Testování neparametrických hypotéz

Testy dobré shody

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

doc. Ing. Petr Seďa, Ph.D.

Co se dnes dozvíte?

- Ověření pravděpodobnostního rozdělení.
- χ^2 test dobré shody.

Ověření pravděpodobnostního rozdělení

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Možnosti ověření pravděpodobnostního rozdělení:

a) **Grafická analýza** pro ověření shody empirického a teoretického rozdělení

b) $\boldsymbol{\chi^2}$ **test dobré shody** s očekávaným rozdělením

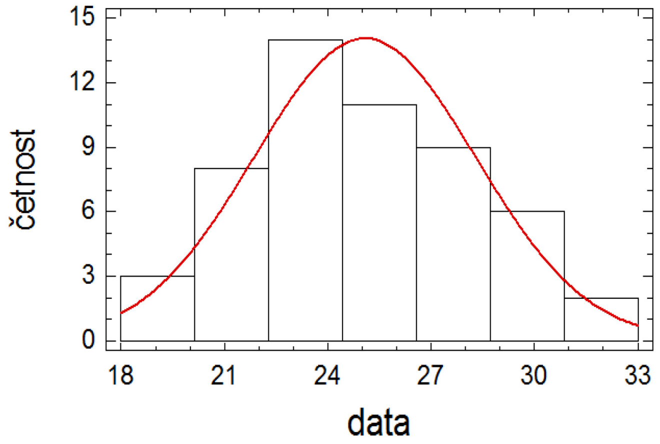
- úplně specifikovaný test,
- neúplně specifikovaný test,

Poznámka: χ^2 čti: chí-kvadrát (česky) nebo „kaj skvére“ (anglicky)

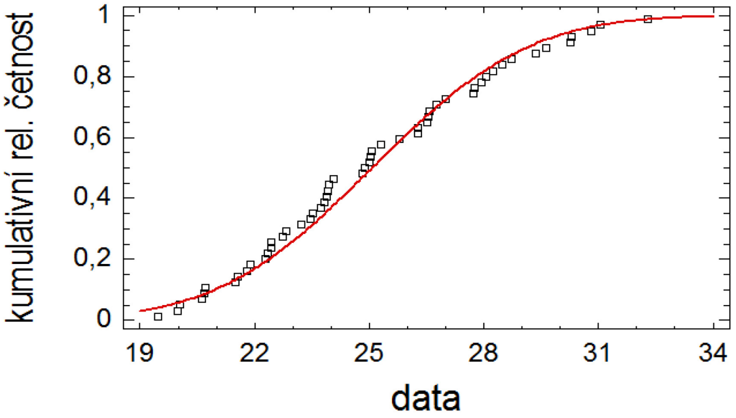
Jak ověřit, zda se empirické rozdělení shoduje s teoretickým?

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Grafické nástroje:



Srovnání histogramu
s teoretickou hustotou pravděpodobností



Srovnání kumulativní relativních četností
(resp. empirické distribuční funkce) s
teoretickou distribuční funkcí

Jak ověřit, zda se empirické rozdělení shoduje s teoretickým?

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Proč statistický test?

Pomocí grafické analýzy můžeme metodou srovnání se standardními modely pouze **odhadnout typ rozdělení**!

Objektivní míru shody dat s teoretickým modelem poskytují tzv. **testy dobré shody.**

Jak ověřit, zda se empirické rozdělení shoduje s teoretickým?

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Testy dobré shody:

H_0: Teoretické a empirické rozdělení se **shoduje**.

H_A: Teoretické a empirické rozdělení se **neshoduje.**

Vybrané testy dobré shody:

a) χ^2 **test dobré shody:**

i) ověření, zda jsou relativní četnosti jednotlivých variant rovny číslům

$\pi_{0_1}; \ldots ; \pi_{0_k}$,

ii) χ^2- test dobré shody s očekávaným rozdělením:

- úplně specifikovaný test,
- neúplně specifikovaný test,

Testy dobré shody

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Ověření shody relativních četností jednotlivých variant:

V nejjednodušším případě použití χ^2 testu lze konečnou populaci roztřídit podle nějakého znaku do ***k* disjunktních skupin** (tzv. variant) a my chceme na základě náhodného výběru ověřit, zda jsou relativní četnosti jednotlivých variant rovny číslům $\pi_{0_1}, \pi_{0_2}, \dots, \pi_{0_k}$.

