Segédanyag a többszörös összehasonlítások témaköréhez

Kemény S., Deák A., Lakné Komka K., Kunovszki P.:
Kísérletek tervezése és értékelése-Kiegészítő kötet alapján

A fejezetbeli példák az alapkötet 9-1. példájához kapcsolódnak.

9-1. példa
Liszthez a sütőipari minőség javítása érdekében háromféle fehérjeizolátumot adtak: kukoricacsírát, búzacsírát és rizscsírát, és mérték az MT technofunkciós tulajdonságot. Összehasonlításképpen adalék nélküli mintát is vizsgáltak. Az adalék fajtája a vizsgált faktor, jelöljük A-val.
A (valóságos elemekből létrehozott) hipotetikus mérési eredményeket mutatja a 9-1. táblázat, mely néhány, a számításokhoz szükséges további adatot is tartalmaz.

9-1. táblázat
1: adalék nélkül
2: kukoricacsíra adalékkal
3: búzacsíra adalékkal
4: rizscsíra adalékkal

Faktorszint	1	2	3	4
y_{ij}	$250^{(1)}$ $259^{(8)}$ $269^{(3)}$	$222^{(12)}$ $228^{(5)}$ $237^{(9)}$	$238^{(2)}$ $247^{(11)}$ $250^{(6)}$	$212^{(10)}$ $225^{(4)}$ $234^{(7)}$
p_i	3	3	3	3
$\bar{y}_{i\cdot}$	259.33	229.00	245.00	223.67
s_i	9.504	7.550	6.245	11.060
s_i^2	90.333	57.003	39.000	122.324

A táblázatban lévő számértékek az y_{ij} értékek, ahol az első index (i) a csoport sorszáma, vagyis azt mutatja, hogy melyik adalékot használták; a második index (j) pedig a csoporton belüli ismétlés sorszámát mutatja. A számok fölött zárójelben lévő kis számok a kísérletek elvégzésének sorrendjét mutatják.

9-4. táblázat: ANOVA-táblázat

Az eltérés forrása	Eltérés-négyzetösszeg	Szabadsági fok	Szórásnégyzet	F_0	p
A hatása (csoportok közötti)	2352.917	3	784.306	10.164	0.0042
Ismétlések (csoportokon belüli)	617.333	8	77.167		
Teljes	2970.250	11			

10. Összehasonlítások egy faktor két vagy több szintjére

A varianciaanalízis csak arra a kérdésre ad választ, hogy van-e különbség a csoportok várható értékei között, vagyis van-e egyáltalán az *A* faktornak hatása. Ha úgy találjuk, hogy van vagyis a nullhipotézist elutasítjuk, nem tudjuk, hogy csak egy csoport különbözik a többitől, vagy mind különböznek egymástól, vagy esetleg bizonyos csoportok egymással a várható értékük szempontjából azonosnak tekinthetők. Erre a kérdésre akkor kapunk választ, ha egy faktor két vagy több szintjére végzünk összehasonlításokat.
Ha a varianciaanalízis során úgy találjuk, hogy nincs szignifikáns hatása a faktornak (vagyis elfogadjuk a nullhipotézist), akkor csak azt mondhatjuk, hogy az adatok alapján nem vagyunk biztosak benne, hogy mindegyik azonos.

10.1. Két szint összehasonlítása

Vegyük például a következő hipotézist:

$$\mathrm{H}_0^1 : \mu_2 = \mu_4 \text{ vagy } \mu_2 - \mu_4 = 0. \tag{10.1}$$

Ez azt jelentené, hogy a 2. és 4. csoport adatai ugyanazon várható értékű sokaságból származnak. A hipotézis teljesülését *t*-próbával vizsgálhatjuk. A próbastatisztika (a kétmintás *t*-próba analógiájára):

$$t_0 = \frac{y_{2\cdot} - y_{4\cdot}}{s_{y_{2\cdot} - y_{4\cdot}}}. \tag{10.2}$$

A nevezőben lévő szórás a két csoport átlaga különbségének szórása. A különbség varianciája:

$$Var(y_{2\cdot} - y_{4\cdot}) = \sigma_e^2 \left(\frac{1}{p_2} + \frac{1}{p_4} \right). \tag{10.3}$$

Ebből a szórásnégyzet:

$$s_{y_{2\cdot} - y_{4\cdot}}^2 = s^2 \left(\frac{1}{p_2} + \frac{1}{p_4} \right). \tag{10.4}$$

A szokásos kétmintás *t*-próbánál s^2-re a két összehasonlítandó csoport szórásnégyzetének egyesítésével kapható értéket veszik, melynek szabadsági foka esetünkben $p_2 + p_4 - 2$ lenne. Itt azonban rendelkezésünkre áll az összes csoportokon belüli szórásnégyzet egyesítésével kapott s_R^2, amelynek szabadsági foka $\sum_i p_i - r$, ez nagyobb, tehát kedvezőbb.

Akkor minősítjük szignifikánsnak a két szint közötti különbséget (a nullhipotézistől való eltérést), ha $|t_0| > t_{\alpha/2}$, vagyis

$$|y_{2\cdot} - y_{4\cdot}| > t_{\alpha/2} s \sqrt{\frac{1}{p_2} + \frac{1}{p_4}}. \tag{10.5}$$

Tehát a faktor két szintje között akkor szignifikáns a különbség, ha a két szinthez tartozó átlagok különbsége nagyobb, mint az $1-\alpha$ valószínűséggel H_0 igazsága esetén várható lenne (a véletlen műveként). Ezt nevezik *LSD* (Least Significant Difference) eljárásnak. A név onnan származik, hogy

$$t_{\alpha/2}s\sqrt{\frac{1}{p_2}+\frac{1}{p_4}}$$

a két átlag között az a legkisebb különbség, amelyet már szignifikánsnak kell minősítenünk.

A próbát úgy is végezhetjük, hogy kiszámítjuk a *p* értékeket (annak valószínűségét, hogy a nullhipotézis fennállása, vagyis a várható értékek egyenlősége esetén ekkora különbségeket tapasztaljunk), és azokra a párokra utasítjuk el a nullhipotézist, amelyekre *p* kisebb a megadott értéknél (pl. 0.05-nél).

A *t*-eloszlás segítségével konfidenciaintervallumot is számolhatunk a várható értékek különbségére:

$$y_{2\cdot}-y_{4\cdot}-t_{\alpha/2}s_R\sqrt{\frac{1}{p_2}+\frac{1}{p_4}}<\mu_2-\mu_4\le y_{2\cdot}-y_{4\cdot}+t_{\alpha/2}s_R\sqrt{\frac{1}{p_2}+\frac{1}{p_4}}\,. \tag{10.6}$$

Tudjuk, hogy a *t*-eloszlású valószínűségi változó négyzete olyan *F*-eloszlású, amelyre a számláló szabadsági foka 1. A kétoldali *t*-próbának egyoldali *F*-próba felel meg. Akkor utasítjuk el a *t*-próbánál a nullhipotézist, ha $|t_0|>t_{\alpha/2}$, ez pedig azt jelenti, hogy *F*-nek a fölső kritikus értéket kell meghaladnia.

