

5. Závislost dvou náhodných veličin různých typů (kategoriální a metrická veličina)

Cílem tématu je správné posouzení a výběr vhodného testu v závislosti na povaze metrické a kategoriální veličiny.

V následující tabulce jsou uvedeny testy, kterým se budeme v naší kapitole věnovat. Testy se liší jednak podle toho, zda má metrická veličina normální či jiné rozdělení pravděpodobnosti, a jednak podle toho, zda kategoriální veličina nabývá jen dvou anebo více hodnot.

Přehled testů		Rozdělení metrické veličiny	
		Normální	jiné než normální
Počet hodnot kategoriální veličiny	2	F-test a následný t-test	Mann-Whitney test
	3 a více	ANOVA test	Kruskal-Wallis test

Jednotlivé testy budeme prezentovat na zkoumání vlivu pohlaví, resp. třídy na výšku a hmotnost. Jako výběrový soubor použijeme datový soubor data_deti_min.

5.1 Krabicové grafy

Před vlastním použitím vhodného testu, je vždy vhodné provést vizualizaci závislosti, kterou získáme pomocí tzv. *krabicového grafu* (boxplot).

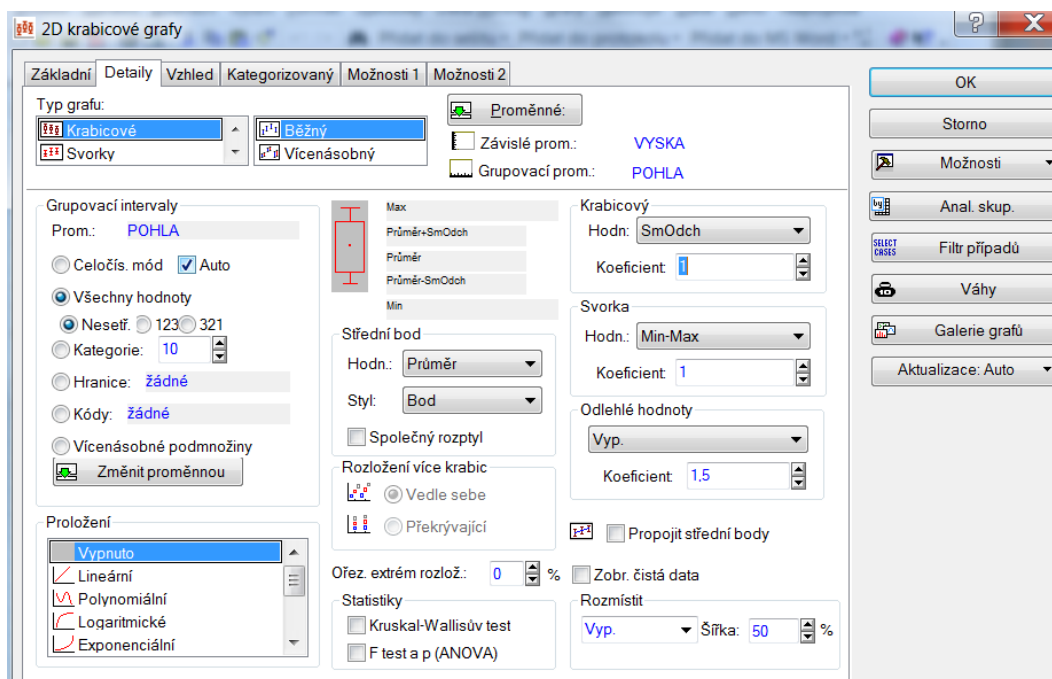
Krabicové grafy nám slouží k lepší představě o možných rozdílech u metrické veličiny v jednotlivých kategoriích (například o rozdílech v tělesné výšce v závislosti na pohlaví). Krabicové grafy naleznete v programu Statistica pod nabídkou Grafy - 2D grafy – Krabicové grafy (viz následující obrázek).

The screenshot shows the Statistica software interface. The 'Grafy' menu is open, and 'Krabicové grafy...' is selected. The background shows a data table with the following columns: TRIDA, VEK, POHLA, ZNAM, VYSKA, HMO TN, and OBLIB. The data rows are numbered 1 to 18.

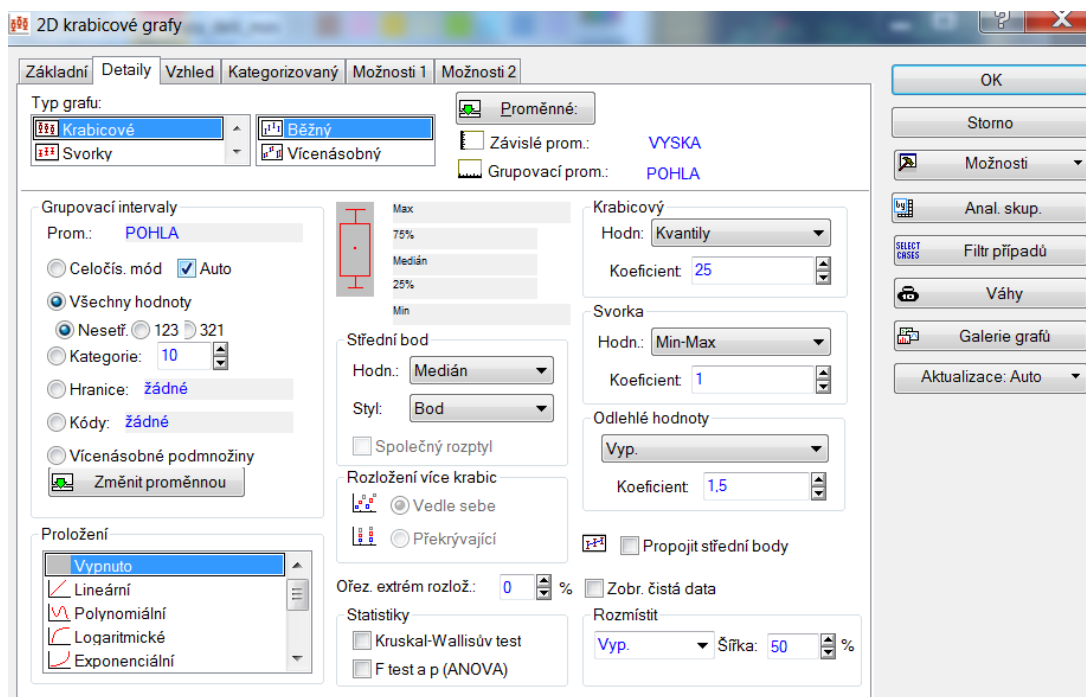
	1	2	3	4	5	6	7
	TRIDA	VEK	POHLA	ZNAM	VYSKA	HMO TN	OBLIB
1	7	13	0	1	160	52	1
2	7	13	0	1	165	57	1
3	8	13	0	1	172	60	1
4	8	13	0	1	169	55	1
5	8	13	0	1	159	51	1
6	7	13	0	1	162	44	1
7	8	13	0	1	156	53	1
8	8	13	0	1	158	49	1
9	7	13	0	1	159	49	1
10	7	13	0	1	174	67	1
11	7	13	0	1	165	51	1
12	8	13	0	2	166	55	1
13	7	13	0	2	160	55	0
14	7	13	0	1	163	50	1
15	8	13	0	1	164	48	1
16	7	13	0	1	156	46	1
17	7	13	0	1	155	47	1
18	7	13	0	1	160	50	1

Po zvolení typu tohoto grafu, se vám nabídne možnost voleb menu (*základní, detaily, vzhled, atd.*), kdy je nutné v menu zvolit *Detaily*, abyste se dostali na stejnou volbu možností jako je na následujících obrázcích. V detailech volíte základní parametry krabicového grafu. Užívají se zejména dva typy krabicových grafů (s průměrem a s mediánem). V následujících dvou obrázcích máte znázorněno, jaké parametry máte volit u grafu pro průměr a jaké pro medián.