Předpoklady testu: Dostatečně velký výběr – tj. v praxi: všechny **očekávané** četnosti větší než 2 a alespoň 80% **očekávaných** četností větších než 5.

Testové kritérium: $G = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$, kde E_i jsou očekávané četnosti a O_i jsou skutečné četnosti.

Rozhodnutí: Pokud je hodnota testového kritéria větší než kvantil $\chi^2_{1-\alpha}$ rozdělení s $k-1$ stupni volnosti, kde k je počet skupin, **zamítáme nulovou hypotézu** o shodě rozdělení.

Příklad 1- test dobré shody (relativní četnosti)

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Bylo provedeno šetření mezi ženami staršími 18 let. Mezi 246 náhodně oslovenými ženami bylo 80 (32,5%) svobodných, 110 (44,7%) vdaných, 30 (12,2%) rozvedených a 26 (10,6%) ovdovělých. Je známo (viz Český statistický úřad), že v ČR je mezi ženami staršími 15 let cca 24,8% svobodných, 49,0% vdaných, 12,6% rozvedených a 13,6% ovdovělých. Lze provedený výběr označit za reprezentativní? Hladina významnosti je 5%.

Stav	svobodná	vdaná	rozvedená	ovdovělá	celkem
očekávaná rel. četnost π_{0_i}	0,248	0,490	0,126	0,136	1,000
pozorovaná (angl. observed) četnost O_i	80	110	30	26	246
očekávaná (angl. expected) četnost E_i	61,0	120,5	31,0	33,5	---

Liší se pozorované a očekávané četnosti statisticky významně?

Příklad 1- test dobré shody (relativní četnosti)

Bylo provedeno šetření mezi ženami staršími 18 let. Mezi 246 náhodně oslovenými ženami bylo 80 (32,5%) svobodných, 110 (44,7%) vdaných, 30 (12,2%) rozvedených a 26 (10,6%) ovdovělých. Je známo (viz Český statistický úřad), že v ČR je mezi ženami staršími 15 let cca 24,8% svobodných, 49,0% vdaných, 12,6% rozvedených a 13,6% ovdovělých. Lze provedený výběr označit za reprezentativní? Hladina významnosti je 5%.

H_0: Provedený výběr **je výběrem z populace**, v níž jsou relativní četnosti jednotlivých variant dány tabulkou:

Stav	svobodná	vdaná	rozvedená	ovdovělá	celkem
očekávaná rel. četnost π_{0_i}	0,248	0,490	0,126	0,136	1,000

H_A: Provedený výběr **není výběrem z populace**, v níž jsou relativní četnosti jednotlivých variant dány výše uvedenou tabulkou.

Příklad 1- test dobré shody (relativní četnosti)

Bylo provedeno šetření mezi ženami staršími 18 let. Mezi 246 náhodně oslovenými ženami bylo 80 (32,5%) svobodných, 110 (44,7%) vdaných, 30 (12,2%) rozvedených a 26 (10,6%) ovdovělých. Je známo (viz Český statistický úřad), že v ČR je mezi ženami staršími 15 let cca 24,8% svobodných, 49,0% vdaných, 12,6% rozvedených a 13,6% ovdovělých. Lze provedený výběr označit za reprezentativní? Hladina významnosti je 5%.