10-1. példa

Vizsgáljuk meg a $H_0^1:\mu_2=\mu_4$ hipotézis érvényességét a 9-1. példa adataira!

$$s^2_{y_{2\cdot}-y_{4\cdot}}=s_R^2\left(\frac{1}{p_2}+\frac{1}{p_4}\right)=77.167\cdot\left(\frac{1}{3}+\frac{1}{3}\right)=51.44\,,$$

$$t_0=\frac{y_{2\cdot}-y_{4\cdot}}{s_{y_{2\cdot}-y_{4\cdot}}}=\frac{229.0-223.67}{\sqrt{51.44}}=0.743\,.$$

A *t*-táblázatból a 8 szabadsági fokhoz tartozó kritikus érték $t_{\alpha/2}(8)=2.306$, tehát a nullhipotézist (hogy a 2. és 4. csoport várható értéke egyenlő) elfogadjuk.
A 95% valószínűséghez tartozó konfidenciaintervallum a μ_2 és μ_4 különbségére:

$$229.0-233.67-2.306\cdot\sqrt{51.44}\sqrt{\frac{1}{3}+\frac{1}{3}}<\mu_2-\mu_4\le 229.0-233.67+2.306\cdot\sqrt{51.44}\sqrt{\frac{1}{3}+\frac{1}{3}}$$

10.2. Az összehasonlítás általánosítása: kontrasztok

Az eddigiekben két csoport átlagát hasonlítottuk össze, arra a kérdésre keresve a választ, hogy van-e különbség a mögöttük álló sokaságok várható értékében. Azt is kérdezhetjük, hogy az első csoport különbözik-e a másik háromtól. Ekkor a nullhipotézis:

$$\mathrm{H}_0^2 : \mu_1 = \frac{\mu_2 + \mu_3 + \mu_4}{3} \text{ vagy } \mu_1 - \frac{\mu_2 + \mu_3 + \mu_4}{3} = 0 . \tag{10.7}$$

A t_0 próbastatisztika számlálójába írandó különbség varianciája:

$$Var\left(y_{1\cdot} - \frac{y_{2\cdot} + y_{3\cdot} + y_{4\cdot}}{3} \right) = \sigma_e^2 \left[\left(\frac{1}{p_1} \right) + \frac{1}{9} \left(\frac{1}{p_2} + \frac{1}{p_3} + \frac{1}{p_4} \right) \right].$$

A (10.7) nullhipotézishez tartozó próbastatisztika:

$$t_0 = \frac{y_{1\cdot} - \dfrac{y_{2\cdot} + y_{3\cdot} + y_{4\cdot}}{3}}{s_R \sqrt{\dfrac{1}{p_1} + \dfrac{1}{9}\left(\dfrac{1}{p_2} + \dfrac{1}{p_3} + \dfrac{1}{p_4} \right)}} . \tag{10.8}$$

Látható, hogy a H_0^1 és a H_0^2 hipotézisek is a csoportok várható értékeinek lineáris kombinációira vonatkoznak. Kézenfekvő, hogy a formulákat a csoportok várható értékeinek bármilyen lineáris kombinációjára általánosítsuk.
Általában egy hipotézis, és ezt jelöljük *k*-adiknak, és *i* a csoport indexe:

$$\mathrm{H}_0^k : \sum_i c_{ik} \mu_i = 0 . \tag{10.9}$$

A c_{ik} együtthatókat kontrasztegyütthatóknak nevezik, ha összegük zérus:

$$\sum_i c_{ik} = 0 . \tag{10.10}$$

A H_0^1 hipotézisre a kontrasztegyütthatók vektora:

$$c_{11} = 0 ,\ c_{21} = 1 ,\ c_{31} = 0 ,\ c_{41} = -1 .$$

A H_0^2 hipotézisre:

$$c_{12} = 1 ,\ c_{22} = c_{32} = c_{42} = -1/3 .$$

A kontrasztegyütthatókkal tulajdonképpen egy új, normális eloszlású változót állítunk elő, ezt nevezik kontrasztnak:

$$C_k = \sum_i c_{ik} y_{i\cdot} . \tag{10.11}$$

A H_0^1 hipotézisre $C_1 = y_{2\cdot} - y_{4\cdot}$, a H_0^2 hipotézisre $C_2 = y_{1\cdot} - \dfrac{y_{2\cdot} + y_{3\cdot} + y_{4\cdot}}{3}$.
A C_k kontraszt még akkor is közelítőleg normális eloszlású lesz, ha az eredeti *y* adatok nem voltak azok, mert átlagaik lineáris kombinációját képezzük, és ezek az $y_{i\cdot}$ átlagok a centrális határeloszlási tétel értelmében már közel normális eloszlásúak.
A C_k kontraszt várható értéke és varianciája:

$$E(C_k) = \sum_i c_{ik}\mu_i \,, \tag{10.12}$$

$$Var(C_k) = \sigma_e^2 \sum_i \frac{c_{ik}^2}{p_i} \,. \tag{10.13}$$

A C_k kontraszt felhasználásával a H_0^k nullhipotézis így is megfogalmazható:

$$\mathrm{H}_0^k:\ E(C_k) = \sum_i c_{ik}\mu_i = 0 \,. \tag{10.14}$$

Például az előbbi H_0^1 és H_0^2 hipotézishez tartozó kontrasztokra:

$$E(C_1) = \mu_2 - \mu_4 \qquad \text{ill.} \qquad E(C_2) = \mu_1 - \frac{\mu_2 + \mu_3 + \mu_4}{3} \,,$$

$$Var(C_1) = \sigma_e^2\left(\frac{1}{p_2} + \frac{1}{p_4}\right) \qquad \text{és} \qquad Var(C_2) = \sigma_e^2\left[\frac{1}{p_1} + \frac{1}{9}\left(\frac{1}{p_2} + \frac{1}{p_3} + \frac{1}{p_4}\right)\right].$$

Mivel C_k normális eloszlású, a következő kifejezés t-eloszlású lesz $\sum_i p_i - r$ szabadsági fokkal:

$$t = \frac{C_k - E(C_k)}{s_{C_k}} \,, \tag{10.15}$$

ahol $s_{C_k}^2 = s_R^2 \sum_i c_{ik}^2 / p_i$ (10.16)

Ha a nullhipotézis igaz, $E(C_k) = \sum_i c_{ik}\mu_i = 0$, így az alkalmazandó próbastatisztika:

$$t_0 = \frac{\sum_i c_{ik} y_{i\cdot}}{s_R \sqrt{\sum_i c_{ik}^2 / p_i}} \,. \tag{10.17}$$

Az *LSD* módszer szerint akkor utasítjuk el a nullhipotézist, ha

$$\left|\sum_i c_{ik} y_{i\cdot}\right| > t_{\alpha/2} s_R \sqrt{\sum_i c_{ik}^2 / p_i} \,. \tag{10.18}$$

Ugyanezen $\mathrm{H}_0^k : \sum_i c_{ik}\mu_i = 0$ hipotézis vizsgálatára F-próbát is végezhetünk. Ehhez lássuk be, hogy ha a nullhipotézis igaz, az

$$S_k = \frac{\left(\sum_i c_{ik} y_{i\cdot}\right)^2}{\sum_i c_{ik}^2 / p_i} \tag{10.19}$$

kontraszt négyzetösszege $\chi^2\sigma_e^2$ eloszlású, szabadsági foka 1. A következő hányados ezért, ha a nullhipotézis igaz, F-eloszlású:

$$F_0 = \frac{S_k}{s_R^2} = \frac{\left(\sum_i c_{ik} y_{i.}\right)^2}{s_R^2 \sum_i c_{ik}^2 / p_i}. \qquad (10.20)$$

A számláló szabadsági foka 1, a nevezőé $\sum_i p_i - r$.

Vegyük észre, hogy ez az F_0 próbastatisztika az előbbi t_0 négyzete.

