11 TESTOVÁNÍ HYPOTÉZ

Čas ke studiu kapitoly: 360 minut

Cíl

Po prostudování tohoto odstavce budete:

- znát základní pojmy a principy testování hypotéz
- znát koncepci klasického testu
- umět rozhodovat pomocí čistého testu významnosti
- umět posoudit chybu při rozhodování
- umět zkonstruovat operativní charakteristiku
- umět používat základní jednovýběrové a dvouvýběrové parametrické testy pro normální rozdělení (z-test, t-test, test relativní četností, test rozptylu + totéž pro dva výběry)
- umět rozhodovat podle párového testu
- umět používat vybrané neparamerické testy (jednovýběrové a dvouvýběrové testy o mediánu (znaménkový, Wilcoxonův), testy o shodě (χ^2-test dobré shody, jednovýběrový a dvouvýběrový Kolmogorovův a Smirnovův test), test závislosti v kombinační tabulce)

Výklad:

11.1 Úvod

Již víme, že pomocí statistické indukce můžeme učinit závěry o populaci na základě výběrového souboru z této populace. V předcházející kapitole jsme se zabývali problémem, jak odhadnout prostřednictvím bodového, popř. intervalového odhadu neznámý parametr populace. V této kapitole budeme konstruovat testy, s jejichž pomocí potvrdíme nebo vyvrátíme nějakou hypotézu o populaci.

Statistické hypotézy (hypotézy o základním souboru (populaci)) můžeme rozdělit do dvou skupin – a to na hypotézy parametrické a hypotézy neparametrické.

Parametrické hypotézy jsou hypotézy o parametrech rozdělení (populace). Můžeme se setkat se třemi typy těchto hypotéz:

a) Hypotézy o parametru jedné populace (o střední hodnotě, mediánu, rozptylu, relativní četnosti…)
b) Hypotézy o parametrech dvou populací (srovnávací testy)
c) Hypotézy o parametrech více než dvou populací (ANOVA …)

Parametrické hypotézy můžeme zapsat jako rovnosti (resp. nerovnosti) mezi testovaným parametrem a jeho předpokládanou hodnotou (např. „$\mu = 100$“, „$\pi \leq 0{,}08$“) nebo jako rovnosti (resp. nerovnosti) mezi testovanými parametry (např. „$\mu_1 = \mu_2 = \mu_3$“, „$\pi_1 > \pi_2$“).
Statistickým hypotézám o jiných vlastnostech populace (tvar rozdělení, závislost proměnných…) se říká **neparametrické hypotézy.**

POZOR!!!

Parametrické testy se říká testům, k jejichž odvození je nutné pro daný výběr specifikovat typ rozdělení (v některých případech i některé parametry tohoto rozdělení). (Nejde tedy obecně o libovolné testy parametrických hypotéz.)

Neparametrické testy se říká testům, k jejichž odvození není nutné pro daný výběr specifikovat typ rozdělení.

11.2 Nulová a alternativní hypotéza

Testováním statistických hypotéz se statistici začali zabývat krátce před vypuknutím druhé světové války. Jeho koncepci vytvořili Jerzy Neyman a E. S. Pearson a dále ji pak rozvinul Abraham Wald.

Testování hypotéz pojali jako rozhodovací proces, v němž proti sobě stojí dvě tvrzení (hypotézy). První z nich – **nulová hypotéza H_0** – představuje určitý rovnovážný stav a bývá vyjádřena rovnosti „=“ (např. $\mu = 100$, $\mu_1 = \mu_2 = \mu_3$ …). Jde o takové tvrzení o populaci, které je bráno jak předpoklad při testování. Oproti ní stavíme tzv. **alternativní hypotézu H_A** . Alternativní hypotéza přestavuje porušení rovnovážného stavu a zapisujeme ji tedy jedním ze tří možných zápisů nerovnosti ($\neq$, <, >). Zvolíme-li alternativní hypotézu ve tvaru „<” nebo „>“, mluvíme o **jednostranné alternativní hypotéze** (např. $\mu < 100$, $\mu > 100$), zvolíme-li alternativní hypotézu ve tvaru „$\neq$”, mluvíme o **oboustranné alternativní hypotéze**.

11.2.1 Výběr vhodné alternativní hypotézy

Při testování hypotéz musíme vždy stanovit jak nulovou, tak i alternativní hypotézu. Nulová hypotéza bývá stanovena jednoznačně (pomocí rovnosti, např. $\mu = 100$). Pro stanovení alternativní hypotézy máme tři možnosti. (např. $\mu < 100$, $\mu > 100$, $\mu \neq 100$). Obsahuje-li zadání problému vedoucího na testování hypotéz vztah jednostranné nerovnosti, volí se jako alternativní hypotéza příslušná jednostranná hypotéza. V ostatních případech volíme oboustrannou alternativní hypotézu. Alternativní hypotéza by měla být v souladu s výběrovým souborem. Pokud tomu tak není, přizpůsobujeme alternativní hypotézu závěrům získaným z výběrového souboru.

Průvodce studiem:

Následující příklady statistických hypotéz by Vám měly pomoci ujasnit si probranou terminologii používanou při testování hypotéz:

1. *Průměrný plat v ČR je 20.200,- Kč.*
   ***Hypotéza:** parametrická, o střední hodnotě*
   ***Populace** (základní soubor): všichni pracující občané ČR → jejich platy*
   ***H_0**: $\mu = 20.200$*
   ***H_A**: $\mu \neq 20.200$ (zadání problému neobsahuje jednostrannou nerovnost)*
   ***Výběrový soubor**: Na průměrný plat zjištěný z výběrového souboru nemáme zvláštní požadavky*

2. *Podpora ODS je vyšší než podpora ČSSD (listopad 2006)*
   ***Hypotéza**: parametrická, srovnání relativních četností dvou populací*
   ***Populace 1**: všichni voliči v ČR → relativní četnost voličů ODS*
   ***Populace 2**: všichni voliči v ČR → relativní četnost voličů ČSSD*
   ***H_0**: $\pi_{ODS} = \pi_{ČSSD}$ ($\pi_1 = \pi_2$)*
   ***H_A**: $\pi_{ODS} > \pi_{ČSSD}$ ($\pi_1 > \pi_2$) (zadání problému obsahuje nerovnost v tomto tvaru)*
   ***Výběrový soubor**: Procentuální zastoupení voličů ODS ve výběru by mělo být větší než procentuální zastoupení voličů ČSSD ve výběru. Pokud tomu tak není, měli bychom použít oboustrannou alternativní hypotézu.*

3. *Mzdy ve strojírenství jsou nižší než mzdy v bankovnictví*
   ***Hypotéza**: parametrická, srovnání středních hodnot dvou populací*
   ***Populace 1**: všichni zaměstnanci ve strojírenství → jejich platy*
   ***Populace 2**: všichni zaměstnanci v bankovnictví → jejich platy*
   ***H_0**: $\mu_{strojírenství} = \mu_{bankovnictví}$ ($\mu_1 = \mu_2$)*

***H*$_A$:** $\mu_{strojírenství} < \mu_{bankovnictví}$ ($\mu_1 < \mu_2$)

(zadání problému obsahuje nerovnost v tomto tvaru)

***Výběrový soubor**: Průměrný plat zjištěný z výběru zaměstnanců ve strojírenství by měl být menší než průměrný plat zjištěný z výběru zaměstnanců v bankovnictví. Pokud tomu tak není, měli bychom použít oboustrannou alternativní hypotézu.*

4. *a) Použití bezpečnostních pásů ovlivňuje úmrtnost při dopravních nehodách*
   b) Použití bezpečnostních pásů snižuje úmrtnost při dopravních nehodách
   ***Hypotéza**: parametrická, srovnání relativních četností dvou populací*
   ***Populace 1**: účastníci dopravních nehod sedící na místech, na nichž je možno používat bezpečnostní pásy – ti, kteří byli připoutáni $\rightarrow$ úmrtnost (v procentech)*
   ***Populace 2**: účastníci dopravních nehod sedící na místech, na nichž je možno používat bezpečnostní pásy – ti, kteří nebyli připoutáni $\rightarrow$ úmrtnost (v procentech)*
   ***H*$_0$:** $\pi_1 = \pi_2$
   ada) ***H*$_A$:** $\pi_1 \neq \pi_2$ *(zadání problému neobsahuje nerovnost)*
   adb) ***H*$_A$:** $\pi_1 < \pi_2$ *(zadání problému obsahuje nerovnost v tomto tvaru)*
   ***Výběrový soubor**: Úmrtnost těch co používají bezpečnostní pásy by měla být menší než úmrtnost těch, co bezpečnostní pásy nepoužívají (ve výběru z účastníku dopravních nehod). Pokud tomu tak není, měli bychom použít oboustrannou alternativní hypotézu.*

5. *Dosažené vzdělání závisí na dosaženém vzdělání otce*
   ***Hypotéza**: neparametrická, testování závislosti proměnných*
   ***Kategoriální proměnná 1**: všichni žijící lidé s ukončeným vzděláním $\rightarrow$ jejich dosažené vzdělání*
   ***Kategoriální proměnná 2**: všichni žijící lidé s ukončeným vzděláním $\rightarrow$ dosažené vzdělání jejich otců*
   ***H*$_0$**: *Dosažené vzdělání nezávisí na dosaženém vzdělání otce („závislost je nulová“)*
   ***H*$_A$**: *Dosažené vzdělání závisí na dosaženém vzdělání otce*

Výklad:

11.3 Chyba I. a II. druhu

Jelikož při rozhodování o nulové hypotéze vycházíme z výběrového souboru, který nemusí dostatečně přesně odpovídat vlastnostem základního souboru, můžeme se při rozhodování dopustit chyby. Při rozhodování mohou nastat situace, které popisuje následující tabulka:

		Výsledek testu	
		Nezamítáme H$_0$	**Zamítáme H$_0$**
Skutečnost	**Platí H$_0$**	Správné rozhodnutí Pravděpodobnost rozhodnutí: $1-\alpha$ (spolehlivost)	**Chyba I. druhu** Pravděpodobnost rozhodnutí: α (hladina významnosti)
	Platí H$_A$	**Chyba II. druhu** Pravděpodobnost rozhodnutí: β	Správné rozhodnutí Pravděpodobnost rozhodnutí: $1-\beta$ (síla testu)

Jestliže nulová hypotéza je ve skutečnosti platná a my ji přesto zamítneme, dopouštíme se **chyby I. druhu**. Pravděpodobnost, že k takovému pochybení dojde nazýváme **hladina významnosti** a označujeme ji α. Platí-li nulová hypotéza a my jsme ji nezamítli, rozhodli jsme správně. Pravděpodobnost tohoto rozhodnutí označujeme $(1-\alpha)$ a nazýváme ji **spolehlivost**. Správným rozhodnutím je rovněž zamítnutí nulové hypotézy v případě, že je platná hypotéza alternativní. Tohoto rozhodnutí se dopouštíme s pravděpodobností $(1-\beta)$, což bývá označováno jako **síla testu**. **Chybou II. druhu** je nezamítnutí nulové hypotézy v případě, že je platná hypotéza alternativní. Pravděpodobnost této chyby je β.

Při testování hypotéz se samozřejmě snažíme minimalizovat obě chyby, tj. dosáhnout vysoké síly testu (nízkého β) při co nejnižší hladině významnosti α. To však není možné, neboť snížením β se zvýší hladina významnosti α a naopak. (Můžeme si obě chyby představit jako na houpačce.) Proto je třeba najít kompromis mezi požadavky na α a β.

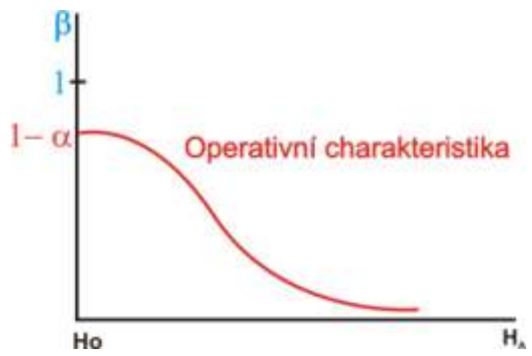
Ve statistice se volí jako rozhodující vstupní parametr testu pravděpodobnost chyby I. druhu – hladina významnosti α. V technických oblastech volíme obvykle 5%-ní nebo 1%-ní hladinu významnosti, pouze ve speciálních případech (lékařské účely) požadavek na pravděpodobnost chyby I. druhu ještě zvyšujeme (volíme ještě nižší α).

Chybu II. druhu snižujeme volbou vhodného testu (pokud máme možnost výběru) popřípadě zvětšením rozsahu výběrového souboru (což je jediný způsob jak snížit β, aniž bychom tím zvýšili α).

11.4 Operativní charakteristika

Pravděpodobnost chyby II. druhu (β, tj. pravděpodobnost, že nezamítneme nulovou hypotézu, přestože je alternativní hypotéza pravdivá) závisí na přesné hodnotě alternativní hypotézy. Dokážeme tedy určit β pro případ, že alternativní hypotéza je přesně specifikovaná. (např. testujeme-li hypotézu, že průměrný plat v ČR je 20.200,- Kč, umíme určit β pro případ, že alternativa je definována ve formě: průměrný plat v ČR je 20.350,- Kč, apod.)

V inženýrských aplikacích se mnohdy setkáváme s tzv. **operativní charakteristikou**, což je závislost pravděpodobnosti chyby II. druhu na přesné specifikaci alternativní hypotézy. Schématické znázornění operativní charakteristiky přináší následující obrázek:



Z obrázku je zřejmé, že vzdaluje-li se alternativa od nulové hypotézy, pravděpodobnost chyby II. druhu (β) klesá.

Místo operativní charakteristiky se mnohdy znázorňuje **křivka síly testu**, tj. závislost síly testu (1-β) na přesné specifikaci alternativní hypotézy (zkráceně se mnohdy označuje pouze jako síla testu (power curve).

Průvodce studiem:

V tomto průvodci se pokusíme o odpovědi na často pokládané otázky.

Proč nepoužíváme pojem „přijímáme nulovou hypotézu"

Testování hypotéz se může provádět různými způsoby. Při každém z nich může být testována hypotéza zamítnuta. Nezamítneme-li ji, znamená to, že prováděným testem jsme ji nemohli zamítnout, nikoliv to, že je správná. Je možné, že nějakým testem se ji zamítnout podaří. Pokud používáme stále přesnější testy a stále docházíme ke stejnému závěru o nezamítnutí nulové hypotézy, můžeme jednat tak, jako by nulová hypotéza byla správná. Nikdy to však nevíme jistě.

Je souvislost mezi testováním parametrických hypotéz a intervalovými odhady?

Ano, pokusme se tuto souvislost objasnit: Spolehlivost testu (1-α), tj. pravděpodobnost, že nezamítneme nulovou hypotézu v případě, že je skutečně platná označuje rovněž pravděpodobnost, že parametr populace leží v příslušném intervalu spolehlivosti. Je tedy zřejmé, že pokud testovaná hodnota parametru leží uvnitř (1-α) intervalu spolehlivosti, můžeme příslušnou nulovou hypotézu nezamítnout na hladině významnosti α. Interval spolehlivosti lze považovat za množinu všech možných (nezamítnutelných) hypotéz.

Příklad:
Vzpomínáte si na řešený příklad o kvalitě disket Sonik a 5M? Zjistili jsme v něm, že rozdíl mezi procentem vadných disket Sonik a 5M leží v intervalu (-1,0%; 2,4%) s 95%-ní spolehlivosti. Chtěli-li bychom testovat, zda diskety Sonik jsou kvalitnější než diskety 5M, mohli bychom (s využitím intervalového odhadu) postupovat takto:

1. *Stanovíme nulovou a alternativní hypotézu:*

H_0: $\pi_{Sonik} = \pi_{5M} \quad (\pi_{Sonik} - \pi_{5M} = 0)$
H_A: $\pi_{Sonik} < \pi_{5M} \quad (\pi_{Sonik} - \pi_{5M} < 0)$

2. *Určíme 95%-ní interval spolehlivosti pro* $(\pi_{Sonik} - \pi_{5M})$

$$P(-1{,}0\ \% < (\pi_{Sonik} - \pi_{5M}) < 2{,}4\ \%) = 0{,}95$$

3. *Určíme, zda testovaná hodnota parametru (v našem případě testovaná hodnota rozdílu parametrů – „0") leží v příslušném intervalu spolehlivosti.*

$$0 \in (-1{,}0\ \%;\ 2{,}4\ \%)$$

4. *Závěr: S 95%-ní spolehlivosti můžeme tvrdit, že kvalita disket Sonik a 5M je stejná (nezamítáme nulovou hypotézu).*

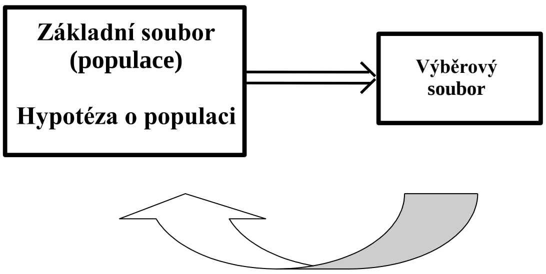
Proč je chyba I. druhu významnější než chyba II. druhu?

V následujícím textu budeme přirovnávat testování hypotéz k principu presumpce neviny. V USA je v soudní praxi při procesech s vrahy pravidlem, že porota rozhoduje o vině obžalovaného. Jde v podstatě o rozhodnutí mezi nulovou hypotézou (nevinen) a alternativní hypotézou (vinen). Chybou I. druhu by bylo uznání obžalovaného vinným, přestože by byl nevinen – došlo by k justičnímu omylu, byl by odsouzen nevinný člověk. Chybou II. druhu by pak bylo osvobození skutečného vraha. Porota se při svém verdiktu musí řídit principem presumpce neviny – "vina musí být prokázána nade vší pochybnost", tzn. minimalizuje chybu I. druhu. Stejně přistupuje k testování hypotéz statistika.

Výklad:

11.5 Princip testování hypotéz

Princip testování hypotéz se dá přirovnat k principu presumpce neviny v soudnictví [Friedrich: Statistika 1, ZČU, Plzeň]. Pokud výběrový soubor ($\underline{X}$) neukáže na (statisticky významný) rozpor s nulovou hypotézou, pak nesmíme nulovou hypotézu zamítnout – podobně jako princip presumpce neviny požaduje, abychom na obžalovaného pohlíželi jako na nevinného do té doby, dokud nepředložíme přesvědčivé důkazy o jeho vině. Statisticky test pak můžeme přirovnat k soudci.



Jsou data konzistentní s hypotézou o populaci ?

Statistický test rozhodne, zda data z výběrového souboru ($\underline{X}$) odpovídají nulové hypotéze. Převedeno do jazyku soudnictví: Soudce rozhodne, zda svědci podali výpověď ve prospěch obhajoby.

Při testování hypotéz se běžně můžeme setkat se dvěma přístupy – klasickým testem a čistým testem významnosti. My se seznámíme obecně s oběma postupy a v dalším textu se pak zaměříme na čistý test významnosti.