Stav	svobodná	vdaná	rozvedená	ovdovělá	celkem
očekávaná rel. četnost π_{0_i}	0,248	0,490	0,126	0,136	1,000
pozorovaná (angl. observed) četnost O_i	80	110	30	26	246
očekávaná (angl. expected) četnost E_i	61,0	120,5	31,0	33,5	---

Ověření předpokladů: Všechny očekávané četnosti jsou větší než 5.

$$x_{obs} = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} = \frac{(80-61{,}0)^2}{61{,}0} + \frac{(110-120{,}5)^2}{120{,}5} + \frac{(30-31{,}0)^2}{31{,}0} + \frac{(26-33{,}5)^2}{33{,}5} = 8{,}53$$

Příklad 1- test dobré shody (relativní četnosti)

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Bylo provedeno šetření mezi ženami staršími 18 let. Mezi 246 náhodně oslovenými ženami bylo 80 (32,5%) svobodných, 110 (44,7%) vdaných, 30 (12,2%) rozvedených a 26 (10,6%) ovdovělých. Je známo (viz Český statistický úřad), že v ČR je mezi ženami staršími 15 let cca 24,8% svobodných, 49,0% vdaných, 12,6% rozvedených a 13,6% ovdovělých. Lze provedený výběr označit za reprezentativní? Hladina významnosti je 5%.

Stav	svobodná	vdaná	rozvedená	ovdovělá	celkem
očekávaná rel. četnost π_{0_i}	0,248	0,490	0,126	0,136	1,000
pozorovaná (angl. observed) četnost O_i	80	110	30	26	246
očekávaná (angl. expected) četnost E_i	61,0	120,5	31,0	33,5	---

Kvantil $\chi^2_{1-\alpha}$ rozdělení s $4-1$ stupni volnosti je roven 7,815 (viz tabulky).

Protože je hodnota testového kritéria větší než hodnota kvantilu, na hladině významnosti 0,05 **zamítáme H_0**, tj. výběr nelze považovat za reprezentativní.

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

χ^2 test dobré shody s očekávaným rozdělením:

H_0: Empirické a teoretické rozdělení se shoduje neboli výběr **pochází z určitého teoretického rozdělení**.

H_A: Empirické a teoretické rozdělení se neshoduje, neboli **není pravda, že výběr pochází z určitého teoretického rozdělení.**

Rozlišujeme:

a) **Úplně specifikovaný test** – v nulové hypotéze jsou specifikovány všechny parametry teoretického rozdělení (např. H_0: Výběr pochází z Poissonova rozdělení se střední hodnotou 5.)

b) **Neúplně specifikovaný test** – v nulové hypotéze nejsou specifikovány všechny parametry teoretického rozdělení (např. H_0: Výběr pochází z Poissonova rozdělení.). V tomto případě musíme nespecifikované parametry teoretického rozdělení odhadnout (bodový odhad) z výběru.

Testy dobré shody

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

χ^2 test dobré shody s očekávaným rozdělením:

Předpoklady testu: Dostatečně velký výběr – tj. v praxi: všechny očekávané četnosti větší než 2 a alespoň 80% očekávaných četností větších než 5.

Testové kritérium: $G = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$, kde E_i jsou očekávané četnosti a O_i jsou skutečné četnosti.

Rozhodnutí: Pokud je hodnota testového kritéria větší než kvantil $\chi^2_{1-\alpha}$ rozdělení s $k - h - 1$ stupni volnosti, kde k je počet skupin a h je počet odhadovaných parametrů teoretického rozdělení, **zamítáme nulovou hypotézu** o shodě rozdělení.

Testy dobré shody

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Je-li teoretické **rozdělení diskrétní**, pak:

1. Stanovíme nulovou a alternativní hypotézu,
2. Označíme pozorované varianty $NV\ x_i$, kde $i = 1, 2, \dots, k$,
3. Určíme pozorované četnosti O_i jednotlivých variant NV $\left(\sum_{i=1}^{k} O_i = n\right)$,
4. Určíme očekávané relativní četnosti π_i jednotlivých variant NV $\left(\pi_i = P_0(x_i)\right)$,
5. Určíme očekávané četnosti E_i jednotlivých variant NV $(E_i = n\pi_i)$,
6. Ověříme předpoklady testu (očekávané četnosti větší než 2, alespoň 80% očekávaných četností větší než 5),
7. Určíme pozorovanou hodnotu testové statistiky $x_{obs} = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$,
8. Pokud je hodnota testového kritéria větší než kvantil $\chi^2_{1-\alpha}$ rozdělení s $k - h - 1$ stupni volnosti, kde k je počet skupin a h je počet odhadovaných parametrů teoretického rozdělení, **zamítáme nulovou hypotézu** o shodě rozdělení,
9. Rozhodneme o výsledku testu.