Számolási egyszerűsítés: minden c_{ik} együtthatót szorozhatunk egy számmal, ekkor a próbastatisztikák kifejezése nem változik (mert a szorzó a számlálóban és a nevezőben is megjelenik). Ez azt jelenti, hogy a $\mu_1 - \frac{\mu_2 + \mu_3 + \mu_4}{3} = 0$ hipotézis helyett azt is írhatjuk, hogy $3\mu_1 - (\mu_2 + \mu_3 + \mu_4) = 0$, utóbbihoz a kontrasztegyütthatók: 3, –1, –1, –1. Általában lehet olyan szorzót találni, hogy minden c_{ik} együttható egész szám legyen, ez pedig kényelmes. Az új kontrasztegyütthatókra is teljesül, hogy $\sum_i c_{ik} = 0$, tehát ténylegesen kontrasztegyütthatóknak tekinthetők.

10.3. Többszörös összehasonlítások, kontrasztok

Sokszor nem egyetlen, hanem több nullhipotézist vizsgálunk ugyanazon kísérletsorozat eredményeiből. Ezeket vizsgálhatjuk egyenként, vagy együttesen. Előbb az egyenkénti vizsgálatot tárgyaljuk.

10.3.1. Ortogonális hipotézisek

Az összehasonlításoknál a hipotézisek lehetnek függetlenek, vagy más néven ortogonálisak, ekkor az egyik hipotézis igaz volta független a másik hipotézis igazságától. Például a $\mu_1 = \mu_2$ és a $\mu_2 = \mu_3$ hipotézisektől nem független a $\mu_1 = \mu_3$ hipotézis, mert ha az első kettő igaz, a harmadiknak is teljesülnie kell.

Az ortogonális (független) hipotézisek száma véges. Ha r csoportunk van (a liszt-példánál $r = 4$), r független összehasonlítást tehetünk. Ezek közül $r - 1$ a különbségekre vonatkozik, 1 a teljes átlagra.

A kontrasztok statisztikai értelemben ortogonálisak, ha $\sum_i c_{ik} c_{il} / p_i = 0$ minden $k \neq l$-re. Logikailag akkor áll fenn az ortogonalitás, ha minden $k \neq l$-re $\sum_i c_{ik} c_{il} = 0$. A logikai és a statisztikai ortogonalitás csak akkor esik egybe, ha a csoportokban az ismétlések p száma azonos.

Bár a hipotézisek függetlenek, a próbák nem teljesen azok, mert ugyanazon s_R^2 szerepel a nevezőjükben. Ha véletlenül az túl nagy vagy túl kicsi, mindegyik elvégzett próbát érinti.

Például a $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ összetett hipotézisnek a következő logikailag független hipotézisek felelhetnek meg:

$$H_0^1 : \mu_1 = \mu_2 \text{ vagy } \mu_1 - \mu_2 = 0,$$

$$H_0^2 : \mu_3 = \mu_4 \text{ vagy } \mu_3 - \mu_4 = 0,$$

$$H_0^3 : \mu_1 + \mu_2 = \mu_3 + \mu_4 \text{ vagy } \mu_1 + \mu_2 - \mu_3 - \mu_4 = 0.$$

(Az első és második nullhipotézist a harmadikba helyettesítve éppen az eredeti összetett hipotézis – mindegyik csoport várható értékének egyenlősége – adódik).

Ezek a hipotézisek akkor lesznek statisztikai értelemben is ortogonálisak, ha az ismétlések p száma minden csoportban azonos.

Az e hipotézisekhez tartozó kontrasztegyütthatókat adja a 10-1. táblázat. Egy-egy oszlop az egy-egy nullhipotézishez tartozó kontrasztegyütthatókat adja. A H_0^1 hipotézisre (k = 1, vagyis a c_{i1} oszlopból) a kontrasztegyütthatók vektora: 1, –1, 0, 0. A táblázatra teljesül a logikai ortogonalitás, vagyis $\sum_i c_{ik} c_{il} = 0$ minden $k \neq l$-re.

10-1. táblázat

		H_0^1	H_0^2	H_0^3			
i	μ_i	c_{i1}	c_{i2}	c_{i3}	$c_{i1}\,c_{i2}$	$c_{i1}\,c_{i3}$	$c_{i2}\,c_{i3}$
1	μ_1	1	0	1	0	1	0
2	μ_2	–1	0	1	0	–1	0
3	μ_3	0	1	–1	0	0	–1
4	μ_4	0	–1	–1	0	0	1
	Σ	0	0	0	0	0	0

10-2. példa

Vizsgáljuk meg, hogy a H_0^1, H_0^2 és H_0^3 nullhipotézisek teljesülnek-e a 9-1. példa adataira!

Számoljuk ki az egyes hipotézisekhez tartozó F_0 próbastatisztika-értékeket!

A 10-2. táblázat foglalja össze a három hipotézishez tartozó kontrasztegyütthatók, a kontrasztok, valamint az F_0 próbastatisztika kiszámításához szükséges kifejezések értékeit. A táblázat egy-egy sora tartozik egy-egy nullhipotézishez!

10-2. táblázat

k	c_{ik}	μ_1	μ_2	μ_3	μ_4	$C_k = \sum_i c_{ik} y_{i\cdot}$	$\sum_i c_{ik}^2 / p$	$S_k = \dfrac{p\left(\sum_i c_{ik} y_{i\cdot}\right)^2}{\sum_i c_{ik}^2}$	ν	$F_0^k = \dfrac{S_k}{s_R^2}$	p
1	c_{i1}	1	–1	0	0	30.33	2/3	1380.167	1	17.886	0.0029
2	c_{i2}	0	0	1	–1	21.33	2/3	682.667	1	8.847	0.0178

3	c_{i3}	1	1	–1	–1	19.66	4/3	290.083	1	3.759	0.0885
								2352.917			

A 9-1. példából $s_R^2 = 77.165$. A táblázat utolsó oszlopa az F-próba p értékét mutatja, vagyis annak valószínűségét, hogy az illető (k-adik) nullhipotézis fennállása esetén ekkora, vagy ennél szélsőségesebb próbastatisztika-érték adódjék. Azokat a nullhipotéziseket utasítjuk el, amelyekre p kisebb a megadott értéknél (pl. 0.05-nél).

Vegyük észre, hogy a három egy szabadsági fokú S_k négyzetösszeg összege (2352.917) éppen egyenlő a 9-4. táblázatban az A faktor (az adalék) hatására kapott 3 szabadsági fokú négyzetösszeggel.

Az Olvasó meggyőződhet róla, hogy a következő részhipotézisek ugyancsak logikailag ortogonálisak (tehát függetlenek), és együttesen ugyancsak a csoportok várható értékének azonosságát jelentik:

$$\mathrm{H}_0^4 : \mu_1 = \frac{\mu_2 + \mu_3 + \mu_4}{3} \qquad \text{vagy} \qquad 3\mu_1 - (\mu_2 + \mu_3 + \mu_4) = 0,$$

$$\mathrm{H}_0^5 : \mu_4 = \frac{\mu_2 + \mu_3}{2} \qquad \text{vagy} \qquad 2\mu_4 - (\mu_2 + \mu_3) = 0,$$

$$\mathrm{H}_0^6 : \mu_2 = \mu_3 \qquad \text{vagy} \qquad \mu_2 - \mu_3 = 0.$$

A 10-3. táblázat foglalja össze a három újabb hipotézisre végzett statisztikai próba eredményeit.