11.5.1 Klasický test

Klasický test se skládá z několika kroků:

1. **Formulace nulové a alternativní hypotézy**

2. **Volba testové statistiky (testového kritéria) T(X̲)** – jde o funkci výběru, která vyjadřuje sílu platnosti nulové hypotézy ve srovnání s hypotézou alternativní. Pro další krok testu musíme znát rovněž rozdělení testové statistiky při platnosti H_0 (**nulové rozdělení) $\mathbf{F_0(x)}$** - $F_0(x) = P(T(\underline{X}) < x \mid H_0)$

3. **Sestrojení kritického oboru a oboru přijetí** – jde o rozdělení prostoru všech možných hodnot testové statistiky (S) na dva podprostory: **obor přijetí** (A) obsahující hodnoty testové statistiky svědčící pro přijetí nulové hypotézy a **kritický obor** (C) obsahující hodnoty svědčící pro zamítnutí nulové hypotézy. Je zřejmé, že $A \cup C = S$; $A \cap C = \emptyset$. Hranice mezi kritickým oborem a oborem přijetí se nazývá **kritická hodnota testu.**

   Konstrukce kritického oboru: Kritický obor bude tak velký, aby pravděpodobnost, že testová statistika leží v kritickém oboru za předpokladu platnosti nulové hypotézy, byla rovna hladině významnosti α.

   $$P\left(T\left(\underline{X}\right) \in C \middle| H_0\right) = \alpha$$

   Jinými slovy: Pravděpodobnost, že hodnota testové statistiky bude ležet v oblasti svědčící pro zamítnutí nulové hypotézy, přestože je nulová hypotéza platná, má být rovna předem zvolené hodnotě α. Jazykem soudnictví: Svědci (výběr) podají falešné svědectví v neprospěch obhajoby (nulové hypotézy) s pravděpodobností α (tady se projevuje rozpor mezi principem testování hypotéz a principem presumpce neviny – soudce nemůže α stanovit a ani jej pro konkrétní případ nezná).

   Známe-li nulové rozdělení testové statistiky T(X̲), není obtížné pro dané α stanovit kritický obor:

   a) Je-li **alternativní hypotéza** ve tvaru **„<“** (ve prospěch alternativy svědčí nízké hodnoty testové statistiky), pak je kritický obor vymezen jako:

   $$\mathbf{C \leq T_{\alpha}}$$

   b) Je-li **alternativní hypotéza** ve tvaru **„>“** (ve prospěch alternativy svědčí vysoké hodnoty testové statistiky), pak je kritický obor vymezen jako:

   $$\mathbf{C \geq T_{1-\alpha}}$$

   c) Je-li **alternativní hypotéza** ve tvaru **„≠“** (ve prospěch alternativy svědčí extrémně nízké nebo extrémně vysoké hodnoty testové statistiky), pak je kritický obor vymezen jako:

   $$\left(C \leq T_{\frac{\alpha}{2}}\right) \vee \left(C \geq T_{1-\frac{\alpha}{2}}\right)$$

4. **Výpočet pozorované hodnoty testové statistiky T(X̲) - x_{OBS}**

   Předcházející kroky jsme mohli podniknout v rámci přípravy testu, nyní již musíme mít k dispozici výběrový soubor a pomocí něj určit konkrétní hodnotu testové statistiky T(X̲) (x_{OBS}). Při tomto výpočtu předpokládáme platnost nulové hypotézy.

5. **Formulace závěru testu** – každý test vede ke dvěma možným výsledkům:

   a) Leží-li testová statistika v kritickém oboru ($x_{OBS} \in C$), pak **zamítáme nulovou hypotézu ve prospěch alternativní hypotézy**

   b) Leží-li testová statistika v oboru přijetí (tzn. neleží v kritickém oboru - $x_{OBS} \notin C$), pak **nulovou hypotézu nezamítáme.**

11.5.2 Čistý test významnosti

Čistý test významnosti zodpovídá otázku, zda získaný náhodný výběr X̲ je či není extrémní s ohledem na nějakou testovanou hypotézu o populaci (zda zjištěné údaje podporují nulovou hypotézu). Oproti klasickému testu nepotřebuje čistý test významnosti znát hladinu významnosti jako vstupní údaj. Jeho výsledek nám umožňuje rozhodnout na jakých hladinách významnosti můžeme nulovou hypotézu zamítnout (resp. nezamítnout).

Čistý test významnosti se skládá z následujících kroků (první dva kroky se shodují s klasickým testem):

1. **Formulace nulové a alternativní hypotézy**

2. **Volba testové statistiky (testového kritéria) T(X̲)** – jde o funkci výběru, která vyjadřuje sílu platnosti nulové hypotézy ve srovnání s hypotézou alternativní. Pro další krok testu musíme znát rovněž rozdělení testové statistiky při platnosti H_0 (**nulové rozdělení) $F_0(x)$ -**
   $F_0(x) = P(T(\underline{X}) < x \mid H_0)$

3. **Výpočet pozorované hodnoty testové statistiky x_{OBS} a výpočet statistiky p-value (p-hodnota)**

   Výpočet pozorované hodnoty testové statistiky t je stejný jako v případě klasického testu.

   Je zřejmé, že čím nižší hladinu významnosti α (čím vyšší spolehlivost) zvolíme, tím širší obor přijetí dostaneme a opačně - čím vyšší hladinu významnosti α (čím nižší spolehlivost) zvolíme, tím užší obor přijetí dostaneme. Při určité hladině významnosti tedy kritická hodnota (hranice mezi oborem přijetí a kritickým oborem) splyne s hodnotou testového statistiky. Tato hodnota hladiny významnosti se nazývá **p-value**. P-value je tedy nejnižší hladina významnosti na níž můžeme nulovou hypotézu zamítnout a zároveň nejvyšší hladiny významnosti na níž se již nulová hypotéza nezamítá.

   Pozorovanou hodnotu statistiky p-value vypočteme podle jedné ze tří možných definic v závislosti na tvaru alternativní hypotézy (je nutné aby alternativní hypotéza korespondovala s výběrovým souborem).

1. H_A ve tvaru „<“: $p - value = F_0(x_{OBS})$

Tuto definici použijeme v případech, kdy pozorovaná data svědčí o tom, že testová statistika by mohla nabývat menších hodnot nežli jsou hodnoty odpovídající nulovému rozdělení. P-value je pak pravděpodobnost, že testovaný parametr populace bude nanejvýš tak velký jako skutečně zjištěný příslušný parametr výběru, bude-li H_0 pravdivá.

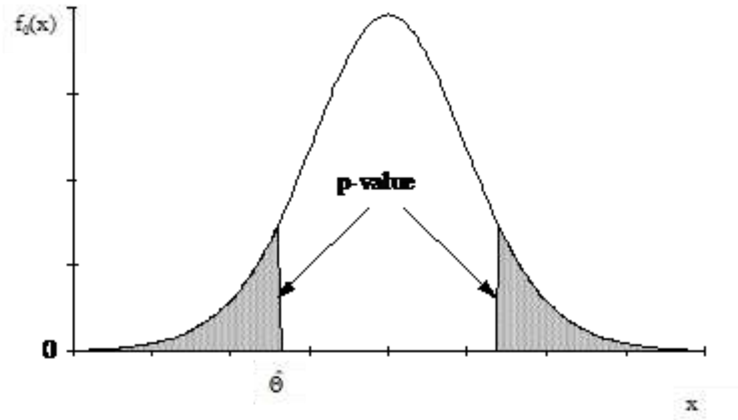
2. H_A ve tvaru „>“: $p - value = 1 - F_0(x_{OBS})$

Tuto definici použijeme v případech, kdy pozorovaná data svědčí o tom, že testová statistika by mohla nabývat vyšších hodnot nežli jsou hodnoty odpovídající nulovému rozdělení. P-value je pak pravděpodobnost, že testovaný parametr populace bude alespoň tak velký jako skutečně zjištěný příslušný parametr výběru, bude-li H_0 pravdivá.

3. H_A ve tvaru „≠“: $p - value = 2 \cdot \min\{F_0(x_{OBS}); 1 - F_0(x_{OBS})\}$

Tuto definici použijeme v případech, kdy pozorovaná data svědčí o tom, že testová statistika by mohla nabývat buď větších nebo menších hodnot nežli jsou hodnoty odpovídající nulovému rozdělení. **Tuto definici však můžeme používat pouze v případech, kdy nulové rozdělení je symetrické** (tzn. nelze použít např. při testování rozptylu). P-value je pak dvojnásobná vzhledem k jednostranným testům.

Následující obrázek znázorňuje p-value pro tuto definici pomocí plochy pod křivkou hustoty nulového rozdělení. Na základě známé geometrické interpretace distribuční funkce je zřejmé, že pro první definici by se dalo p-value ilustrovat jako levá vyšrafovaná plocha v tomto obrázku a pro druhou definici lze p-value schematicky znázornit jako pravou vyšrafovanou plochu.



4. Rozhodnutí na základě p-value

P-value nám říká jaká je **minimální hladina významnosti na níž bychom při daném výběrovém souboru mohli nulovou hypotézu zamítnout**. (např. Je-li p-value = 0,006 pak to znamená, že nulovou hypotézu můžeme zamítnout na hladinách významnosti 0,006 a vyšších, jinak řečeno: nulovou hypotézu můžeme zamítnout se spolehlivostí nejvýše 0,994. Zvolíme-li si spolehlivost testu vyšší než 0,994, p-value nesvědčí pro zamítnutí nulové hypotézy.)

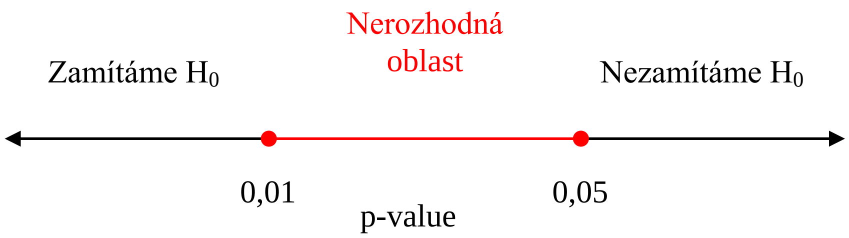
Je zřejmé, že čím menší je p-value, tím silnější je výpověď náhodného výběru proti nulové hypotéze. Ale jak malé musí být p-value, aby empirická výpověď byla dostatečně silná k zamítnutí nulové hypotézy? Výsledek testu obecně závisí na zvolené hladině významnosti α:

Rozhodnutí:

$\alpha > p-value$	Zamítáme H_0 ve prospěch H_A
$\alpha < p-value$	Nezamítáme H_0

Obecně rozhodujeme o zamítnutí nulové hypotézy na základě následujícího schématu, které je založeno na nejběžněji používaných hladinách významnosti (0,01 a 0,05).

$p-value < 0{,}01$	Zamítáme H_0
$0{,}01 < p-value < 0{,}05$	Nedokážeme rozhodnout a většinou doporučujeme opakovat test s větším rozsahem výběru (to vede ke zpřesnění)
$p-value > 0{,}05$	Nezamítáme H_0



V následujících testech budeme používat výhradně čistý test významnosti.

11.6 Test hypotézy o střední hodnotě

Tento typ testu můžeme použít v případě, že populace má normální rozdělení.

ad1.) Volba nulové a alternativní hypotézy

H_0: $\mu = \mu_0$

H_A: $\mu < \mu_0$ [1)]

$\mu > \mu_0$ [2)]

$\mu \neq \mu_0$ [3)]

Volba nulové hypotézy je zřejmá, u alternativy máme tři možnosti. Volba vhodné alternativy je při čistém testu významnosti dána hodnotou příslušné výběrové statistiky, tj. průměru. Je-li průměr jednoznačně nižší než testována hodnota μ_0, volíme alternativu ve tvaru [1)]. Je-li průměr jednoznačně vyšší než testována hodnota μ_0, volíme alternativu ve tvaru [2)]. Pohybuje-li se průměr v blízkosti μ_0, volíme alternativu ve tvaru [3)].

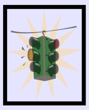
ad2.) Volba testové statistiky

Volba vhodné testové statistiky závisí na tom, zda známe či neznáme směrodatnou odchylku σ. (Srovnejte s postupem při určování intervalového odhadu pro střední hodnotu.) Zároveň si určíme i příslušné nulové rozdělení.

Známe σ: $$T(\underline{X}) = Z = \frac{\overline{X} - \mu}{\sigma} \cdot \sqrt{n} \rightarrow N(0;1)$$

Neznáme σ: $$T(\underline{X}) = T_{n-1} = \frac{\overline{X} - \mu}{s} \cdot \sqrt{n} \rightarrow t_{n-1}$$

Dále pak pokračujeme podle obecného schématu čistého testu významnosti.



Řešený příklad:

Byly naměřeny následující hodnoty IQ (výsledky testu inteligence) pro 10 vybraných účastníků inteligenčního testu (účastníky byli studenti posledního ročníku základní školy):

65	98	103	77	93
102	102	113	80	94

Předpokládejme, že náhodný výběr pochází z normálního rozdělení se směrodatnou odchylkou σ = 15. Ověřte čistým testem významnosti hypotézu, že střední hodnota IQ studentů závěrečného ročníku ZŠ je rovna 100.

Řešení:
Chceme testovat střední hodnotu přičemž známe směrodatnou odchylku. Předpoklad normality základního souboru byl splněn, můžeme tedy přistoupit k testu:

Vstupní data: $\sigma = 15$

Výběr: $$\overline{X} = \frac{65 + 98 + \ldots + 94}{10} = 92{,}7$$

$n = 10$

Stanovení nulové a alternativní hypotézy:

H_0: $\mu = 100$

H_A: $\mu < 100$

(protože výběr ukazuje na to, že střední hodnota by mohla být nižší než 100 – (92,7 < 100))

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = Z = \frac{\overline{X} - \mu}{\sigma} \cdot \sqrt{n} \rightarrow N(0;1)$$

Výpočet hodnoty testové statistiky – x_{OBS}:

$$x_{OBS} = Z_{H_0} = \frac{\overline{X} - \mu_0}{\sigma} \cdot \sqrt{n} = \frac{92{,}7 - 100}{15} \cdot \sqrt{10} = -1{,}54$$

Výpočet p-value:

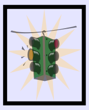
H_A: $\mu < 100$ $\Rightarrow$ $p - value = F_0(x_{OBS})$

$$p - value = \Phi(-1{,}54) = 1 - \Phi(1{,}54) = 1 - 0{,}938 = 0{,}062$$

(tzn. nulovou hypotézu můžeme zamítnou na hladině významnosti 0,062 a nižších)

Rozhodnutí:

$p - value > 0{,}05$ $\Rightarrow$ Nezamítáme nulovou hypotézu, tj. zamítáme alternativu, tj. nelze tvrdit, že IQ studentů závěrečného ročníku ZŠ je nižší než 100.



Řešený příklad:

Výrobce garantuje, že jím vyrobené žárovky mají životnost v průměru 1.000 hodin. Aby útvar kontroly zjistil, zda tomuto konstatování odpovídá i v daném období vyrobená a expedovaná část produkce, vybral z připravené dodávky náhodně 50 žárovek a došel k závěru, že průměrná doba životnosti je 1050 hodin a směrodatná odchylka doby životnosti pak 100 hodin. Ověřte čistým testem významnosti, zda nedošlo ke zlepšení kvality žárovek.

Řešení:

Měřítkem kvality žárovek je jejich střední životnost. Chceme tedy testovat střední hodnotu přičemž směrodatnou odchylku neznáme. Předpokládejme, že životnost žárovek podléhá normálnímu rozdělení.

Vstupní data: Výběr:

$\overline{X} = 1050\ hodin$

$s = 100\ hodin$

$n = 50$

Stanovení nulové a alternativní hypotézy:

H_0: $\mu = 1000$ (rovnovážný stav, střední životnost se nezměnila)

H_A: $\mu > 1000$

(výběr ukazuje na to, že střední životnost by mohla být vyšší než 1000 – (1150 > 1000))

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = T_{n-1} = \frac{\overline{X} - \mu}{s} \cdot \sqrt{n} \to t_{n-1}$$

Výpočet hodnoty testové statistiky – $\mathbf{x_{OBS}}$:

$$x_{OBS} = T_{n-1_{H_0}} = \frac{\overline{X} - \mu_0}{s} \cdot \sqrt{n} = \frac{1050 - 1000}{100} \cdot \sqrt{50} = 3{,}54$$

Výpočet p-value:

H_A: $\mu > 1000$ $\Rightarrow$ $p - value = 1 - F_0(x_{OBS})$

$p - value = 1 - F_0(3{,}54)$

$F_0(3{,}54) > 0{,}9995$ $\quad$ viz. Tabulka 2

(Studentovo rozdělení, 49 stupňů volnosti)

$p - value < 0{,}0005$

Rozhodnutí:

$p - value < 0{,}01 \quad \Rightarrow$

Zamítáme nulovou hypotézu ve prospěch alternativní, tj. lze tvrdit, že kvalita žárovek se zlepšila.

Průvodce studiem:

Pro zájemce o srovnání klasického testu a čistého testu významnosti uvádíme řešení jednoho z výše uvedených příkladů pomocí klasického testu:

Byly naměřeny následující hodnoty IQ (výsledky testu inteligence) pro 10 vybraných účastníků inteligenčního testu (účastníky byli studenti posledního ročníku základní školy):

65	98	103	77	93
102	102	113	80	94

Předpokládejme, že náhodný výběr pochází z normálního rozdělení se směrodatnou odchylkou σ = 15. Ověřte čistým testem významnosti hypotézu, že střední hodnota IQ studentů závěrečného ročníku ZŠ je rovna 100.

Řešení:
Chceme testovat střední hodnotu přičemž známe směrodatnou odchylku. Předpoklad normality základního souboru byl splněn, můžeme tedy přistoupit k testu:

Vstupní data: $\sigma = 15$

Výběr: $\overline{X} = \dfrac{65 + 98 + \ldots + 94}{10} = 92{,}7$

$n = 10$

Stanovení nulové a alternativní hypotézy:

H_0: $\mu = 100$
H_A: $\mu < 100$
(protože výběr ukazuje na to, že střední hodnota by mohla být nižší než 100 – (92,7 < 100))

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = Z = \frac{\overline{X} - \mu}{\sigma} \cdot \sqrt{n} \rightarrow N(0;1)$$

Výpočet hodnoty testové statistiky – x_{OBS}:

$$x_{OBS} = Z_{H_0} = \frac{\overline{X} - \mu_0}{\sigma} \cdot \sqrt{n} = \frac{92{,}7 - 100}{15} \cdot \sqrt{10} = -1{,}54$$

Až do této chvíle se postupy obou typů testu neliší. V klasickém testu však místo p-value určujeme kritický obor.

Stanovení kritického oboru C:

H_A: $\mu < 100$ $\Rightarrow$ $C \leq T_{\alpha}$

Tzn. v tuto chvíli se musíme rozhodnou na jaké hladině významnosti (s jakou spolehlivosti) budeme test provádět. Pro hladinu významnosti 5%:

$C \leq T_{0,05}$
$C \leq z_{0,05}$
$C \leq z_{0,05}$
$C \leq -z_{0,95}$
$C \leq -1{,}645$ (viz. Tabulka 1)

Rozhodnutí:

$x_{OBS} \notin C$ $(-1{,}54 > -1{,}645)$ $\Rightarrow$ x_{OBS} neleží v kritickém oboru, tzn. že leží v oboru přijetí ($x_{OBS} \in A$) $\Rightarrow$

Nezamítáme nulovou hypotézu, tj. zamítáme alternativu, tj. nelze tvrdit, že IQ studentů závěrečného ročníku ZŠ je nižší než 100.

