Testování hypotéz

1 Jednovýběrové testy

90/2 – odhad času

V podmínkách naprostého odloučení má voják prokázat schopnost orientace v čase. Úkolem vojáka je provést odhad časového intervalu 1 hodiny bez hodinek a odeslat signál. Vyhodnocení signálů dává u 1 vojáka během 20 hodin tyto výsledky: 1,16 1,26 1,77 1,15 1,19 0,93 0,87 1,26 1,27 1,31 1,11 0,73 1,25 1,37 1,45 1,08 0,98 0,83 1,17 1,54. Předpokládejte, že odhady časového intervalu mají normální rozdělení.

a) Zjistěte, zda voják v daných podmínkách odhaduje správně hodinový interval. Test proveďte na hladině významnosti $\alpha = 0{,}05$.
b) Sestrojte oboustranný interval spolehlivosti pro střední hodnotu odhadů s rizikem $\alpha = 0{,}05$ a komentujte srovnání s výsledkem testu.

Řešení:

a) $n = 20$, $\bar{x} \doteq 1{,}184$, $s \doteq 0{,}248$, $\alpha = 0{,}05$

Hypotéza a alternativa

$$\text{H: } \mu = 1 \quad \rightarrow \quad \text{A: } \mu \neq 1$$

Testové kritérium

$$t = \frac{\bar{x} - \mu_0}{s}\sqrt{n} = \frac{1{,}184 - 1}{0{,}248}\sqrt{20} \doteq 3{,}317$$

Kritický obor

$$W_{0,05}: \qquad |\, t \,| \geq t_{0,975}(19)$$
$$|\, 3{,}317 \,| \geq 2{,}093$$

tzn., že hodnota testového kritéria patří do kritického oboru, nulovou hypotézu H na hladině významnosti $\alpha = 0{,}05$ zamítáme, platí alternativní hypotéza. S 95% spolehlivostí lze tvrdit, že odhad hodinového intervalu není správný.

b) Oboustranný interval spolehlivosti pro střední hodnotu

$$\bar{x} - t_{0,975}(19)\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{0,975}(19)\frac{s}{\sqrt{n}}$$
$$1{,}184 - 2{,}093\frac{0{,}248}{\sqrt{20}} < \mu < 1{,}184 + 2{,}093\frac{0{,}248}{\sqrt{20}}$$
$$1{,}068 < \mu < 1{,}300$$

Hodnota 1 nepatří do 95% intervalu spolehlivosti pro střední hodnotu. Na základě toho můžeme říct, že odhad hodinového intervalu není s pravděpodobností 95 % správný. Výsledky získané na základě testu a intervalu spolehlivosti jsou stejné. (Odpovídající intervaly spolehlivosti je možné používat při testování hypotéz.)

Provedený test a intervalový odhad lze snadno provést v našem excelovském pracovním sešitu STAT1: Otevřeme si list **1V – normální** (jednovýběrový problém – předpoklad normální rozdělení) a v horní části listu vybereme proměnnou *s90p2 čas*. Ve žlutých buňkách se zobrazí jednoduchý výstup popisné statistiky – hodnoty n, $\bar{x}$, s a s^2 (20; 1,184; 0,248 a

0,062). V této části také vložíme hladinu významnosti α, v našem případě 0,05. V zelených buňkách se budou zobrazovat jednotlivé výsledky statistických analýz – viz obr. 1.

V části 1 se zobrazí bodové odhady parametrů normálního rozdělení, tj. odhad střední hodnoty $\hat{\mu} = \bar{x} = 1{,}184$ a odhad rozptylu $\hat{\sigma}^2 = s^2 = 0{,}062$. Navíc se zobrazí i odhad směrodatné odchylky $\hat{\sigma} = s = 0{,}248$ a odhad směrodatné chyby odhadu střední hodnoty estSE = $s/\sqrt{n}$ = 0,055.

Ve 2. části si můžeme vložit zvolenou velikost přípustné chyby Δ a dostaneme požadovaný minimální rozsah souboru potřebný k tomu, aby velikost přípustné chyby nepřekročila s danou pravděpodobností stanovenou mez. Např. pro zvolenou přípustnou chybu $\Delta = 0{,}1$ dostaneme minimální rozsah výběru $n = 27$.

