Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Júlia Klačanská

Statistické testy ekvivalence

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Arnošt Komárek, Ph.D.
Studijní program: Matematika, Finanční matematika

2010

Chcela by som poďakovať vedúcemu práce za cenné rady a pripomienky a Michaele Tichej za kolegiálnu výpomoc.

Prehlasujem, že som svoju bakalársku prácu napísala samostatne a výhradne s použitím citovaných prameňov. Súhlasím so zapožičaním práce a jej zverejňovaním.

V Prahe dňa 30.7.2010 Júlia Klačanská

Obsah

Názov práce: Štatistické testy ekvivalencie
Autor: Júlia Klačanská
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedúci bakalárskej práce: RNDr. Arnošt Komárek, Ph.D.
e-mail vedúceho: arnost.komarek@mff.cuni.cz

Abstrakt: Predložená práca pojednáva o štatistických testoch ekvivalencie. Tieto testy sú odpoveďou na problém nemožnosti potvrdenia platnosti nulovej hypotézy pri testovaní pomocou klasických testov. Práca sa bude zaoberať dvomi prístupmi odvodenia testov ekvivalencie. Prvý je založený na princípe rozdelenia problému ekvivalencie na dva jednostranné problémy riešené prostredníctvom klasických testov. V druhom prístupe je hľadaný optimálny test, teda test ktorý je najsilnejší zo všetkých testov na zvolenej hladine. Tento test je následne odvodený pre konkrétne vybrané problémy ekvivalencie. Práca je doplnená o ilustračné príklady z oblasti medicíny, v ktorej majú štatistické testy ekvivalencie veľké uplatnenie.

Kľúčové slová: štatistika, testovanie hypotéz, problém ekvivalencie, optimálny test

Title: Statistical tests for equivalence
Author: Júlia Klačanská
Department: Department of Probability and Mathematical Statistics
Supervisor: RNDr. Arnošt Komárek, Ph.D.
Supervisor's e-mail address: arnost.komarek@mff.cuni.cz

Abstract: The present thesis studies statistical tests for equivalence. These tests are the solution to a problem of inability to prove the null hypothesis by using the traditional tests. There will be shown two methods for yielding tests for equivalence. The first one is based on replacing the problem of equivalence with two one-sided problems solved by traditional tests. In the second method there is derived optimal test, the test which has the greatest power among all possible tests of a given size. Afterwards, there will be derived optimal tests for chosen specific situations. In the thesis there are examples which ilustrate use of the statistical tests for equivalence in the medicine.

Keywords: statistics, testing hypotheses, equivalence problem, optimal test

Úvod

Prvý článok o testoch ekvivalencie vyšiel v šesťdesiatych rokoch dvadsiateho storočia. Záujem o tieto testy však vzrástol až po tom, čo FDA (Food and Drug Administration of the U.S. - Úrad pre kontrolu potravín a liečiv USA) vydal opatrenia pre generické lieky. Tieto opatrenia považovali za dostatočnú podmienku pre zavedenie generických liekov do predaja to, že testy ekvivalencie potvrdili ich podobnosť s pôvodným liekom vzhľadom k základným farmakokinetickým vlastnostiam. Dané lieky boli potom považované za bioekvivalentné. Ďalším možným využitím testov ekvivalencie v medicíne sú štúdie porovnávajúce rôzne druhy liečob nejakého ochorenia.

Cieľom tejto bakalárskej práce je uviesť čitateľa do problematiky štatistických testov ekvivalencie a ukázať mu, ako vyzerajú vybrané všeobecne známe testy, ak namiesto klasického problému riešia problém ekvivalencie. Tému som si vybrala, pretože je pomerne zaujímavá a existuje jej veľké uplatnenie v praktickom živote.

Štatistické testy ekvivalencie sú východiskom v situácii, keď klasický test nezamietne nulovú hypotézu, ale nedokáže ju ani potvrdiť. Tieto testy sa zaoberajú situáciou, keď alternatívna hypotéza špecifikuje postačujúco malé okolie odhadovaného parametru, ktoré znamená ekvivalenciu porovnávaných rozdelení.

Práca ukazuje základné možné prístupy k riešeniu problému ekvivalencie a ich použitie v konkrétnej testovacej situácii.

V prvej kapitole sa budeme zaoberať dôvodom pre vznik testov ekvivalencie a ukážeme, ako všeobecný test ekvivalencie vyzerá. V druhej kapitole bližšie rozoberieme testy ekvivalencie založené na princípe inklúzie intervalov spoľahlivosti, ktoré sú jednou z možností, ako problém ekvivalencie riešiť. Ďalšou možnosťou sú optimálne testy, ktorými sa bude zaoberať tretia kapitola. V nej položíme základy, ktoré využijeme v štvrtej kapitole pri odvodení testov v konkrétnych vybraných situáciách.

Práca je založená najmä na štúdiu knihy profesora Welleka, ktorá je celá venovaná danej problematike. Ďalej som použila časť knihy profesora Dupača a profesorky Huškovej, ktorá sa zaoberá Neymanovou-Pearsonovou lemou.

Kapitola 1

Štatistické testy ekvivalencie

Táto kapitola sa zaoberá motiváciou pre vznik testov ekvivalencie a všeobecným zápisom problému, ktorý riešia.

Problém

Uvažujme náhodný výber z rodiny rozdelení $\{F_\Theta : \theta \in \mathbf{R}\}$, pričom θ je parameter, ktorý je mierou odlišnosti zahrnutých pravdepodobnostných rozdelení. Chceme ukázať, že θ sa nachádza v dostatočnej blízkosti od zadanej hodnoty θ_0, ktorú θ nadobúda jedine v prípade ekvivalencie (v praktickom zmysle) zahrnutých pravdepodobnostných rozdelení.

Riešenie pomocou klasického testu

Testujeme

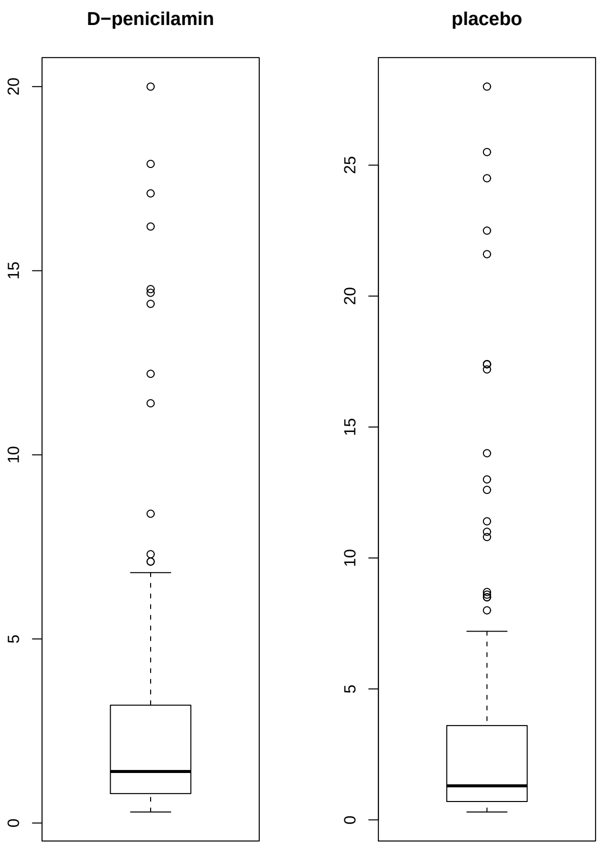
$$H_0 : \theta = \theta_0 \qquad \text{vs.} \qquad H_1 : \theta \neq \theta_0. \tag{1.1}$$

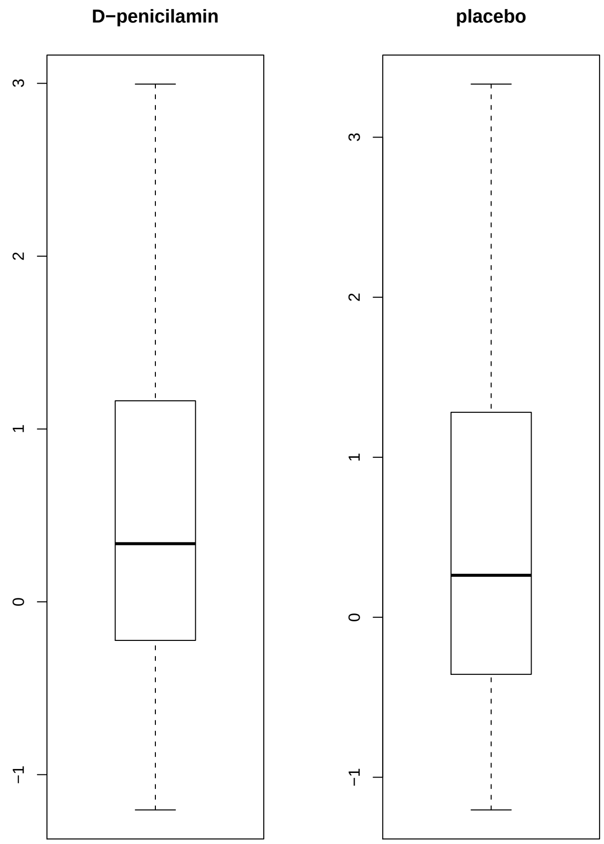
Príklad. V Spojených štátoch amerických prebehol v sedemdesiatych rokoch dvadsiateho storočia výskum zaoberajúci sa vplyvom lieku D-penicilamínu na pacientov s cirhózou pečene. Z tristodvanásť zúčastnených pacientov bolo náhodne vybraných stopäťdesiatosem, ktorí boli liečení pomocou D-penicilamínu. Ostatným bolo podávané placebo. Mimo iných ukazovateľov bol následne sledovaný vývoj hladiny bilirubínu a prítomnosť pavúčikovitých név na koži. Všetky dáta, s ktorými budeme ďalej pracovať, boli zozbierané na začiatku štúdie.
Vzhľadom na to, že úspešná randomizácia by mala na začiatku štúdie viesť k ekvivalencii oboch skupín pacientov vzhľadom k ľubovoľnému ukazovateľu, použijeme

tieto dáta k ilustrácii použitia uvažovaných testov ekvivalencie. Menovite sa budeme zaoberať testovaním ekvivalencie logaritmov hladiny bilirubínu a prítomnosti pavúčikovitých név na koži u oboch skupín pacientov na začiatku štúdie.
Medzi pacientmi, ktorí sa zúčastnili tohoto výskumu, bolo tridsaťšesť mužov a dvestosedemdesiatšesť žien. Vek týchto pacientov sa pohyboval od 26 až do 78 rokov. Ďalšie vybrané popisné štatistiky pre vek pacientov sú nasledujúce:

dolný kvartil ... 42 rokov
medián ... 50 rokov
horný kvartil ... 57 rokov.

Po zistení popisných štatistík a krabičkových grafov pre hladinu bilirubínu u oboch skupín pacientov sme zvážili, že dané rozdelenia nemôžeme považovať za normálne. Z tohoto dôvodu sme sa rozhodli namiesto ekvivalencie hladín bilirubínu testovať ekvivalenciu logaritmov hladiny bilirubínu, ktorých rozdelenie bolo viac podobné normálnemu.
Najprv vidíme krabičkové grafy pre hladiny bilirubínu u oboch skupín pacientov, potom sú zobrazené krabičkové grafy pre logaritmy hladiny bilirubínu pre obe skupiny pacientov.





Popisné štatistiky pre logaritmy hladiny bilirubínu sú nasledujúce

skupina pacientov, ktorým bude podávaný D-penicilamín:
minimum ... -1.204
dolný kvartil ... -0.223
medián ... 0.337
horný kvartil ... 1.163
maximum ... 2.996,

skupina pacientov, ktorým bude podávané placebo:
minimum ... -1.204
dolný kvartil ... -0.323
medián ... 0.262
horný kvartil ... 1.281
maximum ... 3.332.

Prítomnosť pavúčikovitých név bola zistená na koži štyridsiatich piatich pacientov v jednej ako aj v druhej skupine.

Príklad 1. Chceme porovnať logaritmy hladiny bilirubínu dvoch skupín pacientov (jednej bude podávaný D-penicilamín a druhej placebo) na začiatku štúdie.

Máme dva náhodné výbery $(X_1, \dots, X_{158})$ a $(Y_1, \dots, Y_{154})$, kde X_i je hodnota logaritmu hladiny bilirubínu u pacienta, ktorý bude liečený D-penicilamínom, a Y_j je hodnota logaritmu hladiny bilirubínu pacienta, ktorý bude dostávať placebo. Predpokladáme, že $X_i \sim \mathrm{N}(\mu_X, \sigma^2)$ a $Y_j \sim \mathrm{N}(\mu_Y, \sigma^2)$. Vzhľadom na náhodný výber pacientov môžeme predpokladať nezávislosť X_i a Y_j pre každé i a j.

