

UNIVERZITA PALACKÉHO V OLOMOUCI  
PŘÍRODOVĚDECKÁ FAKULTA

**DIPLOMOVÁ PRÁCE**

Obecné metody mnohonásobného porovnávání



**Katedra matematické analýzy a aplikací matematiky**

Vedoucí diplomové práce: **doc. RNDr. Eva Fišerová, Ph.D.**

Vypracovala: **Bc. Hana Zatloukalová**

Studijní program: N1103 Aplikovaná matematika

Studijní obor: Aplikace matematiky v ekonomii

Forma studia: prezenční

Rok odevzdání: 2016

## BIBLIOGRAFICKÁ IDENTIFIKACE

**Autor:** Bc. Hana Zatloukalová

**Název práce:** Obecné metody mnohonásobného porovnávání

**Typ práce:** Diplomová práce

**Pracoviště:** Katedra matematické analýzy a aplikací matematiky

**Vedoucí práce:** doc. RNDr. Eva Fišerová, Ph.D.

**Rok obhajoby práce:** 2017

**Abstrakt:** Tato diplomová práce je zaměřena na obecné metody mnohonásobného porovnávání. V teoretické části jsou tyto metody podrobně představeny a s nimi i několik souvisejících pojmů. V praktické části jsou potom metody aplikované na reálná data a je zde uvedena i simulační studie. Zpracování dat bylo provedeno ve statistickém softwaru R.

**Klíčová slova:** mnohonásobné porovnávání, Boferroniho metoda, Šidákova metoda, Holm-Bonferroniho metoda, Hochbergova metoda, Benjamini - Hochbergova metoda, Benjamini - Yekutieliho metoda, software R

**Počet stran:** 92

**Počet příloh:** 7

**Jazyk:** český

## BIBLIOGRAPHICAL IDENTIFICATION

**Author:** Bc. Hana Zatloukalová

**Title:** General methods for multiple comparisons

**Type of thesis:** Masters's

**Department:** Department of Mathematical Analysis and Application of Mathematics

**Supervisor:** doc. RNDr. Eva Fišerová, Ph.D.

**The year of presentation:** 2017

**Abstract:** This thesis deals with general methods for multiple comparisons. In the theoretical part are described the methods and also some related terms. In the practical part are these methods applied on real data and then is done a simulation study. Statistical software R was used for data processing.

**Key words:** multiple comparisons, Bonferroni method, Šidák method, Holm-Bonferroni method, Hochberg method, Benjamini - Hochberg method, Benjamini - Yekutieli method, software R

**Number of pages:** 92

**Number of appendices:** 7

**Language:** Czech

### **Prohlášení**

Prohlašuji, že jsem diplomovou práci napsala samostatně, za vedení paní doc. RNDr. Evy Fišerové, Ph.D. a že jsem v seznamu použité literatury uvedla všechny zdroje, ze kterých jsem při psaní práce čerpala.

V Olomouci dne .....

.....

podpis

## **Poděkování**

Ráda bych tímto poděkovala své vedoucí diplomové práce, paní doc. RNDr. Evě Fišerové, Ph.D., za odbornou spolupráci a za čas, který mi věnovala při konzultacích. Také bych chtěla poděkovat Mgr. Kamile Fačevicové, Ph.D., která mi rovněž věnovala svůj čas a konzultovala se mnou danou problematiku. Poděkování si také zaslouží má rodina, která mne ve studiu podporovala.

# Obsah

Úvod	8
<b>1 Úvodní pojmy</b>	<b>10</b>
1.1 Chyby prvního a druhého druhu . . . . .	10
1.2 Neupravená a upravená $p$ -hodnota . . . . .	17
1.3 Jednokrokové a vícekové metody . . . . .	18
1.4 Volné a omezené kombinace . . . . .	19
<b>2 Konstrukční metody</b>	<b>20</b>
2.1 Test sjednocení průniku . . . . .	20
2.2 Test průniku sjednocení . . . . .	22
2.3 Princip uzávěru . . . . .	23
2.4 Princip rozdělování . . . . .	28
<b>3 Metody kontrolující FWER</b>	<b>31</b>
3.1 Bonferroniho metoda . . . . .	32
3.2 Jednokroková Šidákova metoda . . . . .	33
3.3 Vícekové Šidákova metoda . . . . .	34
3.4 Holm-Bonferroniho metoda . . . . .	35
3.5 Simesův test . . . . .	38
3.6 Hochbergova metoda . . . . .	40
<b>4 Metody kontrolující FDR</b>	<b>43</b>
4.1 Benjamini-Hochbergova metoda . . . . .	43
4.2 Benjamini-Yekutieliho metoda . . . . .	44
<b>5 Přehled metod</b>	<b>46</b>
<b>6 Reálné příklady</b>	<b>50</b>
6.1 Příklad 1 - váhy kuřat . . . . .	51
6.2 Příklad 2 - mladí muži . . . . .	59
6.3 Příklad 3 - regrese . . . . .	65
<b>7 Simulace</b>	<b>70</b>

<b>Závěr</b>	<b>82</b>
<b>Literatura</b>	<b>83</b>
<b>Příloha A</b>	<b>85</b>
<b>Příloha B</b>	<b>86</b>
<b>Příloha C</b>	<b>87</b>
<b>Příloha D</b>	<b>89</b>
<b>Příloha E</b>	<b>90</b>
<b>Příloha F</b>	<b>91</b>
<b>Příloha G</b>	<b>92</b>

# Úvod

Tématem mé diplomové práce jsou metody mnohonásobného porovnávání. Na mnohonásobném porovnávání je založeno mnoho vědeckých experimentů. Typickým příkladem je využití například v lékařství (porovnávání několika léčebných postupů s nějakým kontrolním léčebným postupem), v zemědělství (porovnávání sklizně určitého druhu plodiny v závislosti na typu hnojiva), při chovu hospodářských zvířat (zjišťování, zda má způsob výživy vliv na hmotnostní přírůstky určitého druhu zvířat) a v mnoha dalších oblastech lidského života. Existují různé metody, které lze k tomuto testování využít. Hlavním cílem celé práce je vytvořit teoretický základ vybraných metod mnohonásobného porovnávání, následně tuto teorii aplikovat na reálná data a provést simulační studii. Pro zpracování praktické části jsem si vybrala program R. Ze začátku nebylo snadné se v tomto softwaru naučit pracovat, ale ve výsledku jsem si díky tomu výpočty značně zjednodušila.

Práce je rozdělena do sedmi kapitol. První kapitola obsahuje obecné pojmy, které je potřeba si připomenout, aby byl čtenář schopen snáze pochopit samotné metody. Je to tedy takový teoretický úvod. Druhá kapitola je věnovaná konstrukčním metodám a seznámí čtenáře s různými postupy, kterých se využívá v metodách mnohonásobného porovnávání. Třetí a čtvrtá kapitola obsahuje základ celé práce. Zde jsou popsány vybrané metody mnohonásobného porovnávání a podrobné postupy testování. Protože teorie je hodně a vzorce pro jednotlivé metody jsou podobné, přišlo mi vhodné v páté kapitole uvést stručný souhrn všech metod a vzorců, aby měl čtenář možnost vidět, jak jsou si metody podobné a v čem se naopak liší. Zároveň tento přehled metod může sloužit jako stručné opakování



a přehledná pomůcka při výpočtech. V šesté kapitole jsou pak uvedeny praktické příklady. Data pro výpočty jsem převzala jednak z datové knihovny programu R, ale využila jsem také data získaná z vlastního průzkumu. Poslední kapitola celé práce je věnovaná simulační studii, jejímž cílem je porovnání jednotlivých metod pomocí silofunkce.

V práci jsou také uvedeny kódy, které byly využity k výpočtům. Čtenář si tak může sám vše v softwaru R vyzkoušet.

# Kapitola 1

## Úvodní pojmy

Cílem této kapitoly je seznámit se se základními pojmy, které je potřeba znát, abychom se mohli dále zabývat metodami mnohonásobného porovnávání. Nejprve se zaměříme na chyby, kterých se nejčastěji při testování můžeme dopustit. V další části si vysvětlíme, co znamená neupravená a upravená  $p$ -hodnota, jaký je rozdíl mezi jednokrokovými a vícekrokovými testovacími postupy a seznámíme se i s dalšími důležitými pojmy.

Tato kapitola byla sepsána pomocí zdrojů [2, 3, 4, 5, 7, 9, 10, 13, 14].

