésünk és azt szeretnénk eldönteni, hogy ez a feltételezés megfelel-e a valóságnak. A **hipotézisvizsgálat** nem más, mint **egy, a sokaságra (vagy sokaságokra) vonatkozó feltételezés vizsgálata minta (vagy minták) alapján**. A becslésekkel ellentétben itt már **van egy előzetes elképzelésünk a sokasági jellemzőről, amelyet a hipotézisvizsgálat eszközeivel megvizsgálunk, hogy teljesül-e.**

Egyszerűbben megfogalmazva tehát megtörténhet, hogy van egy elképzelésünk a valóságról és azt várjuk, hogy a minta is a mi elképzelésünket mutatja. Ehhez össze kell számszerűen hasonlítani a „minta” és az elképzelésünk eltérését. Ehhez szükségünk van egy próbafüggvényre és a próbafüggvény mintán felvett értékének kiszámítására. A „minta” és az elképzelésünk eltéréséről a próbafüggvény mintán felvett értéke segítségével el kell döntenünk, hogy ez jelentős (szignifikáns)-e,

vagyis döntést kell hoznunk az állításunkról. Ha az elképzelésünk és a „minta” ugyanaz, akkor az eltérés (próbafüggvény) értéke 0, azonban minél nagyobb a különbség annál nagyobb a próbafüggvény eltérése nullától.

A hipotézisvizsgálathoz szükségünk van a hipotézisünk statisztikai megfogalmazására, ez nem más, mint a nullhipotézis. **A nullhipotézis jelölése H_0 , tagadása pedig az alternatív hipotézis, amelynek jelölése H_1 . A nullhipotézisben mindig van egyenlőség, tehát megfogalmazásakor háromféle jelölést alkalmazhatunk: $=$, \geq vagy \leq . Az alternatív hipotézis nem tartalmaz egyenlőséget, így ennek megfogalmazásakor csak a \neq , $<$ vagy $>$ jelölések használhatók.** Egy példában megfogalmazva:

Férfiak és nő várható élettartamát vizsgáljuk. A nullhipotézis és az alternatív hipotézis a következők:

$$H_0: \mu_{\text{férfi}} = \mu_{\text{nő}}$$

$$H_1: \mu_{\text{férfi}} \neq \mu_{\text{nő}}$$

A fenti példában a nullhipotézisünk az, hogy a férfiak és nők várható élettartama szignifikánsan nem különbözik, vagyis azonosnak tekinthető. Az alternatív hipotézis ennek az ellenkezőjét állítja, azaz hogy a férfiak és a nők várható élettartama szignifikánsan különbözik, vagyis nem tekinthető egyenlőnek.

1. Hipotézisvizsgálat lépései

1. A nullhipotézis és az alternatív hipotézis megfogalmazása.
2. A rendelkezésre álló információkat figyelembe véve egy próbafüggvény választása.
 - Ez gyakorlatilag egy statisztika (képlet) kiválasztását jelenti.
 - A mintából kiszámított érték(ek)et hasonlítja össze a tesztelt sokasági paraméter feltételezett (hipotetikus) értékével.
3. A próbafüggvény mintán felvett értékeinek kiszámítása
 - Behelyettesítés a kiválasztott próbafüggvény képletébe.
 - A 0-hoz közeli szignifikanciaszint kiválasztása, és a próbafüggvény értékkészletének **elfogadási és kritikus tartományra** bontása.

Az elfogadási tartomány attól függően, hogy mi a nullhipotézisünk, kétféle formát vehet fel: beszélhetünk kétoldali (two-tailed) és egyoldali próbákról (one-tailed). Kétoldali próba esetén a nullhipotézisünk egyenlőséget feltételez (tehát $=$ jelet tartalmaz), az alternatív hipotézis pedig egyenlőtlenséget feltételez (\neq). Ilyen esetben az elfogadási tartományt mindkét oldalról korlátozni kell, tehát egy szimmetrikus intervallumot kell felállítanunk, attól függően, milyen próbafüggvénnyel dolgozunk. Egymintás próba esetén a nullhipotézist a \geq vagy \leq jelölések valamelyikével állítjuk fel, és ilyenkor az elfogadási tartományunkat csak egy oldalról kell korlátozni, a másik irányban a $-\infty$ vagy ∞ szerepelnek. Az SPSS minden esetben kétoldali próbát végez el, és nem tünteti fel közvetlenül az elfogadási tartományt, csak magát a döntést a nullhipotézis elvetéséről vagy elfogadásáról.

