

## One-Way ANOVA – jednoduchá analýza rozptylu

---

### ANOVA

Jednoduchá analýza rozptylu (ANOVA, Analysis of Variance) je statistická metoda, která se užívá k testování hypotézy o shodě středních hodnot dané číselné proměnné v několika nezávislých skupinách. ANOVA je zobecněním T-testu pro dva nezávislé výběry, umožňuje však porovnávat více skupin. Skupiny jsou definovány na základě kategorií jedné kategorizované proměnné (faktoru) a mohou mít různý rozsah. Užití testu má smysl pouze tehdy, pracujeme-li s výběrovými daty a neznáme skutečné střední hodnoty rozdělení, ze kterých data pocházejí.

Analýzu rozptylu aplikujeme například při zjišťování, které z faktorů významně ovlivňují sledovanou proměnnou.

ANOVA vychází z následujících předpokladů:

- pozorování jsou mezi sebou navzájem nezávislá
- skutečné hodnoty a chyby jsou navzájem nezávislé
- výběry pocházejí z normálního rozdělení
- výběry jsou navzájem nezávislé (mají prázdný průnik)
- ve skupinách jsou stejné rozptyly.

Procedura rovněž poskytuje výsledek Levenova testu pro ověření shody rozptylů ve skupinách. Jestliže předpoklad není splněn, jsou k dispozici alternativní testy (Welsch, Brown-Forsythe), jejichž užití je v této situaci vhodnější. Statistika  $F$  je přitom poměrně robustní vzhledem k odchylkám od tohoto předpokladu v případě, že je velikost skupin stejná nebo téměř stejná. Pokud se však velikosti skupin i rozptyly ve skupinách liší, je vhodnější užít Welschův test.

Analýza rozptylu je poměrně robustní vzhledem k odchylkám od normálního rozdělení, je však nutné především ověřit, zda je rozdělení symetrické a nevyskytují se v něm extrémní hodnoty.

Pro podrobnější porovnání průměrů existují dva typy testů – kontrasty a Post Hoc testy. Kontrast je lineární funkce průměrů, která vyjadřuje hypotézu složitější komparace (průměry zde vystupují váženě). Tyto hypotézy jsou obvykle formulovány před začátkem experimentu. Je-li faktorem ordinální proměnná, je dále k dispozici test pro zjišťování, zda v datech existuje trend závislé proměnné vzhledem k uspořádaným hodnotám faktoru. Můžeme tedy například testovat, zda existuje lineární vztah mezi průměrným platem a stupněm vzdělání respondenta.

Post Hoc testy naopak poskytují dodatečnou informaci v případě, že zjistíme mezi průměry statisticky významný rozdíl a chceme konkrétně vědět, které skupiny se od sebe odlišují.

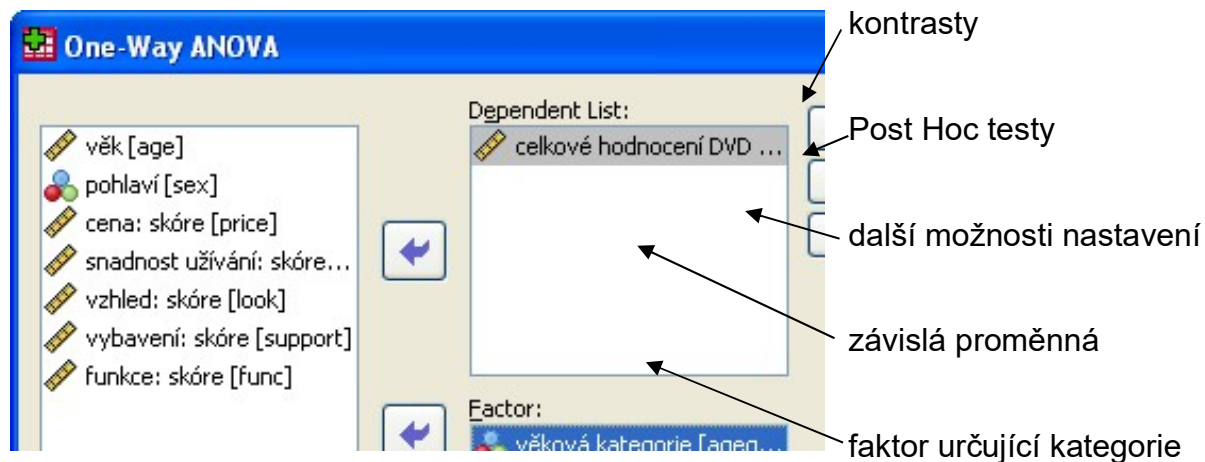
## Listy procedur IBM SPSS Statistics

Pro vícerozměrnou analýzu rozptylu (tj. pro případ, že jsou skupiny definovány na základě kategorií více proměnných a zajímají nás rovněž vzájemné interakce) je určena nabídka *Analyze, General Linear Model, Univariate*.