### 1.1. Chyby prvního a druhého druhu

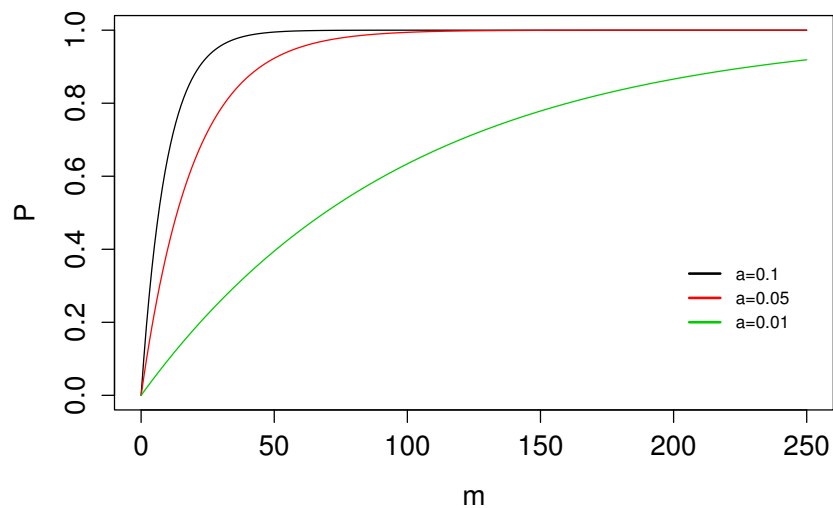
V této části se seznámíme s chybami, kterých se nejčastěji dopouštíme při mnohonásobném testování. Předpokládejme nejprve, že chceme otestovat jednoduchou nulovou hypotézu  $H^0$  proti alternativní hypotéze  $H^A$ . Předpokládejme dále, že k testování využijeme test, který je založený na testovací statistice  $T$ . Její realizaci pak značíme  $t$ . Označme písmenem  $W$  množinu všech hodnot testovací statistiky  $T$ , při kterých budeme  $H^0$  zamítat. Množinu  $W$  nazýváme *kritický obor*. Pro daný kritický obor  $W$  zamítneme nulovou hypotézu  $H^0$ , jestliže  $t \in W$  a nezamítneme  $H^0$ , jestliže  $t \notin W$ . *Chyba prvního druhu* potom nastane, když  $t \in W$ , tedy nulovou hypotézu  $H^0$  zamítneme, ale  $H^0$  platí. *Chyba druhého druhu* nastane když  $t \notin W$ , tedy nulovou hypotézu nezamítneme, ale přitom je  $H^0$  chybná. Kritický obor  $W$  vybíráme tak, abychom omezili pravděpodobnost chyby

prvního druhu nějakým pevně zvoleným malým číslem  $\alpha$ ,  $0 < \alpha < 1$ . Platnou nulovou hypotézu  $H^0$  pak zamítáme nejvýše s pravděpodobností  $\alpha$ . Číslu  $\alpha$  říkáme *hladina testu* a je to pravděpodobnost, že zamítneme nulovou hypotézu, ačkoliv daná hypotéza platí. Hladina testu se volí zpravidla  $\alpha = 0,05$  nebo  $\alpha = 0,01$ .

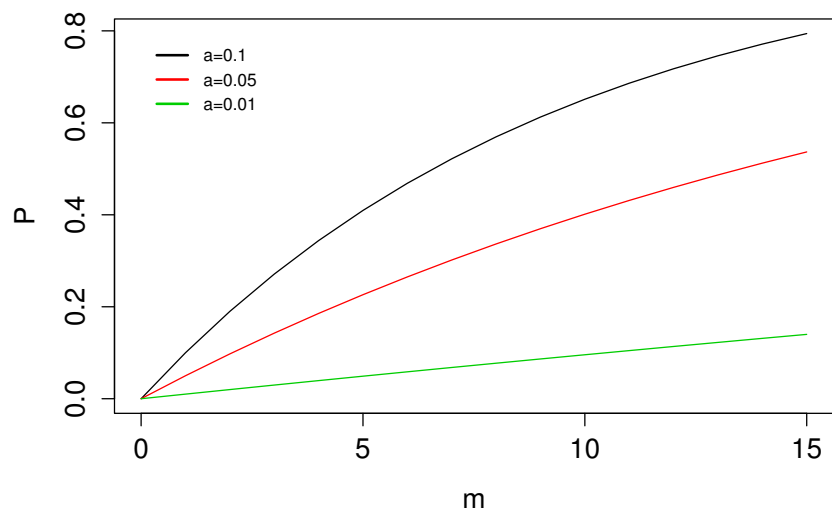
Předpokládejme nyní, že chceme otestovat dvojici hypotéz  $H_1^0$  a  $H_2^0$  na hladině významnosti  $\alpha = 0,05$ . Nechť nulová hypotéza  $H_i^0$  říká, že užívání léku  $i$ , kde  $i = 1, 2$ , nemá na lidský organismus lepší dopad než placebo. Dále předpokládejme, že obě hypotézy  $H_1^0$  a  $H_2^0$  jsou pravdivé. Pak je pravděpodobnost nezamítnutí obou hypotéz, za předpokladu nezávislosti, rovna číslu  $(1 - \alpha)^2 = 0,9025$ . Doplňující pravděpodobnost nesprávného zamítnutí aspoň jedné nulové hypotézy je pak  $1 - (1 - \alpha)^2 = 2\alpha - \alpha^2 = 0,0975$ . To je značně větší než počáteční hladina významnosti  $\alpha = 0,05$ . Obecně, jestliže testujeme  $m$  nulových hypotéz, tak pravděpodobnost dopuštění se nejméně jedné chyby prvního druhu je  $1 - (1 - \alpha)^m$ .

Grafy znázorňující pravděpodobnost dopuštění se nejméně jedné chyby prvního druhu pro  $m = 1, \dots, 250$  a  $\alpha = 0,01, 0,05$  a  $0,10$  jsou uvedeny na Obrázku 1.1 a Obrázku 1.2. Je jasné, že tato pravděpodobnost se pro dostatečně velké hodnoty  $m$  rychle blíží k 1. Jinak řečeno, při velkém počtu testovaných hypotéz a bez žádných úprav pro mnohonásobné testování, se rozhodovatel dopustí chyby prvního druhu skoro jistě. Přesné hodnoty těchto pravděpodobností jsou uvedeny v Tabulce 1.1.

Nechť  $m$  značí počet testovaných nulových hypotéz  $H_1^0, \dots, H_m^0$  a  $M = \{1, \dots, m\}$  množinu indexů těchto hypotéz. Nechť  $M_0 \subseteq M$  značí množinu pravdivých hypotéz a  $m_0$  počet těchto pravdivých hypotéz. Dále nechť náhodné veličiny  $U$  a  $S$  značí počet správných rozhodnutí o hypotézách. Veličina  $U$  značí počet nezamítnutých správných nulových hypotéz a veličina  $S$  značí počet zamítnutých chybných hypotéz. Náhodná veličina  $V$  značí počet chyb prvního druhu a náhodná veličina  $T$  značí počet chyb druhého druhu. Náhodná veličina  $N$  pak značí celkový počet nezamítnutých hypotéz a náhodná veličina  $R$  značí celkový počet zamítnutých nulových hypotéz. Vše je názorně uvedeno v Tabulce 1.2. Čísla  $m$  a  $m_0$  jsou neměnná, kde  $m$  je počet testovaných hypotéz a  $m_0$  je neznámý



Obrázek 1.1: Pravděpodobnost dopuštění se nejméně jedné chyby prvního druhu pro různé počty testovaných hypotéz  $m$  a hladiny významnosti  $\alpha = 0,1$ ,  $\alpha = 0,05$ ,  $\alpha = 0,01$ .



Obrázek 1.2: Pravděpodobnost dopuštění se nejméně jedné chyby prvního druhu pro různé počty testovaných hypotéz  $m$  a hladiny významnosti  $\alpha = 0,1$ ,  $\alpha = 0,05$ ,  $\alpha = 0,01$  - detail pro malé počty testovaných hypotéz.

$m$	$\alpha = 0,1$	$m$	$\alpha = 0,05$	$m$	$\alpha = 0,01$
1	0,1	1	0,05	1	0,01
2	0,19	2	0,0975	2	0,0199
3	0,271	3	0,142625	3	0,029701
4	0,3439	4	0,18549375	4	0,03940399
5	0,40951	5	0,226219063	5	0,04900995
6	0,468559	6	0,264908109	6	0,058519851
7	0,5217031	7	0,301662704	7	0,067934652
8	0,56953279	8	0,336579569	8	0,077255306
9	0,612579511	9	0,36975059	9	0,086482753
10	0,65132156	10	0,401263061	10	0,095617925
11	0,686189404	11	0,431199908	...	...
12	0,717570464	12	0,459639912	83	0,565768673
13	0,745813417	13	0,486657917	84	0,570110986
14	0,771232075	14	0,512325021	85	0,574409877
15	0,794108868	15	0,53670877	86	0,578665778
16	0,814697981	16	0,559873331	87	0,58287912
17	0,833228183	17	0,581879665	88	0,587050329
18	0,849905365	18	0,602785682	89	0,591179826
...	...	...	...	...	...
43	0,989224736	89	0,989591195	458	0,989978814
44	0,990302263	90	0,990111635	459	0,990079026
45	0,991272036	91	0,990606054	460	0,990178236
46	0,992144833	92	0,991075751	461	0,990276453
47	0,99293035	93	0,991521963	462	0,990373689
48	0,993637315	94	0,991945865	463	0,990469952
49	0,994273583	95	0,992348572	464	0,990565252
50	0,994846225	96	0,992731143	465	0,9906596
51	0,995361602	97	0,993094586	466	0,990753004
52	0,995825442	98	0,993439857	467	0,990845474

Tabulka 1.1: Pravděpodobnost dopuštění se nejméně jedné chyby prvního druhu při testování  $m$  hypotéz.

počet platných nulových hypotéz. Realizace těchto náhodných veličin budeme značit  $u, v, t, s, n, r$ . Platí, že  $R = V + S$  a  $N = m - R = U + T$ . Pokud testujeme hypotézy jednotlivě na hladině  $\alpha$ , pak  $R(\alpha)$  je rostoucí funkcí proměnné  $\alpha$ .

Standardním přístupem k testování jednoduchých hypotéz ( $m = 1$ ) je výběr vhodného testu, který bude kontrolovat pravděpodobnost chyby prvního druhu na předem stanovené hladině významnosti  $\alpha$ . Při mnohonásobném testování je

situace obtížnější a využívají se při něm některé z následujících veličin.