Obrázek – Nastavení parametrů u krabicového grafu pro průměr

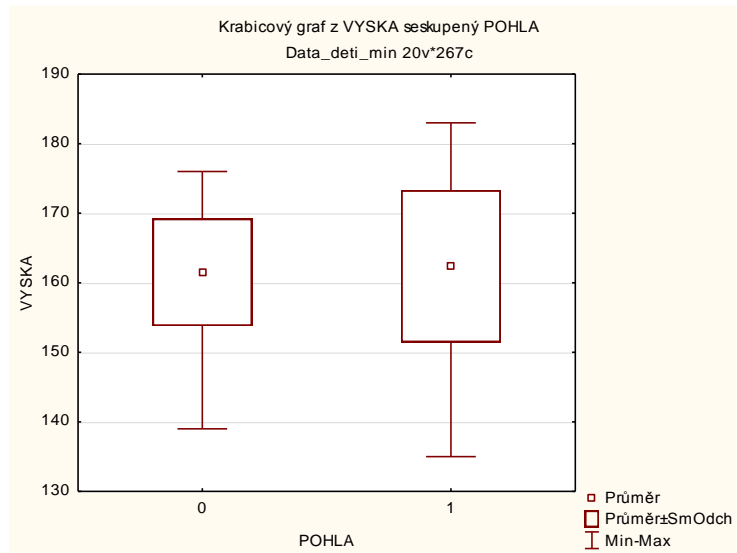


Obrázek – Nastavení parametrů u krabicového grafu pro medián

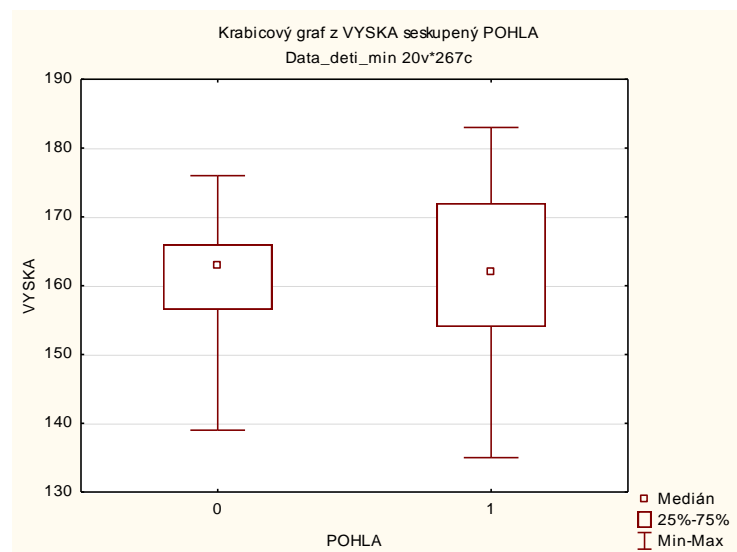


V našem případě jsme zjišťovali vliv pohlaví na výšku žáků. V případě, že jste volili parametry podle obrázků, pak získáváte tyto grafy.

Graf – Vliv pohlaví na výšku – průměrné hodnoty



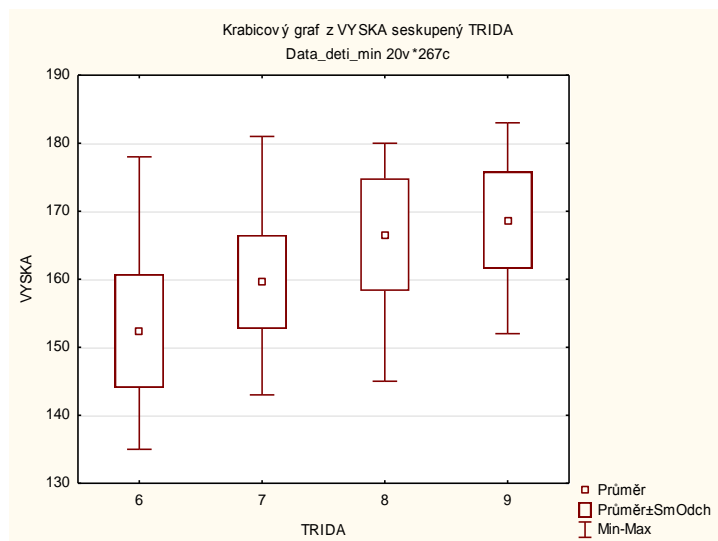
Graf – Vliv pohlaví na výšku – mediánové hodnoty



Prvním náhledem (podle grafů) můžeme získat předběžný závěr, že pohlaví nemá výrazný vliv na výšku (uvědomme si, že se jedná o výběrový soubor dětí z 2. stupně ZŠ, základním souborem tedy nejsou všichni lidé, ale populace jen této věkové skupiny).

Stejný postup je i v případě volby vyššího počtu kategorií než dvou jako u pohlaví. Například v případě třídy, kdy nám třída oproti pohlaví (0, 1) nabývá 4 hodnot (6, 7, 8, 9) tak jak prezentuje následující graf.

Graf – Vliv třídy na výšku u chlapců – průměrné hodnoty



Zde můžeme prvním náhledem (podle grafu) získat předběžný závěr, že třída má vliv na výšku.

5.2 Volba vhodného testu

5.2.1 Zjištění normality dat

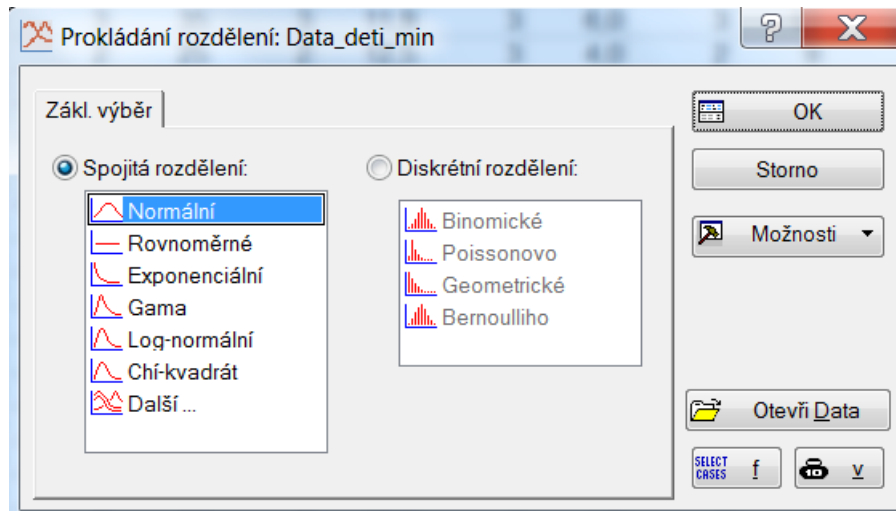
Před každým testem je nutné provést test normality metrické veličiny. Testy normality byly obsahem již tématu 3, avšak přesto se normalitě dat budeme v krátkosti věnovat i v této kapitole. Testem normality sledujeme, zda má metrická veličina normální či jiné pravděpodobnostní rozdělení.

V našem souboru (data_deti_min) jsme provedli test normality u výšky a hmotnosti. Výsledky uvádíme v následujících tabulkách a grafech.

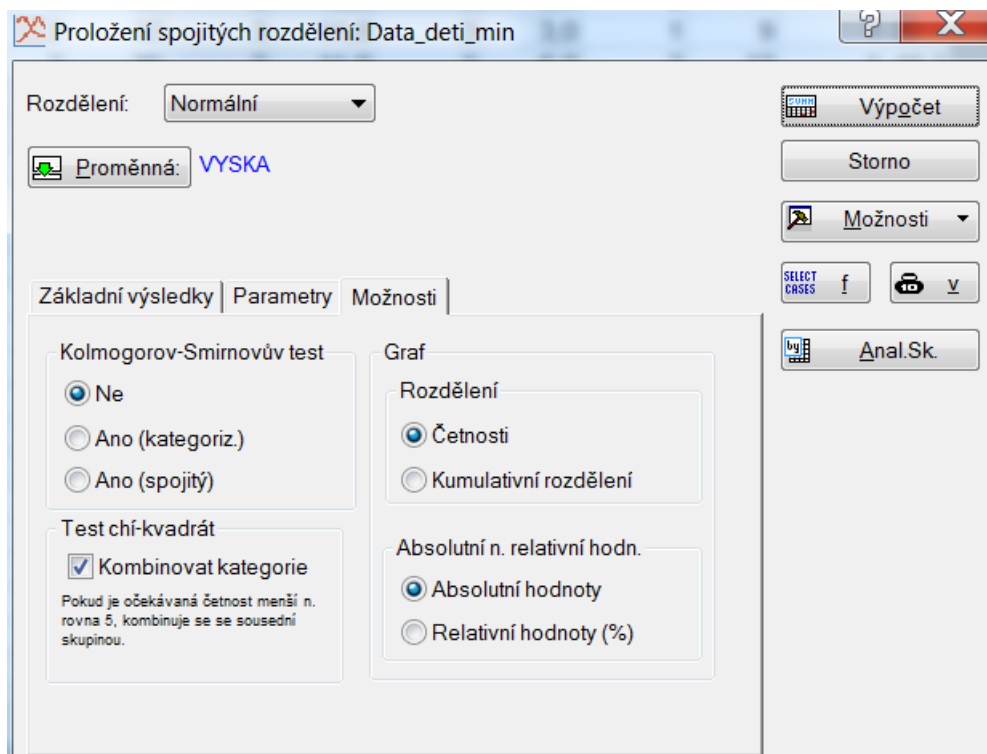
V základním menu zvolíte možnost *Statistiky - prokládání rozdělení*, tak jak je naznačeno na následujícím obrázku.

	1	2	3	4	5	10	11	12	13	14	15	16	17	18					
	TRIDA	VEK	POHLA	ZNAM	VYS	SKOKD	BODYS	LEHSE	BODYL	CLUNB	BODYC	CVYTR	BODYV	SUMA					
1	7	13	0	1		137	1	25	2	12,5	3	4,0	2	8					
2	7	13	0	1		172	3	37	3	10,8	5	6,5	3	14					
3	8	13	0	1		157	2	32	3	11,8	3	3,0	1	9					
4	8	13	0	1		180	3	35	3	11,9	3	6,0	3	12					
5	8	13	0	1		155	2	25	2	12,5	3	4,0	2	9					
6	7	13	0	1		140	1	31	3	12,8	2	3,0	1	7					
7	8	13	0	1		140	1	37	3	12,7	2	2,0	1	7					
8	8	13	0	1		145	2	37	3	11,5	4	5,0	2	11					
9	7	13	0	1		190	4	47	4	10,1	5	1,0	1	14					
10	7	13	0	1		153	2	29	2	12,3	3	4,5	2	9					
11	7	13	0	1		205	5	50	5	10,2	5	7,5	4	19					
12	8	13	0	2		170	3	33	3	12,5	3	6,0	3	12					
13	7	13	0	2	160	55	0	0	0	1	145	2	20	1	13,0	2	2,0	1	6
14	7	13	0	1	163	50	1	1	1	0	145	2	35	3	11,6	4	5,0	2	11
15	8	13	0	1	164	48	1	1	1	1	178	3	27	3	11,6	3	7,0	4	12

Po potvrzení se dostanete na stejnou tabulku jako je na následujícím obrázku, ve které zvolíte normální typ rozdělení a potvrdíte jej OK.