Ve 3. části jsou uvedené intervalové odhady parametru μ. Zde je právě uvedený i oboustranný interval (1,068; 1,300), který jsme v části b) ručního výpočtu také dostali.

Konečně 4. část je určená pro testování hypotéz o střední hodnotě. Nejprve vložíme hodnotu $\mu_0 = 1$ a dostaneme hodnotu testového kritéria $t = 3{,}317$. Dále si mezi nabídnutými alternativami s ohledem na náš řešený problém vybereme alternativu $\mu \neq \mu_0$, tj. $\mu \neq 1$. V řádku tabulky odpovídajícím této alternativě dostaneme následující informace: Hodnota testového kritéria padla do kritického oboru ($t \in W_{0,05}$); hodnota testového kritéria 3,317 překročila kritickou hodnotu – tou je v našem případě Studentův kvantil $t_{0,975}(19) = 2{,}093$; p-hodnota = 0,004 je menší než naše hladina významnosti 0,05. To všechno vede k jedinému závěru, který je zde uvedený také: hypotéza H se zamítá a alternativa A se přijme.

Tento závěr se bude interpretovat zcela shodně jako při ručním zpracování: S 95% spolehlivostí není odhad hodinového intervalu správný.

Výběr z normálního rozdělení

vyber proměnnou s90p2 čas

n	$\bar{x}$	s	s^2
20	1,184	0,248	0,062

α
0,05

1. Bodové odhady parametrů

$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\sigma}$	estSE
1,184	0,062	0,248	0,055

2. Velikost výběru

přípustná chyba Δ	p-st	min.n
0,1	95%	27

3. Intervalové odhady pro střední hodnotu

p-st	oboustranný		Δ	dolní	horní	Δ
95%	1,068	1,300	0,116	1,088	1,280	0,096

4. Testy hypotéz o střední hodnotě

$\mu_0 = 1$ $t = 3{,}317$

spol.	H	A	$t \in W_\alpha$	krit. hodn.	p-hodnota	H	A
95%	$\mu = \mu_0$	$\mu \neq \mu_0$	ano	2,093	0,004	zamítá se	se přijme
		$\mu > \mu_0$	ano	1,729	0,002	zamítá se	se přijme
		$\mu < \mu_0$	ne	-1,729	0,998	nezamítá se	x

Obr. 1: Příklad 90/2 odhad času řešený ve STAT1 – list 1V-normální

2 Dvouvýběrové testy

97/12 – hmotnost sýru

Vážením jsme získali údaje o přesné hmotnosti balíčků sýrů automaticky balených po 250 g, náhodně vybraných před a po seřízení automatu. Údaje o hmotnostech [v gramech]
před seřízením: 243,2 244,8 253,1 247,5 251,0 251,7 254,0 252,5 252,8 250,1
247,3 250,9 253,2 252,7 251,8 245,5
po seřízení: 250,4 250,2 251,1 248,9 249,9 250,2 252,4 250,8
Na 5% hladině významnosti ověřte, zda se seřízením automatu nezměnila nastavená úroveň hmotnosti. Předpokládejte normální rozdělení hmotnosti balíčků.

Řešení:
Před seřízením:
$n_1 = 16$, $\bar{x} \doteq 250{,}131$, $s_1^2 \doteq 11{,}465$, $s_1 \doteq 3{,}386$, $\alpha = 0{,}05$.
Po seřízení:
$n_2 = 8$, $\bar{y} \doteq 250{,}488$, $s_2^2 \doteq 1{,}024$, $s_2 \doteq 1{,}012$.

