Porovnání empirického a teoretického rozložení

Osnova:

- testy dobré shody pro diskrétní a spojité rozložení při úplně i neúplně specifikovaném problému
- jednoduchý test pro exponenciální a Poissonovo rozložení

Motivace

Možnost použití statistických testů je podmíněna nějakými předpoklady o datech. Velmi často je to předpoklad o typu rozložení, z něhož získaná data pocházejí. Mnoho testů je založeno na předpokladu normality.

Opomíjení předpokladů o typu rozložení může v praxi vést i ke zcela zavádějícím výsledkům, proto je nutné věnovat tomuto problému patřičnou pozornost.

Testy dobré shody pro diskrétní a spojité rozložení

Testujeme hypotézu, která tvrdí, že náhodný výběr $X_1, \ldots, X_n$ pochází z rozložení s distribuční funkcí $\Phi(x)$.

a) Je-li distribuční funkce spojitá, pak data rozdělíme do r třídicích intervalů $\langle u_j, u_{j+1})$, j = 1, ..., r. Zjistíme absolutní četnost n_j j-tého třídicího intervalu a vypočteme pravděpodobnost p_j, že náhodná veličina X s distribuční funkcí $\Phi(x)$ se bude realizovat v j-tém třídicím intervalu. Platí-li nulová hypotéza, pak $p_j = \Phi(u_{j+1}) - \Phi(u_j)$.

b) Má-li distribuční funkce nejvýše spočetně mnoho bodů nespojitosti, pak místo třídicích intervalů použijeme varianty $x_{[j]}$, j = 1, …, r. Pro variantu $x_{[j]}$ zjistíme absolutní četnost n_j a vypočteme pravděpodobnost p_j, že náhodná veličina X s distribuční funkcí $\Phi(x)$ se bude realizovat variantou $x_{[j]}$. Platí-li nulová hypotéza, pak $p_j = \Phi(x_{[j]}) - \lim_{x \to x_{[j]}^-} \Phi(x) = P(X = x_{[j]})$.

Testová statistika: $K = \sum_{j=1}^{r} \frac{(n_j - np_j)^2}{np_j}$. Platí-li nulová hypotéza, pak $K \approx \chi^2(r-1-p)$, kde p je počet odhadovaných parametrů daného rozložení. (Např. pro normální rozložení p = 2, protože z dat odhadujeme střední hodnotu a rozptyl.) Nulovou hypotézu zamítáme na asymptotické hladině významnosti α, když testová statistika $K \geq \chi^2_{1-\alpha}(r-1-p)$. Aproximace se považuje za vyhovující, když teoretické četnosti $np_j \geq 5$, j = 1, ..., r.

Upozornění: Hodnota testové statistiky K je silně závislá na volbě třídicích intervalů. Navíc při nesplnění podmínky $np_j \geq 5$, j = 1, ..., r je třeba některé intervaly resp. varianty slučovat, což vede ke ztrátě informace.

Příklad: Testování shody empirického a teoretického rozložení při úplně specifikovaném problému

Byl zjišťován počet poruch určitého zařízení za 100 hodin provozu ve 150 disjunktních 100 h intervalech. Výsledky měření:

Počet poruch za 100 hodin provozu	0	1	2	3	4 a víc
Absolutní četnost	52	48	36	10	4

Na asymptotické hladině významnosti 0,05 testujte hypotézu, že náhodný výběr $X_1, \ldots, X_{150}$ pochází z rozložení Po(1,2).

Řešení:
Pravděpodobnost, že náhodná veličina s rozložením Po(λ), kde λ = 1,2 bude nabývat hodnot 0, 1, ..., 4 a víc je

$$p_j = \frac{\lambda^j}{j!}e^{-\lambda} = \frac{1{,}2^j}{j!}e^{-1{,}2}, j = 0, 1, 2, 3, \quad p_4 = 1 - (p_0 + p_1 + p_2 + p_3).$$

Výpočty potřebné pro stanovení testové statistiky K uspořádáme do tabulky.

j	n_j	p_j	np_j	$(n_j - np_j)^2/ np_j$
0	52	0,301	150.0,301=45,15	1,039
1	48	0,361	150.0,361=54,15	0,698
2	36	0,217	150.0,217=32,55	0,366
3	10	0,087	150.0,087=13,05	0,713
4	4	0,034	150.0,034=5,1	0,237

Podmínky dobré aproximace jsou splněny, všechny teoretické četnosti jsou větší než 5.
K = 1,039 + 0,698 + 0,713 + 0,237 = 3,053, r = 5, $\chi^2_{0,95}(4)$ = 9,488. Protože 3,053 < 9,488, nulovou hypotézu nezamítáme na asymptotické hladině významnosti 0,05.