Hypotéza	$H^0$ nezamítáme	$H^0$ zamítáme	Celkem
$H^0$ je správná	U	V	$m_0$
$H^0$ je chybná	T	S	$m - m_0$
Celkem	N	R	$m$

Tabulka 1.2: Chyby prvního a druhého druhu v mnohonásobném testování hypotéz.

**Definice 1.1.1** Veličina *PCER* (*per-comparison error rate neboli chyba na jednu hypotézu*) je definovaná jako podíl očekávaného počtu chyb prvního druhu a počtu testovaných hypotéz  $m$ :

$$PCER = \frac{E(V)}{m}.$$

Jestliže je každá z  $m$  hypotéz testovaná zvlášť na předem stanovené hladině významnosti  $\alpha$ , potom je  $PCER = \alpha m_0 / m \leq \alpha$ . Nicméně v mnoha aplikacích není kontrola veličiny *PCER* na předem stanovené hladině významnosti  $\alpha$  vhodná.

**Definice 1.1.2** Veličina *PFER* (*per-family error rate neboli chyba na skupinu hypotéz*) je definovaná jako očekávaný počet chyb prvního druhu:

$$PFER = E(V).$$

**Definice 1.1.3** Veličina *FWER* (*familywise error rate neboli chyba přes skupinu hypotéz*) je definovaná jako pravděpodobnost dopuštění se nejméně jedné chyby prvního druhu:

$$FWER = P(V > 0).$$

Veličina *FWER* vyjadřuje pravděpodobnost, že alespoň jedna hypotéza ze všech testovaných hypotéz je chybně zamítnuta. Zatímco veličina *PFER* počítá

se skupinou hypotéz jako s jedním celkem, tak veličina  $FWER$  bere v úvahu postupně všechny podmnožiny celé skupiny hypotéz. Místo kontroly pravděpodobnosti chyby prvního druhu pro každý test na předem stanovené hladině významnosti  $\alpha$ , je pak kontrolována  $FWER$  na hladině významnosti  $\alpha$ . Pro  $m = 1$  je  $FWER$  pravděpodobnost chyby prvního druhu. Veličina  $FWER$  je nejběžněji využívanou veličinou v mnohonásobném testování. Využívá se zejména tam, kde je počet porovnávání rozumný.

Pokud je počet nulových hypotéz  $m$  příliš velký, může být kontrola veličiny  $FWER$  na hladině  $\alpha$  příliš přísná. V takovém případě je možné uvažovat obecnější verzi veličiny  $FWER$ , a to  $gFWER$ .

**Definice 1.1.4** Veličina  $gFWER$  (*generalized familywise error rate neboli zobecněná chyba přes skupinu hypotéz*) je definovaná jako pravděpodobnost dopuštění se více než  $k$  chyb prvního druhu, kde číslo  $k$  je předem stanovené:

$$gFWER = P(V > k).$$

Jestliže je celkový počet nulových hypotéz  $m$  velký, tak může být malý počet  $k$  chyb prvního druhu akceptován. Kontrola veličiny  $gFWER$  na hladině významnosti  $\alpha$  je tedy méně přísná než kontrola veličiny  $FWER$  na hladině významnosti  $\alpha$ . Metody kontrolující  $gFWER$  mohou povolit malý počet chyb prvního druhu za předpokladu, že tento počet je menší nebo roven předem určenému číslu  $k$ .

**Definice 1.1.5** Veličina  $FDR$  (*false discovery rate neboli podíl chybně zamítnutých hypotéz*) je definovaná jako střední hodnota poměru chybně zamítnutých nulových hypotéz a celkového počtu zamítnutých hypotéz ( $Q = V/R$ ):

$$FDR = E(Q).$$

Jestliže  $R > 0$ , pak  $Q = V/R$ . Pro  $R = 0$  položíme  $Q = 0$ . Potom veličinu  $FDR$  můžeme vypočítat následovně:

$$FDR = E(Q) = E\left(\frac{V}{R} \mid R > 0\right) \cdot P(R > 0) + 0 \cdot P(R = 0) = E\left(\frac{V}{R} \mid R > 0\right) \cdot P(R > 0).$$

Uveďme si nyní krátký ilustrativní příklad, ve kterém si budeme demonstrovat užití veličiny  $FDR$ . Uvažujme nějaký nový přístup k léčbě hepatitidy a chceme ho porovnávat se stávajícím léčebným postupem. Nechť nulová hypotéza  $H^0$  říká, že nový postup není lepší než stávající. Tyto dva léčebné postupy budeme srovnávat pomocí  $m$  kritérií (výše jaterních enzymů, koncentrace protilátek, atd.). Původní nulová hypotéza se nám tedy rozdělí na  $m$  dílčích nulových hypotéz, které si označíme  $H_i^0$ ,  $i = 1, \dots, m$ . Jestliže nový postup bude lepší než stávající, pak nastane více zamítnutí jednotlivých hypotéz než přijmutí. Kontrola veličiny  $FWER$  by nám tedy v tomto případě nebyla nic platná, protože kontroluje počet chybně zamítnutých hypotéz  $V$ . Spíše by se nám líbilo, kdybychom uměli kontrolovat poměr mezi počtem chybně zamítnutých hypotéz  $V$  a celkovým počtem zamítnutých hypotéz  $R$ , tj. kontrola veličiny  $FDR$  na hladině  $q$ , kde  $q \leq \alpha$ .

Obecně platí, že  $PCER \leq FWER \leq PFER$ . Metody kontrolující  $PFER$  jsou obecně konzervativnější a vedou k méně zamítnutým hypotézám než ty metody, které kontrolují  $FWER$  nebo  $PCER$ . Metody kontrolující  $FWER$  jsou pak konzervativnější než metody, které kontrolují  $PCER$ . Předpokládejme nyní, že každá hypotéza  $H_j$  je testována na hladině  $\alpha_j$  a rozhodnutí o této hypotéze závisí výhradně na zvoleném testu. Pokud  $m = m_0$ ,  $PCER$  je průměr hladin významnosti  $\alpha_j$  a  $PFER$  je součet hladin významnosti  $\alpha_j$ .  $FWER$  je pak funkce, která nezávisí jen na  $\alpha_j$ , ale také na sdružené distribuci testovacích statistik  $T_j$ . Můžeme tedy psát

$$\begin{aligned} PCER &= \frac{\alpha_1 + \dots + \alpha_m}{m} \leq \max(\alpha_1, \dots, \alpha_m) \leq \\ &\leq FWER \leq PFER = \alpha_1 + \dots + \alpha_m. \end{aligned}$$

Zaměřme se nyní na vztah mezi veličinami  $FWER$  a  $FDR$ .  $FDR$  také závisí na sdružené distribuci testovacích statistik. Jestliže platí, že  $m_0 = m$ , pak  $FWER = FDR$  a jestliže je  $m_0 < m$ , pak  $FWER \geq FDR$ . Obecně pak platí, že  $PCER \leq FDR \leq FWER$ .

Z výše uvedených vlastností je zřejmé, že jestliže určitá metoda kontroluje veličinu  $FWER$ , pak kontroluje i  $FDR$ . Pokud bychom předpokládali platnost



všech nulových hypotéz ( $m = m_0$ ), pak by byl podíl  $\frac{V}{R}$  roven jedné. Tomu se dá vyhnout tím, že budeme předpokládat, že alespoň jedna nulová hypotéza není platná.

Obecně, veličiny *FWER* a *PFER* rostou s rostoucím  $m$ , zatímco *PCER* zůstává stejná. Pokud je  $m = m_0$ ,  $FDR = FWER$  a s rostoucím  $m$  také rostou. Nicméně pokud podíl  $m_0/m$  klesá,  $FDR$  zůstává relativně stabilní a přibližuje se k *PCER*. Pro větší  $m$  (v tisících) se chyby často ustálí a vykazují podobné chování.

Rozlišujeme *silnou kontrolu* a *slabou kontrolu* chyby prvního druhu. Silná kontrola se týká kontroly chyby prvního druhu každé kombinace pravdivých a nepravdivých nulových hypotéz, tj. každé podmnožiny  $M_0 \subseteq \{1, \dots, m\}$  pravdivých nulových hypotéz. Slabá kontrola potom znamená kontrolu chyby prvního druhu v případě, že jsou všechny nulové hypotézy pravdivé, tj. když  $m_0 = m$ . Obecně není moc pravděpodobné, že budou všechny hypotézy pravdivé a slabá kontrola je proto neuspokojivá.

## 1.2. Neupravená a upravená $p$ -hodnota

Výpočet  $p$ -hodnot je běžným úkolem při testování jednorozměrných hypotéz. V mnohonásobném testování je žádoucí počítat takzvané *upravené  $p$ -hodnoty*.

Klasická  $p$ -hodnota je definovaná jako nejmenší hladina, pro kterou bychom hypotézu ještě zamítli. Jestliže máme vysloveny předpoklady o rozdělení pravděpodobností zkoumané náhodné veličiny  $X$  a pokud je formulovaná nulová a alternativní hypotéza, zvolíme vhodnou výběrovou funkci  $T = T(X_1, \dots, X_n)$ , které říkáme testovací statistika. Hodnotu testovací statistiky značíme  $t = T(\mathbf{x})$ .  $P$ -hodnota potom také vyjadřuje pravděpodobnost spočítanou za platnosti nulové hypotézy, že dostaneme právě hodnotu  $t = T(\mathbf{x})$  nebo hodnotu, která ještě více odporuje testované hypotéze. Nulovou hypotézu  $H^0$  zamítáme na hladině testu  $\alpha$ , jestliže je  $p$ -hodnota menší nebo rovna číslu  $\alpha$ . Pokud je  $p$ -hodnota větší než  $\alpha$ , pak  $H^0$  na uvedené hladině významnosti nelze zamítnout.