Dostanete následující tabulku:



Zde zvolíte metrickou proměnnou, kterou chcete testovat (v našem případě výška) a zadáte buď možnost výpočet – pozorované a očekávané rozdělení, kdy je výsledkem tabulka (viz následující tabulka), popřípadě zadáte graf pozorovaného a očekávaného rozdělení a získáte histogram (následující graf).

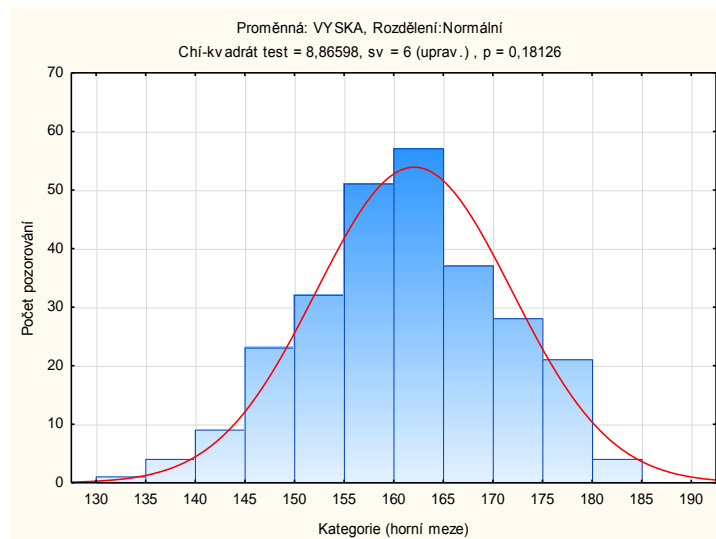
Nulová hypotéza u testu normality zní: **Veličina má v základním souboru normální rozdělení.**

Jak je známo z předchozích témat, pro výsledek testu je podstatná uvedená hodnota p , kterou nalezneme jak v tabulce tak i grafu. V našem případě je hodnota p u výšky rovna 0,18126, což je vyšší než hladina významnosti $\alpha = 0,05$, nulovou hypotézu o normalitě tedy nezamítáme. U hmotnosti je hodnota $p = 0,00001$ a v tomto případě nulovou hypotézu o normalitě zamítáme.

Tabulka - test normality u výšky

Proměnná: VYSKA, Rozdělení: Normální (Data_deti_min) Chi-kvadrát = 8,86598, sv = 6 (uprav.) , p = 0,18126									
Horní hranice	Pozor. Četn.	Kumul. Pozor.	% Pozor.	Kumul. % Pozor.	Očekáv. Četnosti	Kumulativ. Očekáv.	Procent Očekáv.	Kumul. % Očekáv.	Pozorované - Očekáv.
<= 135	1	1	0,37453	0,3745	0,82078	0,8208	0,30741	0,3074	0,17922
140	4	5	1,49813	1,8727	2,58348	3,4043	0,96760	1,2750	1,41652
145	9	14	3,37079	5,2434	7,81511	11,2194	2,92701	4,2020	1,18489
150	23	37	8,61423	13,8577	18,3948	29,6142	6,88946	11,0915	4,60515
155	32	69	11,9850	25,8427	33,6927	63,3070	12,6190	23,7105	-1,69278
160	51	120	19,1011	44,9438	48,0277	111,3348	17,9879	41,6984	2,97221
165	57	177	21,3483	66,2921	53,2825	164,6173	19,9560	61,6544	3,71748
170	37	214	13,8576	80,1498	46,0065	210,6238	17,2309	78,8853	-9,00653
175	28	242	10,48689	90,6367	30,91633	241,5402	11,5791	90,4645	-2,91633
180	21	263	7,86517	98,5019	16,16838	257,7086	6,05557	96,5201	4,83162
185	4	267	1,49813	100,0000	6,57987	264,2884	2,46437	98,9844	-2,57987
< Nekonečno	0	267	0,00000	100,0000	2,71158	267,0000	1,01557	100,0000	-2,71158

Graf – test normality u výšky

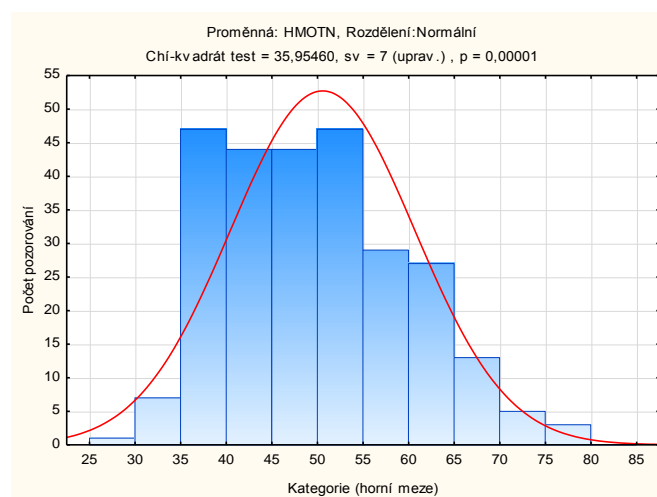


Hodnota $p = 0,18126$, nulovou hypotézu: „Výška má v základním souboru normální rozdělení“ nezamítáme.

Tabulka - test normality u hmotnosti

Horní hranice	Proměnná: HMOTN, Rozdělení: Normální (Data_deti_min) Chí-kvadrát = 35,95460, sv = 7 (uprav.) , p = 0,00001								
	Pozor. Četn.	Kumul. Pozor.	% Pozor.	Kumul. % Pozor.	Očekáv. Četnosti	Kumulativ. Očekáv.	Procent Očekáv.	Kumul. % Očekáv.	Pozorované - Očekáv.
<= 30	1	1	0,3745	0,3745	5,5786	5,5786	2,0893	2,0894	-4,5786
35	7	8	2,6217	2,9963	10,893	16,4719	4,0799	6,1693	-3,8933
40	47	55	17,6030	20,5993	22,997	39,4694	8,6132	14,7825	24,0025
45	44	99	16,4794	37,0787	38,188	77,6580	14,302	29,0854	5,8114
50	44	143	16,4794	53,5581	49,882	127,540	18,682	47,7679	-5,8824
55	47	190	17,6030	71,1610	51,254	178,795	19,196	66,9645	-4,2550
60	29	219	10,8614	82,0225	41,428	220,223	15,516	82,4809	-12,4286
65	27	246	10,1123	92,1348	26,340	246,5648	9,8654	92,3464	0,6592
70	13	259	4,8689	97,0037	13,173	259,7382	4,9338	97,2802	-0,1735
75	5	264	1,8726	98,8764	5,1817	264,9199	1,9407	99,2210	-0,1817
80	3	267	1,1236	100,0000	1,6029	266,5229	0,6003	99,8213	1,3971
< Nekonečno	0	267	0,0000	100,0000	0,4771	267,0000	0,1787	100,000	-0,4771

Graf – test normality u hmotnosti



Hodnota $p=0,00001$, nulovou hypotézu: „Hmotnost má v základním souboru normální rozdělení“ zamítáme.

5.2.2 Počet hodnot kategoriální veličiny

Zjištění počtu hodnot kategoriální veličiny je již velmi jednoduché. U pohlaví nabývá kategoriální veličina pouze dvou hodnot (chlapci, dívky), u třídy čtyř hodnot (6.třída, 7.třída, 8.třída, 9. třída).

5.2.3 Shrnutí

Závěrem shrneme zjištěné skutečnosti. V úvodní tabulce jsme prezentovali typy testů, kterými se budeme v našem téma zabývat. Po ověření normality výšky a hmotnosti a zjištění počtu kategorií u pohlaví a třídy upravíme úvodní tabulku do následující podoby.

Volba testů v našem zkoumání		Rozdělení metrické veličiny	
		Výška - Normální	Hmotnost - jiné než normální
Počet hodnot kategoriální veličiny	Pohlaví - 2	F-test a následný t-test	Mann-Whitney test
	Třída - 4	ANOVA test	Kruskal-Wallis test

Z tabulky je patrné, které testy jsou v našem případě vhodné pro testování.

5.3 F-test a následný t-test

Metrická veličina má normální rozdělení a kategoriální veličina nabývá jen dvou hodnot.

Hodnoty metrické veličiny rozdělíme podle hodnot kategoriální veličiny do dvou výběrů. Předpokládáme, že jeden výběr pochází z rozdělení $No(\mu_1; \sigma_1^2)$ a druhý z rozdělení $No(\mu_2; \sigma_2^2)$.

Rozhodnutí, zda hodnota kategoriální veličiny ovlivňuje hodnoty metrické veličiny, učiníme podle testu nulové hypotézy $H_0 : \mu_1 = \mu_2$ (střední hodnoty obou rozdělení se sobě rovnají). Pro tento test máme ale k dispozici dva *t-testy*, a to:

Dvouvýběrový t-test s rovností rozptylů

a *Dvouvýběrový t-test s nerovností rozptylů*.