Výpočet pomocí systému STATISTICA:

Načteme datový soubor poruchy.sta. Proměnná POCET obsahuje počet poruch, proměnná CETNOST pak absolutní četnosti zjištěného počtu poruch.

Statistiky – Prokládání rozdělení – Diskrétní rozdělení – Poissonovo – OK – Proměnná POCET – klikneme na ikonu se závažím – Proměnná vah CETNOST – Stav Zapnuto – OK – záložka Parametry - Lambda 1,2 - Výpočet.

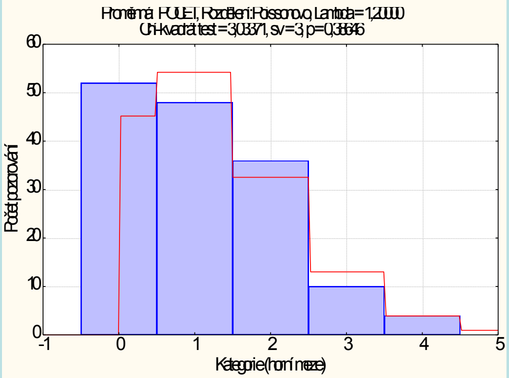
	Proměnná: POCET, Rozdělení:Poissonovo, Lambda = 1,200 (poruchy.s Chí-kvadrát = 3,03371, sv = 3, p = 0,38646							
Kategorie	Pozorov. Četnos	Kumulat Pozorov.	Procen Pozorov.	Kumul. Pozorov.	Očeká Četno	Kumula Očeká	Proce Očeká	Kumul. Očeká
<= 0,000	5	5	34,66	34,6	45,17	45,1	30,11	30,1
1,00000	4	10	32,00	66,6	54,21	99,3	36,14	66,2
2,00000	3	13	24,00	90,6	32,52	131,9	21,68	87,9
3,00000	1	14	6,666	97,3	13,01	144,9	8,67	96,6
< Nekone	4	15	2,666	100,0	5,06	150,0	3,37	100,0

V záhlaví výstupní tabulky je uvedena hodnota testového kritéria (3,03371), počet stupňů volnosti = 3 a p-hodnota (0,38646). Nulová hypotéza se tedy nezamítá na asymptotické hladině významnosti 0,05.

Počet stupňů volnosti 3 však neodpovídá tomu, že známe parametr λ, ve skutečnosti je počet stupňů volnosti 4. Proto pro výpočet p-hodnoty otevřeme nový datový soubor o jedné proměnné a jednom případu.

Do Dlouhého jména napíšeme =1-IChi2(3,03371;4). Dostaneme p-hodnotu 0,5522.

Pro vytvoření grafu se vrátíme do Proložení diskrétních rozložení – Základní výsledky – Graf pozorovaného a očekávaného rozdělení



V grafu jsou patrné určité rozdíly mezi hodnotami pravděpodobnostní a četnostní funkce, ale tyto rozdíly nejsou příliš velké.

Příklad: Testování shody empirického a teoretického rozložení při neúplně specifikovaném problému

V tabulce jsou roztříděny fotbalové zápasy určité soutěže podle počtu vstřelených branek.

Počet branek	0	1	2	3	4 a víc
Počet zápasů	19	30	17	10	8

Na hladině významnosti 0,05 testujte hypotézu, že jde o výběr z Poissonova rozložení.

Výpočet pomocí systému STATISTICA:

Načteme datový soubor branky.sta. Proměnná POCET obsahuje počet vstřelených branek, proměnná CETNOST pak počet zápasů, v nichž bylo dosaženo zjištěného počtu branek.