Nechť  $\theta$  je parametr, který chceme testovat. Oboustranná alternativa od-

povídá nulové hypotéze  $H^0 : \theta = \theta_0$  proti alternativní hypotéze  $H^A : \theta \neq \theta_0$ . Levostranná alternativa odpovídá nulové hypotéze  $H^0 : \theta \geq \theta_0$  proti alternativní hypotéze  $H^A : \theta < \theta_0$ . Pravostranná alternativa odpovídá nulové hypotéze  $H^0 : \theta \leq \theta_0$  proti alternativní hypotéze  $H^A : \theta > \theta_0$ . Pro oboustrannou alternativu je  $p$ -hodnota daná vztahem  $p = 2\min\{P(T \leq t|H^0), P(T \geq t|H^0)\}$ , pro levostrannou alternativu je  $p = P(T \leq t|H^0)$  a pro pravostrannou alternativu je  $p = P(T \geq t|H^0)$ .

Tyto marginální  $p$ -hodnoty  $p$  jsou v tomto textu označovány jako neupravené  $p$ -hodnoty. *Upravená  $p$ -hodnota*  $q_i$  je potom definovaná jako nejmenší hladina významnosti, pro kterou se ještě jednoduchá hypotéza  $H_i^0$ ,  $i \in M$ , pro danou metodu mnohonásobného porovnávání zamítá. Například v případě kontroly veličiny *FWER*

$$q_i = \inf\{\alpha \in (0, 1) : H_i^0 \text{ je zamítnuta na hladině } \alpha\},$$

pokud takové číslo  $\alpha$  existuje, jinak  $q_i = 1$ . Pokud tedy kontrolujeme *FWER* na hladině významnosti  $\alpha$ , tak skupinu nulových hypotéz  $H_i^0$  zamítáme, jestliže  $q_i \leq \alpha$ . Upravené  $p$ -hodnoty pro jednotlivé metody mnohonásobného porovnávání získáme přepočtem pomocí daných vzorců uvedených u jednotlivých metod.

### 1.3. Jednokrokové a vícekové metody

Jednou z možností, jak rozdělit metody mnohonásobného porovnávání, je dělení na *jednokrokové* a *vícekové metody*. Pro jednokrokové metody je charakteristické, že při zamítnutí nebo nezamítnutí jednotlivých nulových hypotéz se nebere v úvahu rozhodnutí o jakékoliv jiné hypotéze. Tudíž nezáleží na pořadí, v jakém jsou hypotézy testovány a rozhodnutí o nulových hypotézách se provádí v jednom kroku. Jednokrokové metody jsou jednoduché, avšak pro kontrolu veličiny *FWER* bývají příliš konzervativní. Jednokrokové metody jsou obecně méně spolehlivé než vícekové metody.

Na rozdíl od jednokrokových metod, zamítnutí nebo nezamítnutí nulových hypotéz u vícekových metod, může záviset na rozhodnutí o jiných hypotézách.

Víceřádkové metody jsou dále rozděleny na *sestupné* a *vzestupné metody*. Obě skupiny metod předpokládají uspořádanou posloupnost hypotéz  $H_{(1)}^0 \prec \dots \prec H_{(m)}^0$ , kde symbol  $\prec$  značí uspořádání hypotéz podle statistické významnosti, tj. podle odpovídajících  $p$ -hodnot. Sestupné metody začínají testováním první z uspořádaných hypotéz  $H_{(1)}^0$  (odpovídající nejmenší neupravené  $p$ -hodnotě) a pokračují další hypotézou v pořadí, dokud dochází k zamítání hypotéz. Metoda končí, když najdeme první hypotézu, kterou nelze zamítnout a potom nelze zamítnout ani ostatní hypotézy s vyšším pořadím. Pokud je tedy například hypotéza  $H_{(i)}^0$  první, která se nezamítá, tak předchozí hypotézy  $H_{(1)}^0, \dots, H_{(i-1)}^0$  se zamítají. Vzestupné metody naopak začínají testováním hypotézy  $H_{(m)}^0$  a pokračují další hypotézou v pořadí, tedy v opačném směru než sestupné metody, dokud se nezamítá. Tyto metody končí prvním zamítnutím a zamítají se i ostatní hypotézy s nižším pořadím. Jestliže je tedy první zamítnutou hypotézou hypotéza  $H_{(i)}^0$ , tak se zamítají všechny hypotézy  $H_{(1)}^0, \dots, H_{(i)}^0$ .

## 1.4. Volné a omezené kombinace

Skupina nulových hypotéz  $H_i^0, i \in M$ , splňuje *podmínku volné kombinace*, jestliže pro libovolnou množinu  $I \subseteq M$  platí, že hypotézy  $H_i^0, i \in I$  jsou pravdivé a zbytek jsou nepravdivé hypotézy. Jinak hypotézy  $H_1^0, \dots, H_m^0$  splňují *podmínku omezené kombinace*.

Oba pojmy si objasníme v následujících jednoduchých příkladech. Podmínka volné kombinace: uvažujme srovnání dvou nových postupů (např. léčebných) s kontrolním postupem ( $m = 2$ ). Potom nastane nějaký ze tří případů: žádný, jeden nebo oba nové postupy jsou lepší než kontrolní postup. Podmínka omezené kombinace: uvažujme srovnání všech párových kombinací tří léčebných postupů pomocí  $\theta_1, \theta_2$  a  $\theta_3$  ( $m = 3$ ). V tomto případě ne všechny kombinace nulových a alternativních hypotéz jsou možné. Například jestliže  $\theta_1 \neq \theta_2$ , pak nemohou být současně pravdivé případy, kdy  $\theta_1 = \theta_3$  a  $\theta_2 = \theta_3$ .

# Kapitola 2

## Konstrukční metody

Ke konstrukci metod mnohonásobného porovnávání lze využít různé postupy, jako například *testy sjednocení průniku*, ke konstrukci metod pro průnik několika jednoduchých hypotéz nebo *testy průnik sjednocení*, ke konstrukci metod pro sjednocení několika jednoduchých hypotéz. Dále pak *princip uzávěru* a *princip rozdělování*, které rozšiřují princip testu sjednocení průniku a umožňují vyslovit závěry pro jednotlivé jednoduché hypotézy.

Tato kapitola byla sepsána pomocí zdrojů [4, 14].

### 2.1. Test sjednocení průniku

Test sjednocení průniku je historicky první konstrukční metoda pro metody mnohonásobného porovnávání. Například předpokládejme, že máme k dispozici několik zavlažovacích systémů, které jsou srovnávány s určitým kontrolním zavlažovacím systémem. Je přirozené prohlásit za úspěch, když bude mít nejméně jeden ze srovnávaných zavlažovacích systémů lepší výsledky než kontrolní systém. Jestliže  $H_i^0$  značí jednoduché hypotézy, které říkají, že zavlažovací systém  $i$  a kontrolní systém mají stejný dopad, pak chceme správně zamítnout libovolnou (nejméně však jednu) nesprávnou hypotézu  $H_i^0$ .

Nyní dáme tomuto problému formální podobu. Mějme skupinu nulových hypotéz  $H_i^0$  a k nim alternativní hypotézy, které budeme značit  $H_i^A$ ,  $i \in M$ . Zajímá nás testování průniku nulových hypotéz  $H_i^0$ , tedy  $H = \bigcap_{i \in M} H_i^0$ . Možný přístup

je použit testovací statistiky  $T_i, i \in M$ , a zamítnout  $H$ , jestliže libovolná testovací statistika  $T_i$  překročí její odpovídající kritickou hodnotu  $c_i$ . Celková oblast zamítnutí je pak sjednocením jednotlivých oblastí zamítnutí,  $\bigcup_{i \in M} \{t_i > c_i\}$ . Tato konstrukce vede k testům sjednocení průniku, které testují průnik skupiny nulových hypotéz  $H_i^0$ , oproti sjednocení skupiny alternativních hypotéz  $H_i^A$ , tj.

$$H = \bigcap_{i \in M} H_i^0$$

oproti

$$A = \bigcup_{i \in M} H_i^A.$$

Nutno poznamenat, že testy sjednocení průniku uvažují celkový průnik nulových hypotéz  $H$  bez možnosti učinit samostatné závěry o jednoduchých hypotézách  $H_1^0, \dots, H_m^0$ . To znamená, že jestliže hypotézu  $H$  testem sjednocení průniku zamítneme, stále zbývá otázka, které z jednoduchých hypotéz  $H_i^0$  by měly být zamítnuty. Tento nedostatek může být vyřešen několika způsoby, jako například aplikací principu uzávěru nebo principu rozdělování.

Důležitou skupinu testů sjednocení průniku tvoří *max-t testy*. Nechť máme hypotézy  $H_1^0, \dots, H_m^0$  a k nim jednotlivé testovací statistiky  $T_1, \dots, T_m$ . Bez ztráty na obecnosti předpokládejme, že velké hodnoty  $T_i$  vedou k zamítnutí  $H_i^0$ . Přírodným postupem je potom uvažovat maximum z jednotlivých testovacích statistik  $T_i$ , což vede přímo k max-t testu

$$T_{max} = \max\{T_1, \dots, T_m\}.$$

Celkovou nulovou hypotézu  $H$  potom zamítáme pouze tehdy, když  $T_{max} \geq c$ , kde konstanta  $c$  je vybírána tak, aby docházelo ke kontrole chyby prvního druhu na hladině významnosti  $\alpha$ , tedy  $P(T_{max} \geq c | H) = \alpha$ . Kritická hodnota  $c$  je počítána ze sdružené distribuční funkce náhodných proměnných  $T_1, \dots, T_m$ . Určení konstanty  $c$  je často obtížné a někdy dokonce nemožné, pokud sdružená distribuční funkce není známá. Pomocí max-t testu je konstruováno mnoho známých metod mnohonásobného porovnávání, jako například Bonferroniho test, Dunnettův test nebo Tukeyho test.