4. Döntés a nullhipotézis helyességének elfogadásáról vagy elvetéséről. Ez a próbafüggvény mintán felvett értékeinek összehasonlítását jelenti a kritikus értékkel. **A nullhipotézis elvetése maga után vonja az alternatív hipotézis elfogadását.**

2. Döntések során elkövethető hibák

	H₀ megfelel a valóságnak	H₀ nem felel meg a valóságnak
H₀-t elfogadjuk	helyes döntés (1- α)	másodfajú hiba (β)
H₀-t elvetjük	elsőfajú hiba (α)	helyes döntés (1- β)

Előfordulhat, hogy a nullhipotézis helyességéről hozott döntésünk nem lesz feltétlenül helyes, ettől függően pedig különböző típusú hibákat követhetünk el. Előfordulhat, hogy a nullhipotézis helyes, viszont a próbafüggvény mintán felvett értéke mégis a kritikus tartományba esik (tehát az elfogadási tartományunkon kívülre). Ekkor a nullhipotézist annak ellenére el fogjuk vetni, hogy az a valóságban helyes. Ezt a hibát **elsőfajú hibának** nevezzük. Ez a fajta hibás döntés α valószínűséggel fordulhat elő.

Előfordulhat az is, hogy a nullhipotézis nem felel meg a valóságnak, a próbafüggvény mintán felvett értéke mégis az elfogadási tartományba esik. Ekkor annak ellenére nem fogjuk elvetni a nullhipotézist, hogy nem igaz. Ezt a döntési hibát **másodfajú hibának** nevezzük, elkövetésének valószínűségét pedig β -val szokás jelölni.

Amennyiben a valóságnak megfelelő nullhipotézist elfogadjuk, illetve a valóságnak nem megfelelő nullhipotézist elvetjük, akkor helyes döntést hozunk, melyek előfordulásainak valószínűségei a hibás döntések meghozatalának komplementerei, helyes nullhipotézis elfogadásának valószínűsége $1-\alpha$, helytelen nullhipotézis elvetésének valószínűsége pedig $1-\beta$.

3. Döntés szoftverrel

Miután felállítottuk a nullhipotézisünket, a hipotézisvizsgálat többi lépését szoftverek segítségével is elvégezhetjük. Ehhez miután kiválasztottuk a megfelelő eljárást, és változókat, a szoftver elvégzi a vizsgálatot. A statisztikai szoftverek hipotézisvizsgálat során, az outputon megadnak egy értéket, a p -érték, p -value, Sig. jelölések valamelyikével. Ez az érték az első fajú hiba elkövetésének valószínűségét jelenti. Ez alapján az érték alapján kell meghoznunk a döntést, hogy elfogadjuk, vagy elvetjük a nullhipotézist. Ha vizsgálataink során például 5 százalékos szignifikanciaszintet használunk, akkor amennyiben a kapott érték 0,05-nál kisebb, akkor a nullhipotézist – ötszázalékos szignifikanciaszint mellett – elvetjük.

3.1. Próbák típusai

- **Egymintás:** egy változó várható értékét hasonlítjuk össze egy feltételezett értékkel. Attól függően, hogy ismert-e az alapsokaság szórása (σ) vagy sem, Z vagy T próbát végezhetünk el, amelyek számoló képletei az alábbiak:

$$\text{Alapsokaság szórása ismert: } Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \quad \text{Alapsokaság szórása nem ismert: } T = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Az SPSS automatikusan azt feltételezi, hogy mintával rendelkezünk, így az alapsokaság szórását nem ismerjük, ezért az a t-próbát fogja futtatni.

