Testování hypotéz

Testování hypotéz jsou klasické statistické úsudky založené na nějakém apriorním předpokladu. Vyslovíme-li předpoklad o hodnotě neznámého parametru nebo o zákonu rozdělení sledované náhodné veličiny, vyslovíme tak **statistickou hypotézu**. Ověřování, zda hypotéza platí či nikoliv, je předmětem testování, které provádíme na základě nějakého výběru (měření, pozorování).

Test statistické hypotézy H proti alternativní hypotéze A je pravidlo, podle něhož na základě náhodného výběru rozhodneme mezi dvěma tvrzeními - sledovanou hypotézou H a alternativní hypotézou A. Výsledkem našeho rozhodování je buď zamítnutí hypotézy H ve prospěch alternativy A či její nezamítnutí. Skutečnost, že hypotézu nezamítáme, neznamená že naměřená data tuto hypotézu potvrzují, ale pouze to, že ji nevyvracejí.

Ve většině programů je testovaná hypotéza označovaná jako **nulová hypotéza H_0 a alternativní hypotéza H_1 nebo H_A .**

Při testování hypotéz se můžeme dopustit chyby dvěma způsoby: Buď zamítneme hypotézu, která platí - to je **chyba prvního druhu α** - nebo naopak tuto hypotézu nezamítneme i když je nesprávná - v tomto případě se jedná o **chybu druhého druhu β**. Při konstrukci testu požadujeme, aby pravděpodobnost chyby 1. druhu byla menší nebo rovna danému číslu **α,** kterému říkáme **hladina významnosti testu.** Obvykle volíme $\alpha = 0,05$ nebo 0,1.

Testování probíhá tak, že vypočítáme hodnotu testové statistiky, porovnáme ji s kritickými hodnotami, odpovídajícími hladině významnosti α, a rozhodneme o zamítnutí či nezamítnutí hypotézy H.

Při **testování pomocí statistických programů** se používá jiný postup: Spočte se hodnota testové statistiky a k ní nejmenší kritický obor, při kterém bychom ještě mohli na základě této hodnoty zamítnout hypotézu H_0 proti dané alternativě. Hladina významnosti, odpovídající tomuto kritickému oboru, se nazývá **minimální hladina významnosti (p-hodnota)**.

Pokud je $p > \alpha$, pak hypotézu H_0 nezamítáme.V opačném případu, kdy $p \leq \alpha$, pak hypotézu H_0 zamítáme.

Příklad 1

O hodinových výdělcích taxikářů dvou měst máme k dispozici údaje ze dvou náhodných výběrů, které jsou uvedeny v tabulce 1.

Proveďte základní statistické vyhodnocení a ověřte na 5% hladině významnosti tvrzení, že dlouhodobý průměr výdělků taxikářů je v obou městech stejný.

Tab.1 Hodinové výdělky taxikářů

1. město:

68	133	144	106	154	175	141	148	75	50	130	151
199	134	183	137	127	101	157	119	112	115	88	168
142	103	135	115	195	133	105	82	78	143	85	85
179	124	113	97	80	84	135	99	116	133	118	200
145	165	123	155	131	98	148	44	125	82	110	111

2. město:

148	127	174	132	125	139	158	140	108	146	125	154	167
132	128	127	111	134	111	118	105	150	109	112	114	133
115	81	132	112	148	162	124	159	198	134	134	157	108
158	73	137	154	138	168	151	136	117	104	141	171	148
145	151	140	113	147	146	105	141	128	167	152	131	