Jestliže k zamítnutí hypotézy  $H_i^0$  vedou malé hodnoty  $T_i$ , musí být bráno namísto maxima minimum z jednotlivých testovacích statistik  $T_i$ , tj.  $T_{min} = \min\{T_1, \dots, T_m\}$  a kritickou hodnotu  $c$  vybíráme tak, aby  $P(T_{min} \leq c | H) = \alpha$ .

## 2.2. Test průniku sjednocení

Uvažujme příklad z oblasti vývoje léčiv. Mezinárodní směrnice nařizují, že kombinovaná léčba (užívání dvou a více léků zároveň) musí mít před uvedením na trh prokazatelně prospěšnější dopad oproti užívání jednotlivých léčiv samostatně. V porovnání s testem sjednocení průniku je nyní požadováno, že všechny nulové hypotézy bez prospěšného dopadu jsou zamítnuty, abychom mohli prohlásit, že kombinovaná léčba má prospěšný efekt.

Mějme skupinu nulových hypotéz  $H_i^0$  a k nim alternativní hypotézy  $H_i^A$ , kde  $i \in M$ . Zajímá nás testování sjednocení nulových hypotéz  $H_i^0$ , tj.  $H = \bigcup_{i \in M} H_i^0$ . Možným přístupem je opět použít testovací statistiku  $T_i, i \in M$ , přičemž k zamítnutí nulové hypotézy vedou velké hodnoty  $T_i$ . Celková oblast zamítnutí je pak průnikem jednotlivých oblastí zamítnutí,  $\bigcap_{i \in M} \{t_i > c_i\}$ . Tato konstrukce vede k testům průniku sjednocení, které testují sjednocení skupiny nulových hypotéz  $H_i^0$ , oproti průniku skupiny alternativních hypotéz  $H_i^A$ , tj.

$$H' = \bigcup_{i \in M} H_i^0$$

oproti

$$A' = \bigcap_{i \in M} H_i^A.$$

Test průniku sjednocení pak zamítá sjednocení nulových hypotéz  $H'$  na celkové hladině  $\alpha$ , jestliže jsou všechny jednoduché hypotézy  $H_i^0$  zamítnuty jejich lokálním testem. Jestliže mají všechny testovací statistiky  $T_i, i \in M$ , stejnou marginální distribuční funkci, pak test průniku sjednocení zamítá  $H'$ , pouze když  $\min_{i \in M} t_i \geq c$ , kde  $c$  je  $(1 - \alpha)$ -kvantil z příslušné marginální distribuční funkce. V tomto konkrétním případě se testu průniku sjednocení také říká *min-t test*.

Vraťme se nyní k příkladu z oblasti vývoje léčiv, který byl zmíněn v úvodu této části. Předpokládejme, že máme  $m$  testovacích statistik, které srovnávají kombinovanou léčbu s léčbou jednotlivými léky. Předpokládejme, že jednotlivé statistiky mají Studentovo rozdělení pravděpodobnosti. Hypotézu  $H'$  zamítáme a tedy usoudíme, že kombinovaná léčba má prospěšný dopad, jestliže je nejmenší hodnota testovací statistiky větší než  $(1 - \alpha)$ -kvantil Studentova rozdělení.

Jestliže jsou některé z nulových hypotéz  $H_i^0$  zamítnuty lokálně, tj. v případě min-testu je pro nějaké  $i$ ,  $t_i > c$ , stále bereme v úvahu sjednocení nulových hypotéz  $H'$  a není možné usuzovat o jednotlivých hypotézách samostatně, neboť by nebyla kontrolována *FWER*.

## 2.3. Princip uzávěru

Testy sjednocení průniku testují celkovou hypotézu  $H$  bez možnosti úsudku o jednotlivých hypotézách  $H_1^0, \dots, H_m^0$ . To znamená, že pokud testem sjednocení průniku zamítneme hypotézu  $H = \bigcap_{i \in M} H_i^0$ , nemůžeme vyslovit žádný závěr o jednotlivých hypotézách  $H_i^0$ . Princip uzávěru je obecná konstrukční metoda, která vede k vícekrokovým testovacím metodám a umožňuje vyslovit závěry o jednotlivých hypotézách  $H_i^0$ .

Abychom mohli popsat princip uzávěru, uvažujme nejprve příklad se dvěma nulovými hypotézami  $H_1^0$  a  $H_2^0$ . Předpokládejme, že chceme zjistit, zda je nějaký ze dvou nových léčebných postupů (například využití dvou nových léků) lepší, než určitý kontrolní (již zavedený) léčebný postup. Nechť  $\mu_j$  značí průměrný dopad léčby  $j$ , kde  $j = 0, 1, 2$  a nula je označení pro kontrolní skupinu. Dále nechť  $\theta_i = \mu_i - \mu_0$ , značí rozdíl průměrného dopadu léčby  $i = 1, 2$  a kontrolní léčby. Parametry  $\theta_i$  jsou parametry, které nás zajímají a výsledné jednoduché hypotézy jsou tedy ve tvaru  $H_i^0 : \theta_i \leq 0$ ,  $i = 1, 2$ . Užitím Bonferroniho testu, který si uvedeme v Kapitole 3, by byla každá hypotéza  $H_i^0$  testována zvlášť na hladině významnosti  $\alpha/2$  a *FWER* by byla kontrolována na hladině významnosti  $\alpha$ . Bonferroniho test může být vylepšen aplikací principu uzávěru.

Nyní budeme považovat hypotézy  $H_i^0$  za podmnožiny parametrického pro-

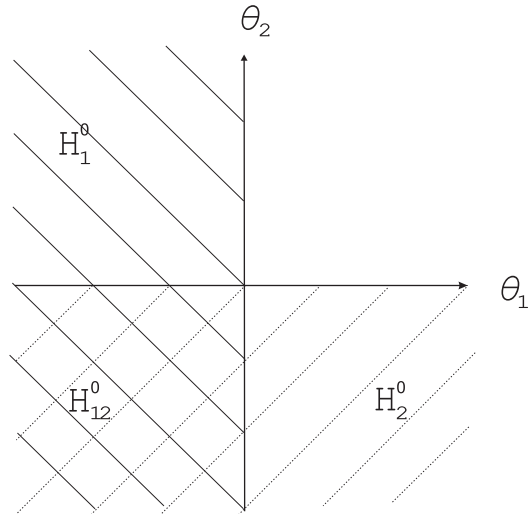
storu, o kterém chceme učinit závěr. Necht'  $\Theta = \mathbb{R}^2$  značí parametrický prostor obsahující  $\theta^*$ , kde  $\theta^* = (\theta_1, \theta_2) \in \Theta$ . Obrázek 2.1 znázorňuje hypotézy  $H_i^0 = \{\theta^* \in \mathbb{R}^2 : \theta_i \leq 0\}$ , kde  $i = 1, 2$ , jako podmnožiny reálné roviny. Je jasné, že dvojice hypotéz  $H_1^0$  a  $H_2^0$  nejsou disjunktní, protože jejich průnik je dán vztahem  $H_{12}^0 = H_1^0 \cap H_2^0 = \{\theta^* \in \mathbb{R}^2 : \theta_1 \leq 0 \wedge \theta_2 \leq 0\}$ , což je dolní levý kvadrant. Testování průniku  $H_{12}^0$  vyžaduje úpravu mnohonásobnosti, která je zahrnuta například u Bonferroniho testu. Bonferroniho test testuje celé sjednocení  $H_1^0 \cup H_2^0$  na hladině významnosti  $\alpha/2$  ( $\alpha/m$ ) a ne pouze průnik hypotéz  $H_{12}^0$ . Nicméně Obrázek 2.1 naznačuje, že zbylé části  $H_1^0 \setminus H_{12}^0$  a  $H_2^0 \setminus H_{12}^0$  mohou být obě testovány na celé hladině významnosti  $\alpha$  bez potřeby dalších úprav mnohonásobnosti. To vede k tomu, že budeme nejprve testovat průnik hypotéz  $H_{12}^0$  pomocí vhodného testu sjednocení průniku. Pokud hypotézu  $H_{12}^0$  zamítneme, pokračujeme testováním hypotéz  $H_1^0$  a  $H_2^0$  na hladině významnosti  $\alpha$ . Nulovou hypotézu  $H_1^0$  zamítneme (zatímco kontrolujeme *FWER* na hladině významnosti  $\alpha$ ) pouze tehdy, když zamítneme obě hypotézy  $H_1^0$  a  $H_{12}^0$  každou na hladině významnosti  $\alpha$ . Stejně tak i hypotézu  $H_2^0$  zamítneme pouze tehdy, když zamítneme obě hypotézy  $H_2^0$  a  $H_{12}^0$ . Pokud však v prvním kroku hypotézu  $H_{12}^0$  nezamítneme, další testování je zbytečné. Tato konstrukční metoda je klíčovou myšlenkou principu uzávěru.

Existují i další možnosti jak znázornit princip uzávěru. Na Obrázku 2.2 jsou znázorněny nulové hypotézy  $H_1^0$  a  $H_2^0$  a jejich průnik  $H_{12}^0$  pomocí Vennova diagramu a na Obrázku 2.3 je diagram, který jasně znázorňuje vztahy mezi testovanými hypotézami. V dolní části jsou dvě základní hypotézy  $H_1^0$  a  $H_2^0$  a jejich průnik  $H_{12}^0$ , je na vrcholu diagramu a testuje se vlastně směrem shora dolů.