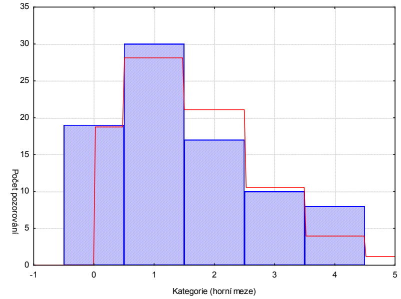
Statistiky – Prokládání rozdělení – Diskrétní rozdělení – Poissonovo – OK – Proměnná POCET – klikneme na ikonu se závažím – Proměnná vah CETNOST – Stav Zapnuto – OK – Výpočet.

Kategorie	Proměnná: POCET, Rozdělení: Poissonovo, Lambda = 1,500 (branky.st Chí-kvadrát = 2,07051, sv = 3, p = 0,55790							
	Pozorov Četnos	Kumulat Pozorov	Procen Pozorov	Kumul. Pozorov	Očeká Četno	Kumula Očeká	Proce Očeká	Kumul. Očeká
<= 0,000	1	1	22,61	22,6	18,74	18,74	22,31	22,3
1,00000	3	4	35,71	58,3	28,11	46,85	33,46	55,7
2,00000	1	6	20,23	78,5	21,08	67,94	25,10	80,8
3,00000	1	7	11,90	90,4	10,54	78,48	12,55	93,4
< Nekone	8	8	9,52	100,0	5,51	84,00	6,56	100,0

V tomto případě je parametr λ Poissonova rozložení neznámý, je odhadnut pomocí výběrového průměru a odhad činí 1,5.

Dále je v záhlaví výstupní tabulky uvedena hodnota testového kritéria (Chí kvadrát = 2,07051), počet stupňů volnosti $r - p - 1 = 5 - 1 - 1 = 3$ a p-hodnota (0,5578). Nulová hypotéza se tedy nezamítá na asymptotické hladině významnosti 0,05.

Pro vytvoření grafu se vrátíme do Proložení diskrétních rozložení – Základní výsledky – Graf pozorovaného a očekávaného rozdělení.



Poznámka k testu dobré shody: Tento test může být použit i v těch případech, kdy rozložení, z něhož daný náhodný výběr pochází, neodpovídá nějakému známému rozložení (např. exponenciálnímu, normálnímu, Poissonovu, ...), ale je určeno intuitivně nebo na základě zkušenosti.

Příklad: Ve svých pokusech pozoroval J.G. Mendel 10 rostlin hrachu a na každé z nich počet žlutých a zelených semen. Výsledky pokusu:

číslo rostliny	1	2	3	4	5	6	7	8	9	10
počet žlutých semen	25	32	14	70	24	20	32	44	50	44
počet zelených semen	11	7	5	27	13	6	13	9	14	18
celkem	36	39	19	97	37	26	45	53	64	62

Z genetických modelů vyplývá, že pravděpodobnost výskytu žlutého semene by měla být 0,75 a zeleného 0,25. Na asymptotické hladině významnosti 0,05 testujte hypotézu, že výsledky Mendelových pokusů se shodují s modelem.

Řešení:
Výpočty potřebné pro stanovení testové statistiky K uspořádáme do tabulky.

j	n_j	p_j	np_j	$(n_j - np_j)^2 / np_j$
1	25	0,75	36.0,75=27	0,148148
2	32	0,75	39.0,75=29,25	0,258547
:	:	:	:	:
10	44	0,75	62.0,75=46,5	0,134409

$K = 0{,}148148 + 0{,}258547 + ... + 0{,}134409 = 1{,}797495$, $r = 10$, $\chi^2_{0,95}(9) = 16{,}9$.
Protože $1{,}797495 < 16{,}9$, nulovou hypotézu nezamítáme na asymptotické hladině významnosti 0,05.

Výpočet pomocí systému STATISTICA:

Načteme datový soubor Mendel hrach.sta. Proměnná celkem obsahuje celkový počet semen, X obsahuje pozorovaný počet žlutých semen a Y vypočítané teoretické četnosti žlutých semen (v našem případě X*0,75).