Testujeme

$$H_0 : \mu_X - \mu_Y = 0 \qquad \text{vs.} \qquad H_1 : \mu_X - \mu_Y \neq 0.$$

Keďže predpokladáme, že obidva náhodné výbery majú normálne rozdelenie, rovnaký rozptyl, a že sú navzájom nezávislé, použijeme dvojvýberový t-test. Ten vyjde s p-hodnotou 0.515, čiže na hladine 0.05 nemôžeme zamietnuť nulovú hypotézu. Z toho však neplynie, že $\mu_X = \mu_Y$. Potvrdenie tejto rovnosti bolo motiváciou vzniku testov ekvivalencie.

Príklad 2. Chceme porovnať výskyt pavúčikovitých név na koži dvoch skupín pacientov s cirhózou pečene (jednej bude podávaný D-penicilamín a druhej placebo) na začiatku štúdie.

Máme dva náhodné výbery $(X_1, \dots, X_{158})$ a $(Y_1, \dots, Y_{154})$, kde X_i má hodnotu 0, ak sa pavúčikovité névy na koži pacienta, ktorý bude liečený D-penicilamínom, nevyskytujú; v opačnom prípade má hodnotu 1. Y_j má hodnotu 0, ak sa pavúčikovité névy na koži pacienta, ktorý dostane placebo, nevyskytujú; v opačnom prípade má hodnotu 1. Predpokladáme, že $X_i \sim \mathrm{Alt}(p_X)$ a $Y_j \sim \mathrm{Alt}(p_Y)$. Vzhľadom na náhodný výber pacientov môžeme predpokladať nezávislosť X_i a Y_j pre každé i a j.

Testujeme

$$H_0 : p_X - p_Y = 0 \qquad \text{vs.} \qquad H_1 : p_X - p_Y \neq 0.$$

Keďže sme predpokladali alternatívne rozdelenie oboch náhodných výberov a ich vzájomnú nezávislosť, použijeme asymptotický dvojvýberový test o proporciách. Ten vyjde s p-hodnotou 0.985, čiže na hladine 0.05 nemôžeme zamietnuť nulovú hypotézu. Opäť sa však nedá tvrdiť, že $p_X = p_Y$.

Vidíme, že nulovú hypotézu nemôžeme pomocou klasického testu potvrdiť. Toto je dôvod pre existenciu testov ekvivalencie.

Riešenie pomocou testu ekvivalencie

Na to, aby sme boli schopní potvrdiť hypotézu $\theta = \theta_0$, potrebujeme testovať

$$H_0 : \theta \neq \theta_0 \qquad \text{vs.} \qquad H_1 : \theta = \theta_0.$$

Ukážeme však, že takto test ekvivalencie nemôžeme vytvoriť.

Keďže sila testu

$$\beta(\theta) = P_\theta[\text{ zamietneme } H_0 \text{ }]$$

je spojitá funkcia parametru θ a chceme, aby pre hladinu testu platilo

$$\sup_{\theta \neq \theta_0} \beta(\theta) = \sup_{\theta \neq \theta_0} P_\theta[\text{ zamietneme } H_0 \text{ }] = \alpha,$$

potom dostaneme, že

$$\beta(\theta_0) = P_{\theta=\theta_0}[\text{ zamietneme } H_0 \text{ }] = \alpha.$$

Teda maximálna dosažiteľná sila testu by bola taktiež α.

Z tohoto dôvodu testujeme

$$H_0 : \theta \leq \theta_0 - \varepsilon_1 \;\text{ alebo }\; \theta \geq \theta_0 + \varepsilon_2 \tag{1.2a}$$

vs.

$$H_1 : \theta_0 - \varepsilon_1 < \theta < \theta_0 + \varepsilon_2, \tag{1.2b}$$

kde $(\theta_0 - \varepsilon_1, \theta_0 + \varepsilon_2)$ predstavuje interval, do ktorého patriace θ považujeme za ekvivalentné s θ_0. ε_1 a ε_2 sú kladné konštanty, ktoré sú zadané spolu s problémom a určujú aké hodnoty z okolia θ_0 sú ešte považované za ekvivalentné s θ_0. Hodnoty týchto konštánt je nutné konzultovať s odborníkom z oboru, z ktorého pochádzajú dáta.

Príklad 1*. Chceme zistiť, či logaritmy hladiny bilirubínu dvoch skupín pacientov (jednej bude podávaný D-penicilamín a druhej placebo) sú na začiatku štúdie ekvivalentné.

Máme dva náhodné výbery $(X_1, \dots, X_{158})$ a $(Y_1, \dots, Y_{154})$, kde X_i je hodnota logaritmu hladiny bilirubínu u pacienta, ktorý bude liečený D-penicilamínom, a Y_j je hodnota logaritmu hladiny bilirubínu pacienta, ktorý bude dostávať placebo. Predpokladáme, že $X_i \sim \mathrm{N}(\mu_X, \sigma^2)$ a $Y_j \sim \mathrm{N}(\mu_Y, \sigma^2)$. Vzhľadom na náhodný výber pacientov môžeme predpokladať nezávislosť X_i a Y_j pre každé i a j.

Budeme testovať

$$H_0 : \mu_X - \mu_Y \leq -\varepsilon_1 \text{ alebo } \mu_X - \mu_Y \geq \varepsilon_2 \quad \text{vs.} \quad H_1 : -\varepsilon_1 < \mu_X - \mu_Y < \varepsilon_2.$$

Príklad 2*. Chceme zistiť, či je výskyt pavúčikovitých név na koži dvoch skupín pacientov s cirhózou pečene (jednej bude podávaný D-penicilamín a druhej placebo) na začiatku štúdie ekvivalentný.

Máme dva náhodné výbery $(X_1, \dots, X_{158})$ a $(Y_1, \dots, Y_{154})$, kde X_i má hodnotu 0, ak sa pavúčikovité névy na koži pacienta, ktorý bude liečený D-penicilamínom, nevyskytujú; v opačnom prípade má hodnotu 1. Y_j má hodnotu 0, ak sa pavúčikovité névy na koži pacienta, ktorý dostane placebo, nevyskytujú; v opačnom prípade má hodnotu 1. Predpokladáme, že $X_i \sim \mathrm{Alt}(p_X)$ a $Y_j \sim \mathrm{Alt}(p_Y)$. Vzhľadom na náhodný výber pacientov môžeme predpokladať nezávislosť X_i a Y_j pre každé i a j.

Budeme testovať

$$H_0 : p_X - p_Y \leq -\varepsilon_1 \text{ alebo } p_X - p_Y \geq \varepsilon_2 \quad \text{vs.} \quad H_1 : -\varepsilon_1 < p_X - p_Y < \varepsilon_2.$$

Keďže nemáme žiadne klinicky zdôvodnené hodnoty konštánt ε_1 a ε_2, tak v ďalších ukážkach budeme používať nami zvolené hodnoty. V prípade párového a dvojvýberového t-testu sme pri určení týchto konštánt využili doporučenia uvedené v tabuľke v [1, strana 12].

Kapitola 2

Princíp inklúzie intervalov spoľahlivosti

Prvou možnosťou, ako pristupovať k testovaniu ekvivalencie, sú testy založené na princípe inklúzie intervalov spoľahlivosti. Tieto testy riešia problém ekvivalencie (1.2a) vs. (1.2b) pomocou testovania dvoch jednostranných hypotéz

$$H_0 : \theta \leq \theta_0 - \varepsilon_1 \qquad \text{vs.} \qquad H_1 : \theta > \theta_0 - \varepsilon_1 \tag{2.1a}$$

a

$$H_0 : \theta \geq \theta_0 + \varepsilon_2 \qquad \text{vs.} \qquad H_1 : \theta < \theta_0 + \varepsilon_2. \tag{2.1b}$$

Odvodenie testu

Navrhnutý test ekvivalencie rozhodne v prospech ekvivalencie (1.2b), tj. zamietne nulovú hypotézu (1.2a), jedine v prípade, že oba testy, ktoré riešia problémy (2.1a) a (2.1b), zároveň zamietnu nulovú hypotézu.
S testovacím problémom (2.1a) súvisí $(1-\alpha)$%-ný jednostranný dolný interval spoľahlivosti pre θ : $(\underline{\theta}, \infty)$. Nulovú hypotézu zamietame, ak $\theta_0 - \varepsilon_1 \notin (\underline{\theta}, \infty) \Leftrightarrow \theta_0 - \varepsilon_1 < \underline{\theta}$.
Pri testovacom probléme (2.1b) odvodíme $(1-\alpha)$%-ný jednostranný horný interval spoľahlivosti pre θ, ktorý je $(-\infty, \overline{\theta})$. Ak $\theta_0 + \varepsilon_2 \notin (-\infty, \overline{\theta}) \quad \Leftrightarrow \quad \theta_0 + \varepsilon_2 > \overline{\theta}$, tak zamietame nulovú hypotézu.
Z tohoto potom vyplýva, že výsledný test ekvivalencie zamietne nulovú hypotézu, ak $\theta_0 - \varepsilon_1 < \underline{\theta} \quad \wedge \quad \theta_0 + \varepsilon_2 > \overline{\theta}$; teda ak $(\underline{\theta}, \overline{\theta}) \in (\theta_0 - \varepsilon_1, \theta_0 + \varepsilon_2)$. Interval spoľahlivosti súvisiaci s týmto testom vznikne prienikom jednostranných intervalov spoľahlivosti $(\underline{\theta}, \infty) \cap (-\infty, \overline{\theta}) = (\underline{\theta}, \overline{\theta})$. Vidíme, že je $(1-2\alpha)$%-ný.

Tvrdenie. *Výsledný test ekvivalencie odvodený na základe predchádzajúcich úvah má hladinu α.*

Dôkaz. Test má hladinu α práve vtedy, keď pravdepodobnosť, že zamietne nulovú hypotézu za podmienky jej platnosti je menšia alebo rovná α.
Nech parameter θ spĺňa nulovú hypotézu (1.2a), teda

1. nech $\theta \leq \theta_0 - \varepsilon_1$. Potom

$$\begin{aligned} P_{H_0}((\underline{\theta}, \overline{\theta}) \in (\theta_0 - \varepsilon_1, \theta_0 + \varepsilon_2)) &= P_{\theta \leq \theta_0 - \varepsilon_1}(\theta_0 - \varepsilon_1 < \underline{\theta} \ \wedge \ \overline{\theta} < \theta_0 + \varepsilon_2)) \\ &= P_{\theta \leq \theta_0 - \varepsilon_1}(\theta_0 - \varepsilon_1 < \underline{\theta}) \\ &= P_{\theta \leq \theta_0 - \varepsilon_1}(\theta < \underline{\theta}) = \alpha \end{aligned}$$

   z (2.1a).

2. nech $\theta \geq \theta_0 + \varepsilon_2$. Potom

$$\begin{aligned} P_{H_0}((\underline{\theta}, \overline{\theta}) \in (\theta_0 - \varepsilon_1, \theta_0 + \varepsilon_2)) &= P_{\theta \geq \theta_0 + \varepsilon_2}(\theta_0 - \varepsilon_1 < \underline{\theta} \ \wedge \ \overline{\theta} < \theta_0 + \varepsilon_2)) \\ &= P_{\theta \geq \theta_0 + \varepsilon_2}(\overline{\theta} < \theta_0 + \varepsilon_2) \\ &= P_{\theta \geq \theta_0 + \varepsilon_2}(\overline{\theta} < \theta) = \alpha \end{aligned}$$

   z (2.1b).

$\square$

Príklad 3. Na riešenie problému ekvivalencie z príkladu 1* použijeme test založený na princípe inklúzie intervalov spoľahlivosti.

Príslušné dva jednostranné problémy sú nasledujúce

$$H_0 : \mu_X - \mu_Y \leq -\varepsilon_1 \qquad \text{vs.} \qquad H_1 : \mu_X - \mu_Y > -\varepsilon_1$$

a

$$H_0 : \mu_X - \mu_Y \geq \varepsilon_2 \qquad \text{vs.} \qquad H_1 : \mu_X - \mu_Y < \varepsilon_2.$$

Keďže predpokladáme normálne rozdelenie oboch náhodných výberov, ich vzájomnú nezávislosť a zhodný rozptyl, budeme oba jednostranné problémy testovať pomocou dvojvýberového t-testu. Zvolíme $\varepsilon_1 = 0.3$ a $\varepsilon_2 = 0.15$.
V prvom prípade dvojvýberový t-test vyjde s p-hodnotou 0.029, v druhom 0.027. Takže na hladine $\alpha = 0.05$ oba testy zamietnu nulovú hypotézu, teda test ekvivalencie zamietne nulovú hypotézu a my môžeme potvrdiť, že platí $-0.3 < \mu_X - \mu_Y < 0.15$.