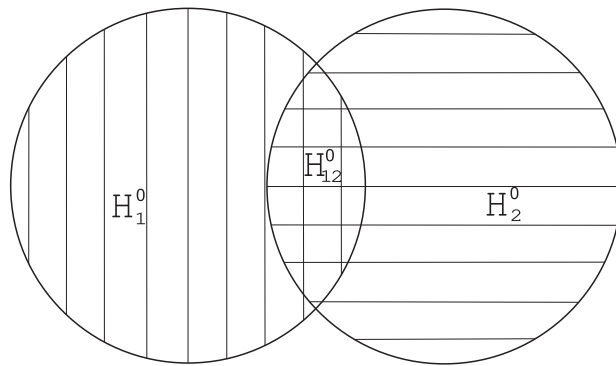
Nyní předpokládejme, že chceme testovat  $m$  hypotéz  $H_i^0, i = 1, \dots, m$  (např. srovnávání  $m$  testovaných skupin se skupinou kontrolní). Princip uzávěru uvažuje všechny průniky hypotéz, které lze sestavit ze základní množiny hypotéz. Každý průnik je testován na lokální hladině významnosti  $\alpha$ . Pro různé hypotézy mohou být použity různé testy sjednocení průniku. Základní nulová hypotéza  $H_i^0$  je zamítnuta pouze tehdy, když jsou na lokální hladině významnosti  $\alpha$  zamítnuty



všechny průniky obsahující  $H_i^0$ .



Obrázek 2.1: Znáornění dvou nulových hypotéz  $H_1^0$  a  $H_2^0$  a jejich průniku  $H_{12}^0$  v parametrickém prostoru.



Obrázek 2.2: Znáornění dvou nulových hypotéz  $H_1^0$  a  $H_2^0$  a jejich průniku  $H_{12}^0$  pomocí Vennova diagramu.

$$H_{12}^0 = H_1^0 \cap H_2^0$$

Obrázek 2.3: Schéma principu uzávěru pro dvě nulové hypotézy  $H_1^0$  a  $H_2^0$  a jejich průnik  $H_{12}^0$ .

Metody mnohonásobného porovnávání založeny na principu uzávěru jsou prováděny následovně:

1. Definujeme množinu základních hypotéz  $\mathcal{H} = \{H_1^0, \dots, H_m^0\}$ .
2. Vytvoříme uzávěr

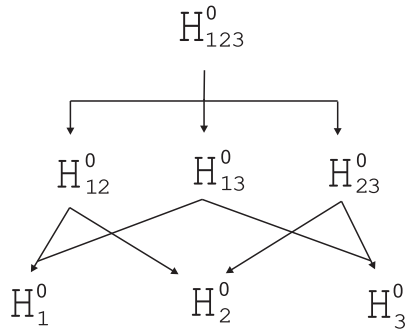
$$\overline{\mathcal{H}} = \left\{ H_I = \bigcap_{i \in I} H_i^0 : I \subseteq \{1, \dots, m\}, H_I \neq \emptyset \right\}.$$

3. Pro všechny neprázdné průniky hypotéz  $H_I \in \overline{\mathcal{H}}$  najdeme vhodný lokální test na hladině  $\alpha$ .
4. Hypotézu  $H_i^0$  zamítneme, jestliže zamítneme všechny hypotézy  $H_I \in \overline{\mathcal{H}}$ ,  $i \in I$ , každou na hladině  $\alpha$ .

Pokud provádíme metody založené na principu uzávěru, upravené  $p$ -hodnoty počítáme následovně. Základní nulová hypotéza  $H_i^0$  je zamítnuta pouze tehdy, když jsou zamítnuty všechny hypotézy  $H_I \in \overline{\mathcal{H}}$ , kde  $i \in I$ , a maximální  $p$ -hodnota z této množiny musí být menší nebo rovna číslu  $\alpha$ . Nechť  $p_I$  značí  $p$ -hodnotu, pro danou hypotézu  $H_I, I \subseteq \{1, \dots, m\}$ . Upravená  $p$ -hodnota pro  $H_i^0$  je definována vztahem

$$q_i = \max_{I: i \in I} \{p_I\}, i = 1, \dots, m.$$

**Příklad 2.3.1** Uvažujme princip uzávěru, znázorněný na Obrázku 2.4, s  $m = 3$  hypotézami. Máme tedy množinu třech hypotéz  $\mathcal{H} = \{H_1^0, H_2^0, H_3^0\}$ . Dále vytvoříme uzávěr množiny všech neprázdných průniků  $\overline{\mathcal{H}}$ , v tomto konkrétním případě  $\overline{\mathcal{H}} = \left\{ H_1^0, H_2^0, H_3^0, H_{12}^0, H_{13}^0, H_{23}^0, H_{123}^0 \right\}$ . Poznamenejme, že  $\overline{\mathcal{H}}$  je uzavřená na operaci průniku. To znamená, že každý průnik  $H_I \cap H_{I'}$ , kde  $H_I, H_{I'} \in \overline{\mathcal{H}}$  leží v  $\overline{\mathcal{H}}$ . Potom pro každou ze sedmi hypotéz najdeme vhodný test. Hypotézu  $H_1^0$  zamítneme, jestliže jsme zamítli hypotézy  $H_{123}^0, H_{12}^0, H_{13}^0$  a  $H_1^0$ , hypotézu  $H_2^0$  zamítneme, jestliže jsme zamítli hypotézy  $H_{123}^0, H_{12}^0, H_{23}^0$  a  $H_2^0$  a hypotézu  $H_3^0$  zamítneme, jestliže jsme zamítli hypotézy  $H_{123}^0, H_{13}^0, H_{23}^0$  a  $H_3^0$ .



Obrázek 2.4: Schéma principu uzávěru pro tři nulové hypotézy  $H_1^0$ ,  $H_2^0$  a  $H_3^0$  a jejich průniků.

Princip uzávěru může být také představen následovně. Připomeňme, že veličina  $FWER$  vyjadřuje pravděpodobnost, že nesprávně zamítneme nejméně jednu hypotézu. Tuto pravděpodobnost chceme kontrolovat na hladině významnosti  $\alpha$  bez ohledu na to, která z množin  $M_0 \subseteq M = \{1, 2, 3\}$  nulových hypotéz se ukáže být správná. Předpokládejme, že se v Příkladě 2.3.1 ukáže být správná hypotéza  $H_{12}^0$ . Potom je  $M_0 = \{1, 2\}$  a chyba prvního druhu nastane, když zamítneme hypotézu  $H_1^0$  nebo  $H_2^0$ .

V případě principu uzávěru je hypotéza  $H_1^0$  zamítnuta pouze tehdy, když jsou zamítnuty všechny hypotézy  $H_{123}^0, H_{12}^0, H_{13}^0$  a  $H_1^0$ . Poznamenejme, že množina pokusů, pro které jsou všechny tyto hypotézy zamítnuty, je podmnožinou množiny pokusů, pro které je zamítnuta hypotéza  $H_{12}^0$ . Matematicky:

$$\{H_1^0 \text{ je zamítnuta pomocí principu uzávěru}\} = \{H_{123}^0 \text{ je zamítnuta}\} \cap \{H_{12}^0 \text{ je zamítnuta}\} \cap \{H_{13}^0 \text{ je zamítnuta}\} \cap \{H_1^0 \text{ je zamítnuta}\} \subseteq \{H_{12}^0 \text{ je zamítnuta}\}.$$

Podobně hypotéza  $H_2^0$  je zamítnuta pouze tehdy, když jsou zamítnuty všechny hypotézy  $H_{123}^0, H_{12}^0, H_{23}^0$  a  $H_2^0$ . Matematicky:

$$\{H_2^0 \text{ je zamítnuta pomocí principu uzávěru}\} = \{H_{123}^0 \text{ je zamítnuta}\} \cap \{H_{12}^0 \text{ je zamítnuta}\} \cap \{H_{23}^0 \text{ je zamítnuta}\} \cap \{H_2^0 \text{ je zamítnuta}\} \subseteq \{H_{12}^0 \text{ je zamítnuta}\}.$$

Protože

$$\{H_1^0 \text{ je zamítnuta pomocí principu uzávěru}\} \subseteq \{H_{12}^0 \text{ je zamítnuta}\}$$

$$\{H_2^0 \text{ je zamítnuta pomocí principu uzávěru}\} \subseteq \{H_{12}^0 \text{ je zamítnuta}\},$$

platí

$$\{\{H_1^0 \text{ je zamítnuta pomocí principu uzávěru}\} \cup \{H_2^0 \text{ je zamítnuta pomocí principu uzávěru}\}\} \subseteq \{H_{12}^0 \text{ je zamítnuta}\},$$

a proto

$$P(\{H_1^0 \text{ je zamítnuta pomocí principu uzávěru}\} \cup \{H_2^0 \text{ je zamítnuta pomocí principu uzávěru}\}) \leq P(H_{12}^0 \text{ je zamítnuta}) = \alpha.$$

Metoda tedy kontroluje *FWER* na hladině menší nebo rovno  $\alpha$ , když je nulová hypotéza  $H_{12}^0$  pravdivá. Stejně tak, když bude pravdivá hypotéza  $H_{13}^0, H_{23}^0$  nebo  $H_{123}^0$ . Odtud vidíme, proč je nutno brát průniky všech hypotéz, když chceme kontrolovat *FWER* na hladině  $\alpha$ .