O tom, který z těchto testů použijeme, se rozhodneme na základě tzv. *F-testu pro rozptyly*, při kterém testujeme nulovou hypotézu $\sigma_1^2 = \sigma_2^2$ (rozptyly obou rozdělení se sobě rovnají). Tento test je tedy jen pomocný a uděláme ho předem. Když tuto hypotézu u F-testu nezamítneme, použijeme první t-test, když ji zamítneme, použijeme druhý t-test.

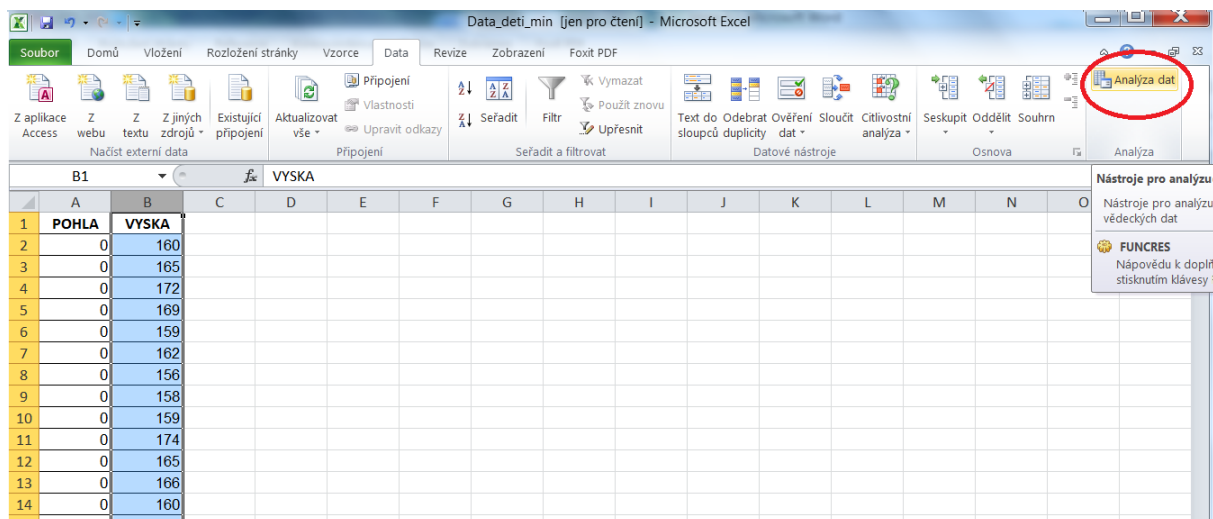
Nulové hypotézy testujeme na hladině významnosti α (obvykle volíme 0,05).

Výsledek testu pak vyhodnotíme podle P-hodnoty obvyklým způsobem:

- Je-li p-hodnota menší než α , pak nulovou hypotézu zamítáme,
- Je-li p-hodnota větší než α , pak nulovou hypotézu nezamítáme.

5.3.1 Provedení F-testu a následného t-testu v Excelu

V Excelu jsou všechny tři výše uvedené testy k dispozici jako nástroje v doplňku Analýza dat. Počítač nám vypočítá hodnoty statistik a uvede je ve formě tabulky. Nejprve je vhodné si data setřídít podle kategoriální proměnné, tedy například dát všechny chlapce za sebou a následně všechny dívky. Zde je vhodné netřídít celý soubor, ale pouze ty veličiny, které potřebujete, tak jako v následujícím obrázku.



Na stejném obrázku pak máte pod menu data nabídku Analyza dat, kdy vám po jejím rozkliknutí vyběhne následující tabulka, s možností volby.

Před vlastním testováním je nejprve vhodné užít nástroj popisná statistika, kterým získáte mimo jiné odhady střední hodnoty a rozptylu u obou souborů (chlapců a dívek). Důvodem je zjištění, který výběrový průměr a výběrový rozptyl je větší.

	<i>Dívky</i>	<i>Chlapci</i>
Stř. hodnota	161,5104	162,386
Chyba stř. hodnoty	0,786676	0,83521
Medián	163	162
Modus	165	160
Směr. odchylka	7,707815	10,92179
Rozptyl výběru	59,41042	119,2854
Špičatost	0,414662	-0,73985
Šikmost	-0,60052	-0,09194

Následně provedeme dvouvýběrový F-test pro rozptyl.

The screenshot shows the Microsoft Excel interface with the 'Data' ribbon active. The spreadsheet contains the following data:

	A	B	C	D	E	F	G	H	I	J
1	POHLA	VYSKA								
2		0								
3		0								
4		0								
5		0								
6		0								
7		0								
8		0								
9		0								
10		0								
11		0								
12		0								
13		0								
14		0								
15		0								

The 'Analyza dat' dialog box is open, showing the following options:

- Anova: jeden faktor
- Anova: dva faktory s opakováním
- Anova: dva faktory bez opakování
- Korelace
- Kovariance
- Popisná statistika
- Exponenciální vyrovnání
- Dvouběrový F-test pro rozptyl** (selected)
- Fourierova analýza
- Histogram

Po potvrzení typu tohoto testu zvolíte do následující tabulky hodnoty výšky u chlapců a dívek (za soubor 1 volíme ten soubor, který má vyšší rozptyl, tedy v našem případě chlapce a za soubor 2 volíme dívky). Současně si navolíte výstupní oblast, kam se vám po potvrzení zhotoví výsledná tabulka.

The screenshot shows the Microsoft Excel interface with the 'Data' ribbon active. The spreadsheet now includes a third column, 'VYSKA', and the 'Dvouběrový F-test pro rozptyl' dialog box is open. The dialog box is configured as follows:

- Vstup:
 - 1. soubor: \$B\$98:\$B\$268
 - 2. soubor: \$B\$2:\$B\$97
- Popisky
- Alfa: 0,05
- Možnosti výstupu:
 - Výstupní oblast: \$K\$4
 - Nový list:
 - Nový sešit

Dvouvýběrový F-test pro rozptyl		
	<i>Soubor 1</i>	<i>Soubor 2</i>
Stř. hodnota	162,386	161,5104
Rozptyl	119,2854	59,41042
Pozorování	171	96
Rozdíl	170	95
F	2,00782	
P(F<=f) (1)	0,000121	
F krit (1)	1,358729	

Z výsledné tabulky je pro nás podstatná hodnota p , podle které v našem případě zamítáme nulovou hypotézu o rovnosti rozptylů a dle toho volíme typ t-testu – t-test s nerovností rozptylů. Pokud by hodnota p byla vyšší než 0,05, volili bychom typ t-testu s rovností rozptylů.

Následně provádíme t-test. Za soubor 1 volíme ten soubor, který má vyšší střední hodnotu – výběrový průměr, tedy v našem případě chlapce a za soubor 2 volíme dívky.

Dvouvýběrový t-test s nerovností rozptylů		
	<i>Soubor 1</i>	<i>Soubor 2</i>
Stř. hodnota	162,386	161,5104
Rozptyl	119,2854	59,41042
Pozorování	171	96
Hyp. rozdíl stř. hodnot	0	
Rozdíl	251	
t Stat	0,763098	
P(T<=t) (1)	0,223061	
t krit (1)	1,650947	
P(T<=t) (2)	0,446121	
t krit (2)	1,96946	

Mimo testového kritéria jsou v tabulce uvedeny dvě hodnoty p a dvě kritické hodnoty. Hodnoty p (1) a t krit (1) se váží k jednostranné alternativní hypotéze, která je ve tvaru: „Jedna střední hodnota je větší než druhá střední hodnota“.

Hodnoty p (2) a t krit (2) se váží k dvoustranné alternativní hypotéze, která je ve tvaru: „Střední hodnoty jsou od sebe různé“.

Obvykle se volí jednostranná alternativa, tedy varianta (1).

Z výsledné tabulky je pro nás podstatná hodnota $p=0,223061$, podle které v našem případě nezamítáme nulovou hypotézu o rovnosti středních hodnot v základním souboru. Věcně to znamená, že v tomto věkovém rozmezí nemá pohlaví vliv na výšku žáků.

5.3.2 Provedení F-testu a následného t-testu ve Statistice

Ve Statistice získáme všechny tři testy najednou: v Základních statistikách volíme T-test pro nezávislé výběry dle skupin.