## Volání procedury v IBM SPSS Statistics

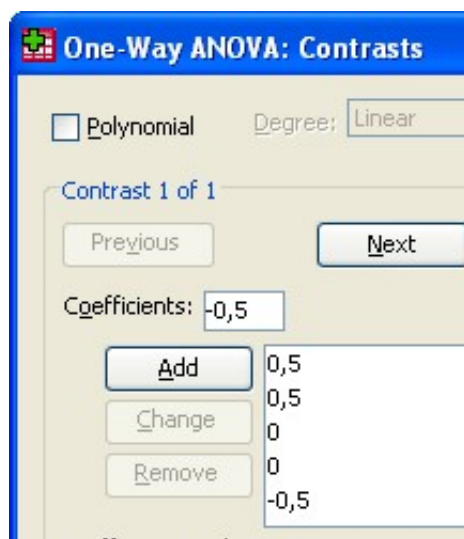
*Analyze* → *Compare Means* → *One-Way ANOVA*

## Nastavení dialogu



- Do pole *Dependent List* přeneseme proměnnou, jejíž střední hodnoty porovnáváme. Uvedeme-li současně více proměnných, provede se test pro každou z nich zvlášť.
- Do pole *Factor* zadáme kategorizovanou proměnnou (faktor), na základě které se data rozdělí do skupin.

## Tlačítka Contrasts



- Při zaškrtnutí políčka *Polynomial* testujeme, zda existuje trend ve vztahu středních hodnot závislé proměnné na faktoru. V rozbalovacím seznamu *Degree* vybereme typ zkoumané závislosti (lineární, kvadratická, kubická, polynom 4. nebo 5. stupně). Tento test má smysl pouze v případě, že je faktorem ordinální proměnná.
- Pro testování uživatelem definovaného typu závislosti se užívá dolní část okna. Do políčka *Coefficients* postupně doplníme vhodné koeficienty kontrastu v pořadí skupin a potvrdíme tlačítkem *Add*. Tyto koeficienty definují lineární kombinaci středních hodnot ve skupinách, která má být za nulové hypotézy rovna nule. Počet koeficientů musí být stejný jako počet skupin (první koeficient se vztahuje k nejnižší skupině, poslední k nejvyšší). Součet koeficientů by měl být nulový. Pokud při zadání uděláme chybu, můžeme ji opravit tlačítky *Change* nebo *Remove*.

### Tlačítko *Post Hoc*

One-Way ANOVA: Post Hoc Multiple Comparisons

Equal Variances Assumed

☐ LSD ☐ S-N-K ☐ Waller-Duncan

☒ Bonferroni ☐ Tukey Type I/Type II Error Ratio: 100

☐ Sidak ☐ Tukey's-b ☐ Dunnett

☐ Scheffe ☐ Duncan Control Category: Last

☐ R-E-G-W F ☐ Hochberg's GT2

☐ R-E-G-W Q ☐ Gabriel

Test

☒ 2-sided ☐ < Control

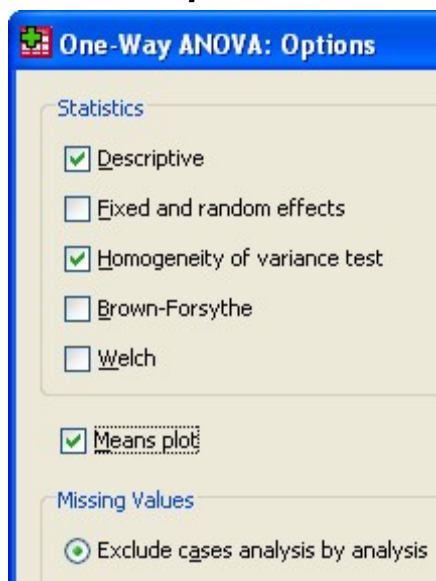
Equal Variances Not Assumed

☐ Tamhane's T2 ☐ Dunnett's T3 ☐ Games-Howell ☐ Dunnett

Tlačítkem *Post Hoc* zadáváme požadavek na výpočet dalších testů pro podrobnější porovnání průměrů. V případě, že zjistíme mezi průměry statisticky významný rozdíl, můžeme zde konkrétně testovat, které průměry se od sebe odlišují. Uvedené testy nabízejí především vzájemné srovnání středních hodnot všech dvojic skupin nebo vytvoření homogenních podmnožin.