Příklad 2 - test dobré shody (diskrétní rozdělení)

Firma MODEL odhaduje počet poruch určitého zařízení během dne pomocí Poissonova rozdělení se střední hodnotou 1,2. Zaměstnanci zaznamenali pro kontrolu skutečné počty poruch celkem ve 150 dnech (výsledky jsou uvedeny v níže uvedené tabulce). Ověřte, zda lze počet poruch daného zařízení během dne skutečně modelovat pomocí Poissonova rozdělení s parametrem $\lambda = 1{,}2$. Hladina významnosti je 5%.

x_i – počet poruch během dne	0	1	2	3	4 a více	celkem
O_i – počet dní, v nichž byl pozorován počet poruch x_i	52	48	36	10	4	150

H_0: Počet poruch daného zařízení během jednoho dne (náhodná veličina X) **má Poissonovo rozdělení** s parametrem $\lambda = 1{,}2$, neboli výběr pochází z Poissonova rozdělení s parametrem $\lambda = 1{,}2$.

H_A: $\neg H_0$.

Příklad 2 - test dobré shody (diskrétní rozdělení)

Firma MODEL odhaduje počet poruch určitého zařízení během dne pomocí Poissonova rozdělení se střední hodnotou 1,2. Zaměstnanci zaznamenali pro kontrolu skutečné počty poruch celkem ve 150 dnech (výsledky jsou uvedeny v níže uvedené tabulce). Ověřte, zda lze počet poruch daného zařízení během dne skutečně modelovat pomocí Poissonova rozdělení s parametrem $\lambda = 1{,}2$. Hladina významnosti je 5%.

x_i – počet poruch během dne	0	1	2	3	4 a více	celkem
O_i – počet dní, v nichž byl pozorován počet poruch x_i	52	48	36	10	4	150
π_i - očekávané pravděpodobnosti výskytu	$P_0(0)$	$P_0(1)$	$P_0(2)$	$P_0(3)$	$P_0(X \geq 4)$	1,000

$P_0(x)$ je pravděpodobnostní funkce Poissonova rozdělení se střední hodnotou 1,2.

Příklad 2 - test dobré shody (diskrétní rozdělení)

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Firma MODEL odhaduje počet poruch určitého zařízení během dne pomocí Poissonova rozdělení se střední hodnotou 1,2. Zaměstnanci zaznamenali pro kontrolu skutečné počty poruch celkem ve 150 dnech (výsledky jsou uvedeny v níže uvedené tabulce). Ověřte, zda lze počet poruch daného zařízení během dne skutečně modelovat pomocí Poissonova rozdělení s parametrem $\lambda = 1{,}2$. Hladina významnosti je 5%.

x_i – počet poruch během dne	0	1	2	3	4 a více	celkem
O_i – počet dní, v nichž byl pozorován počet poruch x_i	52	48	36	10	4	150
π_i - očekávané pravděpodobnosti výskytu	0,301	0,361	0,217	0,087	0,034	1,000
E_i - očekávané četnosti výskytu	45,2	54,2	32,6	13,1	5,1	---

$$x_{obs} = \sum_{i=1}^{5} \frac{(O_i - E_i)^2}{E_i} = \frac{(52 - 45{,}2)^2}{45{,}2} + \cdots + \frac{(4 - 5{,}1)^2}{5{,}1} = 3{,}13$$

Testové kritérium má za předpokladu platnosti H_0 χ^2 rozdělení s 3 $(k - 1 - h)$ stupni volnosti (počet variant $k = 5$, počet odhadovaných parametrů $h = 1$).