Data: Data_deti_min (20s krát 267ř)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	TRIDA	VEK	POHLA	ZNAM	VYSKA	HMOTN	OBLIB	DOVED	SOUTE	SKOKD	BODYS	LEHSE	BODYL	CLUNB	BODYC
1	7	13	0	1	160	52	1	1	0	137	1	25	2	12,5	3
2	7	13	0	1	165	57	1	1	0	179	2	37	3	10,8	5
3	8	13	0	1	172	61	1	1	0	137	1	25	2	11,8	3
4	8	13	0	1	169	59	1	1	0	179	2	37	3	11,9	3
5	8	13	0	1	159	54	1	1	0	137	1	25	2	12,5	3
6	7	13	0	1	162	49	1	1	0	137	1	25	2	12,8	2
7	8	13	0	1	156	50	1	1	0	137	1	25	2	12,7	2
8	8	13	0	1	158	49	1	1	0	137	1	25	2	11,5	4
9	7	13	0	1	159	49	1	1	0	137	1	25	2	10,1	5
10	7	13	0	1	174	61	1	1	0	137	1	25	2	12,3	3
11	7	13	0	1	165	51	1	1	0	137	1	25	2	10,2	5
12	8	13	0	2	166	54	1	1	0	137	1	25	2	12,5	3
13	7	13	0	2	160	50	1	1	0	137	1	25	2	13,0	2
14	7	13	0	1	163	50	1	1	0	137	1	25	2	11,6	4
15	8	13	0	1	164	48	1	1	0	137	1	25	2	11,9	3
16	7	13	0	1	156	49	1	1	0	137	1	25	2	11,9	3
17	7	13	0	1	155	47	1	1	0	137	1	25	2	11,8	3
18	7	13	0	1	160	50	1	1	0	137	1	25	2	12,0	3
19	8	14	0	1	165	52	1	1	0	137	1	25	2	15,9	1
20	8	14	0	2	160	49	1	1	0	137	1	25	2	11,4	4
21	9	14	0	1	156	48	1	1	0	137	1	25	2	12,8	2
22	9	14	0	2	164	56	0	0	1	160	2	26	2	12,8	2
23	9	14	0	2	170	62	1	1	1	179	3	38	3	11,9	3
24	8	14	0	1	175	49	1	1	0	168	3	46	4	11,6	4

Základní statistiky a tabulky: Data_deti_min

Základní výsledky

- Popisné statistiky
- Korelační matice
- t-test, nezávislé, dle skupin**
- t-test, nezávislé, dle proměn.
- t-test, závislé vzorky
- t-test, samost vzorek
- Rozklad & jednofakt. ANOVA
- Rozklad
- Tabulky četností
- Kontingenční tabulky
- Tabulky vícenásob. odpovědí
- Testy rozdílů: r, %, průměry
- Pravděpodobnostní kalkulátor

Možnosti

Otevři data

V nabídce Možnosti volíme Test se samostatnými odhady rozptylu.

Data: Data_deti_min (20s krát 267ř)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	TRIDA	VEK	POHLA	ZNAM	VYSKA	HMOTN	OBLIB	DOVED	SOUTE	SKOKD	BODYS	LEHSE	BODYL	CLUNB	BODYC
1	7	13	0	1	160	52	1	1	0	137	1	25	2	12,5	3
2	7	13	0	1	165	57	1	1	0	179	2	37	3	10,8	5
3	8	13	0	1	172	61	1	1	0	137	1	25	2	11,8	3
4	8	13	0	1	169	59	1	1	0	179	2	37	3	11,9	3
5	8	13	0	1	159	54	1	1	0	137	1	25	2	12,5	3
6	7	13	0	1	162	49	1	1	0	137	1	25	2	12,8	2
7	8	13	0	1	156	50	1	1	0	137	1	25	2	12,7	2
8	8	13	0	1	158	49	1	1	0	137	1	25	2	11,5	4
9	7	13	0	1	159	49	1	1	0	137	1	25	2	10,1	5
10	7	13	0	1	174	61	1	1	0	137	1	25	2	12,3	3
11	7	13	0	1	165	51	1	1	0	137	1	25	2	10,2	5
12	8	13	0	2	166	54	1	1	0	137	1	25	2	12,5	3
13	7	13	0	2	160	50	1	1	0	137	1	25	2	13,0	2
14	7	13	0	1	163	50	1	1	0	137	1	25	2	11,6	4
15	8	13	0	1	164	48	1	1	0	137	1	25	2	11,9	3
16	7	13	0	1	156	49	1	1	0	137	1	25	2	11,9	3
17	7	13	0	1	155	47	1	1	0	137	1	25	2	11,8	3
18	7	13	0	1	160	50	1	1	0	137	1	25	2	12,0	3
19	8	14	0	1	165	52	1	1	0	137	1	25	2	15,9	1
20	8	14	0	2	160	49	1	1	0	137	1	25	2	11,4	4
21	9	14	0	1	156	48	1	1	0	137	1	25	2	12,8	2
22	9	14	0	2	164	56	0	0	1	160	2	26	2	12,8	2
23	9	14	0	2	170	62	1	1	1	179	3	38	3	11,9	3

t-test pro nezávislé vzorky dle skupin: Data_deti_min

Proměnné: Závislé: VYSKA
Grupovací: POHLA

Kód pro skup. 1: 0 Kód pro skup. 2: 1

Základní výsledky | Detailní výsledky | Možnosti

Zobrazit dlouhá jména proměnných

Test se samost odhady rozptylu

Vícerozměrný test (Hotellingovo T²)

p-hodnota pro zvýraznění: .05

Meze spol. pro odhady 95,00 %

Homogenita rozptylů

Leveneův test

Brown-Forsytheův test

Výpočet

Možnosti

Anal skup...

Vážené momenty

SV =

W-1 N-1

ChD vynechána

Celé případy

Párově

Získáme tabulku, na jejímž konci je výsledek F testu, a podle něj vyhodnotíme příslušný t-test. Test s rovností rozptylů má stupeň volnosti (sv) celočíselný (součet rozsahů obou výběrů zmenšený o dvě), test s nerovností rozptylů má stupeň volnosti vyjádřen desetinným číslem. Tabulka se objeví v jednom řádku, pro přehlednost zde uvádíme tři části této tabulky odděleně.

Tato část tabulky se týká F-testu, podle hodnoty p zamítáme nulovou hypotézu o rovnosti rozptylů. Uvedená hodnota p odpovídá dvoustranné alternativě, tj. „rozptyly jsou od sebe různé“. Jednostranné alternativě by odpovídala poloviční hodnota p, tedy $p = 0,000121$ (porovnejte hodnoty s výše uvedeným zpracováním v excelu).

Poč.plat 0	Poč.plat. 1	Sm.odch. 0	Sm.odch. 1	F-poměr Rozptyly	p Rozptyly
96	171	7,707815	10,92179	2,007820	0,000242

Tato část tabulky se týká t-testu s rovností rozptylů (pro nás je tedy neúčinná).

	Průměr 0	Průměr 1	t	sv	p
VYSKA	161,5104	162,3860	-0,694133	265	0,488207

Tato část tabulky se týká t-testu s nerovností rozptylů, vzhledem k hodnotě p, nezamítáme nulovou hypotézu o rovnosti středních hodnot. Hodnota p je opět počítaná pro dvoustrannou alternativu (porovnejte s výsledky v excelu).

	Průměr 0	Průměr 1	t samost odh.rozp	sv	P oboustr.
VYSKA	161,5104	162,3860	-0,763098	251,3832	0,446120

5.4 ANOVA test

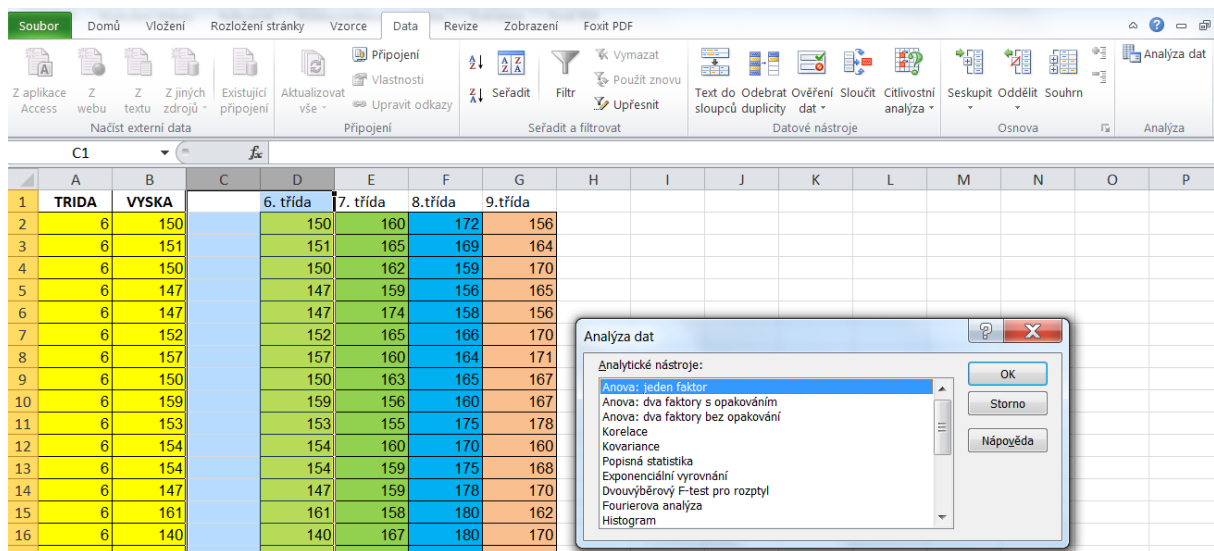
Metrická veličina má normální rozdělení a kategoriální veličina nabývá více než dvou hodnot.

Hodnoty metrické veličiny rozdělíme podle hodnot kategoriální veličiny do tří nebo více výběrů, o nichž předpokládáme, že všechny mají normální rozdělení.

Rozhodnutí, zda hodnota kategoriální veličiny ovlivňuje hodnoty metrické veličiny, učiníme podle testu nulové hypotézy H_0 : střední hodnoty všech rozdělení se sobě rovnají.

5.4.1 Provedení ANOVA-testu v Excelu

V Excelu pro tento test nejprve přeneseme sledované veličiny do nového listu a následně uspořádáme podle hodnot kategoriální veličiny. Jednotlivé podsoubory pak přeneseme do sloupců vedle sebe (sloupce nemusí mít stejnou délku) viz následující obrázek.