Parametre ovplyvňujúce výsledok testu

Existujú situácie, v ktorých testy založené na princípe inklúzie intervalov spoľahlivosti nikdy nebudú schopné zamietnuť nulovú hypotézu. Takéto situácie nastávajú, ak zadané konštanty ε_1 a ε_2 sú príliš malé, rozsah náhodných výberov je nedostatočný, alebo ak testujeme na príliš nízkej hladine. Ako tieto parametre ovplyvňujú výsledok testu, ukážeme na nasledujúcom príklade.

Majme $(X_1, \dots, X_m)$ náhodný výber z rozdelenia F_X, $(Y_1, \dots, Y_n)$ náhodný výber z rozdelenia F_Y, ktoré sú navzájom nezávislé.
Model je $\mathcal{F} = \{F_X \sim \text{Alt}(p_X),\ F_Y \sim \text{Alt}(p_Y),\ p_X, p_Y \in (0,1)\}$.

Testujeme

$$H_0 : p_X - p_Y \leq\ -\varepsilon_1 \ \text{ alebo } \ p_X - p_Y \geq \varepsilon_2 \qquad \text{vs.} \qquad H_1 : -\varepsilon_1 < p_X - p_Y < \varepsilon_2.$$

Najprv odvodíme jednostranné intervaly spoľahlivosti súvisiace s doplnkovými testami:

1. Testujeme

   $$H_0 : p_X - p_Y \leq\ -\varepsilon_1 \qquad \text{vs.} \qquad H_1 : p_X - p_Y > -\varepsilon_1.$$

   Odvodíme dolný interval spoľahlivosti pre $p_X - p_Y$:

   z vlastností alternatívneho rozdelenia platí:

   $$\begin{aligned} \text{E}X_i &= p_X \\ \text{var}X_i &= p_X(1 - p_X). \end{aligned}$$

   Podľa Centrálnej limitnej vety

   $$\begin{aligned} \sqrt{m}(\overline{X} - p_X) &\xrightarrow{\mathcal{D}} \text{N}(0, p_X(1 - p_X)),\ m \to \infty \\ \frac{\sqrt{m}(\overline{X} - p_X)}{\sqrt{p_X(1 - p_X)}} &\xrightarrow{\mathcal{D}} \text{N}(0,1),\ m \to \infty, \end{aligned}$$

   kde $\overline{X} = \dfrac{\sum_{i=1}^{m} X_i}{m}$ je výberový priemer.

   Obdobne ukážeme, že

   $$\frac{\sqrt{n}(\overline{Y} - p_Y)}{\sqrt{p_Y(1 - p_Y)}} \xrightarrow{\mathcal{D}} \text{N}(0,1),\ n \to \infty,$$

kde $\overline{Y} = \dfrac{\sum_{i=1}^{n} Y_i}{n}$ je výberový priemer.

Vďaka nezávislosti $\mathbf{X}$ a $\mathbf{Y}$:

$$\frac{\overline{X} - \overline{Y} - (p_X - p_Y)}{\sqrt{\dfrac{p_X(1-p_X)}{m} + \dfrac{p_Y(1-p_Y)}{n}}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1), \ m, n \to \infty.$$

Keďže zo Zákona veľkých čísel

$$\begin{aligned} \overline{X} &\xrightarrow{P} \mathrm{E}X_i = p_X, \ m \to \infty \\ \overline{Y} &\xrightarrow{P} \mathrm{E}Y_i = p_Y, \ n \to \infty, \end{aligned}$$

tak

$$\frac{\sqrt{\dfrac{p_X(1-p_X)}{m} + \dfrac{p_Y(1-p_Y)}{n}}}{\sqrt{\dfrac{\overline{X}(1-\overline{X})}{m} + \dfrac{\overline{Y}(1-\overline{Y})}{n}}} \xrightarrow{P} 1, \ m, n \to \infty.$$

Zo Sluckého vety potom vyplýva, že

$$\frac{\overline{X} - \overline{Y} - (p_X - p_Y)}{\sqrt{\dfrac{p_X(1-p_X)}{m} + \dfrac{p_Y(1-p_Y)}{n}}} \frac{\sqrt{\dfrac{p_X(1-p_X)}{m} + \dfrac{p_Y(1-p_Y)}{n}}}{\sqrt{\dfrac{\overline{X}(1-\overline{X})}{m} + \dfrac{\overline{Y}(1-\overline{Y})}{n}}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1), \ m, n \to \infty.$$

Teda

$$\frac{\overline{X} - \overline{Y} - (p_X - p_Y)}{\sqrt{\dfrac{\overline{X}(1-\overline{X})}{m} + \dfrac{\overline{Y}(1-\overline{Y})}{n}}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1), \ m, n \to \infty.$$

Odtiaľ teda platí

$$P\left[\frac{\overline{X} - \overline{Y} - (p_X - p_Y)}{\sqrt{\dfrac{\overline{X}(1-\overline{X})}{m} + \dfrac{\overline{Y}(1-\overline{Y})}{n}}} < u_{1-\alpha}\right] =$$

$$P\left[\overline{X} - \overline{Y} - u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})}{m} + \frac{\overline{Y}(1-\overline{Y})}{n}} < p_X - p_Y\right] \longrightarrow 1-\alpha, \ m, n \to \infty.$$

Z predchádzajúcej rovnice vidíme, že približný $(1-\alpha)$%-ný jednostranný dolný interval spoľahlivosti pre $p_X - p_Y$ je

$$\left(\overline{X} - \overline{Y} - u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})}{m} + \frac{\overline{Y}(1-\overline{Y})}{n}},\ \infty\right).$$

2. Testujeme

$$H_0 : p_X - p_Y \geq \varepsilon_2 \qquad \text{vs.} \qquad H_1 : p_X - p_Y < \varepsilon_2.$$

Podobne ako v prvom prípade odvodíme horný interval spoľahlivosti:

$$P\left[-u_{1-\alpha} < \frac{\overline{X} - \overline{Y} - (p_X - p_Y)}{\sqrt{\frac{\overline{X}(1-\overline{X})}{m} + \frac{\overline{Y}(1-\overline{Y})}{n}}}\right] =$$
$$P\left[p_X - p_Y < \overline{X} - \overline{Y} + u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})}{m} + \frac{\overline{Y}(1-\overline{Y})}{n}}\right] \longrightarrow 1-\alpha,\ m, n \to \infty.$$

Potom vidíme, že $(1-\alpha)$%-ný jednostranný horný interval spoľahlivosti pre $p_X - p_Y$ je

$$\left(-\infty,\ \overline{X} - \overline{Y} + u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})}{m} + \frac{\overline{Y}(1-\overline{Y})}{n}}\right).$$

Výsledný test ekvivalencie teda zamietne nulovú hypotézu, ak

$$-\varepsilon_1 < \overline{X} - \overline{Y} - u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})}{m} + \frac{\overline{Y}(1-\overline{Y})}{n}}$$

a zároveň

$$\overline{X} - \overline{Y} + u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})}{m} + \frac{\overline{Y}(1-\overline{Y})}{n}} < \varepsilon_2.$$

Pozrime sa bližšie na situáciu, keď $\varepsilon_1 = \varepsilon_2 = \varepsilon$ a $m = n$. Nulovú hypotézu potom zamietame ak

$$|\overline{X} - \overline{Y}| < \varepsilon - u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X}) + \overline{Y}(1-\overline{Y})}{n}},$$

tj. kritický obor má tvar

$$\left\{ (\mathbf{X},\mathbf{Y}):\ |\overline{X}-\overline{Y}| < \varepsilon - u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})+\overline{Y}(1-\overline{Y})}{n}} \right\}.$$

Aby bol kritický obor neprázdna množina, je treba, aby platilo

$$0 < \varepsilon - u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})+\overline{Y}(1-\overline{Y})}{n}}.$$

1. Nech α, n pevné, potom musí platiť

$$\varepsilon > u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})+\overline{Y}(1-\overline{Y})}{n}}.$$

2. Nech α, ε pevné, potom musí platiť

$$n > \left(\frac{u_{1-\alpha}\left(\overline{X}(1-\overline{X})+\overline{Y}(1-\overline{Y})\right)}{\varepsilon}\right)^2.$$

Vidíme, že podmienka je splnená, ak je ε alebo n dostatočne veľké. V prípade, že je ε alebo n malé, kritický obor môže byť prázdny. K obdobnému záveru sa dostaneme, aj keď $m \neq n$.
Z predchádzajúceho zistenia vyplýva, že zvolenie príliš malého ε (teda ak odborník z daného oboru kladie príliš veľké nároky na interval ekvivalencie), môže viesť k testu s nulovou silou.

Príklad 4. Pri riešení problému ekvivalencie z príkladu 2* použijeme test založený na inklúzii intervalov spoľahlivosti.

Keďže predpokladáme, že oba náhodné výbery majú alternatívne rozdelenie a sú navzájom nezávislé, použijeme test, ktorý sme v tejto kapitole odvodili. Zvolíme $\varepsilon_1 = \varepsilon_2 = \varepsilon$. Potom nulovú hypotézu zamietame ak

$$|\overline{X}-\overline{Y}| < \varepsilon - u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})}{m}+\frac{\overline{Y}(1-\overline{Y})}{n}},$$

a aby bol kritický obor neprázdna množina, je treba, aby platilo

$$0 < \varepsilon - u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})}{m}+\frac{\overline{Y}(1-\overline{Y})}{n}}.$$

Budeme testovať na hladine $\alpha = 0.05$. Chceme ukázať závislosť výsledku testu na ε

pri zvolenom α a danom množstve pozorovaní. Postupne budeme robiť test pre rôzne $\varepsilon = 0.01, 0.02, 0.03, ..., 0.1$.

Na základe týchto skutočností je vytvorená nasledujúca tabuľka.

ε	$\lvert\overline{X}-\overline{Y}\rvert$	M	kritický obor	výsledok testu
0.01	0.007	-0.074	prázdna množina	nezamietam H_0
0.02	0.007	-0.064	prázdna množina	nezamietam H_0
0.03	0.007	-0.054	prázdna množina	nezamietam H_0
0.04	0.007	-0.044	prázdna množina	nezamietam H_0
0.05	0.007	-0.034	prázdna množina	nezamietam H_0
0.06	0.007	-0.024	prázdna množina	nezamietam H_0
0.07	0.007	-0.014	prázdna množina	nezamietam H_0
0.08	0.007	-0.004	prázdna množina	nezamietam H_0
0.09	0.007	0.006	neprázdna množina	nezamietam H_0
0.1	0.007	0.016	neprázdna množina	zamietam H_0

,

kde

$$M = \varepsilon - u_{1-\alpha}\sqrt{\frac{\overline{X}(1-\overline{X})}{m} + \frac{\overline{Y}(1-\overline{Y})}{n}}.$$

Kapitola 3

Optimálne testy ekvivalencie

Obmedzíme sa na prípady, kde testová štatistika má presne, prípadne asymptoticky, normálne alebo t-rozdelenie.

Určenie kritického oboru

Pri testovaní klasického problému (1.1) nulovú hypotézu zamietame, ak testová štatistika $T(\mathbf{X})$ náhodného výberu $\mathbf{X}$ spĺňa $T(\mathbf{X}) < C_1$ alebo $T(\mathbf{X}) > C_2$, kde C_1 a C_2 sú kritické hodnoty. Vzhľadom na to, že test ekvivalencie predstavuje akýsi protiklad ku klasickému testu, tak intuitívne by sme mali nulovú hypotézu (1.2a) zamietnuť, ak pre testovú štatistiku platí $C_1 < T(\mathbf{X}) < C_2$. Určíme teraz kritické hodnoty tak, aby hladina testu bola α. Nech θ je parameter spĺňajúci nulovú hypotézu, teda nech $\theta \leq \theta_0 - \varepsilon_1$ alebo $\theta \geq \theta_0 + \varepsilon_2$. Pre rozdelenia, ktorými sa budeme v tejto práci zaoberať, platí, že čím je θ ďalej od dolnej alebo hornej hranice nulovej hypotézy s alternatívou, teda čím viac θ spĺňa nulovú hypotézu, tým nižšia je sila testu pre dané θ. Z toho vyplýva, že ak chceme, aby hladina testu bola α, tak musí platiť, že sila testu pre θ na oboch hraniciach nulovej hypotézy musí byť α. Budeme sa zaoberať prípadom, keď testová štatistika $T(\mathbf{X})$ má spojité rozdelenie a teda môžeme písať:

$$P_{\theta_1}[C_1 < T(\mathbf{X}) < C_2] = \alpha = P_{\theta_2}[C_1 < T(\mathbf{X}) < C_2], \tag{3.1}$$

$\theta_1 = \theta_0 - \varepsilon_1$, $\theta_2 = \theta_0 + \varepsilon_2$ a $C_1, C_2 \in \tau$, $C_1 < C_2$, kde τ je obor hodnôt testovej štatistiky $T(\mathbf{X})$.