- V oddíle *Equal Variances Assumed* jsou k dispozici metody, které předpokládají vzájemnou shodu rozptylů ve skupinách.
- Pro výběr robustnějších metod, které shodu rozptylů nevyžadují, je určen oddíl *Equal Variances Not Assumed*.
- V poli *Significance level* určíme požadovanou hladinu významnosti pro použité testy.

### Tlačítko *Options*



Pomocí tlačítka *Options* si můžeme vyžádat další typy výstupů a zvolit způsob práce s chybějícími hodnotami.

V části *Statistics* označíme požadované statistiky a doprovodné testy:

- *Descriptive* – přehled popisných statistik ve skupinách i celkem (počet případů, průměr, směrodatná odchylka, standardní chyba průměru, 95% interval spolehlivosti, minimum, maximum).
- *Fixed and random effects* – míry pro pevné efekty: směrodatná odchylka, standardní chyba, 95% interval spolehlivosti a míry pro náhodné efekty: standardní chyba, 95% interval spolehlivosti, odhad rozptylu komponent. Za pevný efekt považujeme takový faktor, jehož kategorie jsou v datech zastoupeny kompletně. U náhodného efektu předpokládáme, že kategorie v našich datech představují pouze náhodný výběr z většího počtu možností.
- *Homogeneity of variance test* – Levenův test shody rozptylů ve skupinách.
- *Brown-Forsythe* – výpočet Brown-Forsytheovy statistiky pro testování shody středních hodnot ve skupinách. Tato statistika je vhodnější než statistika  $F$  v případě, že není splněn předpoklad shody rozptylů ve skupinách.
- *Welch* – výpočet Welchovy statistiky pro testování shody středních hodnot ve skupinách. Tato statistika je vhodnější než statistika  $F$  v případě, že není splněn předpoklad shody rozptylů ve skupinách.

*Mean plot* – čárový graf průměrných hodnot analyzované proměnné ve skupinách určených kategoriemi faktoru.

V části *Missing Values* volíme způsob zacházení s chybějícími hodnotami:

- Označíme-li *Exclude cases analysis by analysis*, vyloučí se v každé analýze samostatně pouze nezbytně nutné případy (tj. ty případy, kde chybí hodnota u analyzované proměnné nebo u faktoru). Výhodou tohoto přístupu je využití maximálního možného počtu případů. Nevýhodou mohou být naopak různé počty případů, ze kterých jednotlivé analýzy vycházejí.

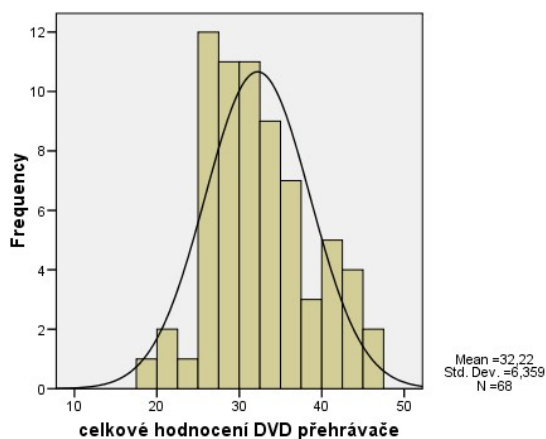
## Listy procedur IBM SPSS Statistics

- Při volbě *Exclude cases listwise* program vyloučí ze všech analýz případy, které mají u některé z analyzovaných proměnných vynechanou hodnotu. Všechny analýzy tedy vycházejí ze stejných případů, je-li však chybějících hodnot větší počet, může se tímto způsobem počet analyzovaných případů výrazně snížit.

## Výstupy

Firma vyvinula nový typ DVD přehrávače. Na základě informací z výzkumu, který provedla, je třeba porovnat, zda respondenti různých věkových kategorií hodnotí přehrávač různě. Datový soubor obsahuje informace o *celkovém hodnocení DVD přehrávače* i o *věkové kategorii respondenta*.