Poté aplikuje test ANOVA (jednofaktorová), který je jako nástroj v doplňku Analýza dat.

Po rozkliknutí zvoleného testu je nutné do vstupní oblasti zadat data ve formě obdélníka, který obsahuje data ze všech čtyř sloupců (může logicky obsahovat u kratších sloupců i prázdná pole).

Získáme následující tabulku:

ANOVA: jeden faktor						
<i>Výběr</i>	<i>Počet</i>	<i>Součet</i>	<i>Průměr</i>	<i>Rozptyl</i>		
Sloupec 1	61	9296	152,3934	69,60929		
Sloupec 2	69	11013	159,6087	47,50639		
Sloupec 3	69	11493	166,5652	68,51407		
Sloupec 4	68	11471	168,6912	50,7241		
ANOVA						
<i>Zdroj variability</i>	<i>SS</i>	<i>Rozdíl</i>	<i>MS</i>	<i>F</i>	<i>Hodnota P</i>	<i>F krit</i>
Mezi výběry	10505,18	3	3501,728	59,55296	2,06E-29	2,638925
Všechny výběry	15464,46	263	58,80024			
Celkem	25969,65	266				

Test je podle P-hodnoty vyhodnocován stejně jako u předchozích testů.

V našem případě je p menší než 0,05, a tedy zamítáme nulovou hypotézu o rovnosti všech čtyř středních hodnot.

Logicky se naskytá otázka, které ze čtyř středních hodnot se liší. Tento problém řeší tzv. post-hoc testy, které si ukážeme v programu Statistica.

5.4.2 Provedení ANOVA-testu ve Statistice

V programu Statistica použijeme Statistiku - Základní statistiky a Rozklad a jednofaktorovou Anovu. Zvolíme sledované veličiny a potvrdíme OK.

The screenshot shows a data table with columns labeled TRIDA, VEK, POHLA, ZNAM, VYSKA, HMOTN, OBLIB, DOVED, SOUTE, SKOKD, BODYD, LEHSE, BODYL, CLUNB, BODYC, CVYTR, BODYV, and SUMA. A dialog box titled "Statistiky dle skupin (rozklad): Data_deti_min" is open, showing settings for the ANOVA test. The dependent variable is "VYSKA" and the grouping variable is "TRIDA". The dialog box includes options for "Proměnné", "Závislé", "Grupovací", "Kódy pro grupovací proměnné", "Vážené momenty", "ChD vynechána", and "Celé případy".

V Základních výsledcích volíme Analýzu rozptylu.

The screenshot shows the same data table as above, but with a dialog box titled "Statistiky dle skupin - výsledek: Data_deti_min" open. The dialog box displays the results of the ANOVA test, including the dependent variable "VYSKA" and the grouping variable "TRIDA". The dialog box includes options for "Základní výsledky", "Popisné statistiky", "ANOVA & testy", "Post-hoc", "Výpočet", "Graf interakcí", "Detailní 2-rozm. tabulky", "Kategoriz. krabicový graf", "Výpočet", "Storno", "Možnosti", and "Anal skup...". The "Analýza rozptylu" option is highlighted with a red circle.

Výsledkem je tato tabulka, podle hodnoty p zamítáme nulovou hypotézu o rovnosti všech středních hodnot.

Proměnná	Analýza rozptylu (Data_deti_min) Označ. efekty jsou význ. na hlad. $p < ,05000$							
	SČ efekt	SV efekt	PČ efekt	SČ chyba	SV chyba	PČ chyba	F	p
VYSKA	10505,18	3	3501,728	15464,46	263	58,80024	59,55296	0,00000

Jestliže nulovou hypotézu zamítneme, tedy konstatujeme, že hodnota kategoriální veličiny ovlivňuje hodnoty metrické veličiny, pak provádíme aposteriorní F-testy a t-testy na jednotlivé dvojice podsouborů (Post-hoc analýza a volba Scheffeova testu).

Data: Data_deti_min (20s krát 267)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
i	TRIDA	VEK	POHLA	ZNAM	VYSKA	HMOTN	OBLIB	DOVED	SOUTE	SKOKD	BODYS	LEHSE	BODYL	CLUNB	BODYC	CVYTR	BODYV	SUMA
1	7	13	0	1	160	52	1	1	0	137	1	25	2	12,5	3	4,0	2	8
2	7	13	0	1	165	57	4	1	0	170	2	27	2	10,9	5	6,5	2	14
3	8	13	0	1	172	60	1	1	0	170	2	27	2	10,9	5	6,5	2	14
4	8	13	0	1	169	55	1	1	0	170	2	27	2	10,9	5	6,5	2	14
5	8	13	0	1	159	51	1	1	0	170	2	27	2	10,9	5	6,5	2	14
6	7	13	0	1	162	44	1	1	0	170	2	27	2	10,9	5	6,5	2	14
7	8	13	0	1	156	53	1	1	0	170	2	27	2	10,9	5	6,5	2	14
8	8	13	0	1	158	49	1	1	0	170	2	27	2	10,9	5	6,5	2	14
9	7	13	0	1	159	49	1	1	0	170	2	27	2	10,9	5	6,5	2	14
10	7	13	0	1	174	67	1	1	0	170	2	27	2	10,9	5	6,5	2	14
11	7	13	0	1	165	51	1	1	0	170	2	27	2	10,9	5	6,5	2	14
12	8	13	0	2	166	55	1	1	0	170	2	27	2	10,9	5	6,5	2	14
13	7	13	0	2	160	55	1	1	0	170	2	27	2	10,9	5	6,5	2	14
14	7	13	0	1	163	50	1	1	0	170	2	27	2	10,9	5	6,5	2	14
15	8	13	0	1	164	48	1	1	0	170	2	27	2	10,9	5	6,5	2	14
16	7	13	0	1	156	46	1	1	0	170	2	27	2	10,9	5	6,5	2	14
17	7	13	0	1	155	47	1	1	0	170	2	27	2	10,9	5	6,5	2	14
18	7	13	0	1	160	50	1	1	0	170	2	27	2	10,9	5	6,5	2	14
19	8	14	0	1	165	52	1	1	0	170	2	27	2	10,9	5	6,5	2	14
20	8	14	0	2	160	42	1	1	0	170	2	27	2	10,9	5	6,5	2	14
21	9	14	0	1	156	45	1	1	0	170	2	27	2	10,9	5	6,5	2	14
22	9	14	0	2	164	56	1	1	0	170	2	27	2	10,9	5	6,5	2	14
23	9	14	0	2	170	62	1	1	0	170	2	27	2	10,9	5	6,5	2	14
24	8	14	0	1	175	49	1	1	0	170	2	27	2	10,9	5	6,5	2	14
25	9	14	0	1	165	61	1	1	0	170	2	27	2	10,9	5	6,5	2	14

Statistiky dle skupin - výsledek: Data_deti_min

ZÁVISLÉ : 1 proměnná: VYSKA

GRUPOV. : 1-TRIDA (4): 6 7 8 9

Základní výsledky | Popisné statistiky | ANOVA & testy | **Post-hoc**

Proměnné: VYSKA

LSD test n. plánovaná porovnání

Schefféův test

Newman-Keulsův test & kritická rozpětí

Duncanův test & kritická rozpětí

Tukeyův HSD

Tukey HSD - nestejná N (Spjotvoll/Stoline)

Úroveň alfa pro kritická rozpětí: .050

p-hodnota pro zvýraznění: .05

Z této výsledné tabulky je patrné, že statisticky významné rozdíly jsou mezi všemi skupinami, vyjma skupiny 3 a 4. Věcně to znamená, že nemůžeme zamítnout hypotézu, že v základní populaci není prokazatelný rozdíl mezi výškami žáků v 8. a 9. třídě.

Scheffeho test; proměn.:VYSKA (Data_deti_min) Označ. rozdíly jsou významné na hlad. $p < ,05000$				
TRID A	{1} M=152,39	{2} M=159,61	{3} M=166,57	{4} M=168,69
6 {1}		0,000005	0,000000	0,000000
7 {2}	0,000005		0,000006	0,000000
8 {3}	0,000000	0,000006		0,453232
9 {4}	0,000000	0,000000	0,453232	

5.5 Mann-Whitney test

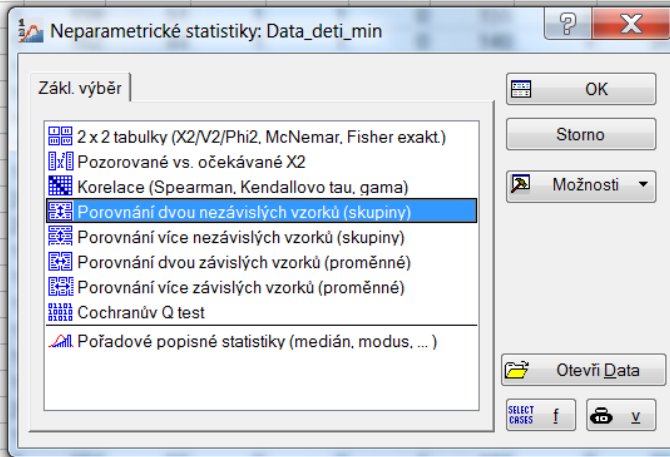
Metrická veličina nemá normální rozdělení a kategoriální veličina nabývá jen dvou hodnot. V našem případě jde o veličiny Hmotnost a Pohlaví.