Kritické hodnoty C_1 a C_2 môžeme určiť numericky ako riešenie predchádzajúcej rovnice na základe nasledujúceho tvrdenia.

Tvrdenie. *Nech pravdepodobnostné rozdelenie testovej štatistiky $T(\boldsymbol{X})$ patrí do ex-*

ponenciálnej triedy rozdelení a nech (C_1', C_2') je ľubovoľný interval, pre ktorý platí $P_{\theta_1}[C_1' < T(\boldsymbol{X}) < C_2'] = \alpha$, ale $P_{\theta_2}[C_1' < T(\boldsymbol{X}) < C_2'] < \alpha$ (alebo $> \alpha$). Potom optimálny kritický obor (C_1, C_2), ktorého kritické hodnoty spĺňajú (3.1), leží napravo (alebo naľavo) od kritického oboru (C_1', C_2').

Príslušný iteračný proces vyzerá nasledovne:

1. zvoľme C_1^0, ktoré je odhadom C_1.
2. Určime C_2^0, tak aby $P_{\theta_1}[C_1^0 < T(\mathbf{X}) < C_2^0] = \alpha$. Teda $C_2^0 = F_{\theta_1}^{-1}[\alpha + F_{\theta_1}(C_1^0)]$, kde F_θ je distribučná funkcia a F_θ^{-1} je kvantilová funkcia testovej štatistiky $T(\mathbf{X})$.
3. Vypočítajme $\alpha_0^2 = P_{\theta_2}[C_1^0 < T(\mathbf{X}) < C_2^0] = F_{\theta_2}(C_2^0) - F_{\theta_2}(C_1^0)$.
4. Za C_1^0 dosaďme buď $\dfrac{\underline{C_1} + C_1^0}{2}$, alebo $\dfrac{C_1^0 + \overline{C_1}}{2}$ podľa toho, či $\alpha_2^0 > \alpha$ alebo $\alpha_2^0 < \alpha$. Interval $(\underline{C_1}, \overline{C_1})$ bol na začiatku iterácie zvolený tak, aby obsahoval C_1^0 a zároveň C_1.
5. Opakujme kroky (1)-(4) až kým $|\alpha_2^0 - \alpha|$ <tolerancia.

Viz [1].

Prípad symetrie

Vidíme, že presné numerické vyčíslenie kritických hodnôt C_1 a C_2 je zložitejšie, ako u klasických testov. Zjednodušenie nastáva v prípade, keď je zaručená symetria:

1. $\theta_1 = -\varepsilon,\ \theta_2 = \varepsilon$, kde $\varepsilon > 0$ a
2. rozdelenie testovej štatistiky $T(\mathbf{X})$ za podmienky $\theta = \varepsilon$ je rovnaké ako rozdelenie testovej štatistiky $-T(\mathbf{X})$ za podmienky $\theta = -\varepsilon$.

Potom platí, že kritický obor optimálneho testu ekvivalencie je symetrický okolo nuly: $(-C, C)$, a ak testová štatistika $T(\mathbf{X})$ má spojité rozdelenie môžeme písať, že C je riešením rovnice

$$P_\varepsilon[|T(\mathbf{X})| < C] = \alpha, \tag{3.2}$$

kde $C \in \tau \cap (0, \infty)$.

Neymanova-Pearsonova lema

V tejto časti na základe Neymanovej-Pearsonovej lemy odvodíme optimálny test pre strednú hodnotu normálneho rozdelenia so známym rozptylom.

Neymanova-Pearsonova lema. *Nech $f_j(\boldsymbol{x})$, $\boldsymbol{x} \in \boldsymbol{R}^n$, $j = 0, 1$, sú hustoty pravdepodobnostných rozdelení a nech pre dané $\alpha \in (0, 1)$ existuje také kladné číslo c, že pre množinu*

$$W^* = \{\boldsymbol{x} \in \boldsymbol{R}^n;\ f_1(\boldsymbol{x}) \geq c f_0(\boldsymbol{x})\}$$

platí

$$\int_{W^*} f_0(\boldsymbol{x}) d\boldsymbol{x} = \alpha.$$

Potom pre ľubovoľnú množinu W, ktorá spĺňa

$$\int_{W} f_0(\boldsymbol{x}) d\boldsymbol{x} \leq \alpha,$$

platí

$$\int_{W^*} f_1(\boldsymbol{x}) d\boldsymbol{x} \geq \int_{W} f_1(\boldsymbol{x}) d\boldsymbol{x}.$$

Dôkaz. Viz [2]. □

Pozn. Nech $\mathbf{X} = (X_1, ..., X_n)$ je náhodný výber z rozdelenia F_θ a testujeme $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$. Potom predchádzajúce tvrdenie nám hovorí, že najsilnejším testom je taký, ktorý má kritický obor W^*, kde $f_0(\mathbf{x})$ a $f_1(\mathbf{x})$ sú hustoty rozdelenia náhodného výberu $\mathbf{X}$ za H_0 a H_1.

Pozn. Nech $\mathbf{X} = (X_1, ..., X_n)$ je náhodný výber z rozdelenia s hustotou $f(\mathbf{x}, \theta)$, $\theta \in \Theta_0 \cup \Theta_1$, a testujeme $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$. Potom optimálny test má kritický obor

$$W^* = \{\mathbf{x} \in \mathbf{R}^n;\ \sup_{\theta \in \Theta_0 \cup \Theta_1} f(\mathbf{x}, \theta) \geq c \sup_{\theta \in \Theta_0} f(\mathbf{x}, \theta)\}, \tag{3.3}$$

pre ktorý platí

$$\forall \theta \in \Theta_0 \qquad \int_{W^*} f(\mathbf{x}, \theta) d\mathbf{x} \leq \alpha, \tag{3.4}$$

ako bolo uvedené v [2].

Príklad. Majme $(X_1, ..., X_n)$ náhodný výber z rozdelenia $F_X \in \mathcal{F}$.
Model je $\mathcal{F} = \{\mathrm{N}(\mu, \sigma^2),\ \mu \in \mathbf{R},\ \sigma^2 > 0 \text{ známa konštanta}\}$.

Testujeme

$$H_0 : \mu \leq \mu_0 - \varepsilon_1 \ \text{ alebo } \ \mu \geq \mu_0 + \varepsilon_2 \qquad \text{vs.} \qquad H_1 : \mu_0 - \varepsilon_1 < \mu < \mu_0 + \varepsilon_2.$$

Určíme kritický obor optimálneho testu podľa predchádzajúcej poznámky:

$$\begin{aligned} \theta &= \mu \\ \Theta_0 &= (-\infty, \mu_0 - \varepsilon_1\rangle \cup \langle \mu_0 + \varepsilon_2, \infty) \\ \Theta_1 &= (\mu_0 - \varepsilon_1, \mu_1 + \varepsilon_2) \\ \Theta &= \Theta_0 \cup \Theta_1 = \mathbf{R} \end{aligned}$$

$$\begin{aligned} f(\mathbf{x}, \mu) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right) \\ &= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2}\right). \end{aligned}$$

Potom

$$\log f(\mathbf{x}, \mu) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

Pre zjednodušenie nasledujúcich výpočtov zadefinujeme funkcie

$$\begin{aligned} l(\mathbf{x}) &= \sup_{\mu \in \Theta} \log f(\mathbf{x}, \mu) \\ l_0(\mathbf{x}) &= \sup_{\mu \in \Theta_0} \log f(\mathbf{x}, \mu) \\ L(\mathbf{x}) &= l(\mathbf{x}) - l_0(\mathbf{x}). \end{aligned}$$

Kritický obor (3.3) môžeme s využitím predchádzajúcich funkcií zapísať ako

$$W^* = \{\mathbf{x} \in \mathbf{R}^n; L(\mathbf{x}) \geq c^*\}.$$

Vzhľadom na to, že $\sum_{i=1}^{n}(x_i - \mu)^2$ je kvadratická funkcia premennej μ s minimom v bode

$$\begin{aligned} \frac{\partial \sum_{i=1}^{n}(x_i - \mu)^2}{\partial \mu} &= 0 \\ \sum_{i=1}^{n} 2(\mu - x_i) &= 0 \\ \mu &= \frac{\sum_{i=1}^{n} x_i}{n} \\ \mu &= \overline{x}, \end{aligned}$$

tak $f(\mathbf{x},\mu)$ nadobúda supremum pre $\mu \in \Theta$ v $\overline{x}$, teda funkcia $l(\mathbf{x})$ je

$$l(\mathbf{x}) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Keďže $\sup_{\mu\in\Theta_0} f(\mathbf{x},\mu)$ závisí na polohe $\overline{x}$, určíme priebeh $l_0(\mathbf{x})$ nasledovne

1. ak $\mu_0 - \varepsilon_1 < \overline{x} \leq \mu_0 + \dfrac{\varepsilon_2 - \varepsilon_1}{2}$, tak

$$l_0(\mathbf{x}) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - (\mu_0 - \varepsilon_1))^2,$$

2. na intervale $\mu_0 + \dfrac{\varepsilon_2 - \varepsilon_1}{2} < \overline{x} < \mu_0 + \varepsilon_2$ je

$$l_0(\mathbf{x}) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - (\mu_0 + \varepsilon_2))^2,$$

3. a pre $\overline{x} \leq \mu_0 - \varepsilon_1$ alebo $\overline{x} \geq \mu_0 + \varepsilon_2$ platí

$$l_0(\mathbf{x}) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Odtiaľ môžeme určiť priebeh funkcie $L(\mathbf{x})$ na predchádzajúcich intervaloch

1.

$$\begin{aligned}
L(\mathbf{x}) &= -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \overline{x})^2 + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - (\mu_0 - \varepsilon_1))^2 \\
&= \frac{1}{2\sigma^2}(\sum_{i=1}^{n}(x_i - \overline{x} + \overline{x} - (\mu_0 - \varepsilon_1))^2 - \sum_{i=1}^{n}(x_i - \overline{x})^2) \\
&= \frac{1}{2\sigma^2}(\sum_{i=1}^{n}(x_i - \overline{x})^2 + n(\overline{x} - (\mu_0 - \varepsilon_1))^2 - \sum_{i=1}^{n}(x_i - \overline{x})^2) \\
&= \frac{n}{2\sigma^2}(\overline{x} - (\mu_0 - \varepsilon_1))^2,
\end{aligned}$$

2.

$$L(\mathbf{x}) = \frac{n}{2\sigma^2}(\overline{x} - (\mu_0 + \varepsilon_2))^2,$$

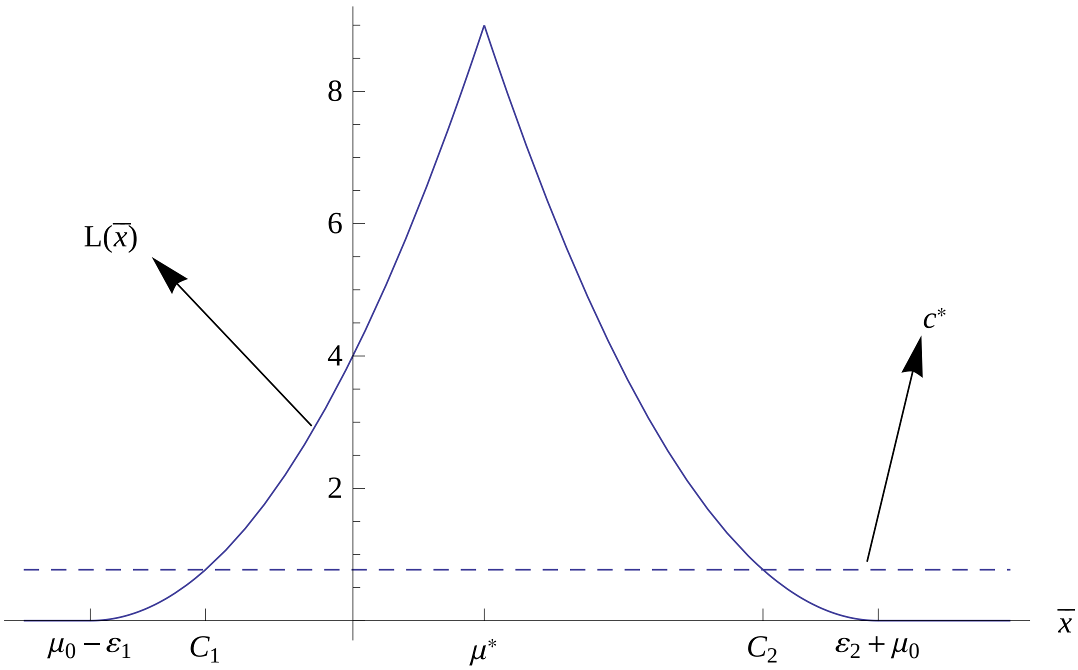
3.

$$L(\mathbf{x}) = 0.$$

Vidíme, že $L(\mathbf{x})$ je funkciou priemeru $\overline{x}$ (testovej štatistiky) a teda môžeme písať $L(\overline{x})$. Ako sme už z Neymanovej-Pearsonovej lemy ukázali, nulovú hypotézu zamietame ak

$$L(\overline{x}) \geq c^*.$$

Danú situáciu znázorňuje nasledujúci graf:



,

kde $\mu^* = \mu_0 + \dfrac{\varepsilon_2 - \varepsilon_1}{2}$.