Princip uzávěru je flexibilní konstrukční metoda, která vyhovuje různým aplikacím. Na principu uzávěru je založeno mnoho běžných metod mnohonásobného porovnávání, například Holm - Bonferroniho metoda.

**Poznámka 2.3.1** Počet operací pro princip uzávěru je obecně v řádu  $2^m$ , kde  $m$  je počáteční počet nulových hypotéz. Často je užitečné najít zkrácené metody, které mohou snížit počet operací a jejichž cílem je rozhodnout o jednotlivých hypotézách, ale ne nezbytně o celém uzávěru. Rozhodnutí plynoucí ze zkrácené metody se shoduje s rozhodnutím plynoucím z principu uzávěru.

## 2.4. Princip rozdělování

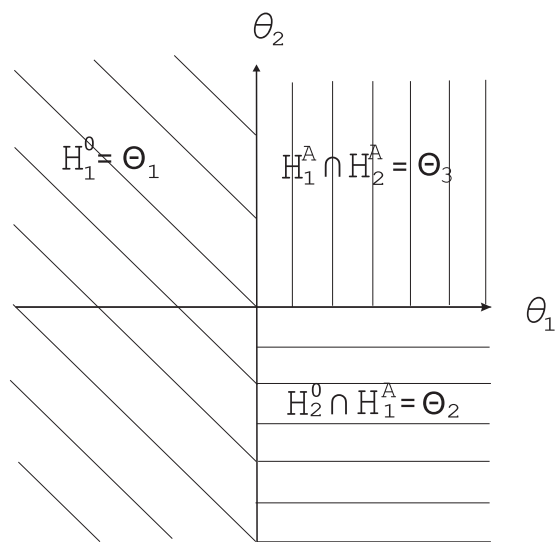
Využitím principu rozdělování můžeme získat silnější metody mnohonásobného porovnávání. Klíčovou myšlenkou principu rozdělování je rozdělit parametrický prostor na disjunktní podmnožiny a pomocí vhodně zvoleného testu takto vzniklé části testovat na hladině významnosti  $\alpha$ . Protože jsou vzniklé části

navzájem disjunktní, pouze jedna z nich obsahuje správný vektor parametrů a může vést k dopuštění se chyby prvního druhu. Proto metody založené na principu rozdělování přísně kontrolují *FWER* na hladině významnosti  $\alpha$ .

Uvažujme znovu problém, týkající se porovnávání dvou nových léčebných postupů s již zavedeným léčebným postupem. Nechť opět  $\theta_i = \mu_i - \mu_0, i = 1, 2$ , značí parametr, který nás zajímá. Dále nechť  $\Theta = \mathbb{R}^2$  značí parametrický prostor s parametrem  $\theta^* = (\theta_1, \theta_2) \in \Theta$ . Hypotézy, které nás zajímají, jsou  $H_i^0 = \{\theta^* \in \mathbb{R}^2 : \theta_i \leq 0\}$ , kde  $i = 1, 2$ . Nechť  $H_i^A = \Theta \setminus H_i^0$  značí odpovídající alternativní hypotézy. Na rozdíl od principu uzávěru, nyní rozdělíme parametrický prostor  $\Theta$  následovně:  $\Theta_1 = H_1^0, \Theta_2 = H_2^0 \cap H_1^A$  a  $\Theta_3 = H_1^A \cap H_2^A$ , viz Obrázek 2.5. Protože  $\Theta_i, i = 1, 2, 3$ , jsou množiny disjunktní a platí, že  $\Theta_1 \cup \Theta_2 \cup \Theta_3 = \Theta$ , představují tyto množiny rozdělení parametrického prostoru  $\Theta$ . Správný vektor parametrů  $\theta^*$  pak leží právě v jedné z těchto množin  $\Theta_i$ . Aplikace testů na každou z těchto podmnožin lokálně na hladině  $\alpha$  vede na metodu mnohonásobného porovnávání, která kontroluje *FWER* na hladině  $\alpha$ . Tato metoda je nejméně tak silná, jako příslušná metoda založená na principu uzávěru, neboť stačí kontrolovat chybu prvního druhu přes  $\Theta_1 = H_1^0$  a menší podprostor  $\Theta_2 \subset H_2^0 \wedge \Theta_2 \neq H_2^0$ .

Interpretace  $\Theta_1$  je stejná jako interpretace  $H_1^0$  (léčebný postup 1 není lepší než kontrolní postup). Interpretace  $\Theta_2$  (léčebný postup 2 není lepší než kontrolní, ale je lepší než léčebný postup 1) se změnila ve srovnání s interpretací  $H_2^0$  (léčebný postup 2 není lepší než kontrolní postup). Existuje mnoho postupů, jak rozdělit parametrický prostor a není vždy jasné, který aplikovat na daný problém. Na Obrázku 2.5 je znázorněn možný příklad rozdělení reálné roviny. Obecně, pro  $m$  hypotéz  $H_1^0, \dots, H_m^0$ , může být princip rozdělování představen následovně:

1. Vybereme vhodné rozdělení  $\{\Theta_l : l \in L\}$  parametrického prostoru  $\Theta$ , pro nějakou množinu indexů  $L$ .
2. Testujeme každou množinu  $\Theta_l$  lokálně na hladině  $\alpha$ .
3. Nulovou hypotézu  $H_i^0$  zamítneme, jestliže jsou zamítnuty všechny  $\Theta_l$ , kde  $\Theta_l \cap H_i^0 \neq \emptyset$ .



Obrázek 2.5: Princip rozdělování pro dvě nulové hypotézy  $H_1^0$  a  $H_2^0$ .

# Kapitola 3

## Metody kontrolující FWER

V této kapitole, která se již věnuje samotným metodám mnohonásobného porovnávání, se zaměříme na metody kontrolující *FWER* a shrneme si nejpoužívanější metody, které jsou konstruovány pomocí neupravených a upravených  $p$ -hodnot.

Nejprve se budeme věnovat metodám, které vycházejí z *Bonferroniho nerovnosti* a to samotné *Bonferroniho metodě*, která je pravděpodobně nejznámější a nejjednodušší jednokrokovou metodou mnohonásobného porovnávání. Ačkoliv existují silnější a spolehlivější metody, Bonferroniho metoda je stále nejpoužívanější díky její jednoduchosti a rovněž kvůli malému počtu předpokladů, které jsou požadovány pro její provedení. Dále se pak zaměříme na vícekrokovou *Holm-Bonferroniho metodu* (tato metoda bývá označována zkráceně jako Holmova metoda) a *Šidákovu metodu*, které vycházejí z Bonferroniho metody. Dále zmíníme *Simesův test* a s ním související vícekrokovou *Hochbergovu metodu*.

V dalším textu předpokládejme, že máme k dispozici neupravené  $p$ -hodnoty  $p_1, p_2, \dots, p_m$ , odpovídající nulovým hypotézám  $H_1^0, \dots, H_m^0$ , kde  $m$  je počet testovaných hypotéz. Necht' máme také stanovenou hladinu významnosti  $\alpha$ .

Tato kapitola byla sepsána pomocí zdrojů [1, 4, 6, 7, 8, 10, 11, 14, 16].

### 3.1. Bonferroniho metoda

Bonferroniho metoda je jednokroková metoda založená na Bonferroniho nerovnosti. Uvažujme  $n$  náhodných jevů  $A_1, \dots, A_n$ . Nechť  $P(A_i)$  je pravděpodobnost, že jev  $A_i$  je pravdivý a  $P(\cup_{i=1}^n A_i)$  je pravděpodobnost, že je pravdivý nejméně jeden z jevů  $A_1, \dots, A_n$ . Bonferroniho nerovností pak rozumíme výraz

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i).$$

Pokud jsou  $A_i$  a  $A_j$  pro všechna  $i, j$  disjunktní, pak se z nerovnosti stává rovnost.

Předpokládejme, že máme k dispozici neupravené  $p$ -hodnoty  $p_1, p_2, p_3, \dots, p_m$ , kde  $m$  je počet uvažovaných hypotéz. Dále nechť  $q_i = \min\{mp_i, 1\}$  jsou upravené  $p$ -hodnoty, kde užití minimum zajistí, že výsledná upravená  $p$ -hodnota  $q_i$  není větší než 1. Označme  $V$  počet chyb prvního druhu a nechť  $M_0 \subseteq M$  značí množinu pravdivých hypotéz,  $m_0 = |M_0|$ . Aplikací Bonferroniho nerovnosti dostaneme výraz

$$P(V > 0) = P\left(\bigcup_{i \in M_0} \{q_i \leq \alpha\}\right) \leq \sum_{i \in M_0} P(q_i \leq \alpha) \leq \alpha \Leftrightarrow P(q_i \leq \alpha) \leq \frac{\alpha}{m}, \forall i \in M_0.$$

Z Bonferroniho nerovnosti vyplývá, že pokud chceme kontrolovat  $FWER$  na hladině nejvýše  $\alpha$ , pak stačí každou nulovou hypotézu testovat na hladině  $\frac{\alpha}{m}$ . Bonferroniho metoda testuje buď pomocí neupravených  $p$ -hodnot a upravených hladin významnosti, nebo pomocí upravených  $p$ -hodnot a neupravené hladiny významnosti.

#### Postup testování pro neupravené $p$ -hodnoty $p_j$ :

1. Nulovou hypotézu  $H_j^0$  zamítneme právě tehdy, když  $p_j \leq \alpha/m$ . Jinak, pokud  $p_j > \alpha/m$ , hypotézu  $H_j^0$  nelze zamítnout.