10-3. táblázat

k	c_{ik}	μ_1	μ_2	μ_3	μ_4	$C_k = \sum_i c_{ik} y_{i\cdot}$	$\sum_i c_{ik}^2 / p$	$S_k = \frac{p\left(\sum_i c_{ik} y_{i\cdot}\right)^2}{\sum_i c_{ik}^2}$	ν	$F_0^k = \frac{S_k}{s_R^2}$	p
4	c_{i4}	3	–1	–1	–1	80.32	9/3+1/3+ +1/3+1/3	1613.361	1	20.907	0.0018
5	c_{i5}	0	1	1	–2	26.66	1/3+1/3+ +4/3	355.555	1	4.6076	0.0641
6	c_{i6}	0	1	–1	0	–16.00	1/3+1/3	384.000	1	4.9762	0.0562
								2352.917			

A három 1 szabadsági fokú S_k négyzetösszeg összege itt is 2352.917.

Általában is igaz, hogy az ortogonális hipotézisekhez tartozó egy szabadsági fokú S_k négyzetösszegek összege ugyanannyi, mint az illető faktor hatására kiszámított egyetlen, $r-1$ szabadsági fokú négyzetösszeg, ha az ismétlések p_i száma a faktor mindegyik i szintjén azonos. Ez formálisan azt jelenti, hogy az ANOVA-táblázat szerint F-próbával az ortogonális hipotéziseket együtt vizsgáljuk, vagy fordítva, az egy hipotézist (hogy a faktornak nincs hatása, a szintek között nincs különbség) három független hipotézisre bontjuk.

Gyakorlatilag (a véletlen ingadozások miatt) előfordulhat, hogy az összetett hipotézist elfogadjuk, miközben, ha egyenként vizsgálnánk a hipotéziseket, valamelyiket közülük elutasítanánk. A fordított eset is lehetséges, vagyis hogy az összetett hipotézist (hogy nincs a szintek között különbség) elutasítjuk, miközben a külön-külön hipotézisek mindegyikét elfogadjuk.

10-3. példa

Vizsgáljunk egy olyan esetet is, amelynél az ismétlések p_i száma nem azonos az A faktor egyes i szintjein (nem kiegyensúlyozott terv)! Az adatokat a 10-4. táblázat tartalmazza.

Először végezzük el a szokásos varianciaanalízist annak vizsgálatára, hogy van-e a faktor 4 szintje között különbség, vagyis van-e a faktornak hatása! A számolásnál ne feledkezzünk meg arról, hogy az ismétlések p_i száma nem azonos az A faktor egyes i szintjein, vagyis a négyzetösszegek kiszámítására a (9.26) és (9.27) képleteket kell alkalmazni!

Az eredményeket a 10-5. táblázatban láthatjuk. Azt találjuk, hogy van különbség, vagyis szignifikáns az A faktor hatása.

10-4. táblázat

Faktorszint	1	2	3	4
y_{ij}	253 259 266	217 233	248 235 263 250	217 229 224
p_i	3	2	4	3
$\bar{y}_{i\cdot}$	259.3	225.0	249.0	223.3
s_i	6.506	11.314	11.460	6.028
s_i^2	42.328	128.007	131.332	36.337

10-5. táblázat

Az eltérés forrása	Eltérés-négyzetösszeg	Szabadsági fok	Szórásnégyzet	F_0	p
A hatása (csoportok közötti)	2712.333	3	904.1111	10.647	0.0036
Ismétlések (csoportokon belüli)	679.333	8	84.91666		
Teljes	3391.666	11			

Itt nem mutatjuk be, de mind a Bartlett-, mind a Levene-próba szerint elfogadjuk a varianciák azonosságára vonatkozó nullhipotézist.

Végezzük most el az első három hipotézis vizsgálatát (10-6. táblázat)!
Ezek a hipotézisek logikai értelemben ortogonálisak, de statisztikai értelemben nem ortogonálisak, mert az ismétlések p száma csoportonként különböző. A felhasználót a logikai ortogonalitás érdekli, ekkor a vizsgált állítások függetlenek.

10-6. táblázat

k	c_{ik}	μ_1	μ_2	μ_3	μ_4	$C_k = \sum_i c_{ik} y_{i\cdot}$	$\sum_i c_{ik}^2 / p_i$	$S_k = \dfrac{\left(\sum_i c_{ik} y_{i\cdot}\right)^2}{\sum_i c_{ik}^2 / p_i}$	ν	$F_0^k = \dfrac{S_k}{s_R^2}$	p
1	c_{i1}	1	–1	0	0	34.3	1/3+1/2	1414.533	1	16.658	0.0035
2	c_{i2}	0	0	1	–1	25.7	1/4+1/3	1129.333	1	13.230	0.0065
3	c_{i3}	1	1	–1	–1	12.0	1/3+1/2+ +1/4+1/3	101.647	1	1.197	0.3058
								2645.513			

Vegyük észre, hogy a három egy szabadsági fokú S_k négyzetösszeg összege (2645.513) itt nem egyenlő a 10-5. táblázatban az *A* faktor (az adalék) hatására kapott 3 szabadsági fokú négyzetösszeggel (2712.333), mert az ismétlések p_i száma nem azonos.

Ha az ismétlések száma különböző, nem igaz, hogy az egyes hipotézisekhez tartozó négyzetösszegek összege kiadná az S_A négyzetösszeget; a vizsgált logikailag ortogonális hipotézisek statisztikai értelemben nem ortogonálisak, vagyis az egyes hipotézisekre elvégzett statisztikai próbák nem függetlenek.

10.3.2. Tervezett összehasonlítások

Az összehasonlítást akkor nevezzük tervezettnek, ha az általa vizsgálandó kérdést már a kísérletek elvégzése előtt is indokoltnak látjuk. Lindman (1992, p. 57) szerint a tervezett összehasonlításokra a következő kritériumok érvényesek:

a) A tervezett összehasonlítások legyenek csakugyan tervezettek, tehát olyanok, amelyeket a kísérletek megkezdése előtt is elhatároznánk. Ha a kísérletek elvégzése után, az eredmények ismeretében jelölünk ki összehasonlításokat, abban elkerülhetetlenül befolyásolnak maguk az adatok, és ez megnöveli az elsőfajú hiba kockázatát.

b) A tervezett összehasonlítások legyenek korlátozott számúak, ezzel is elkerülhetjük, hogy maguk az eredmények keltsenek kedvet az összehasonlításra.

c) Ne válasszunk túlságosan nagy α értéket, vagyis ne legyen nagy az elsőfajú hiba megengedett valószínűsége. Ugyanis ha több összehasonlítást végzünk ugyanazon adatokból, megnő annak esélye, hogy csak a véletlen műveként szignifikánsnak találunk olyan különbségeket, amelyek a valóságban nem léteznek. Ha r tervezett összehasonlítást végzünk, mindegyiket α szinten, annak valószínűsége, hogy egy vagy több vizsgálat eredménye a véletlennek köszönhetően szignifikáns lesz, akár $r\alpha$ nagyságú is lehet (l. 10.4. alfejezet).

d) Ha lehet, ortogonális összehasonlításokat végezzünk.

10-4. példa

A 9-1. példa adataival kapcsolatban már a kísérletek elvégzése előtt természetesen felmerülő kérdés – és így egy tervezett összehasonlítás tárgya –, hogy az adalék nélküli és adalékolt lisztek között van-e különbség. Ez azt igényli, hogy az első mintát hasonlítsuk össze a másik három átlagával. A kontrasztok célszerűen a következők:

–3, 1, 1, 1. Ez a 10-2. példában vizsgált $H_0^4: \mu_1 = \frac{\mu_2 + \mu_3 + \mu_4}{3}$ hipotézis. A különbség nagyon jelentős (p = 0.0018).