Řešený příklad:

Určitý druh lilie dorůstá průměrné výšky 85 cm se směrodatnou odchylkou 10 cm. Skupina 100 těchto lilií byla pěstována za nových, příznivějších podmínek, aby se zjistilo, zda se výška zvýší.

a) Určete mezní hodnotu průměrné výšky tohoto vzorku, za níž bude možno nulovou hypotézu zamítnout na 5%-ní hladině významnosti.
b) Bude-li skutečná průměrná výška těchto 100 rostlin 88cm, jak rozhodneme o nulové hypotéze?
c) Načrtněte operativní charakteristiku.

Řešení:

Ze zadání úlohy usuzujeme, že máme rozhodovat o střední hodnotě výšky rostliny, přičemž známe směrodatnou odchylku populace.

ada) V této části úlohy máme zadánu kritickou hodnotu chyby I. druhu, tj. p-value a máme určit příslušný kritický průměr. Abychom věděli, jakým způsobem určujeme p-value (máme na výběr ze tří možností), musíme nejdříve stanovit nulovou a alternativní hypotézu.

H_0: $\mu = 85$

H_A: $\mu > 85$

$\Rightarrow \quad p\text{-value} = 1 - \mathrm{F}(\mathrm{x}_{\mathrm{OBS}})$

Volba testové statistiky a nulového rozdělení:

$$T(\underline{X}) = Z = \frac{\overline{X} - \mu}{\sigma} \cdot \sqrt{n} \to N(0;1)$$

Výpočet:

$$x_{OBS} = Z_{H_0} = \frac{\overline{X}_{krit} - 85}{10} \cdot \sqrt{100} = \overline{X}_{krit} - 85$$

$$p\text{-value} = 1 - \mathrm{F}(\mathrm{x}_{\mathrm{OBS}})$$

$$0{,}05 = 1 - \Phi\left(\overline{X}_{krit} - 85\right)$$

$$0{,}95 = \Phi\left(\overline{X}_{krit} - 85\right)$$

$$1{,}645 = \overline{X}_{krit} - 85$$

$$\underline{\underline{\overline{X}_{krit} = 86{,}645}}$$

Tzn. překročí-li průměrná výška 100 rostlin 86,6 cm, můžeme nulovou hypotézu na 5%-ní (a vyšší) hladině významnosti zamítnout.

adb) O této otázce můžeme rozhodnout buď na základě výsledku z bodu a) – 88 cm je více než 86,6 cm a proto pro tento průměr můžeme nulovou hypotézu na 5%-ní (a vyšší) hladině významnosti zamítnout – nebo můžeme klasickým způsobem provést čistý test významnosti:

Volba nulové a alternativní hypotézy:

H_0: $\mu = 85$
H_A: $\mu > 85$

Volba testové statistiky a nulového rozdělení:

$$T(\underline{X}) = Z = \frac{\overline{X} - \mu}{\sigma} \cdot \sqrt{n} \rightarrow N(0;1)$$

Výpočet pozorované hodnoty:

$$x_{OBS} = Z_{H_0} = \frac{88 - 85}{10} \cdot \sqrt{100} = 3{,}00$$

Výpočet p-value:

H_A: $\mu > 85$ $\Rightarrow$ $p\text{-value} = 1 - \Phi(3{,}00) < 0{,}003$

Rozhodnutí:

$p\text{-value} < 0{,}01 \Rightarrow$

Zamítáme nulovou hypotézu ve prospěch alternativy, tj. můžeme tvrdit, že lepší podmínky při pěstování tohoto druhu lilií vedly k vyšší výšce rostlin.

adc) Operativní charakteristika je závislosti β na konkrétních hodnotách alternativy (při pevně zvolené hodnotě α). Stanovíme si proto hodnoty pravděpodobnosti chyby II. druhu (β) na několika různých hodnotách alternativy (např. 85,5; 86; 87; 88 cm).

Zvolíme-li α rovno 5%, pak k nezamítnutí nulové hypotézy dojde tehdy, nepřekročí-li průměr hodnotu 86,6 cm (viz. úloha a) – pokud bychom tento výsledek neměli k dispozici, museli bychom kritickou hodnotu průměru určit).

$$\beta = P\left(\overline{X} < 86{,}645 \middle| H_A\right)$$

H_0: $\mu = 85$
H_A: 1) $\mu = 85{,}5$
2) $\mu = 86{,}0$
3) $\mu = 87{,}0$
4) $\mu = 88{,}0$

Volba testové statistiky:

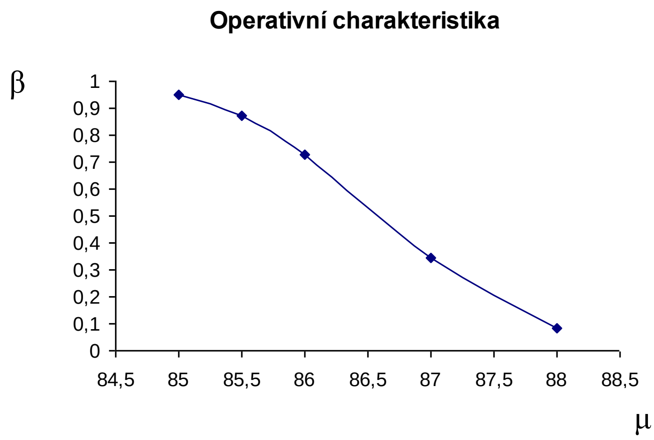
$$Z = \frac{\overline{X} - \mu}{\sigma} \cdot \sqrt{n} \to N(0;1)$$

ad1.) $\beta = P\left(\overline{X} < 86{,}6 \middle| H_A\right) = P\left(Z < \frac{86{,}645 - 85{,}5}{10} \cdot \sqrt{100}\right) = P(Z < 1{,}15) = \Phi(1{,}15) = 0{,}875$

ad2.) $\beta = P\left(\overline{X} < 86{,}6 \middle| H_A\right) = P\left(Z < \frac{86{,}6 - 86{,}0}{10} \cdot \sqrt{100}\right) = P(Z < 0{,}6) = \Phi(0{,}6) = 0{,}726$

ad3.) $\beta = P\left(\overline{X} < 86{,}6 \middle| H_A\right) = P\left(Z < \frac{86{,}6 - 87{,}0}{10} \cdot \sqrt{100}\right) = P(Z < -0{,}4) = 1 - \Phi(0{,}4) = 0{,}345$

ad4.) $\beta = P\left(\overline{X} < 86{,}6 \middle| H_A\right) = P\left(Z < \frac{86{,}6 - 88{,}0}{10} \cdot \sqrt{100}\right) = P(Z < -1{,}4) = 1 - \Phi(1{,}4) = 0{,}081$



Výklad:

11.7 Test hypotézy o rozptylu

Také tento typ testu můžeme použít pouze v případě, že populace má normální rozdělení.

ad1.) Volba nulové a alternativní hypotézy

H_0: $\sigma^2 = \sigma_0^2$

H_A: $\sigma^2 < \sigma_0^2$ 1)

$\sigma^2 > \sigma_0^2$ 2)

Volba nulové hypotézy je zřejmá, u alternativy máme dvě možnosti. Oboustrannou alternativu nemůžeme při čistém testu významnosti volit, neboť rozdělení používané testové statistiky (chí-kvadrát) není symetrické, což znemožňuje výpočet příslušného p-value. Volba vhodné alternativy závisí tentokrát na hodnotě výběrového rozptylu (s^2) Je-li výběrový rozptyl nižší než testována hodnota $(\sigma_0)^2$, volíme alternativu ve tvaru 1). Je-li výběrový rozptyl vyšší než testována hodnota $(\sigma_0)^2$, volíme alternativu ve tvaru 2).

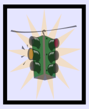
ad2.) Volba testové statistiky

$$T(\underline{X}) = \chi = \frac{(n-1)s^2}{\sigma^2} \rightarrow \chi^2_{n-1}$$

Dále pak pokračujeme podle obecného schématu čistého testu významnosti.

11.8 Test hypotézy o směrodatné odchylce

Chceme-li testovat směrodatnou odchylku, převedeme daný problém na test rozptylu.



Řešený příklad:

Při analýze diferenciace mezd ve velkém podniku bylo zjištěno, že průměrná měsíční mzda činila 9.386,-Kč a směrodatná odchylka mezd 1.562,- Kč. Po rozsáhlých organizačních změnách bylo nutné rychle posoudit, zda došlo ke změnám v diferenciaci mezd. Náhodně bylo vybráno 30 pracovníků a byla zjištěna směrodatná odchylka mezd 1.708,-Kč. Je možné tvrdit, že organizační změny prohloubily diferenciaci mezd?

Řešení:

Měřítkem diferenciace (rozložení) mezd je jejich směrodatná odchylka (resp. rozptyl). Chceme tedy testovat směrodatnou odchylku. Předpokládejme normální rozdělení mezd v podniku.

Vstupní data: Výběr: $s = 1708\ Kč$
$n = 30$

Stanovení nulové a alternativní hypotézy:

H_0: $\sigma = 1562$ (rovnovážný stav, v našem případě počáteční stav)
H_A: $\sigma > 1562$
(výběr ukazuje na to, že směrodatná odchylka by mohla být vyšší než 1562 (1708 > 1562))

Převedení problému na test rozptylu:

H_0: $\sigma^2 = 1562^2$
H_A: $\sigma^2 > 1562^2$

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = \chi = \frac{(n-1)s^2}{\sigma^2} \rightarrow \chi^2_{n-1}$$

Výpočet hodnoty testové statistiky – x_{OBS}:

$$x_{OBS} = \chi_{H_0} = \frac{(n-1)s^2}{\sigma_0^2} = \frac{29 \cdot 1708^2}{1562^2} = 34{,}7$$

Výpočet p-value:

H_A: $\quad \sigma^2 > 1562^2 \quad \Rightarrow$

$$p - value = 1 - F_0(x_{OBS})$$
$$p - value = 1 - F_0(34{,}7)$$
$$0{,}750 < F_0(34{,}7) < 0{,}900 \qquad viz.\,\text{Tabulka 3}$$
$$0{,}100 < p - value < 0{,}250$$

Rozhodnutí:

$p - value > 0{,}05 \quad \Rightarrow$

Nezamítáme nulovou hypotézu, tj. lze tvrdit, že diferenciace mezd se nezvýšila.

Výklad:

11.9 Test hypotézy o relativní četnosti

Také tento typ testu můžeme použít pouze v případě, že populace má normální rozdělení.

ad1.) Volba nulové a alternativní hypotézy

H_0: $\quad \pi = \pi_0$

H_A: $\quad \pi < \pi_0 \ ^{1)}$

$\quad \pi > \pi_0 \ ^{2)}$

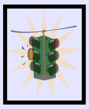
$\quad \pi \neq \pi_0 \ ^{3)}$

Volba nulové hypotézy je zřejmá, u alternativy máme opět tři možnosti. Volba vhodné alternativy je v tomto případě dána hodnotou výběrové relativní četnosti p. Je-li p jednoznačně nižší než testována hodnota π_0, volíme alternativu ve tvaru $^{1)}$. Je-li p jednoznačně vyšší než testována hodnota π_0, volíme alternativu ve tvaru $^{2)}$. Pohybuje-li se p v blízkosti π_0, volíme alternativu ve tvaru $^{3)}$.

ad2.) Volba testové statistiky

$$T(\underline{X}) = P_1 = \frac{p-\pi}{\sqrt{\pi(1-\pi)}} \cdot \sqrt{n} \rightarrow N(0;1)$$

Dále pak pokračujeme podle obecného schématu čistého testu významnosti.



Řešený příklad:

Při volbách do poslanecké sněmovny v červnu 2006 dosáhla ČSSD podpory 30%. Agentura STAT udává, že při průzkumu v prosinci 2006 (1600 respondentů) zjistili pouze 25% podporu této strany. Lze z těchto výsledků usuzovat na klesající podporu ČSSD? Ověřte čistým testem významnosti.

Řešení:

Chceme testovat relativní četnost. Předpokládejme, že relativní četnost podléhá normálnímu rozdělení.

Vstupní data: Výběr: $p = 25\% = 0{,}25$
$n = 1600$

Stanovení nulové a alternativní hypotézy:

H_0: $\pi = 0{,}30$ (rovnovážný stav, podpora ČSSD se nezměnila)
H_A: $\pi < 0{,}30$
(výběr ukazuje na to, že podpora ČSSD by mohla být nižší než 30% – (0,30 < 0,25))

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = P_1 = \frac{p-\pi}{\sqrt{\pi(1-\pi)}} \cdot \sqrt{n} \rightarrow N(0;1)$$

Výpočet hodnoty testové statistiky – x_{OBS}:

$$x_{OBS} = P_{1H_0} = \frac{p-\pi_0}{\sqrt{\pi_0(1-\pi_0)}} \cdot \sqrt{n} = \frac{0{,}25-0{,}30}{\sqrt{0{,}30 \cdot (1-0{,}30)}} \cdot \sqrt{1600} = -4{,}4$$

Výpočet p-value:

H_A: $\pi < 0{,}30$ $\Rightarrow$

$$p-value = F_0(x_{OBS})$$
$$p-value = \Phi(-4{,}4) = 1-\Phi(4{,}4) = 1-1 = 0$$
$$p-value = 0$$

Rozhodnutí:

$p - value < 0{,}01 \qquad \Rightarrow$

Zamítáme nulovou hypotézu, tzn. lze tvrdit, že pokles podpory ČSSD je statisticky významný.

Výklad:

11.10 Test hypotézy o mediánu

V rámci tohoto kurzu se seznámíte s dvěmi neparametrickými testy o mediánu (u těchto testů není nutné dělat žádné předpoklady o rozdělení základního souboru).

11.10.1 Znaménkový test pro medián

Znaménkový test používáme zejména v případech, kdy populace, z níž byl výběr proveden má výrazně zešikmené rozdělení. Jelikož tento test má malou sílu (pravděpodobnost chyby II. druhu je velká ve srovnání s jinými testy), je vhodné mít k dispozici výběr o větším rozsahu.

ad1.) Volba nulové a alternativní hypotézy

H_0: $x_{0,5} = x_{0,5_0}$

H_A: $x_{0,5} < x_{0,5_0}$ $^{1)}$

$x_{0,5} > x_{0,5_0}$ $^{2)}$

$x_{0,5} \neq x_{0,5_0}$ $^{3)}$

Volba nulové hypotézy je zřejmá, u alternativy máme opět tři možnosti. Volba vhodné alternativy je v tomto případě dána hodnotou výběrového mediánu $\tilde{x}$. Je-li $\tilde{x}$ jednoznačně nižší než testována hodnota $x_{0,5_0}$, volíme alternativu ve tvaru $^{1)}$. Je-li $\tilde{x}$ jednoznačně vyšší než testována hodnota $x_{0,5_0}$, volíme alternativu ve tvaru $^{2)}$. Pohybuje-li se $\tilde{x}$ v blízkosti $x_{0,5_0}$, volíme alternativu ve tvaru $^{3)}$.

ad2.) Volba testové statistiky

Pokud medián je $x_{0,5_0}$, potom pravděpodobnost že nějaká pozorovaná hodnota překročí $x_{0,5_0}$ je rovna 0,5. Proto také počet pozorování v náhodném výběru o rozsahu n, které překročí hypotetický medián, bude mít rozdělení binomické s parametry n a 0,5.

Za testovou statistiku volíme tedy v tomto případě:

$$T(\underline{X}) = Y \rightarrow Bi(n;0{,}5),$$

Y … počet pozorování v náhodném výběru o rozsahu n, které překročí $x_{0,5_0}$

Dále pak pokračujeme podle obecného schématu čistého testu významnosti.

11.10.2 Wilcoxnův test pro medián

ad1.) Volba nulové a alternativní hypotézy

Volba nulové a alternativní hypotézy podléhá stejným pravidlům jako u znaménkového testu.

ad2.) Volba testové statistiky

Wilcoxnův test pro testování hypotézy o mediánu je založen na Wilcoxonově statistice, která není závislá na odlehlých pozorováních:

$$T(\underline{X}) = W = \frac{\overline{r^*}}{s_{r^*}} \cdot \sqrt{n} \to N(0;1),$$

kde $y_i = \left| x_i - x_{0,5_0} \right|,$

$r_i = rank(y_i)$

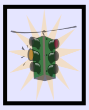
(=pořadí (y_i), nejnižší hodnotě y_i je přiřazena hodnota 1, nejvyšší hodnotě y_i je přiřazena hodnota n , pokud soubor obsahuje několik stejných hodnot, je těmto hodnotám přiřazeno tzv. průměrné pořadí),

$r_i^* = r_i \cdot \mathrm{sgn}\left(x_i - x_{0,5_0}\right)$

(r_i je doplněno znaménkem + nebo – podle toho, zda původní pozorování je větší nebo menší nežli hypotetický medián $x_{0,5_0}$),

$$\overline{r^*} = \frac{\sum_{i=1}^{n} r_i^*}{n}, \qquad s_{r^*} = \sqrt{\frac{\sum_{i=1}^{n} \left(r_i^* - \overline{r}\right)^2}{n-1}}$$

Dále již opět postupujeme známým způsobem.



Řešený příklad:

Byly naměřeny následující hodnoty IQ (výsledky testu inteligence) pro 10 vybraných účastníků inteligenčního testu (účastníky byli studenti posledního ročníku základní školy):

65	98	103	77	93
102	102	113	80	94

Ověřte čistým testem významnosti hypotézu, že medián IQ studentů závěrečného ročníku ZŠ je roven 100.

Řešení:

Ukážeme si řešení pomocí obou výše zmíněných testů hypotéz o mediánu. První krok, tj. stanovení nulové a alternativní hypotézy, je v obou případech stejný.

Vstupní data: Výběr: $\tilde{x} = \frac{94+98}{2} = 96$

$n = 10$

Stanovení nulové a alternativní hypotézy:

H_0: $x_{0,5} = 100$

H_A: $x_{0,5} < 100$

(výběr ukazuje na to, že medián IQ by mohl být nižší než 100)

Znaménkový test

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = Y \to Bi(n;0,5),$$

Y … počet pozorování v náhodném výběru o rozsahu n, které překročí $x_{0,5_0}$

Výpočet hodnoty testové statistiky – x_{OBS}:

65	98	103	77	93
102	102	113	80	94

$x_{OBS} = Y_{H_0} = 4$ (ve výběru jsou 4 hodnoty vyšší než 100)

Výpočet p-value:

H_A: $x_{0,5} < 100$ $\Rightarrow$ $p-value = F_0(x_{OBS})$

$$Y \to Bi(10;0,5)$$

$$p-value = F_0(4) = P(Y < 4) = \sum_{k=0}^{3} \binom{10}{k} \cdot (0,5)^k \cdot (1-0,5)^{10-k}$$

$$p-value = 0,172$$

Rozhodnutí:

$p-value > 0,05$ $\Rightarrow$

Nezamítáme nulovou hypotézu, tj. lze tvrdit, že IQ studentů má medián 100.