Statistiky – Neparametrická statistika – Pozorované versus očekávané χ^2 – OK - Pozorované četnosti X, Očekávané četnosti Y - OK – Výpočet. Dostaneme tabulku:

Přípa	Pozorované vs. očekávané četno Chi-Kvadr. = 1,797495 sv = 9 p = POZN.: Nestejné součty pozor. a			
	pozor X	očeka Y	P - C	(P-O) /O
C:	25,00	27,00	-2,00	0,148
C:	32,00	29,25	2,750	0,258
C:	14,00	14,25	-0,25	0,004
C:	70,00	72,75	-2,75	0,103
C:	24,00	27,75	-3,75	0,506
C:	20,00	19,50	0,500	0,012
C:	32,00	33,75	-1,75	0,090
C:	44,00	39,75	4,250	0,454
C:	50,00	48,00	2,000	0,083
C:	44,00	46,50	-2,50	0,134
Sčt	355,0	358,5	-3,50	1,797

Ve výstupní tabulce najdeme hodnotu testové statistiky (Chi-Kvadr = 1,797495), počet stupňů volnosti (sv = 9) a odpovídající p-hodnotu, kterou porovnáme se zvolenou hladinou významnosti. V našem případě je p-hodnota 0,99428, takže nulová hypotéza se nezamítá na asymptotické hladině významnosti 0,05.

Příklad: Při 60 hodech kostkou jsme dosáhli těchto výsledků: 9 x jednička, 11 x dvojka, 10 x trojka, 13 x čtyřka, 11 x pětka a 6 x šestka. Na asymptotické hladině významnosti 0,05 testujte hypotézu, že kostka je homogenní.

Řešení: n = 60

j	n_j	p_j	np_j	$(n_j - np_j)^2$	$(n_j - np_j)^2/ np_j$
1	9	1/6	10	1	1/10
2	11	1/6	10	1	1/10
3	10	1/6	10	0	0
4	13	1/6	10	9	9/10
5	11	1/6	10	1	1/10
6	6	1/6	10	16	16/10

K = 2,8, r = 6, p = 0, $\chi^2_{0,95}(5) = 11,07$. Protože K < 11,07, H_0 nezamítáme na asymptotické hladině významnosti 0,05.

Výpočet pomocí systému STATISTICA:

Načteme datový soubor kostka.sta. Proměnná X obsahuje pozorované četnosti jednotlivých čísel 1, …, 6 a proměnná Y obsahuje teoretické četnosti (v našem případě 10).

Statistiky – Neparametrická statistika – Pozorované versus očekávané χ^2 – OK - Pozorované četnosti X, Očekávané četnosti Y - OK – Výpočet. Dostaneme tabulku:

	Pozorované vs. očekávané c Chi-Kvadr. = 2,800000 sv = 5			
Přípa	pozorc X	očeka Y	P - C	(P-O) /O
C:	9,00(	10,00	-1,00(	0,100
C:	11,00	10,00	1,00(	0,100
C:	10,00	10,00	0,00(	0,000
C:	13,00	10,00	3,00(	0,900
C:	11,00	10,00	1,00(	0,100
C:	6,00(	10,00	-4,00(	1,600
Sct	60,00	60,00	0,00(	2,800

Ve výstupní tabulce najdeme hodnotu testové statistiky (Chi-Kvadr = 2,8), počet stupňů volnosti (sv = 5) a odpovídající p-hodnotu, kterou porovnáme se zvolenou hladinou významnosti. V našem případě je p-hodnota 0,730786, takže nulová hypotéza se nezamítá na asymptotické hladině významnosti 0,05.

Příklad: Ze záznamů autosalónu byl ve 100 náhodně vybraných dnech zjištěn počet prodaných aut.

Počet prodaných aut za den	0	1	2	3	4	5 a víc
Počet dnů	9	43	29	11	5	3

Na asymptotické hladině významnosti 0,05 testujte hypotézu, že počet prodaných aut za den se řídí Poissonovým rozložením.

Řešení:

Parametr λ Poissonova rozložení neznáme, odhadneme ho pomocí výběrového průměru.