10-5. példa

Szintén már a kísérletek elvégzése előtt föltehető kérdés a 9-1. példa adataira, hogy a háromféle adalék (a háromféle fehérjeizolátum) hatásában van-e különbség. Tehát ez is tervezett összehasonlítás kérdése lehet.

A választ egy olyan varianciaanalízissel adhatjuk meg, amelyben a kontrollcsoport nem szerepel, tehát csak a 3 kérdéses csoport közötti végzünk összehasonlítást.

A számítás eredményét mutatja a 10-7. táblázat. Látjuk, hogy a három adalék közötti különbségből számított eltérés-négyzetösszeg szabadsági foka 2, a nevezőé, minthogy három csoporton belüli 3-3 ismétlésből számított egyesített szórásnégyzet, 6.

Ahogy a többi kontrasztelemzéseknél, az F-próba nevezőjében szereplő szórásnégyzet számításához nemcsak azokat a csoportokat használhatjuk föl, amelyekre a számlálóban az átlagok szerepelnek (vagyis amely csoportokra a kontrasztegyüttható nem zérus), hanem mindegyiket, mert így az összes csoporton belüli ingadozásra kapott szórásnégyzetben rejlő információt hasznosítjuk. Így jártunk el a fejezet bevezetőjében is, a 10-1. példában két csoport átlagának összehasonlítására végzett kétmintás t-próbánál az összes csoporton belüli ingadozásból egyesített s_R^2 szórásnégyzetet használtuk. Ezt itt is megtehetjük, ezzel az F-próba nevezőjében nagyobb (6 helyett 8) szabadsági fokú szórásnégyzetünk lesz (10-8. táblázat), ami előnyös, mert kisebb a másodfajú hiba valószínűsége, vagyis annak kockázata, hogy egy, a valóságban létező hatást nem tudunk kimutatni, mert a szórás elfedi.

10-7. táblázat

Az eltérés forrása	Eltérés-négyzetösszeg	Szabadsági fok	Szórásnégyzet	F_0	p
A hatása (csoportok közötti)	739.556	2	369.778	5.081	0.0511
Ismétlések (csoportokon belüli), a kontroll nélkül	436.667	6	72.778		

10-8. táblázat

Az eltérés forrása	Eltérés-négyzetösszeg	Szabadsági fok	Szórásnégyzet	F_0	p
A hatása (csoportok közötti)	739.556	2	369.778	4.792	0.0428

Ismétlések (csoportokon belüli)	617.333	8	77.167

A példában a két úton kapott következtetés különböző: a nagyobb szabadsági fokú szórásnégyzettel számolva 0.05-os szinten jelentősnek minősítjük a különbséget, a kisebb szabadsági fokú szórásnégyzettel számolva nem. A nagyobb szabadsági fokú nevezővel végzett összehasonlítás több információt tartalmaz, ezért érzékenyebb.

10.3.3. Post hoc összehasonlítások

A latin kifejezés azt jelenti, hogy már a kísérleti adatok birtokában kívánunk összehasonlítást tenni.

10-6. példa

A 9-3. ábra szerint a búzacsíra és a másik két adalék hatása eltér. Az eredmények diktálta kérdés lehet tehát, hogy a látható különbség statisztikailag jelentős-e.
Az ehhez használható kontrasztegyütthatók: 0, 1, –2, 1. Az ANOVA-tábla a 10-9. táblázat.
A különbség szignifikáns, ez összhangban van azzal, hogy a háromféle adalék között jelentős eltérést találtunk.

További, szintén az eredmények láttán felmerülő kérdés, hogy van-e különbség a közel azonos hatású kukoricacsíra és rizscsíra között. A kontrasztegyütthatók: 0, 1, 0, -1. Ez volt a 10-1. példa kérdése, melyre azt a választ adtuk, hogy nincs jelentős különbség.

10-9. táblázat

Az eltérés forrása	Eltérés-négyzetösszeg	Szabadsági fok	Szórásnégyzet	F_0	p
Kontraszt	1345.515	1	1345.515	15.845	0.0041
Ismétlések (csoportokon belüli)	679.333	8	84.917		

10.4. A többszörös összehasonlítások kockázata

Ha egy kísérletsorozat eredményeiből egyszerre nemcsak egy, hanem több összehasonlítást is kívánunk végezni, megnő annak valószínűsége, hogy elsőfajú hibát kövessünk el. Például ha 20 hipotézist vizsgálunk, egyenként 0.05 szignifikancia-szinten, annak valószínűsége, hogy bármelyiküket csak a véletlen műveként szignifikánsnak találjuk, 1/20. Ha több hipotézist vizsgálunk egyszerre, 1/20-nál jóval nagyobb gyakorisággal fordulhat elő, hogy valamelyiket szignifikánsnak találjuk (és elutasítjuk) pusztán a véletlen műveként, amikor mindegyik igaz.

Jelölje α^* annak valószínűségét, hogy az egyik hipotézist elutasítjuk, pedig igaz, tehát elsőfajú hibát vétünk, mert a véletlen ingadozás következtében a próbastatisztika aktuális értéke kívül esik az elfogadási tartományon. Az igaz nullhipotézist $1-\alpha^*$ valószínűséggel fogadjuk el. α^* az egyetlen összehasonlításra vonatkozó elsőfajú hiba-valószínűség (comparisonwise error rate).

Vizsgáljuk először a *független hipotézisek* esetét. Annak esélye, hogy egy döntésnél a k független (ortogonális) hipotézis mindegyikét elfogadjuk, $(1-\alpha^*)^k$. Annak kockázata, hogy a vizsgált k hipotézis közül egyet vagy többet tévesen elutasítsunk, vagyis legalább egy hipotézisnél elsőfajú hibát kövessünk el,

$$\alpha = 1-(1-\alpha^*)^k. \tag{10.21}$$

Tehát egyetlen összehasonlításnál α^* az elsőfajú hiba elkövetésének valószínűsége, α pedig annak valószínűsége, hogy az egész döntésnél, az egész kísérlet adatainak értékelésénél elsőfajú hibát vétsünk, vagyis valamelyik összehasonlításnál tévesen szignifikáns különbséget állapítsunk meg. Ezért α^*-ot összehasonlításonkénti (comparisonwise), α-t az egész kísérletre (döntéscsaládra) vonatkozó hiba-valószínűségnek nevezik (familywise error rate).

Például ha egy kísérlet adataiból 10 független összehasonlítást végzünk, és $\alpha^* = 0.05$, annak valószínűsége, hogy legalább egyet közülük tévesen elutasítsunk (szignifikánsnak minősítsünk), $\alpha = 1-(1-0.05)^{10} = 0.4$. Ezt csökkenthetjük, ha az összehasonlításonkénti α^* kockázatot kisebbre vesszük (ezáltal szélesebbre az elfogadási tartományt). Ekkor azonban a másodfajú hiba valószínűsége megnő, vagyis nagyobb lesz annak kockázata, hogy egy valóságban létező hatást nem mutatunk ki.