Řešení pomocí Statistica 7.0

a) *Základní informace o souboru*

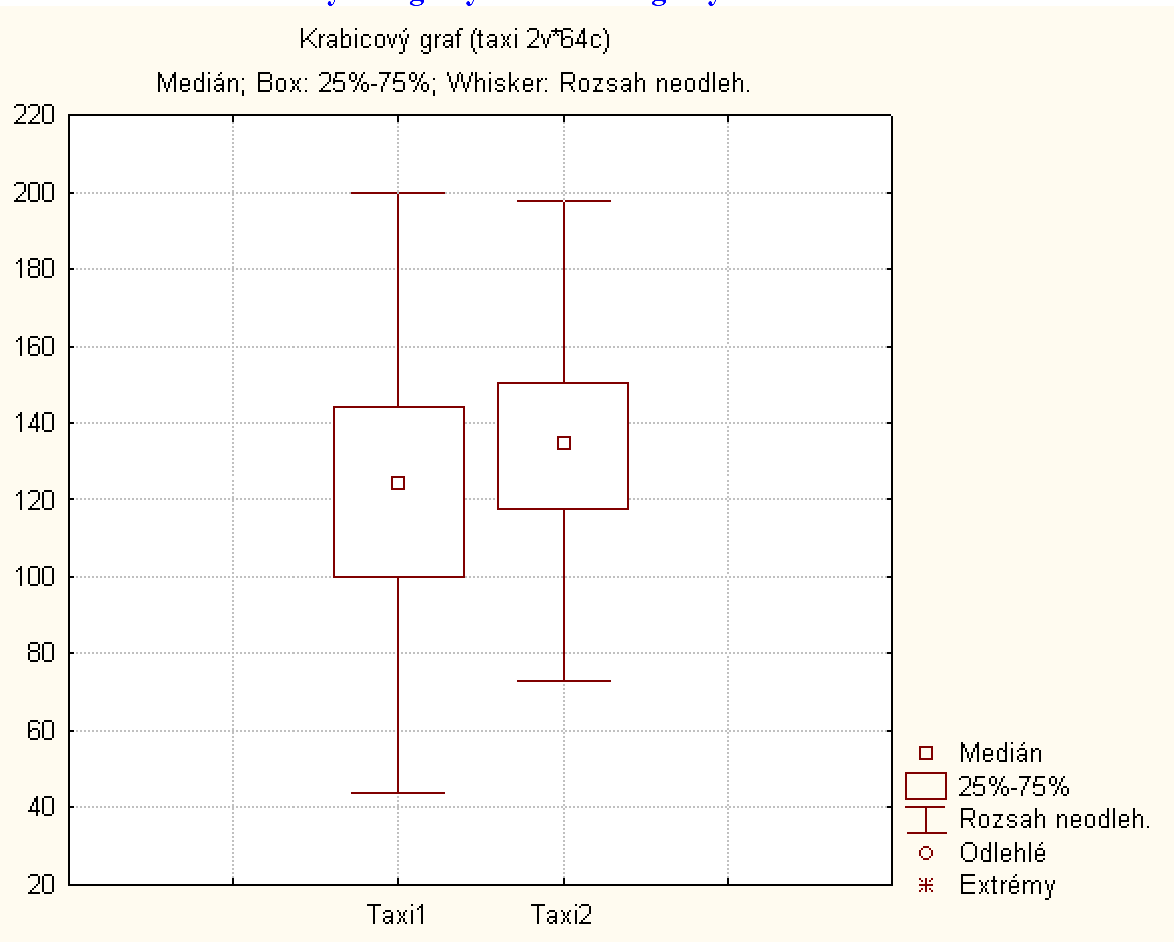
Statistika: Základní statistika a tabulky: Popisné statistiky: Detaily

	N platných	Průměr	Medián	Modus	Min	Max	Rozptyl	Sm. odch.
Taxi1	60	123,85	124,50	133,00	44	200	1223,62	34,98
Taxi2	64	135,20	135,00	Vícenás.	73	198	508,895	22,55