Vidíme, že kritický obor je tvaru (C_1, C_2), ako sme heuristicky odvodili na začiatku kapitoly. Kritické hodnoty C_1 a C_2 zvolíme tak, aby bola splnená podmienka (3.4) pre hustotu pravdepodobnostného rozdelenia priemeru

$$\forall \mu \in \Theta_0 \quad \int_{C_1}^{C_2} f(\overline{x}, \mu) d\overline{x} \leq \alpha.$$

Ak zvolíme konkrétne $\mu \in \Theta_0$, tak potom má priemer $\overline{X}$ rozdelenie $\mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Čím je μ ďalej od jednej alebo druhej hranice nulovej hypotézy s alternatívou, tak tým má integrál $\int_{C_1}^{C_2} f(\overline{x}, \mu) d\overline{x}$ menšiu hodnotu; ako ukážeme v kapitole 4.1. Preto na nájdenie kritických hodnôt stačí splniť podmienku, že na oboch hraniciach nulovej hypotézy s alternatívou, teda pre $\mu = \mu_0 - \varepsilon_1$, $\mu = \mu_0 + \varepsilon_2$ má tento integrál hodnotu α:

$$\int_{C_1}^{C_2} f(\overline{x}, \mu_0 - \varepsilon_1) d\overline{x} = \alpha = \int_{C_1}^{C_2} f(\overline{x}, \mu_0 + \varepsilon_2) d\overline{x}.$$

Tá po uvážení, že $f(\overline{x}, \mu_0 - \varepsilon_1)$ je hustota rozdelenia $\mathrm{N}\left(\mu_0 - \varepsilon_1, \frac{\sigma^2}{n}\right)$ a $f(\overline{x}, \mu_0 + \varepsilon_2)$ je hustota rozdelenia $\mathrm{N}\left(\mu_0 + \varepsilon_2, \frac{\sigma^2}{n}\right)$, vyzerá nasledovne

$$\begin{aligned} \Phi\left(\frac{C_2 - (\mu_0 - \varepsilon_1)}{\frac{\sigma}{\sqrt{n}}}\right) - \Phi\left(\frac{C_1 - (\mu_0 - \varepsilon_1)}{\frac{\sigma}{\sqrt{n}}}\right) &= \alpha \\ &= \Phi\left(\frac{C_2 - (\mu_0 + \varepsilon_2)}{\frac{\sigma}{\sqrt{n}}}\right) - \Phi\left(\frac{C_1 - (\mu_0 + \varepsilon_2)}{\frac{\sigma}{\sqrt{n}}}\right), \end{aligned}$$

kde $\Phi(x)$ je distribučná funkcia normálneho rozdelenia $\mathrm{N}(0, 1)$.

Kapitola 4

Testovanie vybraných konkrétnych problémov

V prvej časti tejto kapitoly odvodíme optimálne testy pre konkrétne vybrané situácie, v druhej časti ukážeme konštrukciu asymptotického testu ekvivalencie pri odvodení jednovýberového a dvojvýberového asymptotického testu o proporciách.

4.1 Test o strednej hodnote normálneho rozdelenia so známym rozptylom

Nech $(X_1, ..., X_n)$ je náhodný výber z rozdelenia $F_X \in \mathcal{F}$.
Model je $\mathcal{F} = \{\mathrm{N}(\mu, \sigma^2),\ \mu \in \mathbf{R},\ \sigma^2 > 0 \text{ známa konštanta}\}$.

Testujeme

$$H_0 : \mu \leq \mu_0 - \varepsilon \ \text{ alebo } \ \mu \geq \mu_0 + \varepsilon \qquad \text{vs.} \qquad H_1 : \mu_0 - \varepsilon < \mu < \mu_0 + \varepsilon.$$

Ako sme si mohli všimnúť v príklade na konci predchádzajúcej kapitoly, je vhodné zvoliť nasledujúcu testovú štatistiku

$$T(\mathbf{X}) = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}},$$

kde $\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ je výberový priemer.

Požadujeme, aby platilo

$$P_{\mu_0-\varepsilon}[C_1 < T(\mathbf{X}) < C_2] = \alpha = P_{\mu_0+\varepsilon}[C_1 < T(\mathbf{X}) < C_2],$$

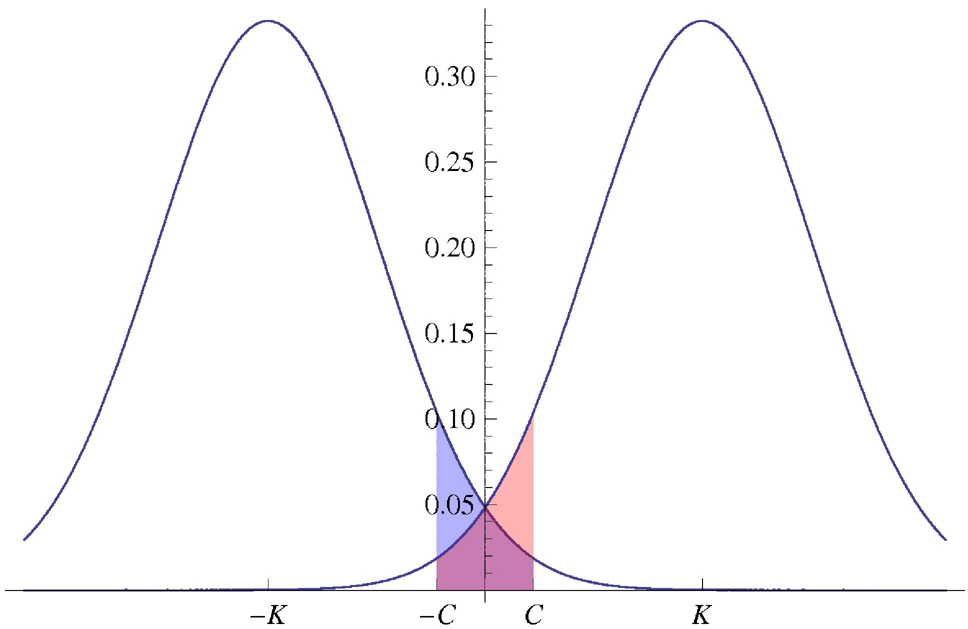
kde (C_1, C_2) je kritický obor.

Určíme rozdelenie $T(\mathbf{X})$ za podmienky $\mu = \mu_0 - \varepsilon$

$$T(\mathbf{X}) = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{\overline{X} - (\mu_0 - \varepsilon) - \varepsilon}{\frac{\sigma}{\sqrt{n}}} = Y - \frac{\varepsilon}{\sigma}\sqrt{n},$$

kde $Y \sim \mathrm{N}(0,1)$. Teda $T(\mathbf{X}) \sim \mathrm{N}\left(-\frac{\varepsilon}{\sigma}\sqrt{n}, 1\right)$.
Analogicky odvodíme, že rozdelenie $T(\mathbf{X})$ za podmienky $\mu = \mu_0 + \varepsilon$ je $\mathrm{N}\left(\frac{\varepsilon}{\sigma}\sqrt{n}, 1\right)$.

Zo symetrie normálneho rozdelenia plynie, že $T(\mathbf{X})$ má za podmienky $\mu = \mu_0 + \varepsilon$ rovnaké rozdelenie ako $-T(\mathbf{X})$ za podmienky $\mu = \mu_0 - \varepsilon$ a to $\mathrm{N}\left(\frac{\varepsilon}{\sigma}\sqrt{n}, 1\right)$. Keďže je zaručená aj symetrickosť intervalu ekvivalencie $(\mu_0 - \varepsilon, \mu_0 + \varepsilon)$, tak kritický obor je tvaru $(-C, C)$, $C > 0$. Môžeme to vidieť na nasledujúcom obrázku.



,

kde $K = \frac{\varepsilon}{\sigma}\sqrt{n}$.

Potom môžeme písať:

$$P_{\mu_0-\varepsilon}[|T(\mathbf{X})| < C] = P_{\mu_0-\varepsilon}[-C < T(\mathbf{X}) < C] =$$

(zo symetrie testovej štatistiky)

$$= P_{\mu_0+\varepsilon}[-C < -T(\mathbf{X}) < C] = P_{\mu_0+\varepsilon}[-C < T(\mathbf{X}) < C] = P_{\mu_0+\varepsilon}[|T(\mathbf{X})| < C].$$

Vidíme, že sa stačí zaoberať jednou hranicou nulovej hypotézy, napr. $\mu_0 + \varepsilon$.

Ukážeme, že pre $\varepsilon^* \geq \varepsilon$ je

$$P_{\mu_0+\varepsilon*}[-C < T(\mathbf{X}) < C] \leq P_{\mu_0+\varepsilon}[-C < T(\mathbf{X}) < C] = \alpha.$$

Vieme, že za podmienky $\mu = \mu_0 + \varepsilon$ má $T(\mathbf{X})$ rozdelenie $\mathrm{N}\left(\frac{\varepsilon}{\sigma}\sqrt{n}, 1\right)$. Za podmienky $\mu = \mu_0 + \varepsilon^*$ platí $T(\mathbf{X}) \sim \mathrm{N}\left(\frac{\varepsilon^*}{\sigma}\sqrt{n}, 1\right)$, kde $\frac{\varepsilon^*}{\sigma}\sqrt{n} \geq \frac{\varepsilon}{\sigma}\sqrt{n}$.

Keďže

$$P_{\mu_0-\varepsilon}[-C < T(\mathbf{X}) < C] = \alpha = P_{\mu_0+\varepsilon}[-C < T(\mathbf{X}) < C]$$

a $\alpha \leq \frac{1}{2}$, tak

$$(-C, C) \in \left(-\frac{\varepsilon}{\sigma}\sqrt{n}, \frac{\varepsilon}{\sigma}\sqrt{n}\right) \in \left(-\frac{\varepsilon^*}{\sigma}\sqrt{n}, \frac{\varepsilon^*}{\sigma}\sqrt{n}\right).$$

Potom môžeme písať

$$P_{\mu_0+\varepsilon^*}[-C < T(\mathbf{X}) < C] = \int_{-C}^{C} f^*(y)\, dy \leq$$

$$\int_{-C}^{C} f(y)\, dy = P_{\mu_0+\varepsilon}[-C < T(\mathbf{X}) < C] = \alpha,$$

kde $f^*(y)$ je hustota $\mathrm{N}\left(\frac{\varepsilon^*}{\sigma}\sqrt{n}, 1\right)$ a $f(y)$ je hustota $\mathrm{N}\left(\frac{\varepsilon}{\sigma}\sqrt{n}, 1\right)$.

Kritickú hodnotu C dostaneme vyriešením nasledujúcej rovnice

$$\begin{aligned}
\alpha &= P_{\mu_0+\varepsilon}[|T(\mathbf{X})| < C] \\
&= P_{\mu_0+\varepsilon}[-C < T(\mathbf{X}) < C] \\
&= P_{\mu_0+\varepsilon}\left[-C < \frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} < C\right] \\
&= P_{\mu_0+\varepsilon}\left[-C < \frac{\overline{X}-(\mu_0+\varepsilon)+\varepsilon}{\frac{\sigma}{\sqrt{n}}} < C\right] \\
&= P_{\mu_0+\varepsilon}\left[-C - \frac{\varepsilon}{\sigma}\sqrt{n} < \frac{\overline{X}-(\mu_0+\varepsilon)}{\frac{\sigma}{\sqrt{n}}} < C - \frac{\varepsilon}{\sigma}\sqrt{n}\right] \\
&= \Phi(C - \frac{\varepsilon}{\sigma}\sqrt{n}) - \Phi(-C - \frac{\varepsilon}{\sigma}\sqrt{n}),
\end{aligned}$$

kde $C > 0$ a $\Phi(x)$ je distribučná funkcia rozdelenia $\mathrm{N}(0,1)$.