$m = \frac{1}{n}\sum_{j=1}^{r} n_j x_j = \frac{1}{100}(0\cdot 9 + 1\cdot 43 + 2\cdot 29 + 3\cdot 11 + 4\cdot 5 + 5\cdot 3) = 1{,}7$. Pravděpodobnost, že náhodná veličina X ~ Po(1,7) bude nabývat hodnot p_j, j = 0,1,2,3,4,5 a víc, je $p_j = \frac{1{,}7^j}{j!}e^{-1{,}7}, j = 0,1,2,3,4, p_5 = 1 - (p_0 + p_1 + p_2 + p_3 + p_4)$

j	n_j	p_j	np_j	$(n_j - np_j)^2$	$(n_j - np_j)^2/ np_j$
0	9	0,1827	18,27	85,9329	4,7035
1	43	0,3106	31,06	142,5636	4,5899
2	29	0,264	26,4	6,76	0,2561
3	11	0,1496	14,96	15,6816	1,0482
4	5	0,0636	6,36	1,8496	0,2908
5	3	0,0296	2,96	0,0016	0,0005

K = 10,8891, r = 6, p = 1, $\chi^2_{0,95}(4) = 9{,}488$. Protože K ≥ 9,488, H_0 zamítáme na asymptotické hladině významnosti 0,05.

Výpočet pomocí systému STATISTICA:

Načteme datový soubor autosalon.sta. Proměnná POCET obsahuje počet prodaných aut, proměnná CETNOST pak počet dnů, v nichž byl prodán zjištěný počet aut.

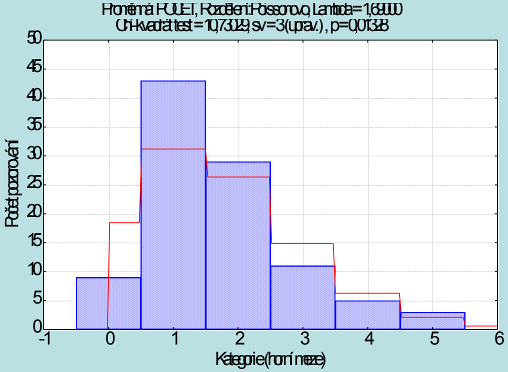
Statistiky – Prokládání rozdělení – Diskrétní rozdělení – Poissonovo – OK – Proměnná POCET – klikneme na ikonu se závažím – Proměnná vah CETNOST – Stav Zapnuto – OK – Výpočet.

	Proměnná: POCET, Rozdělení:Poissonovo, Lambda = 1,69000 Chí-kvadrát = 10,73029, sv = 3 (uprav.) , p = 0,01328							
Kategorie	Pozorov. Četnos	Kumulat Pozorov	Procer Pozorov	Kumul. Pozorov	Očeka Četnc	Kumula Očeká	Proce Očeka	Kumul. Očeká
<= 0,000	9	9	9,000	9,00	18,45	18,45	18,45	18,45
1,00000	43	52	43,00	52,00	31,18	49,63	31,18	49,63
2,00000	29	8	29,00	81,00	26,35	75,98	26,35	75,98
3,00000	1	92	11,00	92,00	14,84	90,83	14,84	90,83
4,00000	5	9	5,000	97,00	6,27	97,10	6,27	97,10
< Nekone	3	10	3,000	100,0	2,898	100,0	2,898	100,0

V záhlaví výstupní tabulky uvedena hodnota testového kritéria (10,73029), počet stupňů volnosti 3 a p-hodnota (0,01328). Nulová hypotéza se tedy zamítá na asymptotické hladině významnosti 0,05.

Vidíme, že nesouhlasí počet stupňů volnosti, měl by být 4. Proto p-hodnotu vypočteme zvlášť. Otevřeme nový datový soubor o jedné proměnné a jednom případu. Do Dlouhého jména napíšeme =1-IChi2(10,73029;4). Dostaneme p-hodnotu 0,0298.

Pro vytvoření grafu se vrátíme do Proložení diskrétních rozložení – Základní výsledky – Graf pozorovaného a očekávaného rozdělení.



V tomto případě jsou patrné značné rozdíly mezi pozorovanými a teoretickými četnostmi.

Jednoduchý test exponenciálního rozložení

Testujeme hypotézu, která tvrdí, že náhodný výběr $X_1, \ldots, X_n$ pochází z exponenciálního rozložení. Označme M výběrový průměr a S^2 výběrový rozptyl tohoto náhodného výběru. Víme, že střední hodnota náhodné veličiny $X \sim Ex(\lambda)$ je $E(X) = 1/\lambda$ a rozptyl je $D(X) = 1/\lambda^2$.

Test založíme na statistice $K = \frac{(n-1)S^2}{M^2}$, která se v případě platnosti H_0 asymptoticky řídí rozložením $\chi^2(n-1)$.