Ha a *hipotézisek nem függetlenek*, annak valószínűségére, hogy legalább egyiket szignifikánsnak találjuk, fölső határt adhatunk (Bonferroni-egyenlőtlenség):

$$\alpha \le k\alpha^*. \tag{10.22}$$

Például, ha a két nem-független hipotézis $\mu_1 = \mu_2$ és $2\mu_1 = 2\mu_2$, és ezeket együtt vizsgáljuk, az elsőfajú hiba valószínűsége megkétszereződik, mert ha az egyiket véletlenül szignifikánsnak találjuk, szükségszerűen a másikat is. Ez a teljes összefüggés szélsőséges esete, ez adja a fölső határt.

Az előbbi 10 (de nem független) összehasonlításra $\alpha \le 10 \cdot 0.05 = 0.5$, ha az összehasonlításonkénti α^* hiba-valószínűség 0.05. Ez azt jelenti, hogy 10 összehasonlítás közül akár 50% valószínűséggel találhatunk legalább egy szignifikánsat, akkor is, ha nincs különbség a csoportok között.

Több eljárás használatos ilyen esetekben a szignifikanciaszintek megválasztására, a megfontolások mellőzése viszont az elsőfajú hiba valószínűségének téves megítéléséhez vezet.

Különösen veszélyes a post hoc vizsgálat, amelynél éppen azok az adatok válnak érdekessé, amelyek eltérnek a többitől. A statisztikai próbák és szignifikanciaszintjeik valószínűségi változókra érvényesek, tehát olyan leendő adatokra, amelyek adott valószínűséggel vehetnek föl bizonyos tartományokban értékeket. Például 20 közül egy jelentősen eltérhet a többitől, pusztán a véletlen miatt, de ez az egy az ismételt kísérleteknél mindig másik lehet. Ha már meglévő adatokkal dolgozunk, az eltérés láttán éppen ennek az eltérőnek a vizsgálatára érzünk indíttatást, és azt kérdezzük, hogy az eltérés szignifikáns-e. A próbánál az elsőfajú hiba megengedett 5%-os valószínűsége azonban csak arra az esetre érvényes, ha a valószínűségi változó még bármilyen értéket fölvehetne, vagyis akkor lenne jogos így vizsgálódnunk, ha nemcsak a gyanús adatokat néznénk. Ha csak a már ránézésre is eltérő adatra végezzük el a próbát, az elsőfajú hiba kockázata jóval nagyobb a deklaráltnál.

Milliken és Johnson (1992) javaslata szerint a következőképpen célszerű eljárni:

1. Végezzünk *F*-próbát annak vizsgálatára, hogy van-e különbség a csoportok várható értéke között!
2. Ha az eltérés 5%-os szinten szignifikáns, végezzünk tervezett összehasonlításokat az LSD-módszerrel (Fisher-féle LSD)! Nemcsak párok, hanem tetszőleges kontrasztok is vizsgálhatók. Ha nem tervezett összehasonlításokat végzünk, vagyis sok kontrasztot vizsgálunk, a Scheffé-próba használandó.
3. Ha az összes várható érték egyenlőségére végzett *F*-próba nem szignifikáns, akkor is végezhetünk tervezett összehasonlításokat, a Bonferroni-módszer alkalmazásával. Ilyenkor a Scheffé-módszer nem jelez különbséget.

Előfordulhat, hogy az *F*-próbával elfogadnánk az összes várható érték egyenlőségére vonatkozó nullhipotézist, egyes párok vizsgálatával azonban szignifikáns különbséget találunk. Ennek oka, hogy a többiek ingadozása elfedi az eltérést. Az is lehetséges, hogy az *F*-próba szignifikáns különbözőséget mutat, páronként mégsem találunk ilyet.

10.4.1. LSD-próba

A többszörös összehasonlításoknál egy-egy kontraszt szignifikanciájára *F*-próbát (vagy azzal ekvivalens *t*-próbát) végzünk.

Láttuk a (10.19) egyenletnél, hogy a *k*-adik kontraszt S_k négyzetösszege, ha a *k*-adik nullhipotézis igaz, $\chi^2\sigma_e^2$ eloszlású, szabadsági foka 1. Az S_k emiatt egyszerre négyzetösszeg és szórásnégyzet is. A próbát úgy végezzük, hogy S_k -ból egy F_0 próbastatisztikát számítunk ki:

$$F_0^k = \frac{S_k}{s_R^2} = \frac{\left(\sum_i c_{ik} y_{i\cdot}\right)^2}{s_R^2 \sum_i c_{ik}^2 / p_i}, \tag{10.23}$$

melyre a szabadsági fokok 1 és $\sum_i p_i - r$.

A $\mathrm{H}_0^k : \sum_i c_{ik}\mu_i = 0$ nullhipotézist akkor utasítjuk el, ha F_0 meghaladja a kritikus értéket. Figyelembe véve, hogy a számláló szabadsági foka 1, az F_0 próbastatisztika négyzetgyöke egy t_0 kifejezés, melynek szabadsági foka $\sum_i p_i - r$. Az *F*-próbával teljesen azonos értékű (és következtetésű) *t*-próbát is végezhetünk. Ennél a $\mathrm{H}_0^k : \sum_i c_{ik}\mu_i = 0$ hipotézist akkor utasítjuk el, ha t_0 meghaladja a kritikus értéket. Egy előre kijelölt α szignifikanciaszinthez (az elsőfajú hiba megengedett α valószínűségéhez) ez azt jelenti, hogy a nullhipotézist elutasítjuk, amennyiben

$$\left|\sum_i c_{ik} y_{i\cdot}\right| \geq t_{\alpha/2} s_R \sqrt{\sum_i c_{ik}^2 / p_i}\ , \tag{10.24}$$

az egyenlőtlenség jobb oldala az a legkisebb eltérés, amelyet már szignifikánsnak minősítünk (Least Significant Difference: *LSD*).

Úgy is eljárhatunk, hogy kiszámítjuk annak p valószínűségét, hogy a különbség elérje vagy meghaladja a ténylegesen talált mértéket:

$$p = P\left(\left|\sum_i c_{ik} y_{i\cdot}\right| \geq t_{p/2} s_R \sqrt{\sum_i c_{ik}^2 / p_i}\right). \tag{10.25}$$

Amennyiben ez a p valószínűség nagyobb az elsőfajú hiba megengedett α valószínűségénél, elvetjük a nullhipotézist, és azt mondjuk, hogy a különbség szignifikáns.

10-7. példa

Hasonlítsunk össze minden lehetséges párt a 10-1. példában!

A négy szintre vonatkozó $\binom{4}{2} = 6$ összehasonlítás egymástól nyilvánvalóan nem független, vagyis nem ortogonális.

Az LSD-próba eredményét mutatja a 10-10. táblázat. A táblázatban a j-edik sor és a k-adik oszlop metszéspontjában annak p valószínűsége van, hogy a j és k csoport (szint) között tapasztalt különbség csak a véletlen ingadozás következménye, vagyis a $\mathrm{H}_0 : \mu_j = \mu_k$ nullhipotézist vizsgáljuk:

$$p = P\left(\left|y_{j\cdot} - y_{k\cdot}\right| \geq t_{p/2} s_R \sqrt{\frac{1}{p_j} - \frac{1}{p_k}}\right)$$

A kontrasztegyütthatók itt: $c_j = 1$, $c_k = -1$, a többi c zérus.

Példaképpen nézzük az 1-2 összehasonlítás p értékének kiszámítását részletesen!

$$y_{1\cdot} = 259.33,\ y_{2\cdot} = 229.00,\ t_{p/2} = \frac{y_{1\cdot} - y_{2\cdot}}{s_R\sqrt{\frac{1}{p_1} - \frac{1}{p_2}}} = \frac{259.33 - 229.00}{\sqrt{77.167}\sqrt{\frac{1}{3} + \frac{1}{3}}} = 4.229.$$

A $\nu_R = 8$ szabadsági fokú t-eloszlású valószínűségi változó a 4.229 értéket fölfelé ill. a –4.229 értéket lefelé (a statisztikai szoftverrel kiszámítva) 0.0029 valószínűséggel haladja meg, ez tehát p értéke.