### Ověření předpokladu normality



One-Sample Kolmogorov-Smirnov Test

		celkové hodnocení DVD přehrávače
N		68
Normal Parameters a,b	Mean	32.22
	Std. Deviation	6.359
Most Extreme Differences	Absolute	.099
	Positive	.099
	Negative	-.069
Kolmogorov-Smirnov Z		.820
Asymp. Sig. (2-tailed)		.511

a. Test distribution is Normal.

b. Calculated from data.

Před užitím analýzy rozptylu je nejprve nutné ověřit, zda jsou splněny předpoklady modelu. K ověření normality použijeme například histogram pro vizuální kontrolu a dále Kolmogorov-Smirnovův neparametrický test (*Analyze, Nonparametric Tests, 1-Sample K-S*). Z posledního řádku tabulky vyplývá, že na 95% hladině spolehlivosti nezamítáme nulovou hypotézu o tom, že výběr pochází z normálního rozložení.

Následující výstupy byly získány pomocí jednoduché analýzy rozptylu.

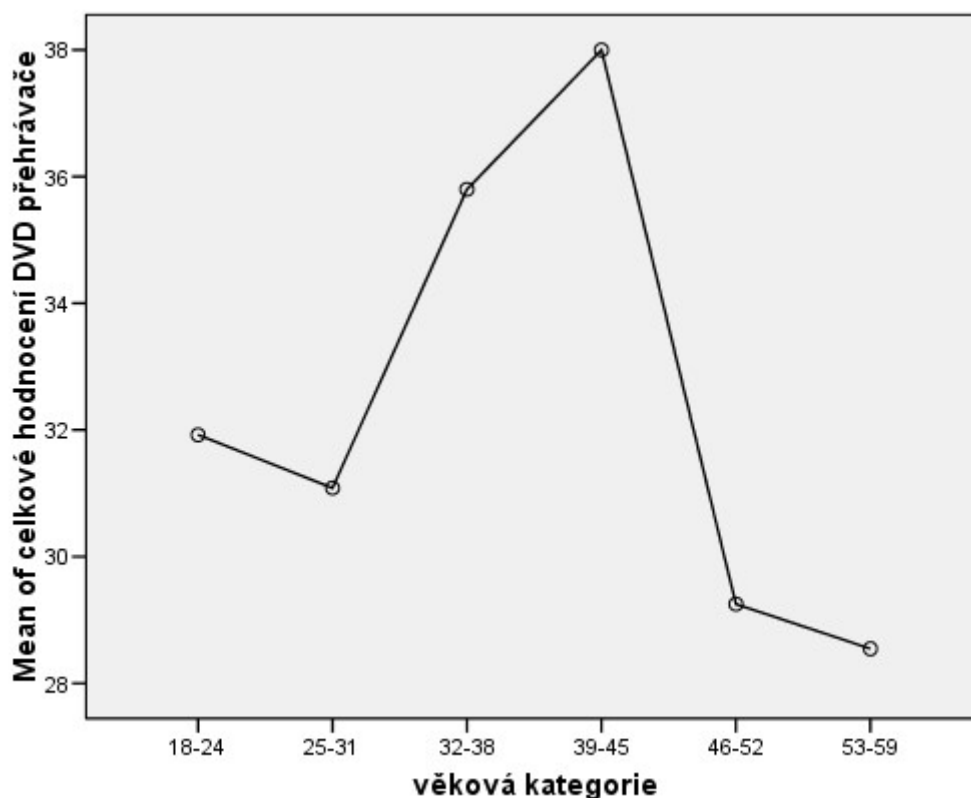
## Popisné statistiky

### Descriptives

celkové hodnocení DVD přehrávače									
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	
					Lower Bound	Upper Bound			
18-24	13	31.92	4.958	1.375	28.93	34.92	26	39	
25-31	12	31.08	5.664	1.635	27.48	34.68	24	40	
32-38	10	35.80	5.308	1.679	32.00	39.60	30	44	
39-45	10	38.00	6.600	2.087	33.28	42.72	28	47	
46-52	12	29.25	6.047	1.746	25.41	33.09	20	41	
53-59	11	28.55	5.298	1.598	24.99	32.10	18	37	
Total	68	32.22	6.359	.771	30.68	33.76	18	47	

Tabulka *Descriptives* poskytuje přehled o průměrech a dalších popisných statistikách celkového hodnocení DVD přehrávače ve skupinách určených věkovými kategoriemi.