Navyše platí:

Tvrdenie. *Ak $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, ..., k$ nezávislé, potom $Y = \sum_{i=1}^{k}\left(\frac{X_i}{\sigma}\right)^2$ má necentrálne χ^2-rozdelenie s k stupňami voľnosti a parametrom necentrality $\lambda = \sum_{i=1}^{k}\left(\frac{\mu_i}{\sigma}\right)^2$.*

Tvrdenie je prevzaté z [6].

Kritickú hodnotu C teda môžeme určiť aj nasledujúcim spôsobom:

$$P_\varepsilon[|T(\mathbf{X})| < C] = P_\varepsilon[T(\mathbf{X})^2 < C^2] = \alpha.$$

Keďže $T(\mathbf{X})$ má za podmienky $\mu = \mu_0 + \varepsilon$ necentrálne χ^2-rozdelenie s jedným stupňom voľnosti a parametrom necentrality $\frac{\varepsilon^2 n}{\sigma^2}$, tak

$$C = \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon^2 n}{\sigma^2}\right)},$$

kde $\chi^2_{1;\alpha}\left(\frac{\varepsilon^2 n}{\sigma^2}\right)$ je α-kvantil necentrálneho χ^2-rozdelenia s jedným stupňom voľnosti a parametrom necentrality $\frac{\varepsilon^2 n}{\sigma^2}$.

Záver:
Hypotézu $H_0 : \mu \leq \mu_0 - \varepsilon$ alebo $\mu \geq \mu_0 + \varepsilon$ zamietame v prospech alternatívy $H_1 : \mu_0 - \varepsilon < \mu < \mu_0 + \varepsilon$, ak

$$|T(\mathbf{X})| < C$$

$$\left|\frac{\overline{X} - \mu_0}{\sigma}\sqrt{n}\right| < \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon^2 n}{\sigma^2}\right)}.$$

4.2 Jednovýberový t-test

Majme $(X_1, ..., X_n)$ náhodný výber z rozdelenia $F_X \in \mathcal{F}$.
Model je $\mathcal{F} = \{\mathrm{N}(\mu, \sigma^2),\ \mu \in \mathbf{R},\ \sigma^2 > 0\}$.

V prípade neznámeho rozptylu sa zameriame na testy ekvivalencie v nasledujúcej špeciálnej situácii:

$$H_0 : \frac{\mu}{\sigma} \leq \theta_1 \text{ alebo } \frac{\mu}{\sigma} \geq \theta_2 \qquad \text{vs.} \qquad H_1 : \theta_1 < \frac{\mu}{\sigma} < \theta_2.$$

Dôvodom je, že pri testovaní všeobecného problému

$$H_0 : \mu \leq \mu_0 - \varepsilon \text{ alebo } \mu \geq \mu_0 + \varepsilon \qquad \text{vs.} \qquad H_1 : \mu_0 - \varepsilon < \mu < \mu_0 + \varepsilon$$

by sme nevedeli určiť rozdelenie testovej štatistiky na hraniciach nulovej hypotézy.

Testová štatistika:

$$T(\mathbf{X}) = \frac{\overline{X}\sqrt{n}}{S_X},$$

kde $\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$ je výberový priemer a $S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$ je výberový rozptyl.

Tvrdenie. *Nech S_X^2 a $\overline{X}$ sú zadefinované ako vyššie, potom sú nezávislé a $\dfrac{S_X^2(n-1)}{\sigma^2}$ má χ^2-rozdelenie s $n-1$ stupňami voľnosti.*

Definícia. *Ak $Z \sim N(0,1)$, $V \sim \chi_k^2$ (χ_k^2 je označenie pre χ^2-rozdelenie s k stupňami voľnosti) nezávislé, potom*

$$T = \frac{Z+\lambda}{\sqrt{\dfrac{V}{k}}}$$

má necentrálne t-rozdelenie s k stupňami voľnosti a parametrom necentrality λ.

Určíme rozdelenie testovej štatistiky:

$$T(\mathbf{X}) = \frac{\overline{X}\sqrt{n}}{S_X} = \frac{\dfrac{\sqrt{n}(\overline{X}-\mu+\mu)}{\sigma}}{\sqrt{\dfrac{\dfrac{S_X^2(n-1)}{\sigma^2}}{n-1}}} = \frac{\dfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma} + \dfrac{\mu\sqrt{n}}{\sigma}}{\sqrt{\dfrac{\dfrac{S_X^2(n-1)}{\sigma^2}}{n-1}}}$$

$$\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \sim \mathrm{N}(0,1), \qquad \frac{S_X^2(n-1)}{\sigma^2} \sim \chi_{n-1}^2 \text{ nezávislé.}$$

Vidíme teda, že $T(\mathbf{X})$ má necentrálne t-rozdelenie s $n-1$ stupňami voľnosti a parametrom necentrality $\dfrac{\sqrt{n}\mu}{\sigma}$.

Kritický obor je tvaru (C_1, C_2) a platí:

$$P_{\theta_1}[C_1 < T(\mathbf{X}) < C_2] = \alpha = P_{\theta_2}[C_1 < T(\mathbf{X}) < C_2].$$

$T(\mathbf{X})$ má za podmienky $\dfrac{\mu}{\sigma} = \theta_i$ necentrálne t-rozdelenie s $n-1$ stupňami voľnosti, parametrom necentrality $\sqrt{n}\theta_i$ a distribučnou funkciou $G_{\sqrt{n}\theta_i}$, $i \in \{1,2\}$. Môžeme písať:

$$G_{\sqrt{n}\theta_1}(C_2) - G_{\sqrt{n}\theta_1}(C_1) = \alpha = G_{\sqrt{n}\theta_2}(C_2) - G_{\sqrt{n}\theta_2}(C_1),$$
$$-\infty < C_1 < C_2 < \infty.$$

Kritické hodnoty C_1, C_2 dostaneme vyriešením predchádzajúcej rovnice.

Odvodenie rozdelenia $T(\mathbf{X})$ za podmienky $\frac{\mu}{\sigma} = \upsilon$:

$$T(\mathbf{X}) = \frac{\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} + \frac{\mu\sqrt{n}}{\sigma}}{\sqrt{\frac{\frac{S_X^2(n-1)}{\sigma^2}}{n-1}}} = \frac{\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} + \sqrt{n}\upsilon}{\sqrt{\frac{\frac{S_X^2(n-1)}{\sigma^2}}{n-1}}}.$$

Odvodenie rozdelenia $-T(\mathbf{X})$ za podmienky $\frac{\mu}{\sigma} = -\upsilon$:

$$-T(\mathbf{X}) = -\frac{\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} + \frac{\mu\sqrt{n}}{\sigma}}{\sqrt{\frac{\frac{S_X^2(n-1)}{\sigma^2}}{n-1}}} = \frac{\frac{\sqrt{n}(-\overline{X}-(-\mu))}{\sigma} + \sqrt{n}\upsilon}{\sqrt{\frac{\frac{S_X^2(n-1)}{\sigma^2}}{n-1}}}.$$

Vďaka symetrii normálneho rozdelenia ide v oboch prípadoch o rovnaké rozdelenie. Ak navyše zvolíme symetrický interval ekvivalencie: $\theta_1 = -\varepsilon,\ \theta_2 = \varepsilon$, potom kritický obor bude tvaru $(-C, C)$ a bude platiť:

$$P_\varepsilon[|T(\mathbf{X})| < C] = P_\varepsilon[-C < T(\mathbf{X}) < C] = \alpha.$$

$T(\mathbf{X})$ má za podmienky $\frac{\mu}{\sigma} = \varepsilon$ necentrálne t-rozdelenie s $n-1$ stupňami voľnosti, parametrom necentrality $\sqrt{n}\varepsilon$ a distribučnou funkciou $G_{\sqrt{n}\varepsilon}$. Môžeme teda písať

$$G_{\sqrt{n}\varepsilon}(C) - G_{\sqrt{n}\varepsilon}(-C) = \alpha,\ 0 < C < \infty.$$

Kritickú hodnotu C dostaneme vyriešením predchádzajúcej rovnice.

Navyše platí:

Tvrdenie. *Ak X_1 má necentrálne χ^2-rozdelenie s k_1 stupňami voľnosti a parametrom necentrality λ, $X_2 \sim \chi^2_{k_2}$ a sú navzájom nezávislé, potom*

$$F = \frac{\frac{X_1}{k_1}}{\frac{X_2}{k_2}}$$

má necentrálne F-rozdelenie s k_1, k_2 stupňami voľnosti a parametrom necentrality λ.

Tvrdenie je prevzaté z [6].

Potom môžeme písať:

$$T(\mathbf{X}) \sim \frac{\mathrm{N}(0,1) + \frac{\sqrt{n}\mu}{\sigma}}{\sqrt{\frac{\chi^2_{n-1}}{n-1}}}$$

$$T(\mathbf{X})^2 \sim \frac{\left[\mathrm{N}(0,1) + \frac{\sqrt{n}\mu}{\sigma}\right]^2}{\frac{\chi^2_{n-1}}{n-1}} \sim \frac{\left[\mathrm{N}\left(\frac{\sqrt{n}\mu}{\sigma}, 1\right)\right]^2}{\frac{\chi^2_{n-1}}{n-1}}$$

$$T(\mathbf{X})^2 \sim \frac{\frac{\text{necentrálne}\chi^2_1\text{-rozdelenie s parametrom necentrality}\left(\frac{\sqrt{n}\mu}{\sigma}\right)^2}{1}}{\frac{\chi^2_{n-1}}{n-1}}.$$

Teda $T(\mathbf{X})^2$ má necentrálne F-rozdelenie s 1, $n-1$ stupňami voľnosti a parametrom necentrality $\frac{n\mu^2}{\sigma^2}$.

Kritickú hodnotu C teda môžeme určiť aj nasledujúcim spôsobom:

$$P_\varepsilon[|T(\mathbf{X})| < C] = P_\varepsilon[T(\mathbf{X})^2 < C^2] = \alpha.$$

$T(\mathbf{X})$ má za podmienky $\frac{\mu}{\sigma} = \varepsilon$ necentrálne F-rozdelenie s 1, $n-1$ stupňami voľnosti a parametrom necentrality $n\varepsilon^2$. Potom

$$C = \sqrt{\mathrm{F}_{1,n-1;\alpha}(n\varepsilon^2)},$$

kde $\mathrm{F}_{1,n-1;\alpha}(n\varepsilon^2)$ je α-kvantil necentrálneho F-rozdelenia s 1, $n-1$ stupňami voľnosti a parametrom necentrality $n\varepsilon^2$.

4.3 Párový t-test

Majme $(\mathbf{Q}_1, \dots, \mathbf{Q}_n)$ náhodný výber z dvojrozmerného rozdelenia $F_{\mathbf{Q}}$, kde

$$\mathbf{Q}_i = \begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \; i \in \{1, \dots, n\}.$$

Urobíme transformáciu dát:

$$\begin{aligned} D_i &= X_i - Y_i, \quad i \in \{1, \dots, n\} \\ \mathrm{E}D_i &= \mu_D \\ \mathrm{var}D_i &= \sigma_D^2. \end{aligned}$$

Predpokladáme:
nech $(D_1, \dots, D_n)$ je náhodný výber z rozdelenia F_D.
Model je $\mathcal{F} = \{F_D \sim \mathrm{N}(\mu_D, \sigma_D^2), \; \mu_D \in \mathbf{R}, \; \sigma_D^2 > 0\}$.

Testujeme

$$H_0 : \frac{\mu_D}{\sigma_D} \leq \theta_1 \;\text{ alebo }\; \frac{\mu_D}{\sigma_D} \geq \theta_2 \qquad \text{vs.} \qquad H_1 : \theta_1 < \frac{\mu_D}{\sigma_D} < \theta_2.$$

Ďalej postupujeme ako pri jednovýberovom t-teste.

Príklad 5. Chceme zistiť, či logaritmy hladiny bilirubínu dvoch skupín pacientov (jednej bude podávaný D-penicilamín a druhej placebo) sú na začiatku štúdie ekvivalentné.