Nejprve provedeme test o shodě rozptylů.

$$\text{H: } \sigma_1^2 = \sigma_2^2 \quad \rightarrow \quad \text{A: } \sigma_1^2 \neq \sigma_2^2$$

Testové kriterium

$$F = \frac{s_1^2}{s_2^2} = \frac{11{,}465}{1{,}024} \doteq 11{,}196$$

Kritický obor

$$W_{0,05}\text{: } F \leq F_{0,025}(15;\,7) \ \vee\ F \geq F_{0,975}(15;\,7)$$
$$11{,}196 \leq 0{,}304 \ \vee\ 11{,}196 \geq 4{,}568$$

Jelikož $11{,}196 > 4{,}568$, tzn., že hodnota testového kriteria patří do kritického oboru, hypotézu o shodě rozptylů na hladině významnosti 0,05 tedy zamítáme. V dalších výpočtech budeme předpokládat, že rozptyly obou výběrů jsou různé.

Nyní přistoupíme k testu o shodě středních hodnot.

$$\text{H: } \mu_1 = \mu_2 \quad \rightarrow \quad \text{A: } \mu_1 \neq \mu_2$$

Testové kriterium

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{250{,}131 - 250{,}488}{\sqrt{\dfrac{11{,}465}{16} + \dfrac{1{,}024}{8}}} \doteq -0{,}388$$

Kritický obor

$$W_{0,05}\text{: } |\, t\, | \geq t_{0,975}(\nu^*), \text{ kde } \nu^* = [k^*] \text{ a}$$

$$k^* \approx \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{1}{n_1 - 1}\left(\dfrac{s_1^2}{n_1}\right)^2 + \dfrac{1}{n_2 - 1}\left(\dfrac{s_2^2}{n_2}\right)^2} = \frac{\left(\dfrac{11{,}465}{16} + \dfrac{1{,}024}{8}\right)^2}{\dfrac{1}{15}\left(\dfrac{11{,}465}{16}\right)^2 + \dfrac{1}{7}\left(\dfrac{1{,}024}{8}\right)^2} \doteq 19{,}504.$$

Protože $v^* = [19{,}504] = 19$ (funkce [x] znamená celou část argumentu, např. [3,8] = 3), potom

$$W_{0,05}: \qquad |\,t\,| \geq t_{0,975}(19)$$
$$|-0{,}388\,| \geq 2{,}093$$
$$0{,}388 \geq 2{,}093$$

Protože tato nerovnost neplatí, znamená to, že hodnota testového kriteria nepaří do kritického oboru a hypotézu o shodě středních hodnot nemůžeme na hladině významnosti 0,05 zamítnout. Změna úrovně hmotnosti před a po seřízení automatu tedy nebyla prokázána.

Také tuto úlohu můžeme pohodlně řešit v našem excelovském pracovním sešitu STAT1: Otevřeme si list **2V – normální** (dvouvýběrový problém – předpoklad normální rozdělení) a v horní části listu vybereme proměnné *s97p12 sýry-před* a *s97p12 sýry-po*. Ve žlutých buňkách se zobrazí jednoduché výstupy popisné statistiky obou souborů – hodnoty n_1, $\bar{x}$, s_1 a s_1^2 resp. n_2, $\bar{y}$, s_2 a s_2^2 (16; 250,131; 3,386 a 11,465 resp. 8; 250,488; 1,012 a 1,024). V této části také vložíme hladinu významnosti α, v našem případě 0,05. V zelených buňkách se budou zobrazovat jednotlivé výsledky statistických analýz – viz obr. 2.

Dva nezávislé výběry z normálního rozdělení

výber proměnné: s97p12 sýry-před | s97p12 sýry-po

1. výběr

n_1	$\bar{x}$	$s_1(x)$	$s_1^2(x)$
16	250,131	3,386	11,465

2. výběr

n_2	$\bar{y}$	$s_2(y)$	$s_2^2(y)$
8	250,488	1,012	1,024

α
0,05

1. Testy hypotéz o shodě rozptylů

F = 11,195

spol.	H	A	F ∈ W_α	kritická hodnota		p-hodnota	H	A	
95%	$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	ano	0,304	4,568	0,003	zamítá se	se přijme	heteroskedasticita
		$\sigma_1^2 > \sigma_2^2$	ano	3,511		0,002	zamítá se	se přijme	
		$\sigma_1^2 < \sigma_2^2$	ne	0,369		0,998	nezamítá se	x	