• **Kétmintás**

- **párosított:** két, egymással párba állítható változót hasonlítunk össze. A teszt során a két változó páronként vett különbségeit vizsgáljuk (amelyet d_i -vel jelölünk), a teszt próbafüggvénye az alábbi:

$$T = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}$$

- **független:** két egymástól független csoport tulajdonságait hasonlítjuk össze. Attól függően, hogy az alapsokaság szórása ismert-e, vagy sem, illetve, hogy teljesül-e a varianciahomogenitás, azaz a szórásnégyzetek egyezősége feltételezhető-e, három próbafüggvény közül lehet választani:

Alapsokaság szórása ismert:
$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Alapsokaság szórása nem ismert, a varianciák egyezősége nem feltételezhető:

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Alapsokaság szórása nem ismert, a varianciák egyezősége nem feltételezhető:

$$T = \frac{\bar{x}_1 - \bar{x}_2}{s_c \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

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

Az SPSS automatikusan mintát feltételez, amelyből nem derül ki az alapsokaság szórása, így t-próbafüggvénnyel számol. Azt, hogy a kettő közül melyik t-próbafüggvényt választja, az határozza meg, hogy a változók szórásnégyzetei azonosnak tekinthetők-e vagy sem.

3.2. Várható érték tesztek SPSS segítségével

3.2.1. Egymintás próbák

Egymintás próbák végezhetőek annak vizsgálatára, hogy a vizsgált változó a sokaságban felvehet-e egy adott értéket. Egymintás próbával vizsgálható a változó várható értéke (sokasági átlag), sokasági szórása és sokasági arány.

Az SPSS-ben a változók várható értékeinek vizsgálatára van lehetőség, amelyet az **Analyze/Compare Means/One-Sample T-Test** menüpontból érünk el.

A teszt elvégzésének alkalmazási feltételei:

1. Nagy mintával rendelkezünk (legalább 100 elem). Vagy:
2. 30-100 elem közötti mintával rendelkezünk, és ne legyen erős baloldali aszimmetria. Az aszimmetria vizsgálatát az **Analyze/Descriptive Statistics/Explore** menüpontban végezhetjük el. A kimeneten a **Skewness** értékét kell figyelembe venni, ami a ferdeség mérőszáma.

Értelmezése megegyezik az Általános statisztika I. kurzuson tárgyalt aszimmetria mutatókkal: pozitív előjel esetén baloldali, negatív előjel esetén pedig jobboldali aszimmetriáról beszélhetünk, amennyiben a Skewness értéke +1 vagy nagyobb pozitív értéket vesz fel, akkor beszélhetünk erős baloldali aszimmetriáról, amely azt is jelenti, hogy az egymintás várható érték tesztet nem végezhetjük el az adott változón.

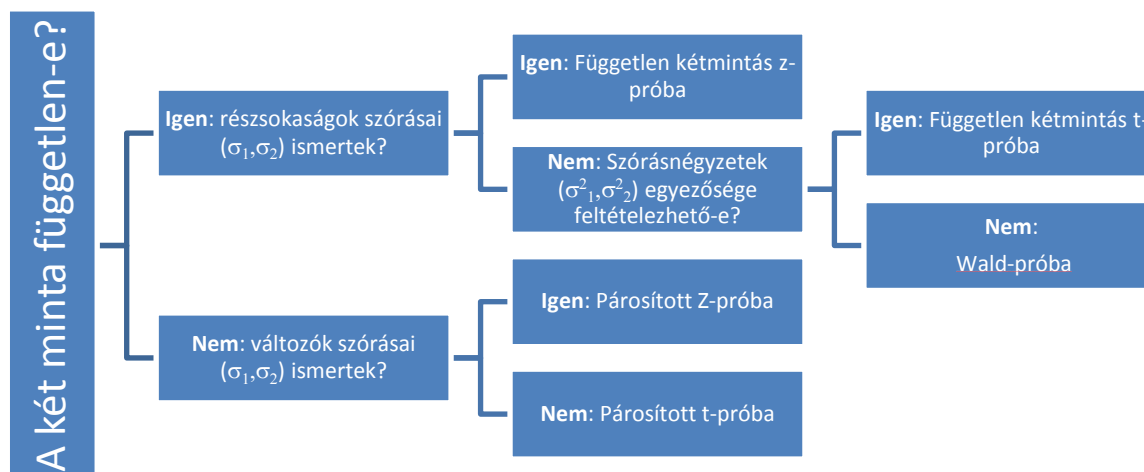
3. Kis minta esetén a változó normális eloszlású legyen (nagy minták esetén feltételezhető a normális eloszlás). A normális eloszlás teljesülését az **Analyze/Descriptive Statistics/Explore** menüpontban végezhetjük el, a **Plots** menüponton belül kiválasztva a **Normality plots with tests** lehetőséget. Normális eloszlás vizsgálatának nullhipotézise, hogy teljesül a normális eloszlás, alternatív hipotézise pedig, hogy nem teljesül. **A teszt elvégzése után a Shapiro-Wilk teszt Sig. értékének figyelembe vételével hozhatunk döntést, ha Sig. > 0,05, akkor elfogadjuk a nullhipotézist, tehát teljesül a normális eloszlás, ha pedig Sig. ≤ 0,05, akkor elvetjük a nullhipotézist, tehát nem teljesül a normális eloszlás,** ami egyben azt is jelenti, hogy az egymintás várható érték tesztet nem végezhetjük el az adott változón.