Wilcoxnův test

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = W = \frac{\overline{r^*}}{s_{r^*}} \cdot \sqrt{n} \to N(0;1),$$

Výpočet hodnoty testové statistiky – x_{OBS}:

Vstupní data postupně transformujeme na proměnnou r^* a z ní vypočteme hodnotu testové statistiky $\left(x_{0,5_0} = 100\right)$:

IQ	Seřazené hodnoty IQ	$y_i = \lvert x_i - x_{0,5_0} \rvert$	$r_i = rank(y_i)$	$r_i^* = r_i \cdot \mathrm{sgn}(x_i - x_{0,5_0})$
93	65	35	10	-10
94	77	23	9	-9
77	80	20	8	-8
80	93	7	6	-6
103	94	6	5	-5
113	98	2	2	-2
98	102	2	2	2
102	102	2	2	2
65	103	3	4	4
102	113	13	7	7

- Nejnižší hodnota y_i je 2. 2 se vyskytuje na 1., 2. a 3. pořadí, proto bude všem těmto hodnotám y_i přiřazeno pořadí 2 ($= \frac{1+2+3}{3}$).
- Např.: $\mathrm{sgn}(65-100) = -1$
  $\mathrm{sgn}(102-100) = 1$

$$\overline{r^*} = \frac{\sum_{i=1}^{n} r_i^*}{10} = -2{,}5, \qquad s_{r^*} = \sqrt{\frac{\sum_{i=1}^{n}\left(r_i^* - \overline{r}\right)^2}{9}} = 6{,}0$$

$$x_{OBS} = W_{H_0} = \left(\frac{\overline{r^*}}{s_{r^*}} \cdot \sqrt{n}\right)_{H_0} = \frac{-2{,}5}{6{,}0} \cdot \sqrt{10} = -1{,}32$$

Výpočet p-value:

H_A: $x_{0,5} < 100 \quad \Rightarrow \quad p - value = F_0(x_{OBS})$

$$p - value = \Phi(-1{,}32) = 1 - \Phi(1{,}32) = 1 - 0{,}907 = 0{,}093$$

Rozhodnutí:

$p - value > 0{,}05 \qquad \Rightarrow$

Nezamítáme nulovou hypotézu, tj. lze tvrdit, že IQ studentů má medián 100.

Výklad:

Následující skupina testů patří mezi **testy o shodě úrovně ve dvou souborech**. Výběr testů bude záviset nejen na srovnávaném parametru, ale také na tom, zda výběry z jednotlivých souborů jsou závislé či nezávislé.

Jako **nezávislé** považujeme takové výběry, kdy příslušné dvojice nejsou fyzicky spjaty, tj. netýkají se stejných prvků (tlak krve u mužů a u žen …).

Jako **závislé** označujeme naopak ty výběry, kdy příslušné dvojice jsou fyzicky spjaty, tj. týkají se stejných prvků pozorovaných za různých podmínek (tlak krve u skupiny osob – před zátěží a po zátěži …).

Testy o shodě úrovně ve dvou souborech pro závislé výběry se nazývají **párové testy**. (Testování vlivu nějakého experimentálního faktoru nebo srovnávání vlivu dvou různých faktorů na jednom měřeném empirickém objektu).

11.11 Test hypotézy o shodě dvou středních hodnot

Jde o jeden z nejpoužívanějších testů, který na základě porovnání dvou **nezávislých** výběrů umožňuje porovnat dvě populace. Nezávislost výběrů bývá v praxi zaručena tím, že každý výběr obsahuje jiné prvky. Také tento test patří mezi parametrické, tj. je založen na předpokladu, že máme výběry z normálního rozdělení.

ad1.) Volba nulové a alternativní hypotézy

H_0: $\quad \mu_1 = \mu_2 \qquad (\mu_1 - \mu_2 = 0)$

H_A: $\quad \mu_1 < \mu_2 \qquad (\mu_1 - \mu_2 < 0)^{1)}$

$\qquad \mu_1 > \mu_2 \qquad (\mu_1 - \mu_2 > 0)^{2)}$

$\qquad \mu_1 \neq \mu_2 \qquad (\mu_1 - \mu_2 \neq 0)^{3)}$

Volba nulové hypotézy je zřejmá, u alternativy máme opět tři možnosti. Volba vhodné alternativy je v tomto případě dána vztahem mezi průměry jednotlivých výběrů. Je-li $\overline{x_1}$ jednoznačně nižší než $\overline{x_2}$, volíme alternativu ve tvaru [1)]. Je-li $\overline{x_1}$ jednoznačně vyšší než $\overline{x_2}$, volíme alternativu ve tvaru [2)]. Pohybuje-li se $\overline{x_1}$ v blízkosti $\overline{x_2}$, volíme alternativu ve tvaru [3)].

ad2.) Volba testové statistiky

Volba vhodné testové statistiky závisí na tom, zda známe či neznáme směrodatné odchylky σ_1 a σ_2. (Srovnejte s postupem při určování intervalového odhadu pro rozdíl středních hodnot.) Zároveň si určíme i příslušné nulové rozdělení.

Známe σ_1, σ_2:
$$T(\underline{X}) = Z_2 = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \to N(0;1)$$

Neznáme σ_1, σ_2:
$$T(\underline{X}) = T_2 = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \to t_{n_1 + n_2 - 2},$$

$$\text{kde } s_p = \sqrt{\frac{(n_1 - 1)s_1^{\,2} + (n_2 - 1)s_2^{\,2}}{n_1 + n_2 - 2}}$$

Dále pak pokračujeme podle obecného schématu čistého testu významnosti.

Řešený příklad:

Tabáková firma TAB prohlašuje, že jejich cigarety mají nižší obsah nikotinu než cigarety NIK. Pro ověření tohoto prohlášení bylo náhodně vybráno z produkce TAB 20 krabiček cigaret (po 20-ti kusech) a v nich bylo zjištěno ($42{,}6 \pm 3{,}7$) mg nikotinu (v jediné cigaretě). Ve 25-ti krabičkách cigaret NIK (po 20-ti kusech) bylo zjištěno ($48{,}9 \pm 4{,}3$) mg nikotinu na cigaretu. Ověřte tvrzení firmy TAB čistým testem významnosti.

Řešení:

Chceme porovnávat střední obsah nikotinu v cigaretách TAB a NIK, směrodatnou odchylku obsahu nikotinu v cigaretách neznáme. Volíme tedy test pro porovnání středních hodnot dvou populací (při neznámých σ) – za předpokladu, že obsah nikotinu v cigaretách podléhá normálnímu rozdělení.

Vstupní data: Výběr 1 – firma TAB:
$$\overline{X_1} = 42{,}6\ mg$$
$$s_1 = 3{,}7\ mg$$
$$n_1 = 20.20 = 400$$

Výběr 2 – firma NIK:
$$\overline{X_2} = 48{,}9\ mg$$
$$s_2 = 4{,}3\ mg$$
$$n_2 = 25.20 = 500$$

Stanovení nulové a alternativní hypotézy:

H$_0$: $\mu_1 = \mu_2$ $(\mu_1 - \mu_2 = 0)$ (rovnovážný stav)
H$_A$: $\mu_1 < \mu_2$ $(\mu_1 - \mu_2 < 0)$
(výběry ukazují na to, že obsah nikotinu v cigaretách TAB je nižší než obsah nikotinu v cigaretách NIK)

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = T_2 = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \to t_{n_1+n_2-2}, \qquad \text{kde } s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$

Výpočet hodnoty testové statistiky – x$_{OBS}$:

Pokud je nulová hypotéza platná, platí, že: $\mu_1 = \mu_2$ $(\mu_1 - \mu_2 = 0)$, proto:

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{399 \cdot (3{,}7)^2 + 499 \cdot (4{,}3)^2}{400+500-2}} = 4{,}0$$

$$x_{OBS} = T_{2_{H_0}} = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)_{H_0}}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(42{,}6 - 48{,}9) - (0)}{4{,}0 \cdot \sqrt{\frac{1}{400} + \frac{1}{500}}} = -23{,}2$$

Výpočet p-value:

H$_A$: $\mu_1 < \mu_2$ $(\mu_1 - \mu_2 < 0)$ $\Rightarrow$ $p - value = F_0(x_{OBS})$
$p - value = F_0(-23{,}2)$
$p - value < 0{,}0005$ viz. Tabulka 2
(Studentovo rozdělení s 898 (=400+500-2) stupni volnosti)

Rozhodnutí:

$p - value < 0{,}01$ $\Rightarrow$

Zamítáme nulovou hypotézu, tj. tvrzení firmy TAB lze považovat za pravdivé.

Výklad:

11.12 Test hypotézy o shodě dvou rozptylů

Opět předpokládejme, že máme dva **nezávislé** výběry z normálního rozdělení.

ad1.) Volba nulové a alternativní hypotézy

H_0: $\sigma_1^2 = \sigma_2^2$

H_A: $\sigma_1^2 < \sigma_2^2$ [1)]

$\sigma_1^2 > \sigma_2^2$ [2)]

Volba nulové hypotézy je zřejmá, u alternativy máme tentokrát pouze dvě možnosti. Oboustrannou alternativu nemůžeme v tomto případě použít, protože výpočet p-value pro oboustrannou alternativu je podmíněn tím, že nulové rozdělení testové statistiky je symetrické. Protože testová statistika používaná pro test shody dvou rozptylů má Fischer-Snedecorovo rozdělení, není tato podmínka splněna. Volba vhodné alternativy je dána vztahem mezi výběrovými rozptyly jednotlivých výběrů. Je-li s_1^2 jednoznačně nižší než, volíme alternativu ve tvaru [1)]. Je-li s_1^2 jednoznačně vyšší než s_2^2, volíme alternativu ve tvaru [2)].

ad2.) Volba testové statistiky

$$T(\underline{X}) = F = \frac{s_1^2}{s_2^2} \rightarrow F(m,n),$$

kde F má Fischer-Snedecorovo rozdělení s m stupni volnosti pro čitatele a n stupni volnosti pro jmenovatele.

Dále pokračujeme podle obecného schématu čistého testu významnosti.

11.13 Test hypotézy o shodě dvou relativních četností

Také tento test bývá často využíván. Opět je zde nutné mít k dispozici dva **nezávislé** výběry z normálního rozdělení.

ad1.) Volba nulové a alternativní hypotézy

H_0: $\pi_1 = \pi_2 \qquad (\pi_1 - \pi_2 = 0)$

H_A: $\pi_1 < \pi_2 \qquad (\pi_1 - \pi_2 < 0)$ [1)]

$\pi_1 > \pi_2 \qquad (\pi_1 - \pi_2 > 0)$ [2)]

$\pi_1 \neq \pi_2 \qquad (\pi_1 - \pi_2 \neq 0)$ [3)]

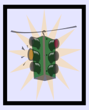
Volba nulové hypotézy je zřejmá, u alternativy máme opět tři možnosti. Volba vhodné alternativy je v tomto případě dána vztahem mezi výběrovými relativními četnostmi jednotlivých výběrů. Je-li p_1 jednoznačně nižší než p_2, volíme alternativu ve tvaru [1)]. Je-li p_1 jednoznačně vyšší než p_2, volíme alternativu ve tvaru [2)]. Pohybuje-li se p_1 v blízkosti p_2, volíme alternativu ve tvaru [3)].

ad2.) Volba testové statistiky

$$T(\underline{X}) = P_2 = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{p \cdot (1-p) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \to N(0;1),$$

kde $p = \dfrac{x_1 + x_2}{n_1 + n_2}$

Dále pak pokračujeme podle obecného schématu čistého testu významnosti.



Řešený příklad:

Byly testovány magnetofony od dvou výrobců – SONIE a PHILL. SONIE prohlašuje, že jejich magnetofony mají nižší procento reklamací. Pro ověření tohoto prohlášení bylo dotazováno několik prodejců magnetofonů a bylo zjištěno, že ze 150 prodaných magnetofonů firmy SONIE bylo v průběhu záruční doby reklamováno 5 výrobků a ze 220 prodaných magnetofonů PHILL bylo v záruční době reklamováno 9 výrobků. Otestujte pravdivost prohlášení firmy SONIE čistým testem významnosti.

Řešení:

Chceme porovnávat procento (relativní četnost) reklamovaných výrobků u obou firem. Volíme tedy test hypotézy a rozdílu mezi podíly (relativními četnostmi).

Vstupní data: Výběr 1 – firma SONIE:

$x_1 = 5$

$n_1 = 150$

$p_1 = \dfrac{5}{150} = 0{,}033$

Výběr 2 – firma PHILL:

$x_2 = 9$

$n_2 = 220$

$p_2 = \dfrac{9}{220} = 0{,}041$

Stanovení nulové a alternativní hypotézy:

H_0: $\pi_1 = \pi_2$ $(\pi_1 - \pi_2 = 0)$ (rovnovážný stav)

H_A: $\pi_1 < \pi_2$ $(\pi_1 - \pi_2 < 0)$

(výběry ukazují na to, že procento reklamovaných výrobků firmy SONIE je nižší než procento reklamovaných výrobků firmy PHILL)

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = P_2 = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{p \cdot (1-p) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \rightarrow N(0;1), \qquad \text{kde } p = \frac{x_1 + x_2}{n_1 + n_2}$$

Výpočet hodnoty testové statistiky – x_{OBS}:

Pokud je nulová hypotéza platná, platí, že: $\pi_1 = \pi_2$ $\quad (\pi_1 - \pi_2 = 0)$, proto:

$$p = \frac{x_1 + x_2}{n_1 + n_2} = \frac{5+9}{150+220} = \frac{14}{370} = 0{,}038$$

$$x_{OBS} = P_{2_{H_0}} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)_{H_0}}{\sqrt{p \cdot (1-p) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0{,}033 - 0{,}041) - (0)}{\sqrt{0{,}038 \cdot (1 - 0{,}038)\left(\frac{1}{150} + \frac{1}{220}\right)}} = -0{,}40$$

Výpočet p-value:

H_A: $\quad \pi_1 < \pi_2 \quad (\pi_1 - \pi_2 < 0) \quad \Rightarrow$

$$p - value = F_0(x_{OBS})$$
$$p - value = \Phi(-0{,}40) = 1 - \Phi(0{,}40)$$
$$p - value = 0{,}345 \qquad viz.\text{Tabulka }1$$

Rozhodnutí:

$p - value > 0{,}05 \qquad \Rightarrow$

Nezamítáme nulovou hypotézu, tj. tvrzení firmy SONIE není oprávněné.

Výklad:

11.14 Test hypotézy o shodě dvou mediánů – Mannův Whitneův test

Jde o další test, který na základě porovnání dvou **nezávislých** výběrů umožňuje porovnat dvě populace. Tento test patří k neparametrickým – nemusíme tedy znát rozdělení populací.

ad1.) Volba nulové a alternativní hypotézy

H_0: $\quad x_{0,5_1} = x_{0,5_2} \quad (x_{0,5_1} - x_{0,5_2} = 0)$

H_A: $\quad x_{0,5_1} < x_{0,5_2} \quad (x_{0,5_1} - x_{0,5_2} < 0)^{1)}$

$$x_{0,5_1} > x_{0,5_2} \qquad (x_{0,5_1} - x_{0,5_2} > 0)^{2)}$$
$$x_{0,5_1} \neq x_{0,5_2} \qquad (x_{0,5_1} - x_{0,5_2} \neq 0)^{3)}$$

Volba nulové hypotézy je zřejmá, u alternativy máme opět tři možnosti. Volba vhodné alternativy je v tomto případě dána vztahem mezi mediány jednotlivých výběrů. Je-li $\tilde{x}_1$ jednoznačně nižší než $\tilde{x}_2$, volíme alternativu ve tvaru [1)]. Je-li $\tilde{x}_1$ jednoznačně vyšší než $\tilde{x}_2$, volíme alternativu ve tvaru [2)]. Pohybuje-li se $\tilde{x}_1$ v blízkosti $\tilde{x}_2$, volíme alternativu ve tvaru [3)].

ad2.) Volba testové statistiky

Volba vhodné testové statistiky závisí na tom, zda známe či neznáme směrodatné odchylky σ_1 a σ_2. (Srovnejte s postupem při určování intervalového odhadu pro rozdíl středních hodnot.) Zároveň si určíme i příslušné nulové rozdělení.

$$T(\underline{X}) = \mathrm{W}_2 = \frac{\overline{r_1} - \overline{r_2}}{s_r\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightarrow \mathrm{N}(0;1)$$

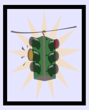
$$s_r = \sqrt{\frac{(n_1 - 1)s_{r_1}^{\ 2} + (n_2 - 1)s_{r_2}^{\ 2}}{n_1 + n_2 - 2}},$$

kde $r_{i_k} = rank(x_i)$

(=pořadí (x_i), nejnižší hodnotě x_i (z obou výběrových souborů) je přiřazena hodnota 1, nejvyšší hodnotě x_i je přiřazena hodnota n , pokud soubor obsahuje několik stejných hodnot, je těmto hodnotám přiřazeno tzv. průměrné pořadí),

$$\overline{r_k} = \frac{\sum_{i=1}^{n} r_{i_k}}{n_k}, \qquad s_{r_k} = \sqrt{\frac{\sum_{i=1}^{n} (r_{i_k} - \bar{r})^2}{n_k - 1}} \qquad (k=1,2)$$

Dále pak pokračujeme podle obecného schématu čistého testu významnosti.



Řešený příklad:

Máme dvě skupiny studentů. První (kontrolní), v níž jsou studenti vyučováni tradičními metodami, a druhá, v níž jsou studenti vyučováni experimentálními metodami. V následujících tabulkách je uvedeno bodové hodnocení vybraných studentů u zkoušky. Na základě srovnání mediánu rozhodněte, zda studenti vyučováni experimentálním metodami dosahují lepších výsledků než studenti s klasickým vyučováním.