Máme dva náhodné výbery $(X_1, \dots, X_{158})$ a $(Y_1, \dots, Y_{154})$, kde X_i je hodnota logaritmu hladiny bilirubínu u pacienta, ktorý bude liečený D-penicilamínom, a Y_j je hodnota logaritmu hladiny bilirubínu pacienta, ktorý bude dostávať placebo. Predpokladáme, že $X_i \sim \mathrm{N}(\mu_X, \sigma^2)$ a $Y_j \sim \mathrm{N}(\mu_Y, \sigma^2)$. Vzhľadom na náhodný výber pacientov môžeme predpokladať nezávislosť X_i a Y_j pre každé i a j.

Pre ilustračné účely spárujeme pacientov podľa veku. Posledné štyri náhodné veličiny prvého výberu vynecháme. Predpokladáme, že $D_i = X_i - Y_i, \; i \in \{1, ..., 154\}$, má rozdelenie $\mathrm{N}(\mu_X, \sigma^2)$ a sú navzájom nezávislé.

Testujeme

$$H_0 : \frac{\mu}{\sigma} \leq -\varepsilon \ \text{ alebo } \ \frac{\mu}{\sigma} \geq \varepsilon \qquad \text{vs.} \qquad H_1 : -\varepsilon < \frac{\mu}{\sigma} < \varepsilon.$$

Budeme testovať na hladine $\alpha = 0.05$ a za ε zvolíme $\varepsilon = 0.25$ (bližšie zdôvodnenie voľby ε je uvedené na konci prvej kapitoly).
Podľa jednovýberového t-testu pre ekvivalenciu je $|T(\mathbf{D})| = 1.242 < C = 2.115$, teda test zamieta nulovú hypotézu.

4.4 Dvojvýberový t-test

Nech $(X_1, \dots, X_m)$ je náhodný výber z rozdelenia F_X, $(Y_1, \dots, Y_n)$ náhodný výber z rozdelenia F_Y a nech sú navzájom nezávislé.
Model je $\mathcal{F} = \{F_X \sim \mathrm{N}(\mu_X, \sigma^2),\ F_Y \sim \mathrm{N}(\mu_Y, \sigma^2),\ \mu_X, \mu_Y \in \mathbf{R},\ \sigma^2 > 0\}$.

Testujeme

$$H_0 : \frac{\mu_X - \mu_Y}{\sigma} \leq -\varepsilon_1 \ \text{ alebo } \ \frac{\mu_X - \mu_Y}{\sigma} \geq \varepsilon_2$$
$$\text{vs.} \qquad H_1 : -\varepsilon_1 < \frac{\mu_X - \mu_Y}{\sigma} < \varepsilon_2, \quad \varepsilon_1, \varepsilon_2 > 0.$$

Testová štatistika:

$$T(\mathbf{X}, \mathbf{Y}) = \sqrt{\frac{m\, n}{m+n}}\, \frac{\overline{X} - \overline{Y}}{S_{X,Y}},$$

kde $\overline{X} = \frac{\sum_{i=1}^m X_i}{m}$, $\overline{Y} = \frac{\sum_{i=1}^n Y_i}{n}$ je výberový priemer a

$$S_{X,Y}^2 = \frac{1}{m+n-2} \left(\sum_{i=1}^m (X_i - \overline{X})^2 + \sum_{i=1}^n (Y_i - \overline{Y})^2 \right).$$

Určíme rozdelenie testovej štatistiky:

$$
\begin{aligned}
T(\mathbf{X},\mathbf{Y}) =& \sqrt{\frac{m\,n}{m+n}}\,\frac{\overline{X}-\overline{Y}}{S_{X,Y}} \\
=& \frac{\sqrt{\dfrac{m\,n}{m+n}}\,\dfrac{(\overline{X}-\mu_X)-(\overline{Y}-\mu_Y)}{\sigma} + \sqrt{\dfrac{m\,n}{m+n}}\,\dfrac{\mu_X-\mu_Y}{\sigma}}{\sqrt{\dfrac{\dfrac{S^2_{X,Y}(m+n-2)}{\sigma^2}}{m+n-2}}}
\end{aligned}
$$

$$
\sqrt{\frac{m\,n}{m+n}}\,\frac{(\overline{X}-\mu_X)-(\overline{Y}-\mu_Y)}{\sigma} \sim \mathrm{N}(0,1),
$$

$$
\frac{S^2_{X,Y}(m+n-2)}{\sigma^2} \sim \chi^2_{m+n-2} \quad \text{nezávislé.}
$$

Teda $T(\mathbf{X},\mathbf{Y})$ má necentrálne t-rozdelenie s $m+n-2$ stupňami voľnosti a parametrom necentrality $\sqrt{\frac{m\,n}{m+n}}\,\frac{\mu_X-\mu_Y}{\sigma}$.

Kritický obor je tvaru (C_1, C_2) a platí:

$$
P_{-\varepsilon_1}[C_1 < T(\mathbf{X},\mathbf{Y}) < C_2] = \alpha = P_{\varepsilon_2}[C_1 < T(\mathbf{X},\mathbf{Y}) < C_2].
$$

$T(\mathbf{X},\mathbf{Y})$ má za podmienky $\frac{\mu_X-\mu_Y}{\sigma} = -\varepsilon_1$ necentrálne t-rozdelenie s $m+n-2$ stupňami voľnosti, parametrom necentrality $-\lambda_1 = -\varepsilon_1\sqrt{\frac{m\,n}{m+n}}$ a distribučnou funkciou $G_{-\lambda_1}$.
$T(\mathbf{X},\mathbf{Y})$ má za podmienky $\frac{\mu_X-\mu_Y}{\sigma} = \varepsilon_2$ necentrálne t-rozdelenie s $m+n-2$ stupňami voľnosti, parametrom necentrality $\lambda_2 = \varepsilon_2\sqrt{\frac{m\,n}{m+n}}$ a distribučnou funkciou G_{λ_2}. Môžeme písať:

$$
\begin{gathered}
G_{-\lambda_1}(C_2) - G_{-\lambda_1}(C_1) = \alpha = G_{\lambda_2}(C_2) - G_{\lambda_2}(C_1), \\
-\infty < C_1 < C_2 < \infty.
\end{gathered}
$$

Kritické hodnoty C_1, C_2 dostaneme vyriešením predchádzajúcej rovnice.

Analogicky ako v jednovýberovom t-teste môžeme ukázať, že rozdelenie $T(\mathbf{X},\mathbf{Y})$ za podmienky $\frac{\mu_X-\mu_Y}{\sigma}=\upsilon$ je rovnaké ako rozdelenie $-T(\mathbf{X},\mathbf{Y})$ za podmienky $\frac{\mu_X-\mu_Y}{\sigma}=-\upsilon$.
Ak navyše zvolíme symetrický interval ekvivalencie: $\varepsilon_1=\varepsilon_2=\varepsilon$, potom kritický obor bude tvaru $(-C,C)$ a bude platiť:

$$P_\varepsilon[|T(\mathbf{X},\mathbf{Y})|<C]=P_\varepsilon[-C<T(\mathbf{X},\mathbf{Y})<C]=\alpha.$$

$T(\mathbf{X},\mathbf{Y})$ má za podmienky $\frac{\mu_X-\mu_Y}{\sigma}=\varepsilon$ necentrálne t-rozdelenie s $m+n-2$ stupňami voľnosti, parametrom necentrality $\lambda=\varepsilon\sqrt{\frac{m\,n}{m+n}}$ a distribučnou funkciou G_λ. Potom platí

$$G_\lambda(C)-G_\lambda(-C)=\alpha,\ 0<C<\infty.$$

Kritickú hodnotu C dostaneme vyriešením predchádzajúcej rovnice.

Navyše platí:
$T(\mathbf{X},\mathbf{Y})^2$ má necentrálne F-rozdelenie s $1,\ m+n-2$ stupňami voľnosti a parametrom necentrality $\frac{m\,n}{m+n}\left(\frac{\mu_X-\mu_Y}{\sigma}\right)^2$; čo bolo ukázané v jednovýberovom t-teste.

Kritickú hodnotu C teda môžeme určiť aj nasledujúcim spôsobom:

$$P_\varepsilon[|T(\mathbf{X},\mathbf{Y})|<C]=P_\varepsilon[T(\mathbf{X},\mathbf{Y})^2<C^2]=\alpha.$$

$T(\mathbf{X},\mathbf{Y})$ má za podmienky $\frac{\mu_X-\mu_Y}{\sigma}=\varepsilon$ necentrálne F-rozdelenie s $1,\ m+n-2$ stupňami voľnosti a parametrom necentrality $\frac{m\,n}{m+n}\varepsilon^2$. Potom

$$C=\sqrt{\mathrm{F}_{1,m+n-2;\alpha}\left(\frac{m\,n}{m+n}\varepsilon^2\right)},$$

kde $\mathrm{F}_{1,m+n-2;\alpha}\left(\frac{m\,n}{m+n}\varepsilon^2\right)$ je α-kvantil necentrálneho F-rozdelenia s $1,\ m+n-2$ stupňami voľnosti a parametrom necentrality $\frac{m\,n}{m+n}\varepsilon^2$.

Príklad 6. Zadanie je rovnaké ako v príklade 1*. Keďže predpokladáme normálne rozdelenie náhodných výberov, ich vzájomnú nezávislosť a rovnaký rozptyl, môžeme použiť dvojvýberový t-test pre ekvivalenciu. Budeme však testovať

$$H_0 : \frac{\mu_X - \mu_Y}{\sigma} \leq -\varepsilon \text{ alebo } \frac{\mu_X - \mu_Y}{\sigma} \geq \varepsilon \quad \text{vs.} \quad H_1 : -\varepsilon < \frac{\mu_X - \mu_Y}{\sigma} < \varepsilon.$$

Testujeme na hladine $\alpha = 0.05$ a za ε zvolíme $\varepsilon = 0.36$ (bližšie zdôvodnenie voľby ε je uvedené na konci prvej kapitoly).
Potom dostaneme, že $|T(\mathbf{X}, \mathbf{Y})| = 0.653 < C = 2.348$, teda zamietame nulovú hypotézu.

4.5 Asymptotický test o proporciách

Nech $(X_1, \dots , X_n)$ je náhodný výber z rozdelenia $F_X \in \mathcal{F}$.
Model je $\mathcal{F} = \{\mathrm{Alt}(p_X),\ p_X \in (0, 1)\}$.

Testujeme

$$H_0 : p_X \leq p_0 - \varepsilon \text{ alebo } p_X \geq p_0 + \varepsilon \quad \text{vs.} \quad H_1 : p_0 - \varepsilon < p_X < p_0 + \varepsilon. \tag{4.1}$$

Z vlastností alternatívneho rozdelenia platí:

$$\begin{aligned} \mathrm{E}X_i &= p_X \\ \mathrm{var}X_i &= p_X(1 - p_X). \end{aligned}$$

Podľa Centrálnej limitnej vety

$$\begin{aligned} \sqrt{n}(\overline{X} - p_X) &\xrightarrow{\mathcal{D}} \mathrm{N}(0, p_X(1 - p_X)),\ n \to \infty \\ \frac{\sqrt{n}(\overline{X} - p_X)}{\sqrt{p_X(1 - p_X)}} &\xrightarrow{\mathcal{D}} \mathrm{N}(0, 1),\ n \to \infty, \end{aligned}$$

kde $\overline{X} = \dfrac{\sum_{i=1}^n X_i}{n}$ je výberový priemer.

Ďalej budeme pokračovať v odvodzovaní asymptotického testu podla postupu uvedenom v [1, strana 43]. Tento postup je založený na výsledkoch uvedených v [5], a za nie príliš veľkých podmienok vedie k testu ekvivalencie, ktorý je asymptotický vzhľadom na hladinu.

Zmienený postup aplikovaný na prípad alternatívneho rozdelenia vyzerá nasledovne: rozptyl $\dfrac{p_X(1-p_X)}{n}$ považujme za známu konštantu a testujme (4.1) pomocou Testu o strednej hodnote normálneho rozdelenia so známym rozptylom, kde $\overline{X}$ je jediná náhodná veličina. Teda:

$$\begin{aligned} \mu &= p_X \\ \mu_0 &= p_0 \\ \sigma^2 &= \frac{p_X(1-p_X)}{n}. \end{aligned}$$

Testujeme:

$$H_0 : \mu \leq \mu_0 - \varepsilon \ \text{ alebo } \ \mu \geq \mu_0 + \varepsilon \qquad \text{vs.} \qquad H_1 : \mu_0 - \varepsilon < \mu < \mu_0 + \varepsilon.$$

Testová štatistika je potom:

$$T = \frac{\overline{X} - \mu_0}{\sigma}.$$

Ako sme odvodili v časti 4.1., hypotézu $H_0 : \mu \leq \mu_0 - \varepsilon$ alebo $\mu \geq \mu_0 + \varepsilon$ zamietame v prospech alternatívy $H_1 : \mu_0 - \varepsilon < \mu < \mu_0 + \varepsilon$, ak

$$|T| < C = \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon}{\sigma}\right)^2},$$

kde $\chi^2_{1;\alpha}\left(\frac{\varepsilon}{\sigma}\right)^2$ je α-kvantil necentrálneho χ^2-rozdelenia s jedným stupňom voľnosti a parametrom necentrality $\left(\frac{\varepsilon}{\sigma}\right)^2$.