2. Testy hypotéz o shodě středních hodnot za předpokladu homoskedasticity $\sigma_1^2 = \sigma_2^2$

t = -0,288 | S = 2,854

spol.	H	A	t ∈ W_α	krit. hodn.	p-hodnota	H	A
95%	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	ne	2,074	0,776	nezamítá se	x
		$\mu_1 > \mu_2$	ne	1,717	0,612	nezamítá se	x
		$\mu_1 < \mu_2$	ne	-1,717	0,388	nezamítá se	x

3. Testy hypotéz o shodě středních hodnot za předpokladu heteroskedasticity $\sigma_1^2 \neq \sigma_2^2$

t = -0,388

spol.	H	A	t ∈ W_α	krit. hodn.	p-hodnota	H	A
95%	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	ne	2,093	0,703	nezamítá se	x
		$\mu_1 > \mu_2$	ne	1,729	0,649	nezamítá se	x
		$\mu_1 < \mu_2$	ne	-1,729	0,351	nezamítá se	x

Obr. 2: Příklad 97/12 hmotnost sýru řešený ve STAT1 – list 2V-normální

V souladu s teorií testování hypotéz o shodě dvou středních hodnot musíme nejprve otestovat shodu obou rozptylů. Výsledky jsou v 1. části listu, hodnota testového kritéria je F = 11,195, a pro alternativu $\sigma_1^2 \neq \sigma_2^2$ překračuje kvantil $F_{0,975}(15; 7)$ = 4,568, také p-hodnota = 0,003 je menší než α = 0,05. Tyto výsledky znamenají, že s 95% pravděpodobností nelze akceptovat shodu rozptylů – homoskedasticitu. Budeme předpokládat neshodu rozptylů – heteroskedasticitu, tento výsledek je v 1. části listu také zobrazený. Ve 2. a 3. části listu řeší STAT1 testy hypotéz o shodě středních hodnot, a to za předpokladu shody (2. část) resp. neshody (3. část) rozptylů. S ohledem na náš výsledek prvního testu – předpoklad heteroskedasticita – použijeme pro další řešení problému 3. část. Hodnota testového kritéria t = –0,388, stupně volnosti ν^* = 19, Studentův kvantil $t_{0,975}(19)$ = 2,093 a p-hodnota = 0,703 vede pro alternativu $\mu_1 \neq \mu_2$ k závěru, že shoda středních hodnot $\mu_1 = \mu_2$ se nezamítá. To prakticky znamená, že změna úrovně hmotnosti před a po seřízení automatu tedy nebyla prokázána.

3 Testy o tvaru rozdělení

Pokud sledujeme reálně jistou náhodnou veličinu prostřednictvím náhodného výběru, potom jednou ze zásadních informací, které budeme při statistické analýze potřebovat, je informace o rozdělení této náhodné veličiny. Přesněji řečeno budeme rozhodovat, zda náš náhodný výběr pochází z normálního rozdělení, nebo zda normální rozdělení jako teoretický model nebude možné akceptovat. I když tuto informaci už můžeme „vysledovat“ z tabulky rozdělení četností resp. z grafu rozdělení četností, korektněji tuto informaci získáme pomocí testů o normalitě – konkrétně pomocí testů o nulové šikmosti a nulové špičatosti resp. C-testu.

V některých reálných situacích může být užitečné ověřit, zda náš výběr nepochází z jiného než normálního rozdělení, např. z Poissonova rozdělení, logaritmicko-normálního rozdělení apod. K tomu slouží χ^2-test dobré shody, kterým lze otestovat shodu dat s jakýmkoliv rozdělením.

99/3 a 91/9 – pneumatiky

Byl proveden test životnosti u 80 kusů pneumatik. Výsledky jsou uvedeny v tabulce.

tisíc km	13	14	15	16	17	18	19
počet	6	22	26	12	9	4	1

a) Vypočítejte koeficienty šikmosti a špičatosti.
b) Pomocí testů o nulové šikmosti a nulové špičatosti ověřte, zda výběr pochází z normálního rozdělení. Použijte hladinu významnosti 0,05 i 0,01.
c) C-testem normality ověřte, zda výběr pochází z normálního rozdělení. Použijte také hladinu významnosti 0,05 i 0,01.