Výběr z první skupiny (klasická výuka)

60	49	52	68	68
45	57	52	13	40
33	30	28	30	48

Výběr z druhé skupiny (experimentální výuka)

38	18	68	84	72
48	36	92	6	54

Řešení:

Volba nulové a alternativní hypotézy

H$_0$: $x_{0,5_1} = x_{0,5_2}$ $\quad \left(x_{0,5_1} - x_{0,5_2} = 0\right)$

H$_A$: $x_{0,5_1} \neq x_{0,5_2}$ $\quad \left(x_{0,5_1} - x_{0,5_2} \neq 0\right)$ $\qquad (\tilde{x}_1 = 48;\ \tilde{x}_2 = 51)$

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = \mathrm{W}_2 = \frac{\overline{r_1} - \overline{r_2}}{s_r\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightarrow \mathrm{N}(0;1)$$

Výpočet hodnoty testové statistiky – x$_{\mathbf{OBS}}$:

x$_\mathbf{i}$	60	49	52	68	68	45	57	52	13	40	33	30	28	30	48
r_{i_1}	19	14	15,5	21	21	11	18	15,5	2	10	7	5,5	4	5,5	12,5

x$_\mathbf{i}$	38	18	68	84	72	48	36	92	6	54
r_{i_2}	9	3	21	24	23	12,5	8	25	1	17

$$\overline{r_1} = \frac{\sum_{i=1}^{n} r_{i_1}}{n_1} = 12{,}1; \qquad s_{r_1} = \sqrt{\frac{\sum_{i=1}^{n}\left(r_{i_1} - \overline{r}\right)^2}{n_1 - 1}} = 6{,}3;$$

$$\overline{r_2} = \frac{\sum_{i=1}^{n} r_{i_2}}{n_2} = 14{,}4; \qquad s_{r_2} = \sqrt{\frac{\sum_{i=1}^{n}\left(r_{i_2} - \overline{r}\right)^2}{n_2 - 1}} = 8{,}9$$

$$s_r = \sqrt{\frac{(n_1 - 1)s_{r_1}{}^2 + (n_2 - 1)s_{r_2}{}^2}{n_1 + n_2 - 2}} = \sqrt{\frac{14\cdot(6{,}3)^2 + 9\cdot(8{,}9)^2}{15 + 10 - 2}} = 7{,}4$$

$$x_{OBS} = W_{2_{H_0}} = \frac{\overline{r_1} - \overline{r_2}}{s_r\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{12{,}1 - 14{,}4}{7{,}4\cdot\sqrt{\frac{1}{15}+\frac{1}{10}}} = (-0{,}76)$$

Výpočet p-value:

$$H_A: \quad x_{0,5_1} \neq x_{0,5_2} \qquad (x_{0,5_1} - x_{0,5_2} \neq 0) \qquad \Rightarrow$$

$$p-value = 2.\min\{F_0(x_{OBS});1-F_0(x_{OBS})\}$$

$$F_0(x_{OBS}) = \Phi(-0{,}76) = 1-\Phi(0{,}76) = 1-0{,}776 = 0{,}224$$

$$1-F_0(x_{OBS}) = 1-\Phi(-0{,}76) = \Phi(0{,}76) = 0{,}776$$

$$p-value = 2.0{,}224 = 0{,}448$$

Rozhodnutí:

$$p-value > 0{,}05 \qquad \Rightarrow$$

Nezamítáme nulovou hypotézu, tzn. nebyl potvrzen vliv typu výuky na výsledky studentů zkoušky.

Výklad:

11.15 Párové výběrové testy

Zopakujme si, že k párovým testům přistupujeme v případech, kdy chceme srovnat úroveň dvou závislých souborů, tj. pokud testujeme vliv nějakého experimentálního faktoru nebo srovnáváme vlivy dvou různých faktorů na jednom měřeném empirickém objektu.

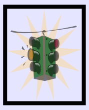
Předpokládejme n měřených jednotek (či objektů), na nichž jsou provedena dvě pozorování, daná různými experimentálními podmínkami (např. působí či nepůsobí nějaký faktor, jehož účinky jsou předmětem šetření). Příkladem může být tepová frekvence srdce před a po nějakém cvičení.

Nechť X_{i0} je počáteční pozorovaná hodnota *i*-tého měřeného objektu (tepová frekvence před cvičením) a X_{i1} následující pozorovaná hodnota (tepová frekvence po cvičení) pro stejný měřený objekt. Nyní můžeme analyzovat tato data a testovat hypotézu, zda existuje rozdíl mezi oběmi pozorováními na bázi výše uvedených dvouvýběrových testů. Avšak tento postup by eliminoval možnost posoudit rozdíly pozorovaných hodnot na týchž měřených objektech.

Mnohem efektivnějším postupem ze statistického hlediska je využít párového charakteru takto získaných dat a vytvořit jednu datovou hodnotu pro každý měřený objekt. V nejjednodušším datovém modelu bude touto hodnotou rozdíl získaných dvou pozorování pro daný *i*-tý měřený objekt. Tímto novým pozorováním je:

$$d_i = X_{i1} - X_{i0}$$

Rozdíly d_i pak mohou být použity pro jednovýběrové testy o tom, zda sledovaný parametr (střední hodnota, medián) d_i je nula, což je ekvivalentní s tím, že neexistují žádné rozdíly mezi experimentálními podmínkami (nebo že zkoumaný faktor je neúčinný).



Řešený příklad:

Máme k dispozici údaje o tepové frekvenci pacientů v klidu a po 10 minutách cvičení. Rozhodněte na základě porovnání středních hodnot a mediánů tepových frekvencí, zda se 10 minutové cvičení projeví na tepové frekvenci pacientů.

Klidová frekvence X_1	42	173	113	115	69	101	94	93	112	67	104	76
Frekvence po cvičení X_2	52	175	147	83	123	119	69	123	82	57	100	89

Řešení:

Zcela zřejmě se jedná o závislé výběry, proto použijeme párové testy.

Klidová frekvence x_1	42	173	113	115	69	101	94	93	112	67	104	76
Frekvence po cvičení x_2	52	175	147	83	123	119	69	123	82	57	100	89
$d = x_2 - x_1$	10	2	34	-32	54	18	-25	30	-30	-10	-4	13

Párový test střední hodnoty:

Vstupní data: Výběr: $\overline{d} = 5{,}0$

$s_d = 26{,}9$

$n = 12$

Stanovení nulové a alternativní hypotézy:

H_0: $\mu = 0$ (rovnovážný stav, cvičení tepovou frekvenci neovlivnilo)

H_A: $\mu > 0$ (výběr ukazuje na to, že cvičení tepovou frekvenci zvýšilo (5 > 0))

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = T_{n-1} = \frac{\overline{X} - \mu}{s} \cdot \sqrt{n} \to t_{n-1}$$

Výpočet hodnoty testové statistiky – x_{OBS}:

$$x_{OBS} = T_{11_{H_0}} = \frac{\bar{d} - \mu_0}{s_d} \cdot \sqrt{n} = \frac{5,0 - 0}{26,9} \cdot \sqrt{12} = 0,64$$

Výpočet p-value:

H$_A$: $\mu > 0 \Rightarrow$ $p-value = 1 - F_0(x_{OBS})$

$p-value = 1 - F_0(0,64)$

$F_0(3,54) < 0,75$ viz. Tabulka 2

(Studentovo rozdělení, 11 stupňů volnosti)

$p-value > 0,25$

Rozhodnutí:

$p-value > 0,05 \quad \Rightarrow$

Nezamítáme nulovou hypotézu, tj. z hlediska střední hodnoty můžeme vliv 10 minutového cvičení považovat za nevýznamný.

Párový test mediánu:

Vstupní data: Výběr: $\tilde{x} = 6,0$

Stanovení nulové a alternativní hypotézy:

H$_0$: $x_{0,5} = 0$ (rovnovážný stav, cvičení tepovou frekvenci neovlivnilo)

H$_A$: $x_{0,5} > 0$ (výběr ukazuje na to, že cvičení tepovou frekvenci zvýšilo (6 > 0))

Znaménkový test:

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = Y \rightarrow Bi(12;0,5),$$

Y … počet pozorování v náhodném výběru o rozsahu n, které překročí $x_{0,5_0}$ (=0)

Výpočet hodnoty testové statistiky – x$_{OBS}$:

d = x$_2$ –x$_1$	10	2	34	-32	54	18	-25	30	-30	-10	-4	13

$x_{OBS} = Y_{H_0} = 7$ (ve výběru je 7 hodnot vyšších než 0)

Výpočet p-value:

H$_A$: $x_{0,5} > 0 \quad \Rightarrow$

$p-value = 1 - F_0(x_{OBS})$

$$Y \to Bi(12;0,5)$$

$$p-value = 1 - F_0(7) = 1 - P(Y < 7) = P(Y \geq 7) = \sum_{k=7}^{12} \binom{12}{k} \cdot (0,5)^k \cdot (1-0,5)^{10-k}$$

$$p-value = 0,387$$

Wilcoxonův test

Volba testového kritéria a stanovení jeho nulového rozdělení:

$$T(\underline{X}) = W = \frac{\overline{r^*}}{s_{r^*}} \cdot \sqrt{n} \to N(0;1),$$

Výpočet hodnoty testové statistiky – x_{OBS}:

Vstupní data postupně transformujeme na proměnnou r^* a z ní vypočteme hodnotu testové statistiky:

$$y_i = \left| x_i - x_{0,5_0} \right| \qquad \left(x_{0,5_0} = 100 \right),$$
$$r_i = rank(y_i),$$
$$r_i^* = r_i \cdot \mathrm{sgn}\left(x_i - x_{0,5_0} \right)$$

d	Seřazené hodnoty d	$y_i = \lvert d_i - 0 \rvert$	$r_i = rank(y_i)$	$r_i^* = r_i \cdot \mathrm{sgn}\left(x_i - x_{0,5_0} \right)$
10	-32	32	10	-10
2	-30	30	8,5	-8,5
34	-25	25	7	-7
-32	-10	10	3,5	-3,5
54	-4	4	2	-2
18	2	2	1	1
-25	10	10	3,5	3,5
30	13	13	5	5
-30	18	18	6	6
-10	30	30	8,5	8,5
-4	34	34	11,5	11,5
13	54	34	11,5	11,5

$$\overline{r^*} = \frac{\sum_{i=1}^{n} r_i^*}{12} = 1,3, \qquad s_{r^*} = \sqrt{\frac{\sum_{i=1}^{n} \left(r_i^* - \overline{r} \right)^2}{11}} = 7,6$$

$$x_{OBS} = W_{H_0} = \left(\frac{\overline{r^*}}{s_{r^*}} \cdot \sqrt{n} \right)_{H_0} = \frac{1{,}3}{7{,}6} \cdot \sqrt{12} = 0{,}59$$

Výpočet p-value:

H_A: $x_{0,5} > 0$ $\Rightarrow$ $p-value = 1 - F_0(x_{OBS})$

$$p-value = 1 - \Phi(0{,}59) = 1 - \Phi(1{,}32) = 1 - 0{,}722 = 0{,}278$$

Rozhodnutí:

Jak pro znaménkový test, tak pro Wilcoxonův test je

$p-value > 0{,}05$ $\Rightarrow$

Nezamítáme nulovou hypotézu, tj. z hlediska mediánu můžeme vliv 10 minutového cvičení považovat za nevýznamný. Blízkost p-value pro t test a pro testy mediánu ukazuje na nepřítomnost odlehlých pozorování.

Výklad:

V následujícím textu se zaměříme na některé z tzv. **testů dobré shody**. V některých případech se můžeme domnívat, že studovaná data (výběr) pocházejí z určitého teoretického rozdělení. Tato domněnka bývá podložena buď informacemi o sledovaném jevu nebo odhadem teoretického rozdělení na základě grafického zobrazení výběrového rozdělení. Náš odhad však nemusí být správný, a proto jej v praxi ověřujeme testem dobré shody (tj. shody mezi výběrovým a teoretickým rozdělením (χ^2 – test dobré shody, Kolmogorovův – Smirnovův test pro jeden výběr, …). Obdobně můžeme ověřit, zda dva nezávislé výběry pocházejí z rozdělení se stejnými distribučními funkcemi (Kolmogorovův – Smirnovův test pro dva výběry).

Z formulace problémů vyplývá, že není třeba rozlišovat jednostranné a oboustranné alternativní hypotézy. Alternativa prostě popírá platnost nulové hypotézy, tj. tvrdí, že rozdělení je jiné než udává nulová hypotéza. Proto **je nutné pro jednotlivé testy určit způsob výpočtu p-value**.

11.16 χ^2 – test dobré shody

ad1.) Volba nulové hypotézy

Test dobré shody se používá nejčastěji pro ověřování těchto hypotéz:

a) H_0: Výběr pochází z populace, v níž jsou relativní četnosti jednotlivých variant rovny číslům $\pi_{0,1}; \pi_{0,2}; \ldots; \pi_{0,k}$ (populace musí být roztříditelná podle nějakého znaku do k skupin)

b) H_0: Výběr pochází z rozdělení určitého typu (např. normální), jehož parametry jsou dány (**úplně specifikovaný model**)

c) H_0: Výběrový soubor pochází z rozdělení určitého typu (např. normální) (**neúplně specifikovaný model** – neověřujeme informace o parametrech rozdělení, parametry modelu odhadujeme)

ad2.) Volba testové statistiky

Jako testovou statistiku volíme statistiku G, která má pro dostatečný rozsah výběru asymptoticky χ^2_{k-h-1} rozdělení:

$$T(\underline{X}) = G = \sum_{i=1}^{k} \frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}} \to \chi^2_{k-h-1},$$

kde **n** je rozsah výběru, **k** je počet variant, **h** je počet odhadovaných parametrů modelu, $\mathbf{n_i}$ jsou skutečné četnosti jednotlivých variant a $\boldsymbol{\pi_{0,i}}$ jsou očekávané relativní četnosti (tj. relativní četnosti, jichž by měly nabýt jednotlivé varianty v případě, že je splněna nulová hypotéza).

$\mathbf{n.\pi_{0,i}}$ jsou tedy očekávané četnosti jednotlivých variant (tj. četnosti, jichž by měly nabýt jednotlivé varianty v případě, že je splněna nulová hypotéza) a $\mathbf{(n_i - n.\pi_{0,i})}$ pak jsou odchylky očekávaných četností od četností skutečných.

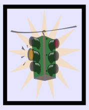
ad3) Předpoklad testu

Za výběr dostatečného rozsahu považujeme výběr, pro nějž platí, že všechny očekávané četnosti jsou vyšší než 5 ($n \cdot \pi_{0,i} > 5$ (i = 1,2, …, k))

Dále postupujeme opět podle obecného postupu při čistém testu významnosti.

ad4) Výpočet p-value

Při tomto testu určujeme p-value jako: $p - value = 1 - F_0(x_{OBS})$



Řešený příklad:

Hodilo se 6000 krát hrací kostkou a zaznamenaly se počty padlých ok...

$\mathbf{x_i}$ **(číslo které padlo)**	1	2	3	4	5	6
$\mathbf{n_i}$ **(četnost jeho výskytu)**	979	1002	1015	980	1040	984

Je možné na základě příslušného testu na hladině významnosti 5% spolehlivě tvrdit, že kostka je "falešná", tj. že pravděpodobnosti všech čísel na kostce nejsou stejné?

Řešení:

Musíme testovat, zda rozdělení „počtu ok“ padlých na kostce je takové, že pravděpodobnosti všech možných hodnot jsou 1/6.
Pro tento test dobré shody doporučujeme použít χ^2 test dobré shody (H_0 je ve tvaru a)):

Volba nulové a alternativní hypotézy

H_0: Pravděpodobnost „počtu ok“ na kostce je dána následující tabulkou:

x_i (číslo které může padnout)	1	2	3	4	5	6
$\pi_{0,i}$ (nulová pravděpodobnost jeho výskytu)	1/6	1/6	1/6	1/6	1/6	1/6

H_A: $\overline{H_0}$**,** tj. pravděpodobnost „počtu ok“ na kostce je jiná než je uvedeno ve výše uvedené tabulce

Volba testové statistiky

Rozsah výběru: n = 6000
Počet variant: k = 6
Počet odhadovaných parametrů: h = 0

$\pi_{0,1} = \pi_{0,2} = \ldots = \pi_{0,6} = 1/6 \quad \Rightarrow n \cdot \pi_{0,1} = n \cdot \pi_{0,2} = \ldots = n \cdot \pi_{0,6} = 1000 \Rightarrow 1000 > 5 \Rightarrow$
Rozsah výběru je dostatečný proto, abychom mohli použít testovou statistiku G

$$T(\underline{X}) = G = \sum_{i=1}^{k} \frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}} \to \chi^2_{k-h-1}$$

Výpočet pozorované hodnoty x_{OBS}:

$$x_{OBS} = T(\underline{X})_{H_0} = G_{H_0} = \sum_{i=1}^{k} \frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}} =$$

$$= \frac{\left(979 - 6000 \cdot \frac{1}{6}\right)^2}{6000 \cdot \frac{1}{6}} + \frac{\left(1002 - 6000 \cdot \frac{1}{6}\right)^2}{6000 \cdot \frac{1}{6}} + \ldots + \frac{\left(984 - 6000 \cdot \frac{1}{6}\right)^2}{6000 \cdot \frac{1}{6}} = 2{,}93$$

Výpočet p-value:

$p - value = 1 - F_0(x_{OBS})$
$F_0(x_{OBS}) = F_0(2{,}93)$
$0{,}250 < F_0(2{,}93) < 0{,}500$ (viz. Tabulka 3, počet stupňů volnosti je 5 (6-1))

$0{,}500 < 1 - F_0(2{,}93) < 0{,}750$

$0{,}500 < p - value < 0{,}750$

Rozhodnutí:

$p - value > 0{,}05 \Rightarrow$ Nezamítáme nulovou hypotézu, tj. nelze tvrdit, že kostka je „falešná“.

Řešený příklad:

Výrobní firma odhaduje počet poruch určitého zařízení během 100 hodin pomocí Poissonova rozdělení s parametrem 1,2. Zaměstnanci zaznamenali pro kontrolu skutečné počty poruch celkem ve 150-ti 100 hodinových intervalech (výsledky jsou uvedeny v tabulce). Ověřte čistým testem významnosti, zda má počet poruch daného zařízení během 100 hodin skutečně Poissonovo rozdělení s parametrem λ=1,2.

x_i – počet poruch během 100 hodin provozu	0	1	2	3	4
n_i - počet pozorování	52	48	36	10	4

Řešení:

Musíme testovat, zda počet poruch daného zařízení během 100 hodin má skutečně Poissonovo rozdělení s parametrem 1,2. Pro tento test dobré shody doporučujeme použít χ^2 test dobré shody (H_0 je ve tvaru b) – tj. jde o **úplně specifikovaný model** (víme jaký má být parametr rozdělení)):

Definujme si náhodnou veličinu X jako počet poruch daného zařízení během 100 hodin provozu.

Volba nulové a alternativní hypotézy

H_0: Počet poruch daného zařízení během 100 hodin (náhodná veličina X) má Poissonovo rozdělení s parametrem 1,2

H_A: $\overline{H_0}$, tj. počet poruch daného zařízení během 100 hodin (náhodná veličina X) nemá Poissonovo rozdělení s parametrem λ=1,2

Volba testové statistiky

Rozsah výběru: n = 150
Počet variant: k = 5
Počet odhadovaných parametrů: h = 0

Pokud platí H_0, pak X (počet poruch během 100 hodin) má Poisoonovo rozdělení se střední hodnotou 1,2 (= λt). Na základě této informace můžeme určit nulové pravděpodobnosti $\pi_{0,i}$.

$$\pi_{0,i} = P(X = x_i) = \frac{(\lambda t)^{x_i}}{x_i!} \cdot e^{-\lambda t} = \frac{(1{,}2)^{x_i}}{x_i!} \cdot e^{-1{,}2}$$

Zároveň si určíme očekávané četnosti.

x_i – **počet poruch během 100 hodin provozu**	0	1	2	3	4
n_i – **počet pozorování**	52	48	36	10	4
$\pi_{0,i}$	0,301	0,361	0,217	0,087	0,034
$n.\pi_{0,i}$ **- očekávané četnosti**	45,2	54,2	32,6	13,1	5,1

Všechny očekávané četnosti jsou větší než 5, tudíž rozsah výběru je dostatečný proto, abychom mohli použít testovou statistiku G

$$T(\underline{X}) = G = \sum_{i=1}^{k} \frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}} \to \chi^2_{k-h-1}$$

Výpočet pozorované hodnoty x_{OBS}:

$$x_{OBS} = T(\underline{X})_{H_0} = G_{H_0} = \sum_{i=1}^{k} \frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}} =$$

$$= \frac{(54-45{,}2)^2}{45{,}2} + \frac{(48-54{,}2)^2}{54{,}2} + \ldots + \frac{(4-5{,}1)^2}{5{,}1} = 3{,}13$$

Výpočet p-value:

H_A: $p-value = 1 - F_0(x_{OBS})$

$F_0(x_{OBS}) = F_0(3{,}13)$

$0{,}250 < F_0(3{,}13) < 0{,}500$ (viz. Tabulka 3, počet stupňů volnosti = 5-0-1 = 4)

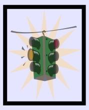
$0{,}500 < 1 - F_0(2{,}93) < 0{,}750$

$0{,}500 < p-value < 0{,}750$

Rozhodnutí:

$p-value > 0{,}05 \Rightarrow$

Nezamítáme nulovou hypotézu, tzn. nemáme námitek proti použití Poissonova rozdělení s parametrem 1,2 pro odhad počtu poruch daného zařízení během 100 hodin provozu (toto rozdělení je vhodným modelem pro počet poruch).