Nulová hypotéza je: oba podsoubory podle hodnot kategoriální veličiny mají stejné distribuční funkce (speciálně tedy mají stejné mediány, stejné kvantily, obecně stejné všechny kvantily).

Test nalezneme ve Statistice v nabídce Neparametrické statistiky, volíme Porovnání dvou nezávislých vzorků (skupiny).

Data: Data_deti_min (20s krát 267f)

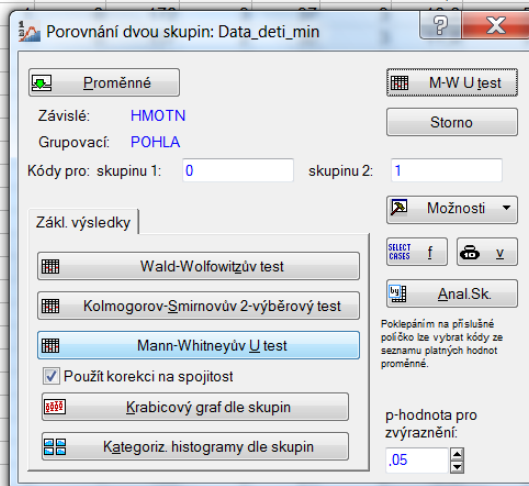
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	TRIDA	VEK	POHLA	ZNAM	VYSKA	HMOTN	OBLIB	DOVED	SOUTE	SKOKD	BODYS	LEHSE	BODYL	CLUNB	BODYC
1	7	13	0	1	160	52	1	1	0	137	1	25	2	12,5	3
2	7	13	0	1	165	57	1	1	0	172	3	37	3	10,8	5
3	8	13	0	1	172	60	1	1	0	157	2	32	3	11,8	3
4	8	13	0	1	169	55	1	1	1	180	3	35	3	11,9	3
5	8	13	0										2	12,5	3
6	7	13	0										3	12,8	2
7	8	13	0										3	12,7	2
8	8	13	0										3	11,5	4
9	7	13	0										4	10,1	5
10	7	13	0										2	12,3	3
11	7	13	0										5	10,2	5
12	8	13	0										3	12,5	3
13	7	13	0										1	13,0	2
14	7	13	0										3	11,6	4
15	8	13	0										3	11,9	3
16	7	13	0										2	11,9	3
17	7	13	0										3	11,8	3
18	7	13	0										2	12,0	3
19	8	14	0										2	15,9	1
20	8	14	0										4	11,4	4
21	9	14	0										3	12,8	2
22	9	14	0	2	164	56	0	0	1	180	2	20	2	12,8	2
23	9	14	0	2	170	62	1	1	1	179	3	38	3	11,9	3
24	8	14	0	1	175	49	1	1	0	168	3	46	4	11,6	4
25	9	14	0	1	165	61	0	1	0	140	1	29	2	13,5	1



Následně zvolíme sledované proměnné a ze Základních výsledků volíme možnost Mann-Whitneyův U test.

Data: Data_deti_min (20s krát 267f)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	TRIDA	VEK	POHLA	ZNAM	VYSKA	HMOTN	OBLIB	DOVED	SOUTE	SKOKD	BODYS	LEHSE	BODYL	CLUNB	BODYC	CVYTR	BODYV
1	7	13	0	1	160	52	1	1	0	137	1	25	2	12,5	3	4,0	2
2	7	13	0	1	165	57	1	1	0	172	3	37	3	10,8	5	6,5	3
3	8	13	0	1	172	60	1	1	0	157	2	32	3	11,8	3	3,0	1
4	8	13	0	1	169	55	1	1	1	180	3	35	3	11,9	3	6,0	3
5	8	13	0	1	159	51	1	1	0	137	1	25	2	12,5	3	4,0	2
6	7	13	0	1	162	44	1	1	0	137	1	25	2	12,5	3	3,0	1
7	8	13	0	1	156	53	1	1	0	137	1	25	2	12,5	3	6,0	3
8	8	13	0	1	158	49	1	1	0	137	1	25	2	12,5	3	2,0	1
9	7	13	0	1	159	49	1	1	0	137	1	25	2	12,5	3	5,0	2
10	7	13	0	1	174	67	1	1	0	137	1	25	2	12,5	3	1,0	1
11	7	13	0	1	165	51	1	1	0	137	1	25	2	12,5	3	4,5	2
12	8	13	0	2	166	55	1	1	0	137	1	25	2	12,5	3	7,5	4
13	7	13	0	2	160	55	0	1	0	137	1	25	2	12,5	3	6,0	3
14	7	13	0	1	163	50	1	1	0	137	1	25	2	12,5	3	2,0	1
15	8	13	0	1	164	48	1	1	0	137	1	25	2	12,5	3	5,0	2
16	7	13	0	1	156	46	1	1	0	137	1	25	2	12,5	3	7,0	4
17	7	13	0	1	155	47	1	1	0	137	1	25	2	12,5	3	5,5	3
18	7	13	0	1	160	50	1	1	0	137	1	25	2	12,5	3	5,0	2
19	8	14	0	1	165	52	1	1	0	137	1	25	2	12,5	3	4,5	2
20	8	14	0	2	160	42	1	1	0	137	1	25	2	12,5	3	1,5	1
21	9	14	0	1	156	45	1	1	0	137	1	25	2	12,5	3	3,5	1
22	9	14	0	2	164	56	0	1	0	137	1	25	2	12,5	3	6,5	3
23	9	14	0	2	170	62	1	1	1	179	3	38	3	11,9	3	3,0	1
24	8	14	0	1	175	49	1	1	0	168	3	46	4	11,6	4	8,5	4
25	9	14	0	1	165	61	0	1	0	140	1	29	2	13,5	1	2,0	1



Získáváme následující tabulku. Test se vyhodnocuje podle p-hodnoty analogicky jako v předchozích případech.

Mann-Whitneyův U Test (w/ oprava na spojitost) (Data_deti_min) Dle proměn. POHLA Označené testy jsou významné na hladině $p < ,05000$									
Proměnná	Sčt poř. skup. 1	Sčt poř. skup. 2	U	Z	p-hodn.	Z upravené	p-hodn.	N platn. skup. 1	N platn. skup. 2
HMOTN	12656,50	23121,50	8000,500	-0,341869	0,732450	-0,342130	0,732253	96	171

Vzhledem k hodnotě p , která je vyšší než 0,05, nemůžeme zamítnout nulovou hypotézu o rovnosti distribučních funkcí hmotnosti u obou pohlaví.

5.6 Kruskal-Wallis test

Metrická veličina nemá normální rozdělení a kategoriální veličina nabývá více než dvou hodnot. V našem případě jde o veličiny Hmotnost a Třída.

Nulová hypotéza je: všechny podsoubory podle hodnot kategoriální veličiny mají stejné distribuční funkce (speciálně tedy mají stejné mediány, stejné kvartily, obecně stejné všechny kvantily).

Test nalezneme ve Statistice v nabídce Neparametrické statistiky, volíme Porovnání více nezávislých vzorků (skupiny).

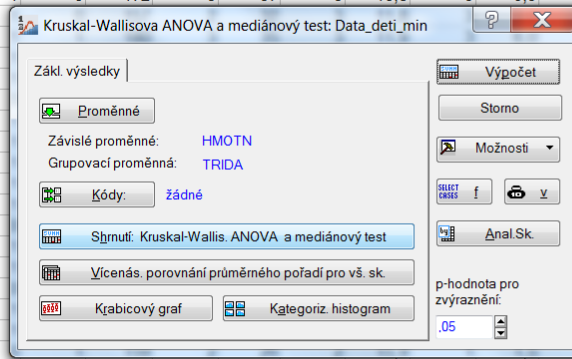
The screenshot shows a spreadsheet with columns labeled TRIDA, VEK, POHLA, ZNAM, VYSKA, HMOTN, OBLIB, DOVED, SOUTE, SKOKD, BODYD, LEHSE, BODYL, CLUNB, and BODYC. A dialog box for non-parametric statistics is open, with the following options listed:

- 2 x 2 tabulky (X²/V₂/Phi², McNemar, Fisher exakt)
- Pozorované vs. očekávané X²
- Korelace (Spearman, Kendallovo tau, gama)
- Porovnání dvou nezávislých vzorků (skupiny)
- Porovnání více nezávislých vzorků (skupiny)**
- Porovnání dvou závislých vzorků (proměnné)
- Porovnání více závislých vzorků (proměnné)
- Cochranův Q test
- Pořadové popisné statistiky (medián, modus, ...)

Následně zvolíme sledované proměnné a ze Základních výsledků volíme možnost Kruskal-Wallis ANOVA a mediánový test.