Az SPSS alapbeállításaként 5%-os szignifikanciaszinttel dolgozik, ami azt jelenti, hogy az elsőfajú hiba elkövetésének valószínűségét 5 százalékban határozza meg (tehát 5% az esélye annak, hogy elvetjük a helyes nullhipotézist). Ahhoz, hogy döntést hozzunk, meg kell vizsgálnunk a kimeneten szereplő Sig. értéket, amely mellett a *two-tailed* felirat utal arra, hogy itt kétoldali próbával van dolgunk, tehát nullhipotézisként azt feltételezzük, hogy az adott változó várható értéke megegyezik a feltételezett értékkel, alternatív hipotézisünk pedig az, hogy nem egyezik meg vele. **Ha a Sig. értéke nagyobb, mint 0,05, akkor elfogadjuk a nullhipotézist, ha pedig kisebb vagy egyenlő, elvetjük a nullhipotézist,** tehát az alternatív hipotézist fogadjuk el. A kimenetről leolvasható az is, hogy a várható érték magasabbnak vagy alacsonyabbnak tekinthető, mint a feltételezett értékünk.

3.2.2. Kétmintás próbák

Kétmintás próbával az egymintás próbákhoz hasonlóan várható érték, sokasági szórás vagy arány vizsgálható, ebben az esetben viszont két csoport esetén. Az SPSS-ben szintén csak várható érték vizsgálatára van lehetőség, azonban kétféleképpen is elvégezhetjük azt, attól függően, hogy ugyanazon változó várható értékét vizsgáljuk két csoport esetén (pl. nők és férfiak átlagéletkora) vagy két változó várható értékét hasonlítjuk össze (pl. diákok matematika és irodalom érdemjegyeinek az átlagát). Előbbi esetén független kétmintás várható érték tesztről beszélhetünk, ugyanis a két vizsgált csoport egymástól független, a vizsgált változó egyezik csak meg, utóbbi esetén pedig párosított kétmintás várható érték tesztről beszélhetünk, ugyanis ugyanazon egyedek esetén vizsgálunk meg két változót (a fenti példában tehát ugyanazon diákok esetén vizsgáljuk meg a jegyeiket).

Annak eldöntésében, hogy melyik vizsgálatot kell elvégeznünk, segíthet az alábbi ábra:



Párosított kétmintás várható érték teszt

Párosított kétmintás várható érték vizsgálatot hajthatunk végre, ha két egymással párba állítható változót szeretnénk összehasonlítani (pl. ugyanazon személyek vizsga előtti és utáni átlagos feszültségi szintjét). Ezt az SPSS-ben az **Analyze/Compare Means/Paired-Samples T-Test** menüpontból érjük el.

A teszt elvégzésének alkalmazási feltételei:

1. Mindkét csoportban nagy mintával rendelkezünk (legalább 100 elem). Vagy:
2. 30-100 elem közötti mintákkal rendelkezünk, és ne legyen erős baloldali aszimmetria. Az aszimmetria vizsgálatát az **Analyze/Descriptive Statistics/Explore** menüpontban végezhetjük el. A kimeneten a **Skewness** értékét kell figyelembe venni, ami a ferdeség mérőszáma. Értelmezése megegyezik az Általános statisztika I. kurzuson tárgyalt aszimmetria mutatókkal: pozitív előjel esetén baloldali, negatív előjel esetén pedig jobboldali aszimmetriáról beszélhetünk, amennyiben a Skewness értéke +1 vagy nagyobb pozitív értéket vesz fel, akkor beszélhetünk erős baloldali aszimmetriáról, amely azt is jelenti, hogy a párosított kétmintás várható érték tesztet nem végezhetjük el az adott változópáron.
3. Kis minták esetén a változók normális eloszlásúak legyenek (nagy minták esetén feltételezhető a normális eloszlás). A normális eloszlás teljesülését az **Analyze/Descriptive Statistics/Explore** menüpontban végezhetjük el, a **Plots** menüponton belül kiválasztva a **Normality plots with tests** lehetőséget. Normális eloszlás vizsgálatának nullhipotézise, hogy teljesül a normális eloszlás, alternatív hipotézise pedig, hogy nem teljesül. A teszt elvégzése után a Shapiro-Wilk teszt Sig. értékének figyelembe vételével hozhatunk döntést, ha Sig. > 0,05, akkor elfogadjuk a nullhipotézist, tehát teljesül a normális eloszlás, ha pedig Sig. ≤ 0,05, akkor elvetjük a nullhipotézist, tehát nem teljesül a normális eloszlás, ami egyben azt is jelenti, hogy a párosított kétmintás várható érték tesztet nem végezhetjük el az adott változópáron.

Az SPSS alapbeállításként 5%-os szignifikanciaszinttel dolgozik, ami azt jelenti, hogy az elsőfajú hiba elkövetésének valószínűségét 5 százalékban határozza meg (tehát 5% az esélye annak, hogy elvetjük a helyes nullhipotézist). Ahhoz, hogy döntést hozzunk, meg kell vizsgálnunk a kimeneten szereplő Sig. értéket, amely mellett a *two-tailed* felirat utal arra, hogy itt kétoldali próbával van

dolgunk, tehát nullhipotézisként azt feltételezzük, hogy az adott változó várható értéke megegyezik a másik változó várható értékével, alternatív hipotézisünk pedig az, hogy nem egyezik meg vele. **Ha a Sig. értéke nagyobb, mint 0,05, akkor elfogadjuk a nullhipotézist, ha pedig kisebb vagy egyenlő, elvetjük a nullhipotézist,** tehát az alternatív hipotézist fogadjuk el. A kimenetről leolvasható az is, hogy mely csoport várható értéke magasabb, és a különbség mértéke is.

Független kétmintás várható érték teszt

Független kétmintás várható érték vizsgálatot hajthatunk végre, ha két egymástól egyértelműen elkülöníthető csoport tulajdonságait szeretnénk összehasonlítani (pl. férfiak és nők átlagos jövedelmét). Ezt az SPSS-ben az **Analyze/Compare Means/Independent-Samples T-Test** menüpontból érjük el.

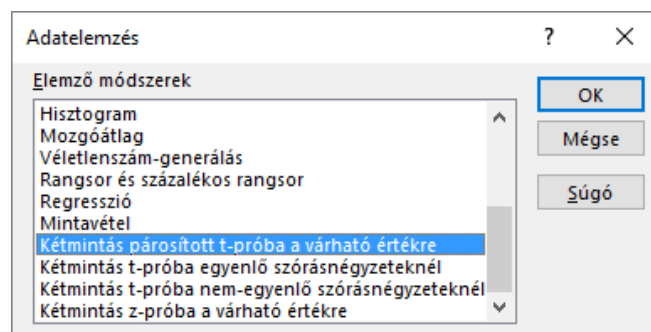
A teszt elvégzésének alkalmazási feltételei:

4. Mindkét csoportban nagy mintával rendelkezünk (legalább 100 elem). Vagy:
5. 30-100 elem közötti mintákkal rendelkezünk, és ne legyen erős baloldali aszimmetria. Az aszimmetria vizsgálatát az **Analyze/Descriptive Statistics/Explore** menüpontban végezhetjük el. A kimeneten a **Skewness** értékét kell figyelembe venni, ami a ferdeség mérőszáma. Értelmezése megegyezik az Általános statisztika I. kurzuson tárgyalt aszimmetria mutatókkal: pozitív előjel esetén baloldali, negatív előjel esetén pedig jobboldali aszimmetriáról beszélhetünk, amennyiben a Skewness értéke +1 vagy nagyobb pozitív értéket vesz fel, akkor beszélhetünk erős baloldali aszimmetriáról, amely azt is jelenti, hogy a független kétmintás várható érték tesztet nem végezhetjük el az adott változópáron.
6. Kis minták esetén a változók normális eloszlásúak legyenek (nagy minták esetén feltételezhető a normális eloszlás). A normális eloszlás teljesülését az **Analyze/Descriptive Statistics/Explore** menüpontban végezhetjük el, a **Plots** menüponton belül kiválasztva a **Normality plots with tests** lehetőséget. Normális eloszlás vizsgálatának nullhipotézise, hogy teljesül a normális eloszlás, alternatív hipotézise pedig, hogy nem teljesül. A teszt elvégzése után a Shapiro-Wilk teszt Sig. értékének figyelembe vételével hozhatunk döntést, ha Sig. > 0,05, akkor elfogadjuk a nullhipotézist, tehát teljesül a normális eloszlás, ha pedig Sig. ≤ 0,05, akkor elvetjük a nullhipotézist, tehát nem teljesül a normális eloszlás, ami egyben azt is jelenti, hogy a független kétmintás várható érték tesztet nem végezhetjük el az adott változópáron.
7. Varianciahomogenitás: a varianciahomogenitás a szórásnégyzetek egyezőségének feltételezését jelenti. Ennek vizsgálatát a független kétmintás t-próba kimenetén levő **Levene teszt** értékeinek figyelembe vételével tehetjük meg. A varianciahomogenitás teszt nullhipotézise, hogy a szórásnégyzetek, vagyis varianciák egyezősége feltételezhető ($\sigma^2_1 = \sigma^2_2$), alternatív hipotézise pedig, hogy a varianciák egyezősége nem feltételezhető ($\sigma^2_1 \neq \sigma^2_2$). A varianciahomogenitás teszt esetén, ha Sig. > 0,05, akkor elfogadjuk a nullhipotézist, tehát a varianciák egyezősége feltételezhető -ekkor a kimenet első sora alapján kell a független kétmintás t próbát elvégezni-, ha pedig Sig. ≤ 0,05, akkor elvetjük a nullhipotézist, tehát a varianciák egyezősége nem feltételezhető -ebben az esetben pedig a kimenet második sorát kell figyelembe venni a független kétmintás t próba elvégzéséhez.

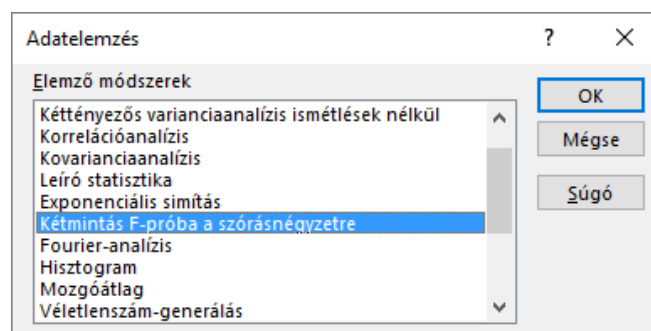
Az SPSS alapbeállításként 5%-os szignifikanciaszinttel dolgozik, ami azt jelenti, hogy az elsőfajú hiba elkövetésének valószínűségét 5 százalékban határozza meg (tehát 5% az esélye annak, hogy elvetjük a helyes nullhipotézist). Ahhoz, hogy döntést hozzunk, meg kell vizsgálnunk a kimeneten szereplő Sig. értéket (azt, hogy a táblázat mely sorának Sig. értékét kell megvizsgálnunk, a varianciahomogenitás teszt alapján dönthetjük el), amely mellett a *two-tailed* felirat utal arra, hogy itt kétoldali próbával van dolgunk, tehát nullhipotézisként azt feltételezzük, hogy az adott csoport várható értéke megegyezik a másik csoport várható értékével, alternatív hipotézisünk pedig az, hogy nem egyezik meg vele. **Ha a Sig. értéke nagyobb, mint 0,05, akkor elfogadjuk a nullhipotézist, ha pedig kisebb vagy egyenlő, elvetjük a nullhipotézist,** tehát az alternatív hipotézist fogadjuk el. A kimenetről leolvasható az is, hogy mely csoport várható értéke magasabb, és a különbség mértéke is.

3.3. Az Excel alkalmazása két mintás próbák esetén

Az **Adatelemzés** bővítményben a kétmintás próbák, mint elemzési módszerek beállítása nagyon hasonló. A párosított kétmintás várható érték tesztet a **Kétmintás párosított t-próba a várható értékre** menüpontból érhetjük el. Nem csak a várható értékek egyezősége tesztelhető, hanem ezek tetszőleges, feltételezett eltérése is. Ezért bemenetként a változók tartománya mellett meg kell adnunk a feltételezett, nem negatív eltérést is. Amennyiben **nulla** értéket adunk meg, úgy a várható értékek egyezőségét teszteljük.