### Graf průměrů ve skupinách



V čárovém grafu průměrů máme k dispozici vizuální srovnání průměrných hodnot celkového hodnocení ve skupinách. Je-li kategorizovaná proměnná ordinální (kategorie jsou uspořádané), získáme zde například informaci o tom, zda je mezi skupinami patrný určitý trend. V našem případě přehrávač průměrně hodnotí o něco výše respondenti ve věku 32–38 a 39–45 let, ve srovnání s ostatními kategoriemi.

### Levenův test shody rozptylů

#### Test of Homogeneity of Variances

celkové hodnocení DVD přehrávače

Levene Statistic	df1	df2	Sig.
.574	5	62	.720

Pomocí *Levenova testu* ověříme, zda je splněn předpoklad shody rozptylů ve skupinách. Tabulka zobrazuje ve sloupcích postupně: testovou statistiku (*Levene Statistic*), stupně volnosti (*df1*, *df2*) a signifikance testu (*Sig.*). Vzhledem k tomu, že hodnota signifikance je větší než 5 %, nezamítáme na 95% hladině spolehlivosti nulovou hypotézu o shodě rozptylů.

### Tabulka analýzy rozptylu

#### ANOVA

celkové hodnocení DVD přehrávače

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	733.274	5	146.655	4.601	.001
Within Groups	1976.417	62	31.878		
Total	2709.691	67			

Analýza rozptylu testuje nulovou hypotézu o rovnosti středních hodnot ve skupinách proti alternativní hypotéze, že existují alespoň dvě skupiny, jejichž střední hodnoty se od sebe liší. Metoda je založena na podílu variability mezi skupinami a variability uvnitř skupin. Z tohoto poměru vychází výpočet testové statistiky *F*, která je dále charakterizována dvěma stupni volnosti.

Tabulka ANOVA obsahuje součty čtverců (tj. míru variability) ve skupinách, mezi skupinami a celkem (*Sum of Squares*), stupně volnosti (*df*), průměrné čtverce, tj. podíl součtu čtverců a stupňů volnosti (*Mean Square*), testovou statistiku (*F*) a dosaženou hladinu významnosti (*Significance*). V našem případě je hodnota signifikance menší než 5 %, a proto na 95% hladině spolehlivosti zamítáme nulovou hypotézu o shodě středních hodnot ve skupinách.

Z tabulky ANOVA tedy vyplývá, že existují rozdíly ve středních hodnotách celkového hodnocení DVD přehrávače ve skupinách. Nyní se pokusíme konkrétně prozkoumat, které skupiny se od sebe odlišují. Procedura umožňuje testovat kromě shody všech středních hodnot také dílčí hypotézy.

## Koeficienty kontrastů

Contrast Coefficients

Contrast	věková kategorie					
	18-24	25-31	32-38	39-45	46-52	53-59
1	.5	.5	0	0	-.5	-.5

Kontrast je lineární funkce průměrů, která vyjadřuje hypotézu složitější komparace. Pomocí kontrastů můžeme testovat, zda střední hodnoty ve skupinách vykazují určitý typ závislosti na kategorizované proměnné (např. lineární závislost), nebo formulovat obecnější hypotézu přímo zadáním koeficientů lineární funkce.

Tabulka *Contrast Coefficients* poskytuje přehled koeficientů požadovaných kontrastů. V tomto případě jsme formulovali jednu hypotézu, která je určena koeficienty 0.5, 0.5, 0, 0, -0.5, -0.5, které se postupně vztahují k jednotlivým věkovým kategoriím. Budeme tedy testovat, zda se součet středních hodnot dvou nejnižších věkových kategorií rovná součtu středních hodnot dvou nejvyšších kategorií.

## Kontrasty – testování hypotéz

Contrast Tests

			Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
celkové hodnocení DVD přehrávače	Assume equal variances	1	2.61	1.633	1.596	62	.116
	Does not assume equal	1	2.61	1.594	1.635	42.280	.110

Tabulka *Contrast Tests* obsahuje výsledky testování hypotéz kontrastů. Tabulka je rozdělena na dvě části – pro případ, že je splněn předpoklad o shodě rozptylů ve skupinách (*Assume equal variance*), a pro případ, že předpoklad splněn není (*Does not assume equal*). Zde budeme číst pouze první řádek.