Příklad 2 - test dobré shody (diskrétní rozdělení)

Firma MODEL odhaduje počet poruch určitého zařízení během dne pomocí Poissonova rozdělení se střední hodnotou 1,2. Zaměstnanci zaznamenali pro kontrolu skutečné počty poruch celkem ve 150 dnech (výsledky jsou uvedeny v níže uvedené tabulce). Ověřte, zda lze počet poruch daného zařízení během dne skutečně modelovat pomocí Poissonova rozdělení s parametrem $\lambda = 1,2$. Hladina významnosti je 5%.

x_i – počet poruch během dne	0	1	2	3	4 a více	celkem
O_i – počet dní, v nichž byl pozorován počet poruch x_i	52	48	36	10	4	150
π_i - očekávané pravděpodobnosti výskytu	0,301	0,361	0,217	0,087	0,034	1,000
E_i - očekávané četnosti výskytu	45,2	54,2	32,6	13,1	5,1	---

Kvantil $\chi^2_{1-\alpha}$ rozdělení se 3 stupni volnosti je roven 7,815 (viz tabulky).

Protože je hodnota testového kritéria menší než hodnota kvantilu, na hladině významnosti 0,05 **nezamítáme H_0**, tj. tzn. nemáme námitek proti použití Poissonova rozdělení s parametrem 1,2 pro odhad počtu poruch daného zařízení během jednoho dne.

Testy dobré shody

Je-li teoretické **rozdělení spojité**, pak:

1. Stanovíme nulovou a alternativní hypotézu,
2. pozorované hodnoty NV roztřídíme do k tříd (obvykle volíme $5 \leq k \leq 15$),
3. Určíme pozorované četnosti O_i v jednotlivých třídách $\left(\sum_{i=1}^{k} O_i = n\right)$,
4. Určíme očekávané relativní četnosti π_i v jednotlivých třídách,
5. Určíme očekávané četnosti E_i v jednotlivých třídách $(E_i = n\pi_i)$,
6. Ověříme předpoklady testu (očekávané četnosti větší než 2, alespoň 80%očekávaných četností větší než 5),
7. Není-li předpoklad testu splněn, pokusíme se třídy korigovat (nemusí být stejně velké),
8. Určíme pozorovanou hodnotu testové statistiky $x_{obs} = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$,
9. Pokud je hodnota testového kritéria větší než kvantil $\chi^2_{1-\alpha}$ rozdělení s $k - h - 1$ stupni volnosti, kde k je počet skupin a h je počet odhadovaných parametrů teoretického rozdělení, **zamítáme nulovou hypotézu** o shodě rozdělení,
10. Rozhodneme o výsledku testu.

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Příklad 3 - test dobré shody (spojité rozdělení)

Na křižovatce v Opavě byly v průběhu několika minut měřeny časové odstupy v sekundách mezi průjezdy jednotlivých vozidel. Zjištěné hodnoty těchto odstupů jsou uvedeny v tabulce. Ověřte, zda lze časové odstupy mezi vozidly modelovat pomocí náhodné veličiny s normálním rozdělením. Hladina významnosti je 5%.

2,5	6,8	5,0	9,8	4,0	2,3	4,2	19	8,7	7,7	5,9	5,3	8,4	3,6	9,2
4,3	2,6	13,0	5,4	8,6	4,2	2,9	1,5	1,8	1,6	5,9	8,3	5,2	6,9	5,1
1,3	6,4	6,5	5,7	3,6	4,8	4,0	7,3	24,9	10,6	15,0	5,3	4,0	3,3	6,0
4,6	1,6	1,9	1,5	11,1	4,3	5,5	2,1	2,9	3,0	3,8	1,0	1,5	8,6	4,4
6,8	5,2	3,0	8,0	4,0	4,7	7,3	2,3	1,9	1,9	4,6	6,4	5,3	3,9	2,4
1,2	6,2	4,3	2,6	2,7	2,0	0,8	3,7	6,9	2,8	4,3	4,9	4,1	4,5	4,4
11,9	9,0	5,6	4,8	2,8	2,1	4,3	1,0	1,6	2,5	2,2	1,3	1,8	1,6	3,8
3,1	1,6	4,9	1,8	3,9	3,4	1,6	4,5	5,8	6,9	1,8	2,6	6,8	2,5	1,9
3,1	10,8	1,6	2,0	4,9	11,2	1,6	2,2	3,8	1,1	1,8	1,4			

X … časový odstup mezi průjezdy jednotlivých vozidel.