10-10. táblázat

	1	2	3	4
átlag	259.33	229.00	245.00	223.67
1		0.0029	0.0807	0.0011

2	0.0029		0.0562	0.4784
3	0.0807	0.0562		0.0178
4	0.0011	0.4784	0.0178	

Például az 1-2 összehasonlításhoz tartozó *p* valószínűség (annak valószínűsége, hogy a két csoport átlaga között a talált vagy nagyobb különbség legyen, ha a várható értékek a valóságban megegyeznek), 0.0029, igen kicsi, tehát a különbséget szignifikánsnak minősítjük, vagyis elutasítjuk azt a hipotézist, hogy az 1. és 2. szint között nincs különbség. Annak valószínűsége viszont, hogy a 2. és 4. szint között tapasztalt különbség a véletlen műve, 0.4784, igen nagy, tehát ezt a különbséget nem minősítjük szignifikánsnak.

Eszerint sincs a kukoricacsíra és rizscsíra (2-4) között jelentős különbség, és az adalékolatlan és a rizscsíra (1-3) között sem, minden más pár között azonban van.

Ha csak az adalékoltakat hasonlítjuk egymással (vagyis az adalékolatlan lisztre kapott eredményeket nem használjuk föl), a következtetés hasonló (10-11. táblázat).

10-11. táblázat

	2	3	4
átlag	229.00	245.00	223.67
2		0.0613	0.4729
3	0.0613		0.0221
4	0.4729	0.0221	

Érdemes felfigyelni arra, hogy a *p* értékek ugyanarra a párra nagyobbak, ha csak 3 csoportot hasonlítunk össze. Ennek oka, hogy a reziduális szórásnégyzet szabadsági foka ilyenkor kisebb (8 helyett 6, l. a 10-8. táblázatot).

10.4.2. A Fisher-féle LSD

Először globális *F*-próbát végzünk annak a nullhipotézisnek a vizsgálatára, hogy nincs különbség a csoportok között. Csak ha ezt elutasítjuk, akkor teszünk további összehasonlításokat az egyes kontrasztokra (pl. párokra). Az irodalom szerint ekkor az elsőfajú hiba kockázata lényegesen kisebb, mivel már eleve csak akkor haladunk tovább, ha az *F*-próba különbséget jelzett.

A 10-7. példában (ennek külön említése nélkül) a páronkénti összehasonlítást (az LSD-próbát) már annak tudatában írtuk elő, hogy a 9-1. példában, a varianciaanalízisnél az adalékok közötti különbségeket szignifikánsnak találtuk, vagyis tulajdonképpen Fisher-féle LSD-vizsgálatot végeztünk.

10.4.3. A Bonferroni-eljárás

Ha azt akarjuk, hogy az egész kiértékelésre (tehát az összes hipotézis vizsgálatára) együtt érvényes elsőfajú hiba-valószínűség (familywise error rate) ne haladjon meg egy határt, az egyes vizsgálatokra (az F-próbánál vagy t-próbánál) vonatkozó α^* (comparisonwise error rate) e szerint választandó.

Független hipotézisek esetén:

$$\alpha^* = 1 - (1-\alpha)^{1/k}. \qquad (10.26)$$

Ha a hipotézisek nem függetlenek, vagy legalábbis ennek veszélye fönnáll, az $\alpha^* = \alpha/k$ választás a megfelelő.

A 10-7. példára $\alpha^* = 0.05/6 = 0.0083$.

Ugyan az egyes hipotézisvizsgálatokat $\alpha^* = 0.0083$ szinten végezzük, de az összes vizsgálatra vonatkozó következtetés levonásakor 0.05 szintet nevezünk meg.

A módszer használata akkor javasolható, ha néhány összehasonlítást akarunk csak végezni, és az F-próba nem bizonyult szignifikánsnak.

10-8. példa

A 10-7. példában (10-10. táblázat) a 9-1. példa adataira kiszámítottuk annak valószínűségét, hogy az egyes párok között tapasztalt mértékű (vagy annál nagyobb) különbség a véletlen ingadozás következtében jöjjön létre. Összesen 6 ilyen pár van, tehát ha mindegyik összehasonlítást egyszerre végezzük el, 6 összehasonlítást teszünk ($\mu_1 = \mu_2$, $\mu_1 = \mu_3$, $\mu_1 = \mu_4$, $\mu_2 = \mu_3$, $\mu_2 = \mu_4$, $\mu_3 = \mu_4$). Ezek az összehasonlítások nem függetlenek, tehát ha a teljes vizsgálatnál az elsőfajú hiba megengedett α valószínűsége 0.05, egy-egy összehasonlításra $\alpha^* = 0.05/6 = 0.0083$. Akkor utasítunk el egy nullhipotézist a Bonferroni-módszer szerint (minősítünk két csoportot egymástól szignifikánsan különbözőnek), ha a talált p értékre $p < 0.0083$. Eszerint a 10-10. táblázat alapján az 1-2 és 1-4 párok szignifikánsan különböznek, a többi nem. Az LSD-módszerrel a 3-4 pár is szignifikánsan különbözőnek bizonyult, a Bonferroni-módszerrel nem.

10.4.4. A Holm-eljárás

A szakirodalom szerint a Bonferroni-eljárás túlságosan konzervatív, vagyis nehezebben utasítja el a nullhipotéziseket.

A szekvenciális eljárások kevésbé konzervatívak, de kicsivel bonyolultabbak. Például a Holm-eljárás döntési szabálya a következő:

1. Vizsgáljuk meg egyenként a nullhipotéziseket, és számítsuk ki a p értékeket (annak valószínűségét, hogy az illető nullhipotézis igazsága esetén a talált vagy annál szélsőségesebb próbastatisztika-értéket kapunk)! Rendezzük sorba a p értékek szerint a próbákat, kezdve a legkisebbel!
2. Ha $p_{(1)} < \alpha^*$, utasítsuk el az első nullhipotézist, és folytassuk a vizsgálatot a következővel, ha nem, fogadjuk el az első nullhipotézist, és hagyjuk abba a vizsgálatot!
3. Ha folytatjuk, ugyanezt az utat kövessük a második, harmadik stb. nullhipotézisre.

Természetesen ezt a módszert csak akkor alkalmazhatjuk, ha a próbákat statisztikai programmal végezzük, és megkapjuk a p értékeket.

10-9. példa

A Holm-módszerrel először a legkisebb *p* értékre vizsgálódunk, az a 10-10. táblázat szerint az 1-4 párra vonatkozik (0.0011). Mivel $0.0011 < 0.0083$, az 1. és 4. csoport egyenlőségére vonatkozó nullhipotézist elutasítjuk, és folytatjuk a vizsgálatot. A második legkisebb *p* érték az 1-2 párra vonatkozik (0.0029). Mivel $0.0029 < 0.0083$, az 1. és 2. csoport egyenlőségére vonatkozó nullhipotézist is elutasítjuk, és folytatjuk a vizsgálatot. A harmadik legkisebb *p* érték a 3-4 párra vonatkozik (0.0178). Mivel $0.0178 > 0.0083$, a 3. és 4. csoport egyenlőségére vonatkozó nullhipotézist elfogadjuk, és befejezzük a vizsgálatot, a további nullhipotéziseket (1-3, 2-3, 2-4) is elfogadva.