Tabulka postupně zobrazuje: hodnotu kontrastu, tj. hodnotu dané lineární kombinace průměrů (*Value of Contrast*) a její standardní chybu (*Std. Error*), testovou statistiku pro testování hypotézy kontrastu (*t*), stupně volnosti (*df*) a oboustrannou dosaženou hladinu významnosti (*Sig. (2-tailed)*). Na základě significance rozhodujeme o zamítnutí nulové hypotézy. Zde je hodnota significance větší než 5 %, proto na 95% hladině spolehlivosti nezamítáme nulovou hypotézu.



## Tabulka mnohonásobných srovnání

### Multiple Comparisons

Dependent Variable: celkové hodnocení DVD přehrávače

Bonferroni

(I) věková kategorie	(J) věková kategorie	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
18-24	25-31	.840	2.260	1.000	-6.06	7.74
	32-38	-3.877	2.375	1.000	-11.13	3.37
	39-45	-6.077	2.375	.194	-13.33	1.17
	46-52	2.673	2.260	1.000	-4.23	9.57
	53-59	3.378	2.313	1.000	-3.68	10.44
25-31	18-24	-.840	2.260	1.000	-7.74	6.06
	32-38	-4.717	2.417	.834	-12.10	2.66
	39-45	-6.917	2.417	.086	-14.30	.46
	46-52	1.833	2.305	1.000	-5.20	8.87
	53-59	2.538	2.357	1.000	-4.66	9.73
32-38	18-24	3.877	2.375	1.000	-3.37	11.13
	25-31	4.717	2.417	.834	-2.66	12.10
	39-45	-2.200	2.525	1.000	-9.91	5.51
	46-52	6.550	2.417	.131	-.83	13.93
	53-59	7.255	2.467	.069	-.28	14.79
39-45	18-24	6.077	2.375	.194	-1.17	13.33
	25-31	6.917	2.417	.086	-.46	14.30
	32-38	2.200	2.525	1.000	-5.51	9.91
	46-52	8.750*	2.417	.009	1.37	16.13
	53-59	9.455*	2.467	.004	1.92	16.99
46-52	18-24	-2.673	2.260	1.000	-9.57	4.23
	25-31	-1.833	2.305	1.000	-8.87	5.20
	32-38	-6.550	2.417	.131	-13.93	.83
	39-45	-8.750*	2.417	.009	-16.13	-1.37
	53-59	.705	2.357	1.000	-6.49	7.90
53-59	18-24	-3.378	2.313	1.000	-10.44	3.68
	25-31	-2.538	2.357	1.000	-9.73	4.66
	32-38	-7.255	2.467	.069	-14.79	.28
	39-45	-9.455*	2.467	.004	-16.99	-1.92
	46-52	-.705	2.357	1.000	-7.90	6.49

\*. The mean difference is significant at the .05 level.

Dalším typem testů pro detailnější porovnání skupin jsou Post Hoc testy. Procedura nabízí široké spektrum metod, které umožňují především provádět párová srovnání nebo nalézt podmnožiny skupin s blízkými hodnotami průměrů.

V tomto případě chceme zjistit, které dvojice průměrů se od sebe významně odlišují. Opakované užití dvouvýběrového T-testu však není vhodné. Pokud bychom každou z hypotéz testovali samostatně na hladině spolehlivosti 95 %, celková spolehlivost složeného výroku nebude 95 %, ale menší. (Například při testování dvou hypotéz o shodě středních hodnot skupin A a B a skupin C a D na 95% hladině spolehlivosti, bude celková spolehlivost  $0.95 \times 0.95 = 0.9025$ .) Z tohoto důvodu je vhodnější užít některou z metod pro mnohonásobná srovnání, které procedura nabízí. Tímto

## *Listy procedur IBM SPSS Statistics*

způsobem zabráníme tomu, aby bylo příliš mnoho dvojic označeno jako významně odlišných.

Zde jsme užili *Bonferroniho* metodu, která vychází z jednoduchého principu: dělí hladinu významnosti počtem testovaných párů. Tabulka *Multiple Comparisons* obsahuje: rozdíly průměrů (*Mean Difference*) a jejich standardní chyby (*Std. Error*), dosažené hladiny významnosti (*Sig.*) a dolní i horní meze intervalů spolehlivosti pro rozdíl (*95% Confidence Interval*). Dvojice skupin, které se od sebe na 95% hladině spolehlivosti významně odlišují, jsou v prvním sloupci označeny hvězdičkou. V našem případě se tedy jedná o dvě významně odlišné dvojice: věkové kategorie 39-45 a 46-52 a dále 39-45 a 53-59.