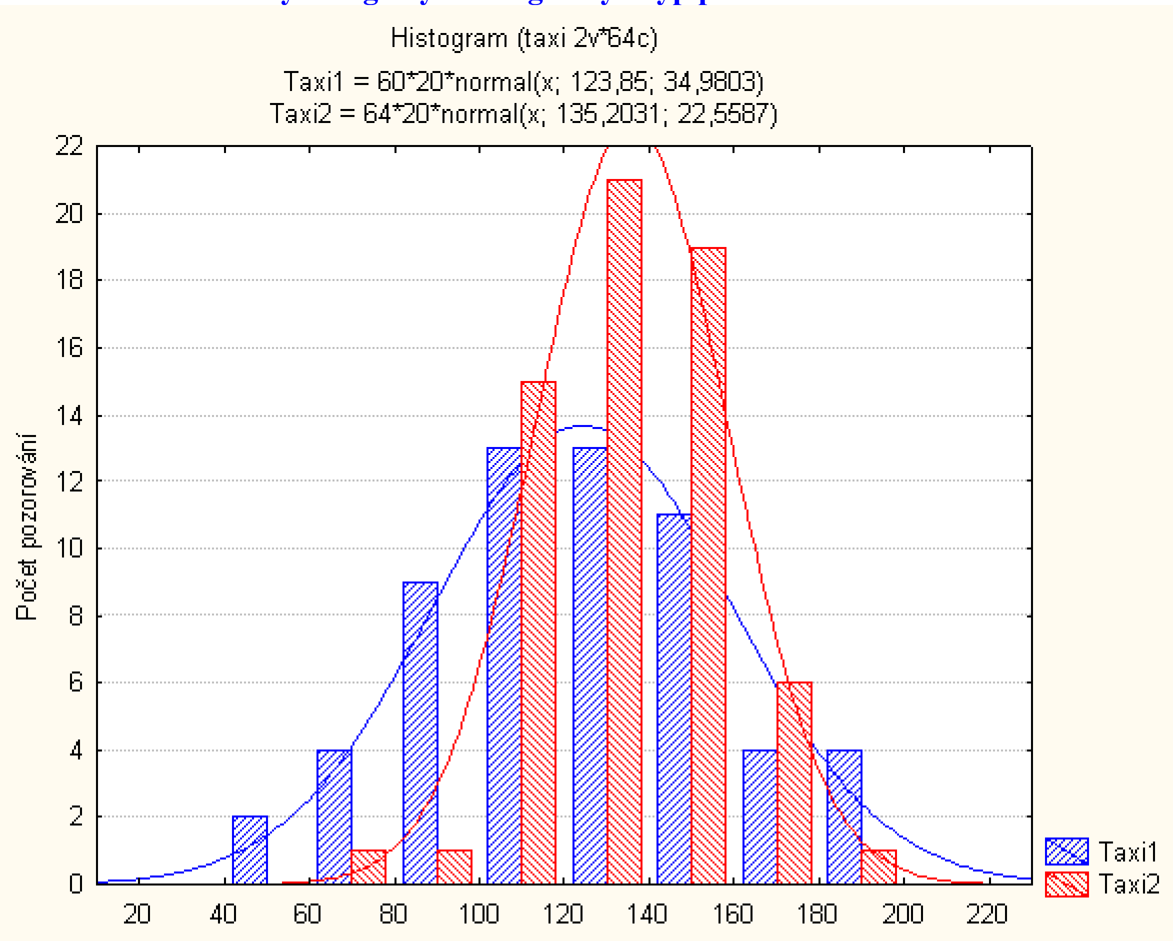
b) *Grafické ověření normality dat*

Protože se při odvozování intervalů spolehlivosti i testů hypotéz předpokládá normální rozdělení výběrového souboru, měli bychom vždy tento předpoklad ověřit. Nesplnění předpokladu normality dat, vede k přibližným a někdy i chybným řešením.

Grafy: 2D grafy: Krabicové grafy: Vícenásobné



Grafy: 2D grafy: Histogramy: Typ proložení normální



c) *Testy normality*

**Statistika: Základní statistika a tabulky: Popisné statistiky:
Normalita: K-S, Shapiro-Wilksův W test**

Tabulka četností:Taxi1 (taxi) K-S d=,04997, p> .20; Lilliefors p> .20 Shapiro-WilksW=,98863, p<,85

Tabulka četností:Taxi2 (taxi) K-S d=,05292, p> .20; Lilliefors p> .20 Shapiro-WilksW=,98626, p<,69

Na základě údajů v tabulce c) nezamítáme hypotézu o normalitě obou výběrů.

d1) *Test rovnosti rozptylů a Test rovnosti středních hodnot*

Statistika: Základní statistika a tabulky: t-test, nezávislé, dle proměnné

T-test pro nezávislé vzorky (taxi) Pozn.: Proměnné byly brány jako nezávislé vzorky

	Průměr	Průměr	Hodnota t	sv	p	t separ.	sv	p
Taxi1 vs. Taxi2	123,8500	135,2031	-2,16124	122	0,032630	-2,13244	99,77278	0,035425

	Poč.plat.	Poč.plat.	Sm.odch.	Sm.odch.	F-poměr	p
Taxi1 vs. Taxi2	60	64	34,98030	22,55869	2,404469	0,000735

V horní části tabulky najdeme opět hodnoty základních výběrových statistik: rozsah výběru, průměr a směrodatnou odchylku. Následuje hodnota testového kritéria = t-statistika = -2,16124, počet stupňů volnosti = 122 a minimální oboustranná pravděpodobnost p = 0,03263 při které ještě zamítáme hypotézu o rovnosti středních hodnot.

d2) V případě, kdy data nepocházejí z normálního rozdělení, použijeme neparametrické testy

!!! POZOR !!! Jiná struktura dat (taxi2.sta)

Statistika: Neparametrické statistiky: porovnání dvou nezávislých vzorků skupiny

Mann-Whitneyův U test (taxi2) Dle proměn. Prom1
Označené testy jsou významné na hladině p <,05000

	Sčt poř.	Sčt poř.	U	Z	Úroveň p
Prom2	3315,000	4435,000	1485,000	-2,17500	0,029631

	Z	Úroveň p	N platn.	N platn.	2*1str.
Prom2	-2,17533	0,029606	60	64	0,029470

V tabulky najdeme opět hodnoty základních výběrových statistik: tentokráte součet pořadí. Následuje hodnota testového kritéria = U = 1485, a minimální oboustranná pravděpodobnost p = 0,029631 při které ještě zamítáme hypotézu o rovnosti středních hodnot.