Řešený příklad:

Na dálnici byly v průběhu několika minut měřeny časové odstupy [s] mezi průjezdy jednotlivých vozidel. Zjištěné hodnoty těchto odstupů jsou v další tabulce:

2,5	6,8	5,0	9,8	4,0	2,3	4,2	1,9	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			

Otestujte čistým testem významnosti, zda lze časové odstupy mezi vozidly považovat za náhodnou veličinu s normálním rozdělením.

Řešení:

Nechť: náhodná veličina X je definována jako časový odstup mezi průjezdy jednotlivých vozidel.

Volba nulové a alternativní hypotézy:

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

Volba testové statistiky:

Pokud se nám podaří splnit předpoklady pro χ^2 test dobré shody ($n \cdot \pi_{0,i} > 5$), můžeme řešit daný problém pomocí tohoto testu (H_0 bude vyjádřená ve verzi c) – **neúplně specifikovaný model**).

- Nejdříve **odhadneme parametry rozdělení** (μ odhadneme průměrem, σ odhadneme výběrovou směrodatnou odchylkou (nejlepší nestranné bodové odhady)):

  Rozsah výběru: n = 132

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{132} x_i}{132} = 4{,}6 \qquad \hat{\sigma} = s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}} = 3{,}3$$

- V dalším kroku musíme **rozdělit data do „rozumného“ počtu intervalů a najít očekávané četnosti pro příslušné intervaly**. Na jejich základě rozhodneme, zda můžeme pro řešení daného problému použít χ^2 test dobré shody.

  Intervaly se volí většinou pouze na základě vlastní úvahy. Snažíme se však dodržovat několik pravidel:

  - Pokud je to možné, dodržujeme konstantní šířku intervalu (třídy)

- Počet intervalů v „rozumných“ mezích. Obvykle se považuje za vhodné volit 5 až 15 intervalů. Počet intervalů nemá být ani příliš malý (vede k hrubému, zjednodušenému pohledu na rozdělení pravděpodobnosti), ani příliš velký (který dělá rozdělení pravděpodobnosti nepřehledným).

- Intervaly nemusí mít stejnou šířku, avšak proto, abychom mohli použít χ^2 test dobré shody, musí být očekávané četnosti pro příslušné intervaly větší než 5.

Pokusíme se tedy rozdělit data do „rozumného“ počtu intervalů, najdeme očekávané četnosti pro příslušné intervaly a pak data přerozdělíme tak, aby byla splněna podmínka pro použití χ^2 testu dobré shody.

Jak spočítat očekávané četnosti?

Očekávané četnosti: $n \cdot \pi_{0,i}$

Očekávané relativní četnosti: $\pi_{0,i}$ určíme jako pravděpodobnosti výskytu náhodné veličiny X na příslušném intervalu (předpokládáme-li platnost H_0, známe rozdělení X (parametry tohoto rozdělení jsme odhadli). Pravděpodobnost, že náhodná veličina s normálním rozdělením ($N(\hat{\mu}; \hat{\sigma}^2)$) leží v i-tém intervalu je:

$$\pi_{0,i} = F(x_i) - F(x_{i-1}),$$

kde x_i je horní hranice intervalu a $x_0 = -\infty$.

Rozdělení do intervalů, příslušné očekávané relativní četnosti a očekávané četnosti

i	Časový interval [s]	Počet pozorování v časovém intervalu	Očekávané relativní četnosti $\pi_{0,i}$	Očekávané četnosti $n.\,\pi_{0,i}$
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	$\langle 5{,}8;\ 6{,}8\rangle$	10	0,106	13,9
12	$\langle 6{,}8;\ 8{,}7\rangle$	12	0,145	19,2
13	$(8{,}7;\ \infty)$	11	0,107	14,1
Součet	x	132	1,000	x

Protože normální náhodná veličina může nabývat libovolné hodnoty z množiny reálných čísel, volíme jsou dva krajní intervaly pro potřeby testu rozšířeny na: $(-\infty; 1{,}5\rangle$, $(8{,}7; \infty)$.

- Platí-li H_0: $X \rightarrow N\left(4{,}6; (3{,}3)^2\right)$

$$\pi_{0,1} = \mathrm{P}(\mathrm{X} \in (-\infty;1{,}5)) = \mathrm{P}(\mathrm{X} < 1{,}5) = \mathrm{F}(1{,}5) = \Phi\left(\frac{1{,}5-4{,}6}{3{,}3}\right) = \Phi(-0{,}94) = 1-\Phi(0{,}94) =$$

$$= 1 - 0{,}826 = 0{,}174$$

$$\vdots$$

$$\pi_{0,13} = \mathrm{P}(\mathrm{X} \in (8{,}7;\infty)) = \mathrm{P}(\mathrm{X} > 8{,}7) = 1-\mathrm{F}(8{,}7) = 1-\Phi\left(\frac{8{,}7-4{,}6}{3{,}3}\right) = 1-\Phi(1{,}24) =$$

$$= 1 - 0{,}893 = 0{,}107$$

- Pohledem na očekávané četnosti zjistíme, že jsme intervaly zvolili poměrně dobře – pouze 2. a 3. intervalu přísluší očekávané četnosti nižší než 5 (to odporuje použitelnosti χ^2 testu dobré shody). Tento nedostatek snadno napravíme tím, že tyto intervaly sloučíme.

i	Časový interval [s]	Počet pozorování v časovém intervalu	Očekávané relativní četnosti $\pi_{0,i}$	Očekávané četnosti *n*. $\pi_{0,i}$
1	$(-\infty; 1{,}5\rangle$	11	0,174	22,9
2	$\langle 1{,}5;\ 2{,}0\rangle$	20	0,041	5,4
3	$\langle 2{,}0;\ 2{,}5\rangle$	10	0,047	6,2
4	$\langle 2{,}5;\ 2{,}9\rangle$	8	0,041	5,4
5	$\langle 2{,}9;\ 3{,}6\rangle$	8	0,078	10,3
6	$\langle 3{,}6;\ 4{,}0\rangle$	10	0,047	6,2
7	$\langle 4{,}0;\ 4{,}4\rangle$	10	0,048	6,3
8	$\langle 4{,}4;\ 4{,}9\rangle$	10	0,060	8,0
9	$\langle 4{,}9;\ 5{,}8\rangle$	12	0,106	14,0
10	$\langle 5{,}8;\ 6{,}8\rangle$	10	0,106	13,9
11	$\langle 6{,}8;\ 8{,}7\rangle$	12	0,145	19,2
12	$(8{,}7;\ \infty)$	11	0,107	14,1
Součet	X	132	1,000	x

- Nyní jsou splněny předpoklady pro použití χ^2 testu dobré shody. Jako testovou statistiku tedy volíme:

$$T(\underline{X}) = G = \sum_{i=1}^{k} \frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}} \to \chi^2_{k-h-1}$$

Výpočet pozorované hodnoty x_{OBS}:

$$x_{OBS} = T(\underline{X})_{H_0} = G_{H_0} = \sum_{i=1}^{k} \frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}} =$$

$$= \frac{(11 - 22{,}9)^2}{22{,}9} + \frac{(20 - 5{,}4)^2}{5{,}4} + \ldots + \frac{(11 - 14{,}1)^2}{14{,}1} = 59{,}7$$

Výpočet p-value:

Počet variant: k = 12
Počet odhadovaných parametrů: h = 2

$p - value = 1 - F_0(x_{OBS})$
$F_0(x_{OBS}) = F_0(59{,}7)$
$F_0(59{,}7) >>> 0{,}999$ (viz. Tabulka 3, počet stupňů volnosti = 12-2-1 =
$1 - F_0(59{,}7) <<< 0{,}001$
$p\text{-}value <<< 0{,}001$

Rozhodnutí:

$p - value <<< 0{,}001 \Rightarrow$ Zamítáme nulovou hypotézu, tzn. že naměřené časové odstupy nelze považovat za výběr z normálního rozdělení.

Výklad:

11.17 Kolmogorovův – Smirnovův test pro 1 výběr

Kolmogorovův – Smirnovův test se používá k ověření hypotézy, že pořízený **výběr pochází z rozdělení se spojitou distribuční funkcí F(x)**. F(x) musí být **úplně specifikovaná**.

Máme-li při ověřování dobré shody mezi empirickým a teoretickým rozdělením k dispozici pouze **výběr malého rozsahu**, dáváme tomuto testu přednost před χ^2 testem dobré shody.

Výhody Kolmogorovova - Smirnovova test oproti χ^2 testu dobré shody:

- větší síla testu $(1-\beta)$
- nemá omezující podmínky

- vychází z jednotlivých pozorování a nikoliv u údajů setříděných do skupin (nedochází ke ztrátě informace obsažené ve výběru)

ad1.) Volba nulové a alternativní hypotézy

H$_0$: $F(x) = F_0(x)$

H$_A$: $\overline{H_0}$

kde F(x) je distribuční funkce rozdělení, z něhož náhodný výběr pochází (teoretická distribuční funkce)

ad2.) Volba testové statistiky $T(\underline{X})$ (včetně nulového rozdělení)

Uvažujme vzestupně uspořádaný náhodný výběr ze spojitého rozdělení: $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$

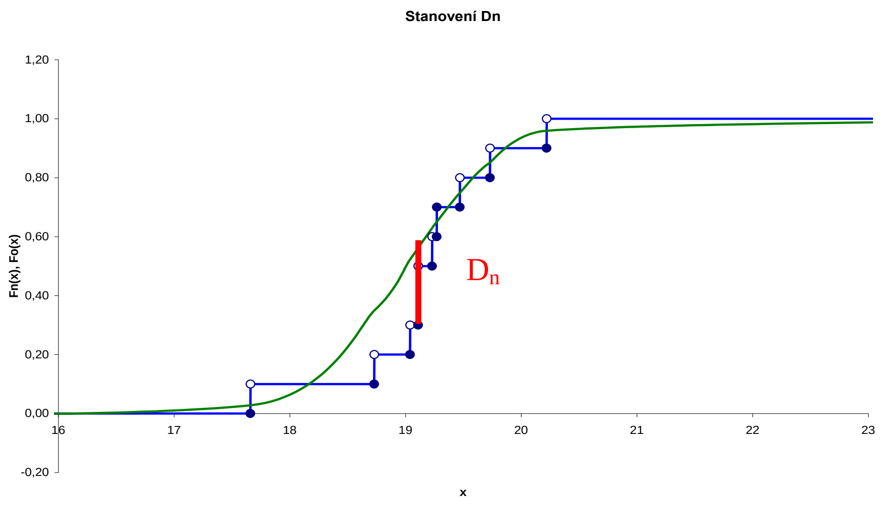
Pak výběrová (empirická) distribuční funkce $F_n(x)$ je dána jako:

$$
\begin{aligned}
F_n(x) &= 0, && x \le x_{(1)} \\
&= \frac{i}{n}, && x_{(i)} < x \le x_{(i+1)} \qquad i = 1,2,\ldots,(n-1) \\
&= 1, && x > x_{(n)}
\end{aligned}
$$

Jako testové kritérium použijeme statistiku D_n, jejíž význačné kvantily jsou tabelovány. Testová statistika D_n je definována jako maximální odchylka teoretické a empirické distribuční funkce.

$$T(\underline{X}) = D_n = \sup_x |F_n(x) - F_0(x)| = \max(D_1^*, D_2^*, \ldots, D_n^*),$$

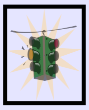
$$\text{kde } D_i^* = \max\left\{\left|F_0(x_i) - \frac{i-1}{n}\right|, \left|\frac{i}{n} - F_0(x_i)\right|\right\} \qquad pro\ i = 1,2,\ldots,n$$



Dále postupujeme standardně podle čistého testu významnosti.

ad4) Výpočet p-value

Při tomto testu určujeme p-value jako: $p-value = 1 - F_0(x_{OBS})$



Řešený příklad:

V tabulce je 10 čísel generovaných jako hodnoty rozdělení N (19; 0,7^2). Ověřte Kolmogorovovým – Smirnovovým testem, zda generované hodnoty pocházejí z předpokládaného rozdělení.

Generované hodnoty x_i	19,732	19,108	19,234	19,038	19,270	19,105	19,473	17,660	20,219	18,727

Řešení:

Volba nulové a alternativní hypotézy:

H$_0$: $F(x) = F_0(x)$, kde $F_0(x)$ je distribuční funkce normálního rozdělení o parametrech μ = 19, σ = 0,7. (neboli: data pocházejí z N (19; 0,7^2))

H$_A$: Data nepocházejí z N (19; 0,7^2)

Volba testové statistiky:

$$T(\underline{X}) = D_n = \sup_x |F_n(x) - F_0(x)| = \max(D_1^*, D_2^*, \ldots, D_n^*)$$

$$\text{kde } D_i^* = \max\left\{\left|F_0(x_i) - \frac{i-1}{n}\right|, \left|\frac{i}{n} - F_0(x_i)\right|\right\} \qquad pro\; i = 1,2,\ldots,n$$

Výpočet pozorované hodnoty x_{OBS}:

Seřazené hodnoty $x_{(i)}$	**Pořadí (i)**	**(i-1)/n**	**i/n**	**$F_0(x_{(i)})$**	**D_i pro i/n**	**D_i pro (i-1)/n**	**D_i^***
17,660	1	0,00	0,10	0,03	0,07	0,03	0,07
18,727	2	0,10	0,20	0,35	0,15	0,25	0,25
19,038	3	0,20	0,30	0,52	0,22	0,32	0,32
19,105	4	0,30	0,40	0,56	0,16	0,26	0,26
19,108	5	0,40	0,50	0,56	0,06	0,16	0,16
19,234	6	0,50	0,60	0,63	0,03	0,13	0,13
19,270	7	0,60	0,70	0,65	0,05	0,15	0,15
19,473	8	0,70	0,80	0,75	0,05	0,05	0,05
19,732	9	0,80	0,90	0,85	0,05	0,05	0,05
20,219	10	0,90	1,00	0,96	0,04	0,06	0,06

$$x_{OBS} = 0{,}32$$

Výpočet p-value:

$$p-value = 1 - F_0(x_{OBS})$$
$$F_0(x_{OBS}) = F_0(0{,}32)$$
$$F_0(0{,}32) < 0{,}9 \qquad \text{(viz. Tabulka 5, n = 10)}$$
$$1 - F_0(0{,}32) > 0{,}1$$
$$p-value > 0{,}1$$

Rozhodnutí:

$p-value > 0{,}1 \Rightarrow$ Nezamítáme nulovou hypotézu, tzn. lze tvrdit, že získaná data podléhají normálnímu rozdělení s parametry $\mu = 19$, $\sigma = 0{,}7$.

Výklad:

11.18 Kolmogorovův – Smirnovův test pro 2 výběry

Tento test se používá k ověření hypotézy, zda dva nezávislé výběry jsou z určitého spojitého rozdělení se stejnou distribuční funkcí.

ad1.) Volba nulové a alternativní hypotézy

H_0: $F^I(x) = F^{II}(x)$

H_A: $F^I(x) \neq F^{II}(x)$,

kde $F^I(x)$, $F^{II}(x)$ jsou předpokládané (teoretické) distribuční funkce prvního, resp. druhého výběru.

ad2.) Volba testové statistiky $T(\underline{X})$ (včetně nulového rozdělení)

Jako testové kritérium použijeme statistiku d_{n_1,n_2}, jejíž význačné kvantily jsou tabelovány.

$$T(\underline{X}) = d_{n_1,n_2} = \left|F_n^I(x) - F_n^{II}(x)\right|,$$

kde $F_n^I(x)$, $F_n^{II}(x)$ jsou empirické (výběrové) distribuční funkce jednotlivých výběru.

$F_n^I(x)$, $F_n^{II}(x)$ konstruujeme stejným způsobem jako při Kolmogorově – Smirnovově testu pro jeden výběr (tzn. stačí když sledujeme rozdíly mezi empirickými distribučními funkcemi v bodech jejich nespojitosti).

Dále postupujeme opět podle standardního postupu čistého testu významnosti.

ad4) Výpočet p-value

Při tomto testu určujeme p-value jako: $p-value = 1 - F_0(x_{OBS})$

11.19 Testy v kontingenční (kombinační) tabulce

Testy nezávislosti v kontingenční tabulce (Contingency Tables, Crosstables) řadíme mezi tzv. analýzu kategoriálních dat (Categorial Data Analysis). Setkáváme se s nimi v ekonomii, personalistice, psychologii, sociologii, marketingu… Abychom se s tímto testem mohli seznámit, seznámíme se nejdříve se základními pojmy v této oblasti.