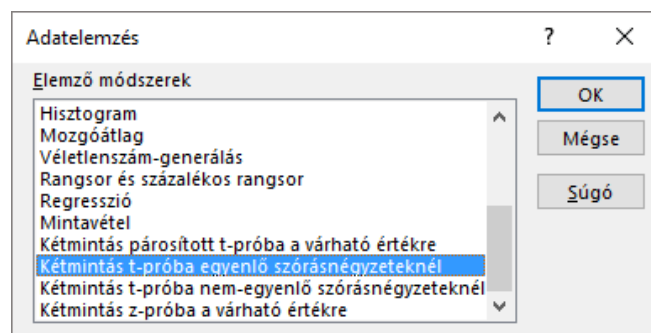
Kétmintás *t*-próba alkalmazásakor, amennyiben nem párosított *t*-próbáról van szó, első lépésben meg kell vizsgálnunk a két változó varianciájának egyezőségét a **Kétmintás F-próba a szórásnégyzetre** funkcióval. Ez alapján történik a háttérben a próbafüggvény kiválasztása.



A kétmintás *F*-próba kimenete az alábbi.

Kétmintás F-próba a szórásnégyzetre		
	Változó 1	Változó 2
Várható érték	3,333333333	3,5
Variancia	2,666666667	3,5
Megfigyelések	6	6
df	5	5
F	0,761904762	
P(F<=f) egyszélű	0,386338214	
F kritikus egyszélű	0,1980069	

Ekkor a kimenet alapján a következő döntést hozhatjuk: mivel $P=0,386 > 0,05$; ezért ötszázalékos szignifikanciaszinten a nullhipotézist elfogadjuk, azaz a két változó variáciája azonosnak tekinthető. Ezután alkalmazhatjuk a **Kétmintás t-próba egyenlő szórásnégyzeteknél** esetet (ellenkező esetben a **Kétmintás t-próba nem egyenlő szórásnégyzeteknél** menüpontot kell kiválasztanunk).



Ennek egy szabványos kimenete az alábbi.

Kétmintás t-próba egyenlő szórásnégyzeteknél		
	Változó 1	Változó 2
Várható érték	3,333333333	3,5
Variancia	2,666666667	3,5
Megfigyelések	6	6
Súlyozott variancia	3,083333333	
Feltételezett átlagos eltérés	0	
df	10	
t érték	-0,164398987	
P(T<=t) egyszélű	0,436346277	
t kritikus egyszélű	1,812461102	
P(T<=t) kétszélű	0,872692554	
t kritikus kétszélű	2,228138842	

A táblázatban egyszélű és kétszélű értékeket is találunk. Ezek az **egyoldali** és a **kétoldali** próbák megfelelői. Számunkra a kétoldali próbák lesznek az érdekesek. Ha tehát kétoldali próbát hajtunk végre, akkor a táblázatban szereplő kétszélű p -érték $=0,87 > 0,05$; ezért ötszázalékos szignifikanciaszinten a nullhipotézist elfogadjuk, azaz a két változó várható értéke azonosnak tekinthető.

SZEGEDI TUDOMÁNYEGYETEM
GAZDASÁGTUDOMÁNYI KAR
KÖZGAZDÁSZ KÉPZÉS
TÁVOKTATÁSI TAGOZAT
LECKESOROZAT
COPYRIGHT © SZTE GTK 2017/2018

A LECKE TARTALMA, ILLETVE ALKOTÓ ELEMEI ELŐZETES,
ÍRÁSBELI ENGEDÉLY MELLETT HASZNÁLHATÓK FEL.

JELLEN TARTALOM
A SZEGEDI TUDOMÁNYEGYETEMEN KÉSZÜLT
AZ EURÓPAI UNIÓ TÁMOGATÁSÁVAL.
PROJEKT AZONOSÍTÓ: EFOP-3.4.3-16-2016-00014

SZÉCHENYI 2020



MAGYARORSZÁG
KORMÁNYA

Európai Unió
Európai Szociális
Alap



BEFEKTETÉS A JÖVŐBE