Závěr příkladu

Protože vypočítaná dvoustranná pravděpodobnost p = 0.0354 < 0,05 = α, zamítáme hypotézu $H_0: \mu_1 - \mu_2 = 0$ o rovnosti středních hodnot a s pravděpodobností 0,95 tvrdíme, že dlouhodobé průměrné výdělky taxikářů z uvedených měst jsou různé.

TESTOVÁNÍ NEZÁVISLOSTI NOMINÁLNÍCH VELIČIN

Test nezávislosti χ^2 (chí-kvadrát), měření síly závislosti

Cramérův koeficient. Tento koeficient nabývá hodnot mezi 0 a 1. Čím blíže je 1, tím je těsnější závislost mezi X a Y, čím blíže je 0, tím je tato závislost volnější

V roce 1950 zkoumali Yule a Kendall barvu očí a vlasů u 6800 mužů.

Barva očí	Barva vlasů			
	světlá	kaštanová	černá	rezavá
modrá	1768	807	180	47
šedá nebo zelená	946	1387	746	53
hnědá	115	438	288	16

Na asymptotické hladině významnosti 0,05 testujte hypotézu o nezávislosti barvy očí a barvy vlasů. Vypočtěte Cramérův koeficient. Simultánní četnosti znázorněte graficky.

Návod: Vytvořte nový datový soubor o 12 případech a třech proměnných (OCI, VLASY, CETNOST). Do proměnné OCI napište varianty barvy očí $x_{[1]} = 1$ (modrá), $x_{[2]} = 2$ (šedá nebo zelená), $x_{[3]} = 3$ (hnědá), přičemž každá varianta se objeví čtyřikrát pod sebou. Do proměnné VLASY napište třikrát pod sebe všechny varianty $y_{[1]} = 1$ (světlá), $y_{[2]} = 2$ (kaštanová), $y_{[3]} = 3$ (černá), $y_{[4]} = 4$ (rezavá).

	oci	**vlasy**	**cet-nost**
1	modrá	světlá	1768
2	modrá	kaštano-vá	807
3	modrá	černá	180
4	modrá	rezavá	47
5	šedozele-ná	světlá	946
6	šedozele-ná	kaštano-vá	1387
7	šedozele-ná	černá	746
8	šedozele-ná	rezavá	53
9	hnědá	světlá	115
10	hnědá	kaštano-vá	438
11	hnědá	černá	288
12	hnědá	rezavá	20

Statistika: Základní statistika/Tabulky: Kontingenční tabulky: Specifikace tabulky (List1-oci, List2-vlasy): OK: v (váhy) četnost: OK: Možnosti: zaškrtněte Pearson & M-L Chi -square, Phi & Cramer's V: **Detailní výsledky: Detailní 2-rozm. tabulky**