H_0: Časové odstupy mezi průjezdy jednotlivých vozidel **mají normální rozdělení**.
H_A: Časové odstupy mezi průjezdy jednotlivých vozidel **nemají normální rozdělení**.

Příklad 3 - test dobré shody (spojité rozdělení)

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Na křižovatce v Opavě byly v průběhu několika minut měřeny časové odstupy v sekundách mezi průjezdy jednotlivých vozidel. Zjištěné hodnoty těchto odstupů jsou uvedeny v tabulce. Ověřte, zda lze časové odstupy mezi vozidly modelovat pomocí náhodné veličiny s normálním rozdělením. Hladina významnosti je 5%.

X … časový odstup mezi průjezdy jednotlivých vozidel.

Neúplně specifikovaný test – normální rozdělení má 2 parametry (μ, σ^2), které je třeba odhadnout.

$$\hat{\mu} = \bar{x} = 4{,}6 \qquad \hat{\sigma}^2 = s^2 = 10{,}9$$

Proč musíme nyní pozorované hodnoty kategorizovat?

Pravděpodobnostní funkce spojité NV je nulová. Pravděpodobnost výskytu spojité NV na určitém intervalu již nulová není!

Příklad 3 - test dobré shody (spojité rozdělení)

Na křižovatce v Opavě byly v průběhu několika minut měřeny časové odstupy v sekundách mezi průjezdy jednotlivých vozidel. Zjištěné hodnoty těchto odstupů jsou uvedeny v tabulce. Ověřte, zda lze časové odstupy mezi vozidly modelovat pomocí náhodné veličiny s normálním rozdělením. Hladina významnosti je 5%.

i	Třídící interval [s]	Empirické četnosti O_i	Očekávané pravd. π_{0_i}
1	$(-\infty; 1{,}5\rangle$	11	0,174
2	$(1{,}5; 1{,}8\rangle$	13	0,024
3	$(1{,}8; 2{,}0\rangle$	7	0,017
4	$(2{,}0; 2{,}5\rangle$	10	0,047
5	$(2{,}5; 2{,}9\rangle$	8	0,041
6	$(2{,}9; 3{,}6\rangle$	8	0,078
7	$(3{,}6; 4{,}0\rangle$	10	0,047
8	$(4{,}0; 4{,}4\rangle$	10	0,048
9	$(4{,}4; 4{,}9\rangle$	10	0,060
10	$(4{,}9; 5{,}8\rangle$	12	0,106
11	$(5{,}8; 6{,}8\rangle$	10	0,106
12	$(6{,}8; 8{,}7\rangle$	12	0,145
13	$(8{,}7; \infty)$	11	0,107
Celkem	-	132	1,000

$P(X \in (-\infty;\ 1{,}5\rangle) = P(X < 1{,}5)$ $= F(1{,}5)$, kde $F(x)$ je distribuční funkce rozdělení $N(4{,}6; 10{,}9)$

Příklad 3 - test dobré shody (spojité rozdělení)

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Na křižovatce v Opavě byly v průběhu několika minut měřeny časové odstupy v sekundách mezi průjezdy jednotlivých vozidel. Zjištěné hodnoty těchto odstupů jsou uvedeny v tabulce. Ověřte, zda lze časové odstupy mezi vozidly modelovat pomocí náhodné veličiny s normálním rozdělením. Hladina významnosti je 5%.