Řešení:
a) Výběrové koeficienty šikmosti a špičatosti určíme v programu STAT1 jako momentové koeficienty a_3 = 0,624 a a_4 = –0,068 – viz obr. 1.

b) Ověření normality je založené na skutečnosti, že normální rozdělení má nulovou šikmost a současně nulovou špičatost: $\alpha_3 = 0 \wedge \alpha_4 = 0$. Proto použijeme tuto nejjednodušší filozofii, která spočívá pouze ve snaze zamítnout nulovou šikmost nebo zamítnout nulovou špičatost. Pokud by se to podařilo, potom prohlásíme, že výběr z normálního rozdělení nepochází. V opačném případě, tedy když nulovou šikmost ani nulovou špičatost nezamítneme, bude možné normální rozdělení jako model pro popis sledované náhodné veličiny akceptovat.

- Nejprve otestujeme nulovou šikmost pro $\alpha = 0{,}05$: užijeme $n = 80$ a $a_3 = 0{,}624$

  H: $\alpha_3 = 0 \;\rightarrow\;$ A: $\alpha_3 \neq 0$

  $$u_3 = \frac{a_3}{\sqrt{D(a_3)}} = \frac{0{,}624}{\sqrt{0{,}0696}} \doteq 2{,}365, \text{ kde } D(a_3) = \frac{6(n-2)}{(n+1)(n+3)} = \frac{6\cdot 78}{81\cdot 83} \doteq 0{,}0696$$

  $W_{0,05}$: $|\,u_3\,| \geq u_{0{,}975}$
  $2{,}365 \geq 1{,}960$ … platí – H se zamítá

  Výběr tedy pochází z rozdělení, které s 95% spolehlivostí vykazuje nenulovou šikmost, to tedy znamená, že normální rozdělení není vhodným modelem pro popis naší náhodné veličiny! V takovém případě test o nulové špičatosti už není potřebné provádět.

- Nyní otestujeme nulovou šikmost pro $\alpha = 0{,}01$: užijeme $n = 80$ a $a_3 = 0{,}624$

  H: $\alpha_3 = 0 \;\rightarrow\;$ A: $\alpha_3 \neq 0$
  výpočet $D(a_3) = 0{,}0696$ a $u_3 = 2{,}365$ se nemění
  $W_{0,01}$: $|\,u_3\,| \geq u_{0{,}995}$
  $2{,}365 \geq 2{,}576$ neplatí – H se nezamítá

  V tomto případě se s 99% spolehlivostí nepodařila prokázat nenulová šikmost. To tedy znamená, že výběr pochází ze symetrického rozdělení, a o normálním rozdělení musíme rozhodnout pomocí testu o nulové špičatosti pro $\alpha = 0{,}01$: užijeme $n = 80$ a $a_4 = -0{,}068$

  H: $\alpha_4 = 0 \;\rightarrow\;$ A: $\alpha_4 \neq 0$

  $$u_4 = \frac{a_4 + \frac{6}{n+1}}{\sqrt{D(a_4)}} = \frac{-0{,}068 + \frac{6}{81}}{\sqrt{0{,}249}} \doteq 0{,}012,$$

  $$\text{kde } D(a_4) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)} = \frac{24\cdot 80\cdot 78\cdot 77}{81^2\cdot 83\cdot 85} \doteq 0{,}249$$

  $W_{0,01}$: $|\,u_4\,| \geq u_{0{,}995}$
  $0{,}012 \geq 2{,}576$ … neplatí – H se nezamítá

  S 99% spolehlivostí se nepodařila prokázat ani nenulová špičatost, to tedy znamená, že na hladině významnosti $\alpha = 0{,}01$ lze normální rozdělení akceptovat jako vhodný model pro popis sledované veličiny.

Dáme-li dohromady naše úvahy, je patrné, že normalita se na hladině významnosti 0,05 zamítá (koeficient šikmosti je nenulový, říkáme také, že je statisticky významný), avšak na hladině významnosti 0,01 je možné považovat data za výběr z normálního rozdělení (oba koeficienty jsou statisticky nevýznamné).

c) C-test normality je založený na skutečnosti, že součet čtverců normovaných veličin u_3 a u_4 má Pearsonovo rozdělení se dvěma stupni volnosti.