Dosadením pôvodných parametrov dostaneme

$$|T| = \left|\frac{\overline{X} - \mu_0}{\sigma}\right| = \left|\frac{\overline{X} - p_0}{\sqrt{\frac{p_X(1-p_X)}{n}}}\right| < \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon^2}{\frac{p_X(1-p_X)}{n}}\right)}. \tag{4.2}$$

Navyše zo Zákona veľkých čísel platí

$$\begin{aligned} \overline{X} &\xrightarrow{P} \mathrm{E}X_i = p_X, \ n \to \infty \\ \overline{X}(1-\overline{X}) &\xrightarrow{P} p_X(1-p_X), \ n \to \infty, \end{aligned}$$

a teda $\overline{X}(1-\overline{X})$ je slabo konzistentný odhad $p_X(1-p_X)$. Týmto odhadom nahradíme všetky $p_X(1-p_X)$, ktoré sa vyskytujú v (4.2) a dostaneme

$$\left| \frac{\overline{X} - p_0}{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}} \right| < \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon^2}{\frac{\overline{X}(1-\overline{X})}{n}} \right)}.$$

Vidíme, že $\chi^2_{1;\alpha}\left(\frac{\varepsilon^2}{\frac{\overline{X}(1-\overline{X})}{n}} \right)$ je náhodná kritická hodnota, avšak Wellek v [5] ukázal, že test s takto definovaným kritickým oborom má asymptoticky hladinu α.

Záver:
Hypotézu $H_0 : p_X \leq p_0 - \varepsilon$ alebo $p_X \geq p_0 + \varepsilon$ zamietame v prospech alternatívy $H_1 : p_0 - \varepsilon < p_X < p_0 + \varepsilon$ ak

$$\left| \frac{\overline{X} - p_0}{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}} \right| < \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon^2}{\frac{\overline{X}(1-\overline{X})}{n}} \right)}.$$

4.6 Dvojvýberový asymptotický test o proporciách

Majme $(X_1, \ldots, X_n)$ náhodný výber z rozdelenia F_X, $(Y_1, \ldots, Y_n)$ náhodný výber z rozdelenia F_Y, ktoré sú navzájom nezávislé.

Model je $\mathcal{F} = \{F_X \sim \text{Alt}(p_X),\ F_Y \sim \text{Alt}(p_Y),\ p_X, p_Y \in (0,1)\}$.

Rovnaký rozsah oboch náhodných výberov sme zvolili kvôli postupu, ktorý bude použitý.

Testujeme

$$H_0 : p_X - p_Y \leq\ -\varepsilon \ \text{ alebo } \ p_X - p_Y \geq \varepsilon \qquad \text{vs.} \qquad H_1 : -\varepsilon < p_X - p_Y < \varepsilon. \tag{4.3}$$

Už sme ukázali v druhej kapitole, že

$$\frac{\sqrt{n}(\overline{X} - \overline{Y} - (p_X - p_Y))}{\sqrt{p_X(1-p_X) + p_Y(1-p_Y)}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1),\ n \to \infty.$$

Ako v predchádzajúcej kapitole, tak aj pri odvodení tohoto testu využijeme postup uvedený v [1, strana 43], ktorý za určitých slabých podmienok vedie k asymptotickému testu ekvivalencie.

Zmienený postup aplikovaný na prípad alternatívneho rozdelenia vyzerá nasledovne: rozptyl $\dfrac{p_X(1-p_X) + p_Y(1-p_Y)}{n}$ považujme za známu konštantu a testujme (4.3) pomocou Testu o strednej hodnote normálneho rozdelenia so známym rozptylom, kde $\overline{X} - \overline{Y}$ je jediná náhodná veličina. Teda:

$$\begin{aligned} \mu &= p_X - p_Y \\ \mu_0 &= 0 \\ \sigma^2 &= \frac{p_X(1-p_X) + p_Y(1-p_Y)}{n}. \end{aligned}$$

Testujeme:

$$H_0 : \mu \leq \mu_0 - \varepsilon \ \text{ alebo } \ \mu \geq \mu_0 + \varepsilon \qquad \text{vs.} \qquad H_1 : \mu_0 - \varepsilon < \mu < \mu_0 + \varepsilon.$$

Testová štatistika je nasledujúca:

$$T = \frac{\overline{X} - \overline{Y} - \mu_0}{\sigma}.$$

Ako sme už odvodili v časti 4.1., hypotézu $H_0 : \mu \leq \mu_0 - \varepsilon$ alebo $\mu \geq \mu_0 + \varepsilon$ zamietame v prospech alternatívy $H_1 : \mu_0 - \varepsilon < \mu < \mu_0 + \varepsilon$ ak

$$|T| < C = \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon}{\sigma}\right)^2},$$

kde $\chi^2_{1;\alpha}\left(\frac{\varepsilon}{\sigma}\right)^2$ je α-kvantil necentrálneho χ^2-rozdelenia s jedným stupňom voľnosti a parametrom necentrality $\left(\frac{\varepsilon}{\sigma}\right)^2$.

Dosadením pôvodných parametrov do predchádzajúceho výrazu dostaneme

$$|T| = \left|\frac{\overline{X} - \overline{Y} - \mu_0}{\sigma}\right| = \left|\frac{\overline{X} - \overline{Y}}{\sqrt{\frac{p_X(1-p_X) + p_Y(1-p_Y)}{n}}}\right| < \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon^2}{\frac{p_X(1-p_X) + p_Y(1-p_Y)}{n}}\right)}. \tag{4.4}$$

Ďalej zo Zákona veľkých čísel platí

$$\begin{aligned} \overline{X} &\xrightarrow{P} \mathrm{E}X_i = p_X,\ n \to \infty \\ \overline{X}(1-\overline{X}) &\xrightarrow{P} p_X(1-p_X),\ n \to \infty \end{aligned}$$

$$\begin{aligned} \overline{Y} &\xrightarrow{P} \mathrm{E}Y_i = p_Y,\ n \to \infty \\ \overline{Y}(1-\overline{Y}) &\xrightarrow{P} p_Y(1-p_Y),\ n \to \infty \end{aligned}$$

$$\overline{X}(1-\overline{X}) + \overline{Y}(1-\overline{Y}) \xrightarrow{P} p_X(1-p_X) + p_Y(1-p_Y),\ n \to \infty,$$

a teda $\overline{X}(1-\overline{X}) + \overline{Y}(1-\overline{Y})$ je slabo konzistentný odhad $p_X(1-p_X) + p_Y(1-p_Y)$. Týmto odhadom nahradíme všetky $p_X(1-p_X)+p_Y(1-p_Y)$, ktoré sa vyskytujú v (4.4) a dostaneme

$$\left|\frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\overline{X}(1-\overline{X}) + \overline{Y}(1-\overline{Y})}{n}}}\right| < \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon^2}{\frac{\overline{X}(1-\overline{X}) + \overline{Y}(1-\overline{Y})}{n}}\right)}.$$

Vidíme, že $\chi^2_{1;\alpha}\left(\frac{\varepsilon^2}{\frac{\overline{X}(1-\overline{X})+\overline{Y}(1-\overline{Y})}{n}}\right)$ je náhodná kritická hodnota, avšak Wellek v [5] ukázal, že test s takto definovaným kritickým oborom má asymptoticky hladinu α.

<u>Záver:</u>
Hypotézu $H_0 : p_X - p_Y \leq -\varepsilon$ alebo $p_X - p_Y \geq \varepsilon$ zamietame v prospech alternatívy $H_1 : -\varepsilon < p_X - p_Y < \varepsilon$ ak

$$\left|\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\overline{X}(1-\overline{X})+\overline{Y}(1-\overline{Y})}{n}}}\right| < \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon^2}{\frac{\overline{X}(1-\overline{X})+\overline{Y}(1-\overline{Y})}{n}}\right)}.$$

Príklad 7. Zadanie je rovnaké ako v príklade 2*. Vzhľadom na to, že predpokladáme alternatívne rozdelenie oboch náhodných výberov a ich vzájomnú nezávislosť, použijeme dvojvýberový asymptotický test o proporciách. Zvolíme $\varepsilon_1 = \varepsilon_2 = \varepsilon$. Pre ilustráciu použitia testu, ktorý sme odvodili v predchádzajúcej časti, vynecháme posledné štyri náhodné veličiny z prvého výberu.

Budeme testovať na hladine $\alpha = 0.05$ a zvolíme $\varepsilon = 0.10$.
Podľa dvojvýberového asymptotického testu je

$$0 = \left|\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\overline{X}(1-\overline{X})+\overline{Y}(1-\overline{Y})}{n}}}\right| < \sqrt{\chi^2_{1;\alpha}\left(\frac{\varepsilon^2}{\frac{\overline{X}(1-\overline{X})+\overline{Y}(1-\overline{Y})}{n}}\right)} = 0.144,$$

teda zamietame nulovú hypotézu.

Záver

V práci som sa zaoberala základmi z oblasti štatistických testov ekvivalencie. Najprv som v prvej kapitole na príkladoch ilustrovala všeobecne známy fakt, že pomocou klasického testu nie sme schopní potvrdiť nulovú hypotézu, ktorý je motiváciou pre vznik testov ekvivalencie. Ďalej som za pomoci knihy profesora Welleka dokázala, že problém ekvivalencie nemôžeme vytvoriť zámenou nulovej hypotézy a alternatívy klasického obojstranného testu. Nakoniec som uviedla ako vyzerá všeobecný problém ekvivalencie.

V druhej kapitole som s pomocou knihy profesora Welleka odvodila jeden z možných prístupov k riešeniu problému ekvivalencie - test založený na princípe inklúzie intervalov spoľahlivosti. Následne som na konkrétnom testovacom probléme ukázala jednu z nevýhod tohoto prístupu - že za určitých podmienok nám dáva test s nulovou silou.

V tretej kapitole som na základe knihy profesora Welleka heuristicky odvodila optimálny test pre ekvivalenciu. Ďalej som s využitím Neymanovej-Pearsonovej lemy na príklade testu o strednej hodnote normálneho rozdelenia so známym rozptylom ukázala, že sú tieto heuristické úvahy správne.

Štvrtá kapitola obsahuje testy pre konkrétne vybrané situácie. V prvej časti tejto kapitoly som s pomocou knihy profesora Welleka odvodila test o strednej hodnote normálneho rozdelenia so známym rozptylom, jednovýberový t-test, párový t-test a dvojvýberový t-test. Posledné dva som doplnila o príklady, kde som tieto testy aplikovala na konkrétne dáta. V druhej časti som na základe postupu uvedenom v knihe od profesora Welleka odvodila jednovýberový a dvojvýberový asymptotický test o proporciách. Posledný zmienený som následne na konkrétnych dátach ukázala.

Vzhľadom na veľký rozsah témy, existuje ešte veľa vecí, ktorými by sa táto práca mohla ďalej zaoberať. Jednou z nich je testovanie ďalších konkrétnych vybraných problémov.

Pri písaní tejto práce ma najviac zaujalo odvodenie konkrétnych testov v štvrtej kapitole, a potom ich následná aplikácia na dáta s využitím štatistického programu R.

Literatúra

[1] Wellek S.: *Testing statistical hypotheses of equivalence*, Chapman and Hall/CRC Press LLC, 2003.

[2] Dupač V., Hušková M.: *Pravděpodobnost a matematická statistika*, Karolinum, Praha, 1999 .

[3] R Development Core Team: *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2010.

[4] Wolfram S.: *The Mathematica Book. 5th edition.*, Wolfram Media Champaign (IL), 2003.

[5] Wellek S. (1996): *A new approach to equivalence assessment in standard comparative bioavailability trials by means of the Mann-Whitney statistic*, Biometrical journal **38**, 695-710.

[6] Hogg R. V., Craig A: *Introduction to Mathematical Statistics*, Prentice Hall, 5th edition, 1995.