Data: Data_deti_min (20s krát 267f)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	TRIDA	VEK	POHLA	ZNAM	VYSKA	HMOTN	OBLIB	DOVED	SOUTE	SKOKD	BODYS	LEHSE	BODYL	CLUNB	BODYC	CVYTR	BODYV	SUMA	VYKON
1	7	13	0	1	160	52	1	1	0	137	1	25	2	12,5	3	4,0	2	8	0
2	7	13	0	1	165	57	1	1	0	172	3	37	3	10,8	5	6,5	3	14	2
3	8	13	0	1	172	60	1	1	0	157	2	32	3	11,8	3	3,0	1	9	1
4	8	13	0	1	169	55	1	1	1	180	3	35	3	11,9	3	6,0	3	12	1
5	8	13	0	1	159	51	1	1	0	155	2	25	2	12,5	3	4,0	2	9	1
6	7	13	0	1	162	44	1	1	0	140	1	31	3	12,8	2	3,0	1	7	0
7	8	13	0	1	156	53	1	0	0	140	1	37	3	12,7	2	2,0	1	6	0
8	8	13	0	1	158	49	1	1	0	145	2	37	3	11,5	4	5,0	2	11	1
9	7	13	0	1	159	49	1	1	1	190	4	47	4	10,1	5	1,0	1	7	0
10	7	13	0	1	174	67	1	1	0	159	2	33	3	10,9	3	4,5	2	9	1
11	7	13	0	1	165	51	1	1	0	173	3	32	3	12,8	2	6,5	3	11	1
12	8	13	0	2	166	55	1	0	0	160	2	26	2	12,8	2	3,0	1	7	0
13	7	13	0	2	160	55	0	0	1	179	3	38	3	11,9	3	6,0	3	12	1
14	7	13	0	1	163	50	1	1	1	168	3	46	4	11,6	4	8,5	4	15	2
15	8	14	0	1	165	52	1	1	0	140	1	29	2	13,5	1	2,0	1	5	0
16	8	14	0	2	160	42	1	1	0	170	3	47	4	11,4	4	3,5	1	12	1
17	9	14	0	1	156	45	1	0	0	173	3	32	3	12,8	2	6,5	3	11	1
18	9	14	0	2	164	56	0	0	1	160	2	26	2	12,8	2	3,0	1	7	0
19	9	14	0	2	170	62	1	1	1	179	3	38	3	11,9	3	6,0	3	12	1
20	8	14	0	1	175	49	1	1	0	168	3	46	4	11,6	4	8,5	4	15	2
21	9	14	0	1	165	61	0	1	0	140	1	29	2	13,5	1	2,0	1	5	0



Podstatný je pro nás Kruskal-Wallisův test, který je nutno zvolit v levém menu, viz následující obrázek.

Data: Data_deti_min (20s krát 267f)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	TRIDA	VEK	POHLA	ZNAM	VYSKA	HMOTN	OBLIB	DOVED	SOUTE	SKOKD	BODYS	LEHSE	BODYL	CLUNB	BODYC	CVYTR	BODYV
1	7	13	0	1	160	52	1	1	0	137	1	25	2	12,5	3	4,0	2
2	7	13	0	1	165	57	1	1	0	172	3	37	3	10,8	5	6,5	3
3	8	13	0	1	172	60	1	1	0	157	2	32	3	11,8	3	3,0	1
4	8	13	0	1	169	55	1	1	1	180	3	35	3	11,9	3	6,0	3
5	8	13	0	1	159	51	1	1	0	155	2	25	2	12,5	3	4,0	2
6	7	13	0	1	162	44	1	1	0	140	1	31	3	12,8	2	3,0	1
7	8	13	0	1	156	53	1	0	0	140	1	37	3	12,7	2	2,0	1
8	8	13	0	1	158	49	1	1	0	145	2	37	3	11,5	4	5,0	2
9	7	13	0	1	159	49	1	1	1	190	4	47	4	10,1	5	1,0	1
10	7	13	0	1	174	67	1	1	0	159	2	33	3	10,9	3	4,5	2
11	7	13	0	1	165	51	1	1	0	173	3	32	3	12,8	2	6,5	3
12	8	13	0	2	166	55	1	0	0	160	2	26	2	12,8	2	3,0	1
13	7	13	0	2	160	55	0	0	1	179	3	38	3	11,9	3	6,0	3
14	7	13	0	1	163	50	1	1	1	168	3	46	4	11,6	4	8,5	4
15	8	14	0	1	165	52	1	1	0	140	1	29	2	13,5	1	2,0	1
16	8	14	0	2	160	42	1	1	0	170	3	47	4	11,4	4	3,5	1
17	9	14	0	1	156	45	1	0	0	173	3	32	3	12,8	2	6,5	3
18	9	14	0	2	164	56	0	0	1	160	2	26	2	12,8	2	3,0	1
19	9	14	0	2	170	62	1	1	1	179	3	38	3	11,9	3	6,0	3
20	8	14	0	1	175	49	1	1	0	168	3	46	4	11,6	4	8,5	4
21	9	14	0	1	165	61	0	1	0	140	1	29	2	13,5	1	2,0	1
22	9	14	0	1	156	47	1	0	0	150	2	35	3	12,5	2	4,0	2
23	9	14	0	1	170	66	1	0	0	162	2	38	3	13,6	1	5,0	2
24	8	14	0	1	171	59	0	1	1	159	2	30	2	12,6	2	3,0	1

PS6* - Kruskal-Wallisova ANOVA založ. na poř.; HMOTN (Data_deti_min)

Kruskal-Wallisova ANOVA založ. na poř.; HMOTN (Data_deti_min)
 Nezávislá (grupovací) proměnná : TRIDA
 Kruskal-Wallisův test: $H(3, N=267) = 99,50189$ $p = 0,000$

Závislá: HMOTN	Kód	Počet platných	Součet pořadí	Prům. Pořadí
6	6	61	3641,00	59,6885
7	7	69	8181,50	118,5725
8	8	69	11467,00	166,1884
9	9	68	12488,50	183,6544

Získáváme následující tabulku. Test se vyhodnocuje podle p-hodnoty analogicky jako v předchozích případech.

Závislá: HMOTN	Kruskal-Wallisova ANOVA založ. na poř.; HMOTN (Data_deti_min) Nezávislá (grupovací) proměnná : TRIDA Kruskal-Wallisův test: $H(3, N=267) = 99,50189$ $p = 0,000$			
	Kód	Počet platných	Součet pořadí	Prům. Pořadí
6	6	61	3641,00	59,6885
7	7	69	8181,50	118,5725
8	8	69	11467,00	166,1884
9	9	68	12488,50	183,6544

Test se vyhodnocuje podle p-hodnoty analogicky jako v předchozích případech. Vzhledem k hodnotě p, která je nižší než 0,05, zamítáme nulovou hypotézu o rovnosti distribučních funkcí hmotnosti u všech tříd.

Post-hoc analýza se provede pomocí volby Vícenásobné porovnání průměrného pořadí pro všechny skupiny.

Data: Data_deti_min (20s krát 267f)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	TRIDA	VEK	POHLA	ZNAM	VYSKA	HMOTN	OBLIB	DOVED	SOUTE	SKOKD	BODYS	LEHSE	BODYL	CLUNB	BODYC	CVYTR	BODYV	SUMA
1	7	13	0	1	160	52	1	1	0	137	1	25	2	12,5	3	4,0	2	8
2	7	13	0	1	165	57	1	1	0	172	3	37	3	10,8	5	6,5	3	14
3	8	13	0	1	172	60	1										1	9
4	8	13	0	1	169	55	1										3	12
5	8	13	0	1	159	51	1										2	9
6	7	13	0	1	162	44	1										1	7
7	8	13	0	1	156	53	1										1	7
8	8	13	0	1	158	49	1										2	11
9	7	13	0	1	159	49	1										1	14
10	7	13	0	1	174	67	1										2	9
11	7	13	0	1	165	51	1										4	19
12	8	13	0	2	166	55	1										3	12
13	7	13	0	2	160	55	0										1	6
14	7	13	0	1	163	50	1										2	11
15	8	13	0	1	164	48	1										4	13
16	7	13	0	1	156	46	1										3	10
17	7	13	0	1	155	47	1										2	10
18	7	13	0	1	160	50	1										2	9
19	8	14	0	1	165	52	1										1	6
20	8	14	0	2	160	42	1	1	0	170	3	47	4	11,4	4	3,5	1	12
21	9	14	0	1	156	45	1	0	0	173	3	32	3	12,8	2	6,5	3	11
22	9	14	0	2	164	56	0	0	1	160	2	26	2	12,8	2	3,0	1	7
23	9	14	0	2	170	62	1	1	1	179	3	38	3	11,9	3	6,0	3	12
24	8	14	0	1	175	49	1	1	0	168	3	46	4	11,6	4	8,5	4	15
25	9	14	0	1	165	61	0	1	0	140	1	29	2	13,5	1	2,0	1	5

Kruskal-Wallisova ANOVA a mediánový test: Data_deti_min

Zákl. výsledky

Proměnné

Závislé proměnné: **HMOTN**

Grupovací proměnná: **TRIDA**

Kódy: 4

Shrnutí: Kruskal-Wallis ANOVA a mediánový test

Vícenás. porovnání průměrného pořadí pro vš. sk

Křabicový graf

Kategoriz. histogram

Výpočet

Storno

Možnosti

Anal.Sk.

p-hodnota pro zvýraznění: .05

Závislá: HMOTN	Vícenásobné porovnání p hodnot (oboustr.); HMOTN (Data_deti_min) Nezávislá (grupovací) proměnná : TRIDA Kruskal-Wallisův test: $H(3, N=267) = 99,50189$ $p = 0,000$			
	6 R:59,689	7 R:118,57	8 R:166,19	9 R:183,65
6		0,000086	0,000000	0,000000
7	0,000086		0,001755	0,000005
8	0,000000	0,001755		1,000000
9	0,000000	0,000005	1,000000	

Z této výsledné tabulky je patrné, že statisticky významné rozdíly jsou mezi všemi třídami, vyjma tříd 8. a 9. Věcně to znamená, že nemůžeme zamítnout hypotézu, že v základní populaci není prokazatelný rozdíl mezi hmotnostmi žáků v 8. a 9. třídě.