- Nejprve otestujeme normalitu pro $\alpha = 0{,}05$: užijeme $u_3 = 2{,}365$ a $u_4 = 0{,}012$
  H: X má normální rozdělení $\rightarrow$ A: X nemá normální rozdělení
  $C = u_3^2 + u_4^2 = 2{,}365^2 + 0{,}012^2 \doteq 5{,}593$

$W_{0,05}$: $C \geq \chi^2_{0,95}(2)$

$5{,}593 \geq 5{,}991$ … neplatí – H se nezamítá

S 95% spolehlivostí se nepodařilo hypotézu o normálním rozdělení zamítnout, a proto budeme normální rozdělení považovat za vhodný model pro popis naší náhodné veličiny.

- Dále otestujeme normalitu pro $\alpha = 0{,}01$: užijeme $u_3 = 2{,}365$, $u_4 = 0{,}012$ a $C = 5{,}593$.
  H: X má normální rozdělení $\rightarrow$ A: X nemá normální rozdělení

  $W_{0,05}$: $C \geq \chi^2_{0,99}(2)$

  $5{,}593 \geq 9{,}210$ … neplatí – H se nezamítá

  S 99% spolehlivostí se také nepodařilo hypotézu o normálním rozdělení zamítnout, a proto budeme i v tomto případě normální rozdělení považovat za vhodný model pro popis naší náhodné veličiny. Rozdíly od normality nejsou tedy na obou hladinách významnosti statisticky významné.

Excelovský pracovní sešit STAT1 nám poskytuje základní informace o normalitě na 3 listech, které jsou určené pro základní zpracování dat: *Popisné charakteristiky*, *Bodové rozdělení* a *Intervalové rozdělení*. Pod tabulkou s popisnými charakteristikami a grafy se nachází část *Ověření normality* – viz obr. 1. Samostatně je zde provedený test o nulové šikmosti, test o nulové špičatosti (závěr o normalitě si musí uživatel udělat sám!) a C-test o normalitě. Na obr. 1 se týkají všechny výstupy našeho řešeného příkladu 91/9 pneumatiky, všechny na hladině významnosti 0,05.

13	Ověření normality		hladina významnosti α : 0,05			
14	test o nulové šikmosti					
15	a_3	$D(a_3)$	u_3	$u_{1-\alpha/2}$	p-hodnota	
16	0,624	0,070	2,366	1,960	0,018	→ nulová šikmost se zamítá
17	test o nulové špičatosti					
18	a_4	$D(a_4)$	u_4	$u_{1-\alpha/2}$	p-hodnota	
19	-0,068	0,249	0,011	1,960	0,991	→ nulová špičatost se nezamítá
20	kombinovaný test o šikmosti a špičatosti: C-test					
21	u_3	u_4	C	$\chi^2_{1-\alpha}(2)$	p-hodnota	
22	2,366	0,011	5,597	5,991	0,061	→ normální rozdělení se nezamítá

Obr. 1: Příklad 91/9 pneumatiky řešený ve STAT1 – list Popisná statistika

101/15 myčka

Po dobu 3 měsíců se v pracovních dnech sledoval počet aut na mycí lince za den.

počet aut	1	2	3	4	5	6	7	8	9	10
počet případů	2	4	8	11	15	12	5	4	2	1

Předpokládejte, že počet aut na myčce má Poissonovo rozdělení. Je tento předpoklad opodstatněný? Použijte χ^2-test dobré shody a řešte na hladině významnosti 0,05.