Kritický obor: $W = \left(0, \chi^2_{\alpha/2}(n-1)\right\rangle \cup \left\langle \chi^2_{1-\alpha/2}(n-1), \infty\right)$.

Jestliže $K \in W$, H_0 zamítáme na asymptotické hladině významnosti α.

Příklad: Byla zkoumána doba životnosti 45 součástek (v hodinách). Zjistili jsme, že průměrná doba životnosti činila m = 99,93 h a rozptyl $s^2 = 7328{,}91\ h^2$. Na asymptotické hladině významnosti 0,05 testujte hypotézu, že daný náhodný výběr pochází z exponenciálního rozložení.

Řešení:

Testová statistika: $K = \frac{(n-1)s^2}{m^2} = \frac{44 \cdot 7328{,}91}{99{,}93^2} = 32{,}292$

Kritický obor: $W = \left(0, \chi^2_{\alpha/2}(n-1)\right\rangle \cup \left\langle \chi^2_{1-\alpha/2}(n-1), \infty\right) = \left(0, \chi^2_{0{,}025}(44)\right\rangle \cup \left\langle \chi^2_{0{,}975}(44), \infty\right) = \left(0; 27{,}575\right\rangle \cup \left\langle 64{,}202; \infty\right)$

Protože se testová statistika nerealizuje v kritickém oboru, hypotézu o exponenciálním rozložení nezamítáme na asymptotické hladině významnosti 0,05.

Jednoduchý test Poissonova rozložení

Testujeme hypotézu, která tvrdí, že náhodný výběr $X_1, ..., X_n$ pochází z Poissonova rozložení. Označme M výběrový průměr a S^2 výběrový rozptyl tohoto náhodného výběru. Víme, že střední hodnota náhodné veličiny $X \sim Po(\lambda)$ je $E(X) = \lambda$ a rozptyl je $D(X) = \lambda$.

Test založíme na statistice $K = \frac{(n-1)S^2}{M}$, která se v případě platnosti H_0 asymptoticky řídí rozložením $\chi^2(n-1)$.

Kritický obor: $W = \left(0, \chi^2_{\alpha/2}(n-1)\right\rangle \cup \left\langle \chi^2_{1-\alpha/2}(n-1), \infty\right)$.

Příklad: Studujeme rozložení počtu pacientů, kteří během 75 dnů přijdou na pohotovost. Osmihodinovou pracovní dobu rozdělíme do půlhodinových intervalů a v každém intervalu zjistíme počet příchozích pacientů:

Počet pacientů	0	1	2	3	4	5	6	7	8	9	10
Pozorovaná četnost	79	188	282	275	196	114	45	10	7	3	1

Na asymptotické hladině významnosti 0,05 testujte hypotézu, že daný náhodný výběr pochází z Poissonova rozložení.

Řešení:

Nejprve musíme vypočítat realizaci výběrového průměru a výběrového rozptylu:

$$m = \frac{1}{1200}\left(0 \cdot 79 + 1 \cdot 188 + \dots + 10 \cdot 1\right) = 2,803$$

$$s^2 = \frac{1}{1199}\left[79 \cdot (0-2,803)^2 + 188 \cdot (1-2,803)^2 + \dots + 1 \cdot (10-2,803)^2\right] = 2,7085$$

$$K = \frac{(n-1)s^2}{m} = \frac{1199 \cdot 2,7085}{2,803} = 1158,57,$$

Kritický obor: $W = \left(0, \chi^2_{\alpha/2}(n-1)\right\rangle \cup \left\langle \chi^2_{1-\alpha/2}(n-1), \infty\right) = \left(0; 1104,93\right\rangle \cup \left\langle 1296,86; \infty\right)$

H_0 nezamítáme na asymptotické hladině významnosti 0,05.

Příklad: V systému hromadné obsluhy byla sledována doba obsluhy 70 zákazníků (v min). Výsledky jsou uvedeny v tabulce rozložení četností:

Doba obsluhy	Počet zákazníků
(0, 3]	14
(3,6]	16
(6,9]	10
(9,12]	9
(12,15]	8
(15,18]	5
(18,21]	3
(21,24]	5

Na asymptotické hladině významnosti 0,05 testujte hypotézu, že daný náhodný výběr pochází z exponenciálního rozložení. Použijte:

a) test dobré shody,

b) jednoduchý test exponenciálního rozložení

Řešení:

Testujeme H_0: náhodný výběr $X_1, \ldots, X_{70}$ pochází z Ex(λ) proti H_1: non H_0.