11.19.1 Základní pojmy

Kontingenční tabulka vzniká setříděním prvků populace podle variant dvou kategoriálních znaků, např. znaku X a znaku Y. Nechť znak X má m variant a znak Y má n variant. Názvy jednotlivých variant znaků X a Y jsou pak uvedeny v hlavičce tabulky a uvnitř tabulky uvádíme četnosti n_{ij}, kde i označuje i-tou variantu znaku X $\left(i \in \langle 1; m \rangle\right)$ a j označuje j-tou variantu znaku Y $\left(j \in \langle 1; n \rangle\right)$. Při práci s kontingenční tabulkou budeme dále používat toto značení:

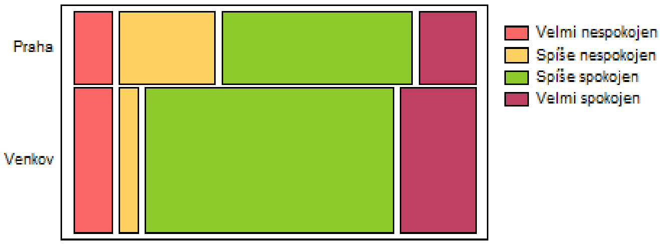
$n_{i.}$ … součet všech četností v i-té řádce
$n_{.j}$ … součet všech četností v j-tém sloupci

Schéma kontingenční tabulky

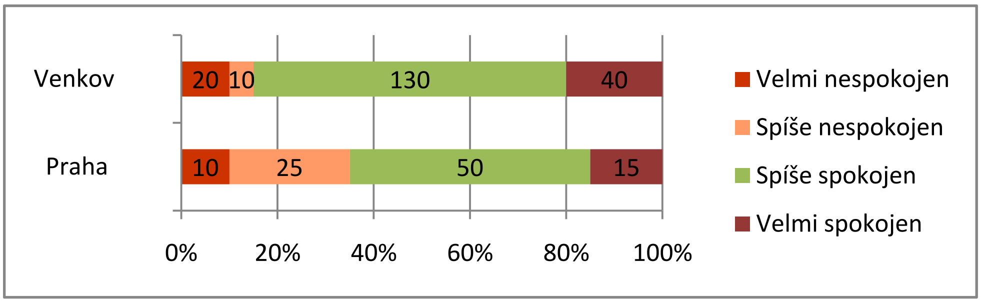
X/Y	Y_1	Y_2	…	Y_n	$\sum_j$
X_1	n_{11}	n_{12}	⋮	n_{1n}	$n_{1.}$
X_2	n_{21}	n_{22}	⋮	n_{2n}	$n_{2.}$
⋮	⋮	⋮	⋮	⋮	⋮
X_m	n_{m1}	n_{m2}	…	n_{mn}	$n_{m.}$
$\sum_i$	$n_{.1}$	$n_{.2}$	…	$n_{.n}$	n

Grafickou obdobou kontingenční tabulky je **mozaikový graf**. Tento graf se skládá z obdélníků, jejichž strany jsou úměrné příslušným marginálním relativním četnostem. Statgraphics Plus konstruuje mozaikový graf tak, že na svislou osu vynáší nezávisle proměnnou (příčina) a na vodorovnou osu závisle proměnnou (důsledek). Pokud by byl mozaikový graf v tomto případě tvořen svislými pruhy (jednotlivé obdélníky stejných barev by měly stejné „vodorovné“ rozměry), znamenalo by to, že sledované proměnné jsou nezávislé. Obdobné vyhodnocení provedeme v případě, kdy statistický software vynáší nezávisle proměnnou na vodorovnou osu (např. JMP-IN). Pak je v případě nezávislosti proměnných mozaikový graf tvořen vodorovnými pásy.

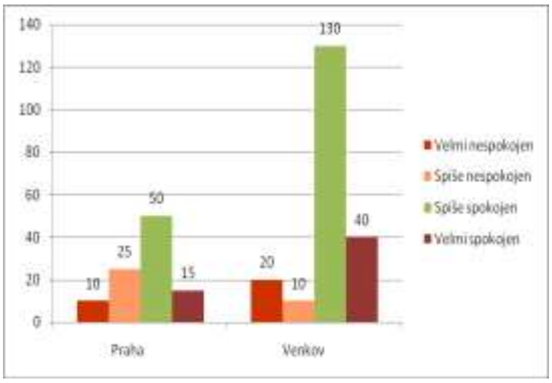
Následující ukázka mozaikového grafu odpovídá datům popisujícím jak jsou zaměstnanci spokojeni v práci v závislosti na umístění podniku.

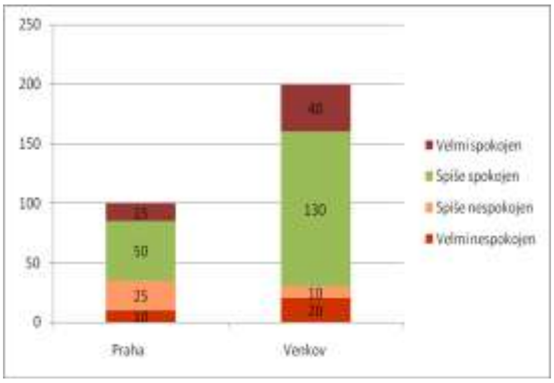


Obdobou mozaikového grafu je **100% skládaný pruhový graf** (např. MS Excel). Od mozaikového grafu se tento graf liší tím, že šířky obou řádků jsou stejné, tzn. že nezohledňuje řádkové marginální relativní četnosti.



Kromě mozaikového grafu se pro prezentaci dat zapsaných v kontingenční tabulce používají **shlukový**, popř. **kumulativní sloupcový graf**.





11.19.2 Testy v kontingenční tabulce

Pro ověření nezávislosti náhodných veličin X a Y (nezávislosti v kombinační tabulce) používáme test, který je založen na **porovnávání empirických** (pozorovaných) **četností s četnostmi teoretickými**, tj. takovými, které bychom očekávali v případě nezávislosti.

Teoretické četnosti označujeme n_{ij}^* a určujeme je jako četnosti odpovídající součinu příslušných marginálních relativních četností (připomeňme si, že v případě, že jsou dvě diskrétní náhodné veličiny nezávislé, pak jejich sdružené pravděpodobnosti jsou rovny součinu příslušných marginálních pravděpodobností).

$$n_{ij}^* = \left(\frac{n_{i.}}{n} \cdot \frac{n_{.j}}{n} \right) \cdot n = \frac{n_{i.} \cdot n_{.j}}{n}$$

11.19.3 X^2 test nezávislosti v kontingenční tabulce

ad1.) Volba nulové a alternativní hypotézy

H_0: Náhodné veličiny v kombinační tabulce jsou nezávislé.
H_A: Náhodné veličiny v kombinační tabulce jsou závislé.

ad2.) Volba testové statistiky $T(\underline{X})$ (včetně nulového rozdělení)

$$T(\underline{X}) = G = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left(\mathrm{n}_{\mathrm{ij}} - n_{ij}^*\right)^2}{n_{ij}^*} \rightarrow \chi^2_{(m-1)(n-1)}$$

Testová statistika G má rozdělení χ^2 s (m-1).(n-1) stupni volnosti. Je zřejmé, že čím bude hodnota testové statistiky G vzdálenější od nuly, tím silněji budou data vypovídat pro zamítnutí nulové hypotézy.

Předpoklad testu:

1. Žádná z očekávaných četností nesmí klesnout pod hodnotu 2.
2. Alespoň 80% očekávaných četností musí být větších než 5.

Další postup je standardní.

ad4) Výpočet p-value

Při tomto testu určujeme p-value jako: $p - value = 1 - F_0(x_{OBS})$

Řešený příklad:

Pro diferencovaný přístup v personální politice potřebuje vedení podniku vědět, zda spokojenost v práci závisí na tom, jedná-li se o pražský závod či závody mimopražské. Výsledky šetření jsou v následující tabulce. Zobrazte data pomocí mozaikového grafu a na základě testu nezávislosti v kombinační tabulce rozhodněte o závislosti spokojenosti v zaměstnání na umístnění podniku.

Stupeň spokojenosti	**Místo**	
	Praha	**Venkov**
Velmi spokojen	15	40
Spíše spokojen	50	130
Spíše nespokojen	25	10
Velmi nespokojen	10	20

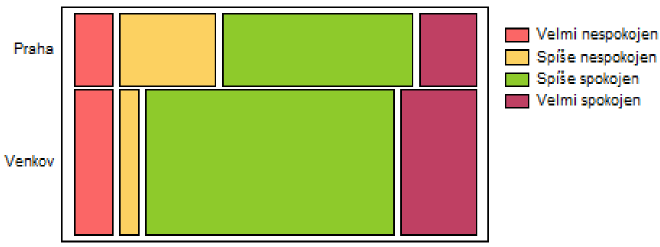
Řešení:

Nejdříve si data znázorníme pomocí mozaikového grafu, k čemuž potřebujeme znát marginální relativní četnosti:

	Velmi nespok	Spíše nespok	Spíše spokoj	Velmi spokoj	Row Total
Praha	10 3,33%	25 8,33%	50 16,67%	15 5,00%	100 33,33%
Venkov	20 6,67%	10 3,33%	130 43,33%	40 13,33%	200 66,67%
Column Total	30 10,00%	35 11,67%	180 60,00%	55 18,33%	300 100,00%

Cell contents:
Observed frequency
Percentage of table

Nyní můžeme sestrojit mozaikový graf. Na svislou osu budeme vynášet nezávisle proměnnou – tj. umístění podniku. Mozaikový graf proto bude tvořen dvěma řadami obdélníků (Praha, Mimo Prahu), přičemž řada odpovídající hodnotě „Praha" bude mít šířku odpovídající 33,33% a řada odpovídající hodnotě „Mimo Prahu" bude mít šířku odpovídající 66,67%. (Tzn., z celkové výšky mozaikového grafu bude řada odpovídající hodnotě „Praha" zabírat 33,33%, …). Závisle proměnná (Stupeň spokojenosti) nabývá 4 hodnot, proto bude každý řádek mozaikového grafu tvořen čtyřmi obdélníky příslušných délek (např. obdélník odpovídající řádku „Praha" a stupni spokojenosti – velmi spokojen bude mít délku odpovídající 15% celkové délky mozaikového grafu).



Všimněte si, že členitost grafu je způsobena zejména odlišný procentem „spíše nespokojených" zaměstnanců.

Rozhodnutí o závislosti provedeme na základě testu nezávislosti v kombinační tabulce.

Volba nulové a alternativní hypotézy:

H_0: Spokojenost v práci **nezávisí** na umístění závodu.
H_A: Spokojenost v práci závisí na umístění závodu.

Volba testové statistiky:

$$T(\underline{X}) = G = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left(n_{ij} - n_{ij}^*\right)^2}{n_{ij}^*} \rightarrow \chi^2_{(m-1)(n-1)}$$

Předpoklady testu:

Nutno ověřit, zda očekávané četnosti neklesly pod 2 a zda alespoň 80% z nich je větších než 5. Nejdříve si tedy z pozorovaných četností určíme četnosti marginální a pomocí nich pak četnosti očekávané.

Výpočet marginálních a očekávaných četností:

Stupeň spokojenosti	**Místo**		**Σ**
	Praha	**Venkov**	
Velmi spokojen	15	40	55
Spíše spokojen	50	130	180
Spíše nespokojen	25	10	35
Velmi nespokojen	10	20	30
Σ	100	200	300

n_i

$n_{.}$

n

Očekávané četnosti n_{ij}^* :

Stupeň spokojenosti	**Místo**	
	Praha	**Venkov**
Velmi spokojen	$\frac{55 \cdot 100}{300} = 18{,}3$	$\frac{55 \cdot 200}{300} = 36{,}6$
Spíše spokojen	$\frac{180 \cdot 100}{300} = 60{,}0$	$\frac{180 \cdot 200}{300} = 120{,}0$
Spíše nespokojen	$\frac{35 \cdot 100}{300} = 11{,}7$	$\frac{35 \cdot 200}{300} = 23{,}4$
Velmi nespokojen	$\frac{30 \cdot 100}{300} = 10{,}0$	$\frac{30 \cdot 200}{300} = 20{,}0$

Všechny očekávané četnosti jsou větší než 5.

Výpočet pozorované hodnoty:

$$x_{OBS} = T(\underline{X})_{H_0} = G = \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{\left(\mathrm{n}_{\mathrm{ij}} - n_{ij}^*\right)^2}{n_{ij}^*} = \frac{(15-18{,}3)^2}{18{,}3} + \frac{(50-60{,}0)^2}{60{,}0} + \ldots + \frac{(20-20{,}0)^2}{20{,}0} = 27{,}0$$

Výpočet p-value:

$m = 4, \quad n = 2 \quad \Rightarrow$ počet stupňů volnosti = $(4-1)\cdot(2-1) = 3$

$p - value = 1 - F_0(x_{OBS})$

$F(27{,}0) >>> 0{,}999$ (viz. Tabulka 3, počet stupňů volnosti = 3)

$1 - F(27{,}0) <<< 0{,}001$

$p - value <<< 0{,}001$

Rozhodnutí:

P- value $< 0{,}01$, proto zamítáme nulovou hypotézu ve prospěch alternativy, tj. spokojenost v práci závisí na umístění závodu.

Výklad:

11.19.4 Yatesova korekce X^2 testu nezávislosti v kontingenční tabulce

V případě, že není splněn předpoklad pro použití X^2 testu nezávislosti v kontingenční tabulce (tzn. že máme extrémně nízké očekávané četnosti), lze použít tzv. Yatesovu korekci. Efektem této korekce je zmenšení hodnoty testového kritéria, což znamená, že je obtížnější zamítnout nulovou hypotézu. Snížíme tak pravděpodobnost chyby I. druhu, chyba II. druhu se však zvýší – test tedy má menší sílu (oproti testu X^2).

Testová statistika:

$$T(\underline{X}) = G = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left(\mathrm{n_{ij}} - n_{ij}^{*} - 0{,}5\right)^2}{n_{ij}^{*}} \rightarrow \chi^2_{(m-1)(n-1)}$$

Výpočet p-value: $p - value = 1 - \chi^2_{(m-1)(n-1)}(x_{OBS})$

11.19.5 Fisherův exaktní test

Pro čtyřpolní tabulku (2x2) lze v případě nízkých očekávaných četností použít Fisherův exaktní test. Tento test považuje marginální četnosti za neměnné, tudíž se předpokládá, že data jsou výběrem z hypergeometrického rozdělení. Určují se pravděpodobnosti výskytu všech možných obměn četností v kontingenční tabulce, které dávají stejné marginální četnosti jako tabulka zjištěných četností. Podrobnější popis tohoto testu naleznete v literatuře věnující se analýze kategoriálních dat.

11.19.6 McNemarův test

McNemarův test je testem shody rozdělení pro čtyřpolní tabulku, pro závislé proměnné (každý respondent přispívá hodnotami do dvou políček). Lze jej použít pouze pro dvě alternativní proměnné se stejnými kódy.

Nulová a alternativní hypotéza:

H_0: Procenta „úspěšností“ jsou u obou veličin stejná.

H_A: Procenta „úspěšností“ nejsou u obou veličin stejná.

Testové kritérium: $T(X) = \frac{(n_{12} - n_{21})^2}{(n_{12} + n_{21})} \rightarrow \chi^2_{(1)}$ (Je-li $(n_{12} + n_{21})/2 \geq 4$)

Výpočet p-value: $p - value = 1 - \chi^2_{(1)}(x_{OBS})$

Řešený příklad:

Byla vybrána skupina 100 řidičů, kteří měli za úkol projet se svými vozidly náročnou uzavřenou trať. Potom po požití alkoholu dostali stejný úkol. Má se zjistit, zda požití alkoholu ovlivňuje pravděpodobnost správného projetí trati. Je tedy třeba rozhodnout, zda se počet úspěšných řidičů před podáním alkoholu (jichž bylo 80) významně liší od počtu úspěšných řidičů po požití alkoholu (jichž pak bylo jen 60). Výsledky experimentu jsou shrnuty v následující tabulce:

Před požitím alkoholu	Po požití alkoholu		Celkem
	Bez chyby	Chybně	
Bez chyby	45	35	80
Chybně	15	5	20
Celkem	60	40	100

Řešení:

Jde o závislé proměnné (stejné osoby prováděly pokus „před“ a „po“), použijeme tedy MCNemarův test.

Nulová hypotéza: Procento „úspěšných“ řidičů nezávisí na podání alkoholu.
Alternativní hypotéza: Procento „úspěšných“ řidičů závisí na podání alkoholu.

Ověření předpokladu testu: $(n_{12}+n_{21})/2=(45+5)/2=25\geq 4$

Výpočet pozorované hodnoty: $x_{OBS}=\frac{(45-5)^2}{(45+5)}=8$

Výpočet p-value:

$0{,}995<\chi^2_{(1)}(8)<0{,}999$

$p-value=1-\chi^2_{(1)}(8)$

$0{,}001<p-value<0{,}005$

Rozhodnutí: Zamítáme nulovou hypotézu, alkohol ovlivňuje „úspěšnost“ řidičů.

Shrnutí:

Pojmem testování statistických hypotéz označujeme rozhodování o pravdivosti **parametrických**, resp. **neparametrických hypotéz** o populaci. V tomto rozhodovacím procesu oproti sobě stojí **nulová** a **alternativní hypotéza** a naším cílem je rozhodnout, zda data z výběrového souboru ($\underline{X}$) odpovídají nulové hypotéze.

Jelikož při rozhodování o nulové hypotéze vycházíme z výběrového souboru, který nemusí dostatečně přesně odpovídat vlastnostem základního souboru, můžeme se při rozhodování dopustit chyby. Při rozhodování mohou nastat situace, které popisuje následující tabulka:

		Výsledek testu	
		Nezamítáme H_0	**Zamítáme H_0**
Skutečnost	**Platí H_0**	Správné rozhodnutí Pravděpodobnost rozhodnutí: $1-\alpha$ (spolehlivost)	**Chyba I. druhu** Pravděpodobnost rozhodnutí: α (hladina významnosti)
	Platí H_A	**Chyba II. druhu** Pravděpodobnost rozhodnutí: β	Správné rozhodnutí Pravděpodobnost rozhodnutí: $1-\beta$ (síla testu)

V inženýrských aplikacích se mnohdy setkáváme s tzv. **operativní charakteristikou**, což je závislost chyby II. druhu na přesné specifikaci alternativní hypotézy. Operativní charakteristika bývá v praxi taktéž nahrazována **křivkou síly testu**, což je závislost (1-β) na přesné specifikaci alternativní hypotézy.

Při testování hypotéz se běžně můžeme setkat se dvěma přístupy – klasickým testem a čistým testem významnosti.

Klasický test se skládá z několika kroků:

1. Formulace nulové a alternativní hypotézy
2. Volba testové statistiky (testového kritéria) T($\underline{X}$)
3. Sestrojení kritického oboru a oboru přijetí
4. Výpočet pozorované hodnoty testové statistiky T($\underline{X}$) - x_{OBS}
5. Formulace závěru testu – každý test vede ke dvěma možným výsledkům

Oproti klasickému testu nepotřebuje čistý test významnosti znát hladinu významnosti jako vstupní údaj. Jeho výsledek nám umožňuje rozhodnout na jakých hladinách významnosti můžeme nulovou hypotézu zamítnout (resp. nezamítnout).

Čistý test významnosti se skládá z následujících kroků:

1. Formulace nulové a alternativní hypotézy
2. Volba testové statistiky (testového kritéria) T($\underline{X}$)
3. Výpočet pozorované hodnoty testové statistiky x_{OBS} a výpočet statistiky p-value
   P-value je tedy nejnižší hladina významnosti na níž můžeme nulovou hypotézu zamítnout a zároveň nejvyšší hladiny významnosti na níž se již nulová hypotéza nezamítá. P-value

vypočteme podle jedné ze tří možných definic v závislosti na tvaru alternativní hypotézy (je nutné aby alternativní hypotéza korespondovala s výběrovým souborem).

a) H_A ve tvaru „<“: $p-value = F_0(x_{OBS})$

b) H_A ve tvaru „>“: $p-value = 1 - F_0(x_{OBS})$

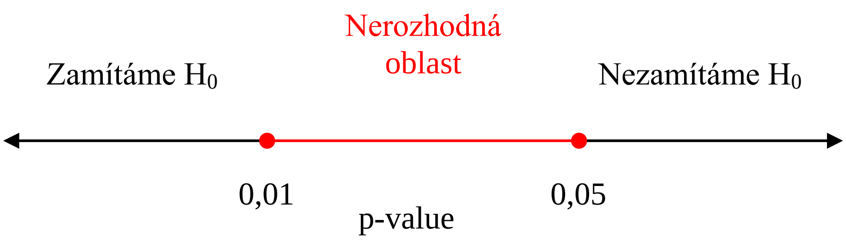
c) H_A ve tvaru „≠“: $p-value = 2 \cdot \min\{F_0(x_{OBS}); 1 - F_0(x_{OBS})\}$

4. Rozhodnutí na základě p-value

Rozhodnutí:

$\alpha > p-value$	Zamítáme H_0 ve prospěch H_A
$\alpha < p-value$	Nezamítáme H_0

Obecně rozhodujeme o zamítnutí nulové hypotézy na základě následujícího schématu, které je založeno na nejběžněji používaných hladinách významnosti (0,01 a 0,05).