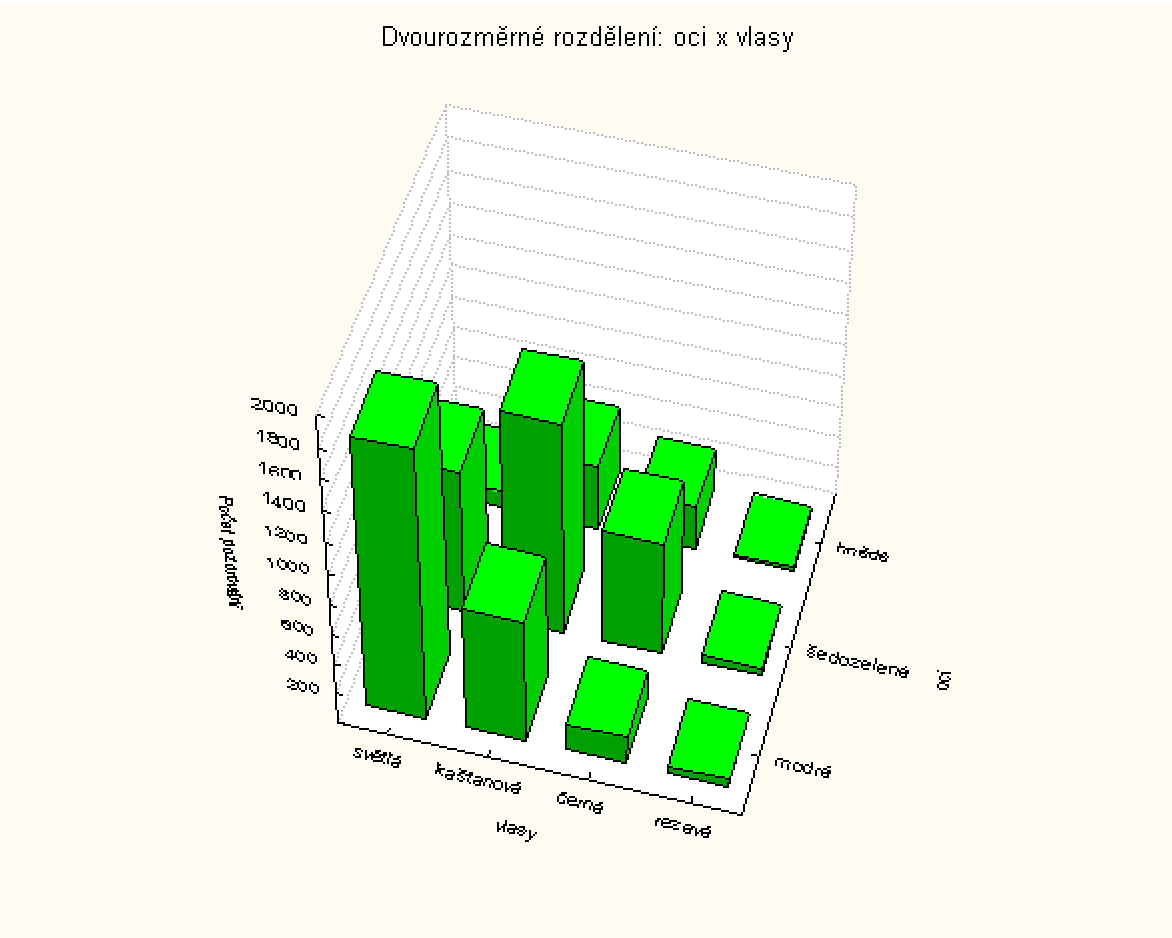
Statist. : oci(3) x vlasy(4) (oci)

	Chí-kvadr.	**sv**	**p**
Pearsonův chí-kv.	1088,934	df=6	p=0,0000

M-V chí-kvadr.	1157,153	df= 6	p=0,000 0
Fí	,4003188		
Kontingenční koefici-ent	,3716458		
Cramér. V	,2830681		

Ve výstupní tabulce najdeme mj. hodnotu testové statistiky (Chi-square = 1088,149) s počtem stupňů volnosti (df = 6) a odpovídající **p-hodnotou** (p = 0,0000) i **Cramérův koeficient (V = 0,283)**. Pro grafické znázornění četností se vraťte do Detailní výsledky –3D histogram. Po vytvoření grafu je nutné manuálně zvětšit rozsah zobrazovaných hodnot na osách x a y.

Pomocí STATISTIKY je možno lehce ověřit splnění podmínek dobré aproximace (tzn., že teoretické četností mají být aspoň v 80% případů větší než 5 a ve zbylých 20% případů nemají klesnout pod 2). Teoretické četnosti se vypočítají tak, že v Options zaškrtneme Oekávané četnosti. V našem případě jsou podmínky dobré aproximace splněny.



Závěr: Existuje závislost mezi barvou očí a barvou vlasů

KORELAČNÍ KOEFICIENT

Pracovník personálního oddělení jistého podniku zkoumá, zda existuje vztah mezi počtem dní absence v práci za rok a věkem pracovníka. Náhodně vybere záznamy o 10 pracovnících a získá údaje o jejich věku (znak X) a počtu dní absence za minulý rok (znak Y).

X	27	61	37	23	46	58	29	36	64	40
Y	15	6	10	18	9	7	14	11	5	8

a) Vypočtěte koeficient korelace r a interpretujte ho.

Řešení:

Základní statistika a tabulky: Korelační matice: 1 seznam proměnné: Možnosti: Zobrazit r, p, N: Souhrn

a) r = -0,9325, p=0,00
Označení korelace jsou významné. Červeně psané hodnoty p jsou statisticky významné.

Závěr: Mezi věkem pracovníka a počtem dnů absence za minulý rok existuje silná nepřímá lineární závislost – čím je pracovník starší, tím méně chybí v práci.