i	Třídící interval [s]	Empirické četnosti O_i	Očekávané pravd. π_{0_i}	Očekávané četnosti E_i
1	$(-\infty; 1,5\rangle$	11	0,174	22,9
2	$(1,5; 1,8\rangle$	13	0,024	3,2
3	$(1,8; 2,0\rangle$	7	0,017	2,3
4	$(2,0; 2,5\rangle$	10	0,047	6,2
5	$(2,5; 2,9\rangle$	8	0,041	5,4
6	$(2,9; 3,6\rangle$	8	0,078	10,3
7	$(3,6; 4,0\rangle$	10	0,047	6,2
8	$(4,0; 4,4\rangle$	10	0,048	6,3
9	$(4,4; 4,9\rangle$	10	0,060	8,0
10	$(4,9; 5,8\rangle$	12	0,106	14,0
11	$(5,8; 6,8\rangle$	10	0,106	13,9
12	$(6,8; 8,7\rangle$	12	0,145	19,2
13	$(8,7; \infty)$	11	0,107	14,1
Celkem	-	132	1,000	-

Všechny očekávané četnosti jsou větší než 2, **sloučíme-li třídy 2 a 3,** bude splněn i silnější předpoklad – očekávané četnosti budou větší než 5.

Příklad 3 - test dobré shody (spojité rozdělení)

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Na křižovatce v Opavě byly v průběhu několika minut měřeny časové odstupy v sekundách mezi průjezdy jednotlivých vozidel. Zjištěné hodnoty těchto odstupů jsou uvedeny v tabulce. Ověřte, zda lze časové odstupy mezi vozidly modelovat pomocí náhodné veličiny s normálním rozdělením. Hladina významnosti je 5%.

i	Třídící interval [s]	Empirické četnosti O_i	Očekávané pravd. π_{0_i}	Očekávané četnosti E_i
1	$(-\infty; 1{,}5\rangle$	11	0,174	22,9
2	$(1{,}5; 2{,}0\rangle$	20	0,041	5,5
3	$(2{,}0; 2{,}5\rangle$	10	0,047	6,2
4	$(2{,}5; 2{,}9\rangle$	8	0,041	5,4
5	$(2{,}9; 3{,}6\rangle$	8	0,078	10,3
6	$(3{,}6; 4{,}0\rangle$	10	0,047	6,2
7	$(4{,}0; 4{,}4\rangle$	10	0,048	6,3
8	$(4{,}4; 4{,}9\rangle$	10	0,060	8,0
9	$(4{,}9; 5{,}8\rangle$	12	0,106	14,0
10	$(5{,}8; 6{,}8\rangle$	10	0,106	13,9
11	$(6{,}8; 8{,}7\rangle$	12	0,145	19,2
12	$(8{,}7; \infty)$	11	0,107	14,1
Celkem	-	132	1,000	-

$$x_{obs} = \sum_{i=1}^{12} \frac{(O_i - E_i)^2}{E_i} = \frac{(11-22{,}9)^2}{22{,}9} + \cdots + \frac{(11-14{,}1)^2}{14{,}1} = 59{,}7$$

Příklad 3 - test dobré shody (spojité rozdělení)

Na křižovatce v Opavě byly v průběhu několika minut měřeny časové odstupy v sekundách mezi průjezdy jednotlivých vozidel. Zjištěné hodnoty těchto odstupů jsou uvedeny v tabulce. Ověřte, zda lze časové odstupy mezi vozidly modelovat pomocí náhodné veličiny s normálním rozdělením. Hladina významnosti je 5%.

Testové kritérium má za předpokladu platnosti H_0 χ^2 rozdělení s 9 $(12-1-2)$ stupni volnosti (počet tříd $k=12$, počet odhadovaných parametrů $h=2$).

Kvantil $\chi^2_{1-\alpha}$ rozdělení s 9 stupni volnosti je roven 16,92 (viz tabulky).

Protože je hodnota testového kritéria větší než hodnota kvantilu, na hladině významnosti 0,05 **zamítáme nulovou hypotézu**, tj. nelze předpokládat, že časové odstupy mezi průjezdy jednotlivých vozidel nemají normální rozdělení.

Literatura

1. Janáček J. *Statistika jednoduše*. Praha: Grada, 2022. **(kapitola 9).**

2. Ramík J. a Čemerková Š. *Statistika A*. Opava, Karviná: SLU, 2000. **(kapitola 9).**

SLEZSKÁ UNIVERZITA
FAKULTA VEŘEJNÝCH POLITIK V OPAVĚ

Děkuji za pozornost.