10.4.5. Scheffé módszere

Az LSD-próbánál a (118.23) szerinti F_0^k próbastatisztika számlálójának szabadsági foka 1, mivel a *k*-adik kontraszt S_k négyzetösszegének 1 a szabadsági foka.

Az elsőfajú hiba kockázatát csökkenthetjük, ha $r-1$ összehasonlítás esetén a számlálóban lévő négyzetösszeg szabadsági fokát 1 helyett $r-1$-nek vesszük, vagyis $r-1$-gyel osztunk, és a táblázatból való kikereséskor vagy a *p* valószínűség számításakor is ezt vesszük alapul:

$$F_{0Sch}^k = \frac{\dfrac{SS_k}{r-1}}{s_R^2} = \frac{\dfrac{\left(\sum_i c_{ik} y_{i\cdot}\right)^2}{r-1}}{s_R^2 \sum_i \dfrac{c_{ik}^2}{p_i}} . \tag{10.27}$$

Akkor utasítjuk el a nullhipotézist, ha

$$\left|\sum_i c_{ik} y_{i\cdot}\right| > \sqrt{(r-1) F_\alpha\left(r-1, \sum_i p_i - 1\right) s_R^2 \sum_i c_{ik}^2 / p_i} . \tag{10.28}$$

Emlékeztetőül, az LSD-eljárásnál akkor utasítjuk el a nullhipotézist, ha

$$\left|\sum_i c_{ik} y_{i\cdot}\right| > \sqrt{F_\alpha\left(1, \sum_i p_i - 1\right) s_R^2 \sum_i c_{ik}^2 / p_i} . \tag{10.29}$$

A (10.29) egyenlet teljesen azonos a (10.24) egyenlettel, ha figyelembe vesszük, hogy $\sqrt{F_\alpha(1,\nu)} = t_{\alpha/2}(\nu)$.

Az LSD-eljárásnál akár *t*-, akár *F*-eloszlással számolhatunk, a Scheffé-módszernél csak *F*-eloszlással.

A Scheffé-módszerrel számított elfogadási tartomány szélesebb lesz. Emiatt csökken a próba ereje, vagyis megnő a másodfajú hiba elkövetésének kockázata, azaz annak veszélye, hogy nem mutatunk ki egy létező különbséget. A módszer nagyszámú összehasonlítás esetén használandó.

10-10. példa

Vessük össze a 9-1. példa adataira az LSD-módszerrel kapott elemzési eredményeket (10-10. táblázat) a Scheffé-próba eredményeivel, megint belevéve a kontrollcsoportot!

Az eredmények a 10-12. táblázatban láthatók. A p értékek itt a következő valószínűséget jelentik:

$$p = P\left\{\left|\sum_i c_{ik} y_{i\cdot}\right| \geq \sqrt{(r-1)F_p\left(r-1, \sum_i p_i - 1\right)s_R^2 \sum_i c_{ik}^2 / p_i}\right\}. \tag{10.30}$$

10-12. táblázat

	1	2	3	4
átlag	259.33	229.00	245.00	223.67
1		0.0195	0.3306	0.0079
2	0.0195		0.2519	0.9041
3	0.3306	0.2519		0.0984
4	0.0079	0.9041	0.0984	

Példaképpen nézzük az 1-2 összehasonlítás p értékének kiszámítását részletesen!

$$y_{1\cdot} = 259.33\,,\ y_{2\cdot} = 229.00\,,\ F = \frac{\dfrac{(y_{1\cdot} - y_{2\cdot})^2}{r-1}}{s_R^2\left(\dfrac{1}{p_1} - \dfrac{1}{p_2}\right)} = \frac{\dfrac{(259.33 - 229.00)^2}{3}}{77.167\left(\dfrac{1}{3} + \dfrac{1}{3}\right)} = 5.961$$

Az F-eloszlású valószínűségi változó (amelynél a számláló szabadsági foka 3, a nevezőé 8), az 5.961 értéket 0.0195 valószínűséggel haladja meg, ez lesz p értéke.

Itt a p értékek nagyobbak, mint az LSD-módszer szerint, vagyis ekkor kisebb biztonságúak a következtetéseink, összhangban azzal, hogy nem túl nagy számú összehasonlításnál a Scheffé-próba gyengébb. Például a 3. és 4. szint közötti különbséget itt nem minősítenénk szignifikánsnak, az LSD-nél pedig annak találtuk.

10.4.6. Összehasonlítás kontrollcsoporttal

A tervezett összehasonlítások között különösen gyakori az ún. kontrollcsoporttal való összevetés. A kérdést úgy szokás föltenni, hogy az egyetlen faktor (a „kezelés") egyes szintjein kapott eredmények különböznek-e a kezeletlen (ún. kontroll-) csoport eredményeitől. A liszt-példában ez úgy fogalmazható meg, hogy az egyes adalékokkal készült lisztek MT tulajdonsága különbözik-e az adalék nélküli lisztétől. A 10-4. példabeli vizsgálatnál a kérdés úgy szólt, hogy az adalékolt lisztek (a háromféle adalékkal átlagosan) különböznek-e az adalékolatlantól.

A kérdés nyilvánvalóan kétmintás t-próbákkal válaszolható meg, de ezek a t-próbák nem lesznek egymástól függetlenek.

10-11. példa

A 9-1. példa adataira vizsgáljuk meg, hogy különböznek-e az egyes adalékokkal készült lisztek az adalékolatlanoktól!
A nullhipotézis az egyes összehasonlításokhoz:

$$H_0 : \mu_1 = \mu_i ,$$

és ezt minden i-edik adalékra el kell végezni (i = 2, 3, 4). A megfelelő próbastatisztika:

$$t_0 = \frac{y_{1\cdot} - y_{i\cdot}}{s_R \sqrt{\frac{1}{p_1} + \frac{1}{p_i}}} .$$

Ehhez a próbához az egyes i-edik adalékokra már kiszámítottuk a p értékeket, ezek a 10-10. táblázat első sorában vannak:

	1	2	3	4
1		0.0029	0.0807	0.0011

Amit itt másképpen kell megadnunk, az az összehasonlításonkénti kritikus p érték (α^*). Mivel a hipotézisek nem függetlenek, egy-egy összehasonlításra az elsőfajú hiba megengedett valószínűsége a Bonferroni-módszer szerint az $\alpha^* = \alpha/k$ képlettel adható meg, ahol k az egy döntésnél elvégzett összehasonlítások száma, itt k = 3.

Ha az itteni három szimultán összehasonlításnál együtt az elsőfajú hiba megengedett valószínűsége 0.05, akkor egyetlen összehasonlításra $\alpha^* = 0.05/3 = 0.0167$, eszerint a 2. és a 4. adalékkal készült liszt lényegesen különbözik az adalék nélkülitől.

A Dunnett-féle (1965) t-próba explicit módon veszi figyelembe, hogy a kétmintás t-próbák nem függetlenek egymástól. Az i-edik és l-edik csoport átlagának a kontrollcsoportétól való $y_{1\cdot} - y_{i\cdot}$ és $y_{1\cdot} - y_{l\cdot}$ eltérése közötti korrelációs együtthatót kiszámolva elvileg élesebb következtetést tesz lehetővé, gyakorlatilag az így kiszámított t ill p értékek nem különböznek a Bonferroni-közelítéssel kiszámítottaktól.