Stručný přehled testových statistik, s nimiž jsme se seznámili

Jednovýběrové parametrické testy

Testovaný parametr	Pozn.	Testová statistika	Nulové rozdělení
Střední hodnota μ	Známe-li σ	$Z = \dfrac{\overline{X} - \mu}{\sigma} \cdot \sqrt{n}$	$N(0;1)$
Střední hodnota μ	Neznáme-li σ	$T_{n-1} = \dfrac{\overline{X} - \mu}{s} \cdot \sqrt{n}$	t_{n-1}
Rozptyl σ^2 (směrodatná odchylka σ)		$\chi = \dfrac{(n-1)s^2}{\sigma^2}$	χ^2_{n-1}
Relativní četnost π		$P_1 = \dfrac{p - \pi}{\sqrt{\pi(1-\pi)}} \cdot \sqrt{n}$	$N(0;1)$

Jednovýběrové neparametrické testy

Testovaný parametr	Pozn.	Testová statistika	Nulové rozdělení
Medián $x_{0,5}$	Znaménkový test, používáme u výrazně zešikmených výběrů většího rozsahu	Y … počet pozorování v náhodném výběru o rozsahu n, které překročí $x_{0,5_0}$	$Bi(n;0,5)$
Medián $x_{0,5}$		$W = \dfrac{\overline{r^*}}{s_{r^*}} \cdot \sqrt{n}$	$N(0;1)$

Dvouvýběrové parametrické testy pro nezávislé výběry

Testované parametry	Pozn.	Testová statistika	Nulové rozdělení
Střední hodnoty μ_1, μ_2	Známe-li σ_1, σ_2	$Z_2 = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$N(0;1)$
Střední hodnoty μ_1, μ_2	Neznáme-li σ_1, σ_2	$T_2 = \dfrac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{s_p \cdot \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$, $s_p = \sqrt{\dfrac{(n_1 - 1)s_1^{\,2} + (n_2 - 1)s_2^{\,2}}{n_1 + n_2 - 2}}$	$t_{n_1 + n_2 - 2}$
Rozptyly σ_1^2, σ_2^2 (směrodatné odchylky σ_1, σ_2)		$F = \dfrac{s_1^2}{s_2^2}$	$F(m,n)$
Relativní četnosti π_1, π_2		$P_2 = \dfrac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{p \cdot (1-p) \cdot \left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$, $p = \dfrac{x_1 + x_2}{n_1 + n_2}$	$N(0;1)$

Dvouvýběrové neparametrické testy

Testované parametry	Pozn.	Testová statistika	Nulové rozdělení
Mediány $x_{0,5_1}, x_{0,5_2}$	Mannův – Whitneův test	$W_2 = \dfrac{\overline{r_1} - \overline{r_2}}{s_r \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $s_r = \sqrt{\dfrac{(n_1 - 1)s_{r_1}^{\ 2} + (n_2 - 1)s_{r_2}^{\ 2}}{n_1 + n_2 - 2}}$	$N(0;1)$

Dvouvýběrové párové testy

Předpokládejme n měřených jednotek (či objektů), na nichž jsou provedena dvě pozorování, daná různými experimentálními podmínkami (např. působí či nepůsobí nějaký faktor, jehož účinky jsou předmětem šetření). Testování provádíme tak, že vytvoříme jednu datovou hodnotu pro každý měřený objekt. V nejjednodušším datovém modelu bude touto hodnotou rozdíl získaných dvou pozorování pro daný *i*-tý měřený objekt. Dané rozdíly pak mohou být použity pro jednovýběrové testy o tom, zda sledovaný parametr je nula, což je ekvivalentní s tím, že neexistují žádné rozdíly mezi experimentálními podmínkami (nebo že zkoumaný faktor je neúčinný).

Testy dobré shody pro jeden výběr

Tyto testy nám umožňují ověřit, zda studovaná data (výběr) pocházejí z určitého teoretického rozdělení.

Test	Podmínky použití	Testová statistika	Nulové rozdělení
χ^2 – test dobré shody	$n \cdot \pi_{0,i} > 5$ (i = 1,2, …, k)	$G = \sum_{i=1}^{k} \dfrac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}}$	χ^2_{k-h-1}
Kolmogorovův – Smirnovův test	F(x) - úplně specifikovaná, výběr může být malého rozsahu	$D_n = \sup_x \lvert F_n(x) - F_0(x) \rvert = \max(D_1^*, D_2^*, \ldots, D_n^*)$ $D_i^* = \max\left\{ \left\lvert F_0(x_i) - \dfrac{i-1}{n} \right\rvert, \left\lvert \dfrac{i}{n} - F_0(x_i) \right\rvert \right\}$ $pro\ i = 1,2,\ldots,n$	tabelováno

Test dobré shody pro dva výběry

Tento test nám umožňuje ověřit, zda studovaná data (výběry) pocházejí ze stejného teoretického rozdělení.

Test	Testová statistika	Nulové rozdělení
Kolmogorovův – Smirnovův test	$d_{n_1,n_2} = \lvert F_n^I(x) - F_n^{II}(x) \rvert$	tabelováno

Test nezávislosti v kontingenční tabulce

Testy nezávislosti v kontingenční tabulce řadíme mezi tzv. analýzu kategoriálních dat. **Kontingenční tabulka** vzniká setříděním prvků populace podle variant dvou kategoriálních znaků. Grafickou obdobou kontingenční tabulky je **mozaikový graf**. Tento graf se skládá z obdélníků, jejichž strany jsou úměrné příslušným marginálním relativním četnostem.

Pro ověření nezávislosti náhodných veličin X a Y (nezávislosti v kombinační tabulce) používáme test, který je založen na **porovnávání empirických** (pozorovaných) **četností s četnostmi teoretickými**, tj. takovými, které bychom očekávali v případě nezávislosti.

Test	Testová statistika	Nulové rozdělení
Test nezávislosti v kontingenční tabulce	$G = \sum_{i=1}^{m} \sum_{j=1}^{\mathrm{n}} \frac{\left(\mathrm{n}_{\mathrm{ij}} - n_{ij}^*\right)^2}{n_{ij}^*}$	$\chi^2_{(m-1)(n-1)}$

Otázky

1. Co je to statistická hypotéza a jaké typy těchto hypotéz znáte?
2. Popište jak můžeme testovat statistické hypotézy pomocí intervalových odhadů?
3. Popište princip klasického testu.
4. Popište princip čistého testu významnosti.
5. Co je to p-value? Jak pomocí něj rozhodujeme při čistém testu významnosti?
6. Jakých chyb se při testování můžete dopustit? Objasněte pojmy: chyba I. druhu, chyba II. druhu, síla testu, hladina významnosti.
7. Co je to operativní charakteristika?
8. V čem spočívá rozdíl mezi parametrickými a neparametrickými testy?
9. Popište jednotlivé testy, s nimiž jste se seznámili (podmínky použití, způsob vytvoření alternativní hypotézy, …)
10. Co je to kontingenční tabulka? Co je to mozaikový graf (popište jeho konstrukci).
11. Popište test závislosti v kontingenční tabulce.

Úlohy k řešení

1. Firma FRIDGER pravidelně přijímá dodávky chladících jednotek pro své chladničky a za posledních 18 měsíců pouze 2% jednotek nedosahovaly požadovaných parametrů. Dodavatel však přešel na novou technologii a fy FRIDGER se obává možného zhoršení dodávek. Proto bylo náhodně vybráno 500 jednotek z následující dodávky a zjištěno, že 21 jednotek nesplňuje požadované parametry.
   a.) Ověřte pomocí 95% intervalu spolehlivosti, zda došlo k zhoršení kvality
   b.) Ověřte pomocí čistého testu významnosti, zda došlo k zhoršení kvality (na 5% hladině významnosti)
   c.) Načrtněte křivku síly testu pro tento případ

2. Firma Modus zjišťovala v roce 2006 názory Čechů na bezpečnost jaderných elektráren. Ze 420 respondentů ve věku od 18 do 30 let považovalo 24% současná bezpečnostní opatření za postačující. Z 510 respondentů ve věku 30 až 50 let považovalo současná bezpečnostní opatření za postačující 34%. Ověřte čistým testem významnosti, zda má věk vliv na odpověď.

3. Výrobní proces produkuje milióny žárovek se střední životnosti 14 000 hodin. Novou technologií byl vyroben vzorek 25 žárovek s průměrnou životností 14 740 hodin a směrodatnou odchylkou 2 000 hodin. Ověřte čistým testem významnosti, zda nová technologie vedla ke zvýšení životnosti žárovek.

4. Majitel rybníka ví z dlouhodobých záznamů, že střední váha kaprů z tohoto rybníka je 1,97 kg. V loňském roce majitel zkoušel nový způsob krmení ryb. Při minulém výlovu byla průměrná váha sta kaprů 1,99 kg se směrodatnou odchylkou 0,21 kg. Ověřte čistým testem významnosti, zda se při novém způsobu krmení:
   a.) váha kaprů změnila
   b.) váha kaprů zvýšila

5. U standardně vyráběného materiálu má mez pevnosti R_m normální rozdělení se střední hodnotou 640,0 MPa a směrodatnou odchylkou 4,5 MPa. Změnou posloupnosti tepelných úprav byl připraven nový materiál (předpokládáme stejný rozptyl), pro nějž bylo naměřeno R_m u deseti vzorků postupně:

   651 639 645 648 650 643 652 640 644 645

   a) Ověřte znaménkovým testem hypotézu, že medián meze pevnosti po změně posloupnosti tepelných úprav je 643.
   b) Ověřte stejnou hypotézu Wilcoxonovým testem
   c) Zvolte pravděpodobnost chyby I. druhu 5% a načrtněte operativní charakteristiku pro test hypotézy o tom, zda došlo ke změně střední hodnoty meze pevnosti.
   Návod: Vypočtěte pravděpodobnost chyby II. druhu pro jednoduché alternativy: $H_A : \mu = \mu_A$. Volte postupně μ_A = 642, 644, 646 MPa.

6. Firma TT udává, že 1% jejich rezistorů nesplňuje požadovaná kritéria. V testované dodávce 1000ks bylo nalezeno 15 nevyhovujících rezistorů. Potvrzuje tento výsledek tvrzení TT? Ověřte čistým testem významnosti.
7. Výrobce garantuje, že jím vyrobené žárovky mají životnost v průměru 1.000 hodin. Aby útvar kontroly zjistil, zda tomuto konstatování odpovídá i v daném období vyrobená a expedovaná část produkce, vybral z připravené dodávky náhodně 50 žárovek a došel k závěru, že průměrná doba životnosti je 950 hodin a směrodatná odchylka doby životnosti pak 100 hodin. Je možné zjištěný rozdíl doby životnosti ve výběru připsat náhodě nebo je známkou nekvality produkce? Ověřte čistým testem významnosti.

8. Představenstvo velké akciové společnosti zvažuje odprodat část akcií zaměstnancům této společnosti. Odhaduje se, že zájem o nákup by mohlo projevit asi 20% z nich. Proto personální útvar připravil předběžný průzkum, v němž oslovil 400 náhodně vybraných pracovníků společnosti, z nichž zájem o nákup akcií projevilo 66 lidí. Je úvaha představenstva reálná? Ověřte čistým testem významnosti.

9. Automat vyrábí pístové kroužky o daném průměru. Výrobce udává, že směrodatná odchylka průměru kroužku je 0,05mm. K ověření této informace bylo náhodně vybráno 80 kroužků a vypočtena směrodatná odchylka jejich průměru 0,04mm. Lze tento rozdíl považovat za významný ve smyslu zlepšení kvality produkce? Ověřte čistým testem významnosti.

10. Byly testovány polovodičové součástky od dvou výrobců – MM a PP. MM prohlašuje, že její výrobky mají nižší procento vadných. Pro ověření tohoto tvrzení bylo z produkce MM náhodně vybráno 200 součástek, z nichž 14 bylo vadných. Podobný experiment byl proveden u firmy PP s výsledkem 10 vadných ze 100 náhodně vybraných součástek.
    a) Otestujte tvrzení firmy MM čistým testem významnosti.
    b) Otestujte tvrzení firmy MM prostřednictvím intervalového odhadu na hladině významnosti 0,05.
    c) Nalezněte 95% interval spolehlivosti pro počet vadných součástek firmy MM.

11. Při analýze diferenciace mezd ve velkém podniku bylo zjištěno, že průměrná měsíční mzda činila 9.386,-Kč a směrodatná odchylka mezd 1.562,- Kč. Po rozsáhlých organizačních změnách bylo nutné rychle posoudit, zda došlo ke změnám v diferenciaci mezd. Náhodně bylo vybráno 30 pracovníků a byla zjištěna směrodatná odchylka mezd 1.708,-Kč. Je možné na 5% hladině významnosti tvrdit, že organizační změny prohloubily diferenciaci mezd?

12. Ropná společnost chce postavit novou čerpací stanici na severním nebo jižním okraji menšího města. Projekt předpokládá, že bude vybrán ten výjezd z města, kde je vyšší intenzita provozu. Na severním výjezdu z města probíhalo šetření během 50 dní a byl zjištěn počet 4.000 projíždějících vozidel (denně, se směrodatnou odchylkou 70 vozidel). Na jižním výjezdu z města bylo za 45 dní zaznamenáno v průměru 3.900 projíždějících vozidel denně (směrodatná odchylka 60 vozidel). Lze rozhodnout, který výjezd je zatíženější? (Volte hladinu významnosti 0,05).

13. Podnik uspořádal školení výpočetní techniky na aplikaci Excel MS Office ’97, jímž jsou vybaveny všechny počítače pracovníků ekonomického oddělení. Pokuste se prokázat, že školení ovlivnilo podíl pracovníků používajících tuto aplikaci pravidelně ve své práci. Výsledky šetření jsou v tabulce.

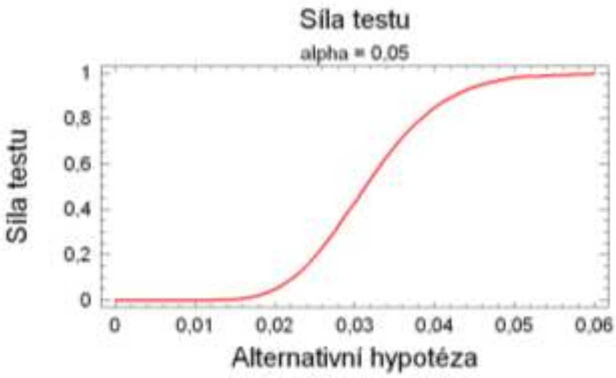
Před školením	**Po školení**	
	Používá	**Nepoužívá**
Používá	28	4
Nepoužívá	23	15

14. V průzkumu veřejného mínění byla sledována závislost mezi názorem na odstoupení vlády a věkem dotazovaných. Určete, zda daná závislost existuje (čistým testem závislosti) a nakreslete mozaikový graf pro tento případ.

Věk. skupina	**Názor na odstoupení vlády**		
	ANO	NE	NEVÍM
Do 20-ti let	20	5	10
(20-35) let	40	5	10
(36-55) let	30	0	20
Nad 55 let	10	30	10

Řešený příklad:

1. a) (0,026; 0,063)
   b) p-value = 0,0014 $\Rightarrow$ na 5% hladině významnosti zamítáme nulovou hypotézu, tj. můžeme říci, že se kvalita chladících zařízení zhoršila
   c)



2. $x_{OBS} = -3{,}33$, p-value = 0,0004 $\Rightarrow$ zamítáme nulovou hypotézu, tj. můžeme tvrdit, že lidé ve věku 18 až 30 let považují jaderné elektrárny za bezpečnější než lidé ve věku 30 až 50 let.

3. $x_{OBS} = 1{,}85$, p-value = 0,038 $\Rightarrow$ zamítáme nulovou hypotézu, tj. můžeme tvrdit, že nová technologie vedla ke zvýšení životnosti žárovek.

4. a) $x_{OBS} = 0{,}95$, p-value = 0,343 $\Rightarrow$ nezamítáme nulovou hypotézu, tj. nemůžeme tvrdit, že nový způsob krmení vedl ke změně hmotnosti kaprů.
   b) $x_{OBS} = 0{,}95$, p-value = 0,172 $\Rightarrow$ nezamítáme nulovou hypotézu, tj. nemůžeme tvrdit, že nový způsob krmení vedl ke zvýšení hmotnosti kaprů.

5. ještě doplním

6. p-value = 0,08 $\Rightarrow$ nezamítáme nulovou hypotézu, tzn. daný výsledek potvrzuje tvrzení firmy TT.

7. $x_{OBS} = -3{,}54$, p-value = 0,00045 $\Rightarrow$ zamítáme nulovou hypotézu, tj. můžeme tvrdit, že zjištěný rozdíl je známkou nekvality produkce.

8. p-value = 0,046 $\Rightarrow$ zamítáme nulovou hypotézu, tj. můžeme tvrdit, že úvaha představenstva není reálná

9. $x_{OBS} = 50{,}56$, p-value = 0,005 $\Rightarrow$ zamítáme nulovou hypotézu, tj. můžeme tvrdit, že došlo ke zlepšení kvality.

10. a) $x_{OBS} = -0{,}90$, p-value = 0,18 $\Rightarrow$ nezamítáme nulovou hypotézu, tj. tvrzení firmy MM nemůžeme považovat za pravdivé.
    b) $0 \in (-0{,}099; 0{,}039) \Rightarrow$ nezamítáme nulovou hypotézu, tj. tvrzení firmy MM nemůžeme považovat za pravdivé.
    c) $(0{,}039; 0{,}115)$

11. $x_{OBS} = 34{,}67$, p-value = 0,22 $\Rightarrow$ nezamítáme nulovou hypotézu, tj. nelze tvrdit, že organizační změny prohloubily diferenciaci mezd.

12. $x_{OBS} = 7{,}43$, p-value = 0,000 $\Rightarrow$ zamítáme nulovou hypotézu, tj. můžeme tvrdit, že severní výjezd je zatíženější

13. $x_{OBS} = 6{,}39$, p-value = 0,0115 $\Rightarrow$ zamítáme nulovou hypotézu, tj. můžeme tvrdit, že školení přineslo požadovaný efekt. (nezapomeňte ověřit použitelnost testu)

14. $x_{OBS} = 71{,}81$, p-value = 0,0000 $\Rightarrow$ zamítáme nulovou hypotézu, tj. můžeme tvrdit, že názor na odstoupení vlády závisí na věku respondenta. (nezapomeňte ověřit použitelnost testu)