Ad a) Nejprve odhadneme parametr λ exponenciálního rozložení: $\hat{\lambda} = \frac{1}{\frac{1}{n}\sum \ldots} = \frac{1}{\frac{1}{70}(14\cdot 1{,}5 + 16\cdot 4{,}5 + \cdots + 5\cdot 22{,}5)} = 0{,}1122$

Pravděpodobnost, že náhodná veličina s rozložením Ex(λ), kde $\lambda = 0{,}1122$ se bude realizovat v intervalu $(u_j, u_{j+1}\rangle$ je

$p_j = \Phi(u_{j+1}) - \Phi(u_j)$, $j = 1, \ldots, r$, kde $\Phi \ldots = \ldots - \ldots$.

Výpočty potřebné pro stanovení testové statistiky K uspořádáme do tabulky.

$(u_j, u_{j+1}\rangle$	$x_{[j]}$	n_j	p_j	np_j
(0, 3]	1,5	14	0,2858	20,0033
(3,6]	4,5	16	0,2041	14,2871
(6,9]	7,5	10	0,1458	10,2044
(9,12]	10,5	9	0,1041	7,2884
(12,15]	13,5	8	0,0744	5,2056
(15,18]	16,5	5	0,0531	3,7181
(18,21]	19,5	3	0,0378	2,6556
(21,24]	22,5	5	0,0271	1,8967

Podmínky dobré aproximace nejsou splněny, sloučíme tedy intervaly (15,18], (18,21] a (21,24].

$(u_j, u_{j+1}\rangle$	$x_{[j]}$	n_j	p_j	np_j	$(n_j - np_j)^2 / np_j$
(0, 3]	1,5	14	0,2858	20,0033	1,8017
(3,6]	4,5	16	0,2041	14,2871	0,2054
(6,9]	7,5	10	0,1458	10,2044	0,0041
(9,12]	10,5	9	0,1041	7,2884	0,4020
(12,15]	13,5	8	0,0744	5,2056	1,5000
(15,24]	19,5	13	0,1181	8,2704	2,7047

Testová statistika K = 1,8017 + … + 2,7047 = 6,6178, r = 6, p = 1, r – p – 1 = 4, $\chi^2_{0,95}(4) = 9{,}4877$.
Testová statistika se nerealizuje v kritickém oboru $W = \langle 9{,}4877, \infty)$, na asymptotické hladině významnosti 0,05 nelze zamítnout hypotézu, že doba obsluhy se řídí exponenciálním rozložením.

Ad b) Jednoduchý test exponenciálního rozložení je založen na statistice $K = \frac{(n-1)S^2}{M^2}$, která se v případě platnosti H_0 asymptoticky řídí rozložením $\chi^2(n-1)$.

Kritický obor: $W = (0, \chi^2_{\alpha/2}(n-1)\rangle \cup \langle \chi^2_{1-\alpha/2}(n-1), \infty)$.

Nejprve musíme vypočítat realizaci výběrového průměru a výběrového rozptylu:

$$m = \frac{1}{70}(14 \cdot 15 + 16 \cdot 45 + \dots + 5 \cdot 225) = 89{,}14[illegible]$$

$$s^2 = \frac{1}{69}\left(19 \cdot (15 - 89{,}143)^2 + 16 \cdot (45 - 89{,}143)^2 + \dots + 5 \cdot (225 - 89{,}143)^2\right) = [illegible],144[illegible]$$

$$K = \frac{(n-1)s^2}{m^2} = \frac{69 \cdot [illegible],11447}{89{,}143^2} = [illegible],5726$$

Kritický obor: $W = (0, \chi^2_{\alpha/2}(n-1)\rangle \cup \langle \chi^2_{1-\alpha/2}(n-1), \infty) = (0, \chi^2_{0,025}(69)\rangle \cup \langle \chi^2_{0,975}(69), \infty) = (0; 47{,}9242\rangle \cup \langle 93{,}8565; \infty)$.

H_0 zamítáme na asymptotické hladině významnosti 0,05.