Řešení:

Odhad parametru lambda provedeme pomocí výběrového průměru (pro Poissonovo rozdělení totiž platí $E(X) = \lambda$ a odhad $\hat{\lambda} = \bar{x} = 5$). Zformulujeme hypotézu a alternativu:

H: X má Poissonovo rozdělení s parametrem $\lambda = 5$

A: X nemá Poissonovo rozdělení s parametrem $\lambda = 5$

Jako testové kriterium použijeme statistiku

$$\chi^2 = \sum_{j=1}^{k} \frac{(n_j - n\pi_j)^2}{n\pi_j},$$

která má při platnosti hypotézy Pearsonovo rozdělení s $\nu = k - c - 1$ stupni volnosti, kde n je rozsah výběrového souboru, k je počet tříd, c je počet neznámých parametrů ověřovaného rozdělení. Potom kritický obor je

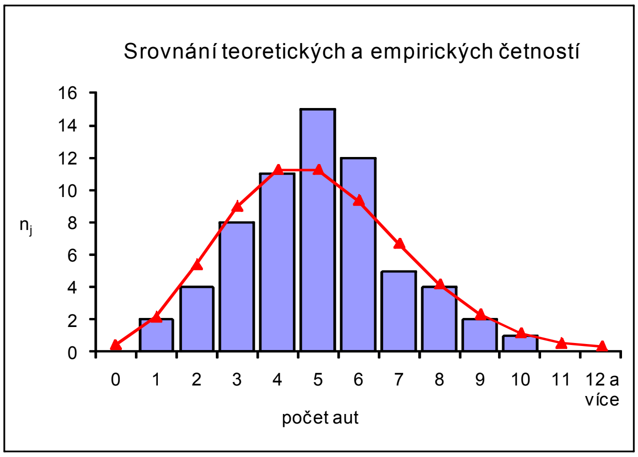
$$W_\alpha = \{\chi^2;\ \chi^2 \geq \chi^2_{1-\alpha}(\nu)\}.$$

x_j	n_j četnosti	π_j hodnoty pravděpodobnostní funkce	$n\pi_j$ teoretické četnosti	sdružené $n\pi_j$	sdružené n_j	$\frac{(n_j - n\pi_j)^2}{n\pi_j}$
0	0	0,00674	0,43136	7,97760	6	0,49024
1	2	0,03369	2,15616			
2	4	0,08422	5,39008			
3	8	0,14037	8,98368	8,98368	8	0,10771
4	11	0,17547	11,23008	11,23008	11	0,00471
5	15	0,17547	11,23008	11,23008	15	1,26556
6	12	0,14622	9,35808	9,35808	12	0,74585
7	5	0,10444	6,68416	6,68416	5	0,42435
8	4	0,06528	4,17792	8,53632	7	0,2765
9	2	0,03627	2,32128			
10 a více	1	0,03183	2,03712			
	64	1,00000	64,00000	64,00000	64	**3,31492**

V tabulce je uveden výpočet testové statistiky. V prvním sloupci je uvedený obor hodnot náhodné veličiny s Poissonovým rozdělením, ve druhém sloupci jsou empirické četnosti. Ve třetím a čtvrtém sloupci jsou pravděpodobnosti (např. z tabulek) a vypočítáme teoretické četnosti. Vzhledem k tomu, že teoretické četnosti v prvních třech a posledních třech třídách jsou menší než 5, provedeme jejich sloučení; sdružené hodnoty jsou uvedené v pátém a šestém sloupci. Sedmý sloupec obsahuje jednotlivé vypočítané hodnoty testového kritéria a jejich součet = hodnota testového kritéria.

Kritický obor pro $\alpha = 0{,}05$ je $\chi^2 \geq \chi^2_{0,95}(5)$, tedy $3{,}315 \geq 11{,}1$ (neplatí). Stupně volnosti určíme ze vztahu $\nu = k - c - 1 = 7 - 1 - 1 = 5$. Protože hodnota testového kriteria nepatří do kritického oboru, testovanou hypotézu, že Poissonovo rozdělení s parametrem $\lambda = 5$ je vhodným modelem pro popis naší náhodné veličiny počet aut na myčce, nemůžeme na hladině významnosti 0,05 zamítnout.

Na obr. 2 je zobrazené srovnání teoretických a empirických četností, ze kterého je vidět, jak empirické četnosti přibližně „kopírují“ teoretický model, což vizuálně také napovídá, že Poissonův model s parametrem $\lambda = 5$ bude možné považovat pro popis naší veličiny jako vhodný. Zobrazený grafický výstup je vytvořený v běžném excelovském prostředí, není součástí programu STAT1.



Obr. 2: Srovnání teoretických a empirických četností