### Postup testování pro upravené $p$ -hodnoty $q_j$ :

1. Nejprve vypočítáme upravené  $p$ -hodnoty  $q_j$ :

$$q_j = \min\{mp_j, 1\}.$$

2. Nulovou hypotézu  $H_j^0$  zamítneme právě tehdy, když  $q_j \leq \alpha$ . Jinak, pokud  $q_j > \alpha$ , hypotézu  $H_j^0$  nelze zamítnout.

Upravené  $p$ -hodnoty pro Bonferroniho metodu můžeme získat pomocí programu R z vektoru neupravených  $p$ -hodnot  $\mathbf{p}$  následovně:

```
p.adjust(p, "bonferroni")
```

**Příklad 3.1.1** Uvažujme  $m = 3$  nulové hypotézy, které chceme otestovat na hladině významnosti  $\alpha = 0,05$ . Nechť  $p_1 = 0,01$ ,  $p_2 = 0,015$  a  $p_3 = 0,005$  jsou neupravené  $p$ -hodnoty. Protože  $p_1, p_2, p_3 < \alpha/3 = 0,016667$ , zamítáme všechny tři hypotézy. Obdobně, když vypočítáme upravené  $p$ -hodnoty  $q_1 = 0,03$ ,  $q_2 = 0,045$  a  $q_3 = 0,015$  a srovnáme je s hladinou významnosti  $\alpha = 0,05$ , dojdeme ke stejnému závěru.

Velkou výhodou Bonferroniho metody je snadnost provedení, avšak je obecně považovaná za metodu velmi konzervativní, tzn., že si udržuje požadovanou hladinu významnosti v celém experimentu a nedovolí, aby pravděpodobnost chyby nekontrolovatelně vzrostla. Tato metoda představuje velmi obecný přístup k testování hypotéz, který lze použít pro jakoukoliv korelační strukturu mezi testovými statistikami. V následující části se seznámíme se zlepšenými metodami, které vycházejí právě z Bonferroniho metody.

## 3.2. Jednokroková Šidáková metoda

Rozlišujeme jednokrokovou Šidákovu metodu a vícekrokovou Šidákovu metodu. Nejprve se zaměříme na jednokrokovou Šidákovu metodu.

Pokud jsou testovací statistiky nezávislé,  $FWER$  je pro  $m$  hypotéz daná vztahem

$$FWER = P(V > 0) = 1 - P(V = 0) = 1 - (1 - \alpha)^m.$$

To je motivací k seznámení se s přístupem Zbyňka Šidáka. Poznamenejme, že Šidákův přístup je spolehlivější než Bonferroniho metoda, ačkoliv v praxi je rozdíl zanedbatelný. Pro  $m_0 = m$  a nezávislé testové statistiky je  $FWER$  přesně  $\alpha$ . Pokud jsou statistiky pozitivně korelovány, je metoda konzervativní. V opačném případě je liberální.

**Postup testování pro neupravené  $p$ -hodnoty  $p_j$ :**

1. Hypotézu  $H_j^0$  zamítáme právě tehdy, když  $p_j \leq 1 - (1 - \alpha)^{1/m}$ .

**Postup testování pro upravené  $p$ -hodnoty  $q_j$ :**

1. Nejprve vypočítáme upravené  $p$ -hodnoty  $q_j$  podle vzorce:

$$q_j = 1 - (1 - p_j)^m.$$

2. Nulovou hypotézu  $H_j^0$  zamítneme právě tehdy, když  $q_j \leq \alpha$ . Jinak, pokud  $q_j > \alpha$ , hypotézu  $H_j^0$  nelze zamítnout.

### 3.3. Vícekroková Šidáková metoda

Nyní se seznámíme s vícekrokovou sestupnou Šidákovou metodou, která rovněž testuje jak pomocí neupravených  $p$ -hodnot a upravených hladin významnosti, tak pomocí upravených  $p$ -hodnot  $q_{(j)}$ . Jak uvidíme v postupech testování, u více-krokové Šidákovy metody, je potřeba vektor neupravených  $p$ -hodnot nejprve uspořádat.

**Postup testování pro neupravené  $p$ -hodnoty  $p_{(j)}$ :**

1. Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti

$$p_{(1)} \leq \dots \leq p_{(m)}.$$

- Postupně testujeme od nejmenší  $p$ -hodnoty a metoda končí prvním nezamítnutím. Tj. nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když  $p_{(i)} \leq (1 - (1 - \alpha)^{1/(m-i+1)})$ ,  $i = 1, \dots, j$ . Nulová hypotéza  $H_{(j)}^0$  je tedy zamítnuta, jestliže jsou zamítnuty všechny hypotézy jí předcházející. Pokud nezamítneme hned první testovanou hypotézu ( $p_{(1)} > \alpha/m$ ), metoda končí a nezamítá se žádná z hypotéz.

**Postup testování pro upravené  $p$ -hodnoty  $q_{(j)}$ :**

- Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti  $p_{(1)} \leq \dots \leq p_{(m)}$ .
- Dále vypočítáme upravené  $p$ -hodnoty  $q_{(j)}$  podle vzorce:

$$q_{(j)} = \max_{k=1, \dots, j} \{1 - (1 - p_{(k)})^{(m-k+1)}\}.$$

- Nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když  $q_{(j)} \leq \alpha$ . Jinak, pokud  $q_{(j)} > \alpha$ , hypotézu  $H_{(j)}^0$  nelze zamítnout.

### 3.4. Holm-Bonferroniho metoda

Holm-Bonferroniho metoda, zkráceně Holmova metoda, představuje metodu mnohonásobného porovnávání, která zlepšuje Bonferroniho přístup. Holmova metoda je vícezkroková sestupná metoda, která spočívá v opakované aplikaci Bonferroniho nerovnosti. Hypotézy testujeme popořadě, dle uspořádaných  $p$ -hodnot.

Holmova metoda testuje také pomocí upravených  $p$ -hodnot  $q_{(j)}$ , které odpovídají neupraveným  $p$ -hodnotám  $p_{(j)}$  a jsou srovnávány s hladinou významnosti  $\alpha$ . Upravená  $p$ -hodnota pro Holmovu metodu je dána výrazem

$$q_{(j)} = \min\{1, \max((m - j + 1)p_{(j)}, q_{(j-1)})\},$$

nebo také

$$q_{(j)} = \max_{k=1, \dots, j} \{\min\{(m - k + 1)p_{(k)}, 1\}\}.$$

Tím, že uvažujeme maxima z množiny minim, zajistíme monotonnost upravených  $p$ -hodnot. To znamená, že upravené  $p$ -hodnoty jsou uspořádány do neklesající posloupnosti  $q_{(1)} \leq q_{(2)} \leq \dots \leq q_{(m)}$  a konkrétní hypotéza může být zamítnuta pouze tehdy, pokud byly zamítnuty všechny hypotézy s menší neupravenou  $p$ -hodnotou.

**Postup testování pro neupravené  $p$ -hodnoty  $p_{(j)}$ :**

1. Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti  $p_{(1)} \leq \dots \leq p_{(m)}$ .
2. Postupně testujeme od nejmenší  $p$ -hodnoty a metoda končí prvním nezamítnutím. Tj. nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když platí, že  $p_{(i)} \leq \alpha / (m - i + 1), i = 1, \dots, j$ . Nulová hypotéza  $H_{(j)}^0$  je tedy zamítnuta, jestliže jsou zamítnuty všechny hypotézy jí předcházející. Pokud nezamítneme hned první testovanou hypotézu,  $p_{(1)} > \alpha / m$ , metoda končí a nezamítá se žádná z hypotéz.

**Postup testování pro upravené  $p$ -hodnoty  $q_{(j)}$ :**

1. Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti  $p_{(1)} \leq \dots \leq p_{(m)}$ .
2. Dále vypočítáme upravené  $p$ -hodnoty  $q_{(j)}$  podle vzorce:

$$q_{(j)} = \max_{k=1, \dots, j} \{ \min \{ (m - k + 1)p_{(k)}, 1 \} \}.$$

3. Nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když  $q_{(j)} \leq \alpha$ . Jinak, pokud  $q_{(j)} > \alpha$ , hypotézu  $H_{(j)}^0$  nelze zamítnout.

Upravené  $p$ -hodnoty pro Holm-Bonferroniho metodu můžeme získat pomocí programu R z vektoru neupravených  $p$ -hodnot následovně:

```
p.adjust(p, "holm")
```

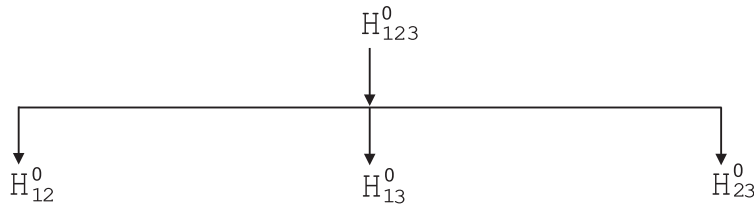
**Příklad 3.4.1** Uvažujme znovu  $p$ -hodnoty  $p_1 = 0,01, p_2 = 0,015$  a  $p_3 = 0,005$ , odpovídající nulovým hypotézám  $H_1^0, H_2^0$  a  $H_3^0$ . Nejprve uspořádáme  $p$ -hodnoty podle velikosti a začneme testováním té nejmenší, tedy  $p_{(1)} = 0,005 = p_3$ . Protože  $p_{(1)} = 0,005 < 0,016667 = \alpha/3$ , tak  $H_{(1)}^0 = H_3^0$  zamítáme na hladině významnosti  $\alpha = 0,05$ . Dále testujeme  $p_{(2)} = 0,01 < 0,025 = \alpha/2$ , a tedy i hypotéza  $H_{(2)}^0 = H_1^0$  se zamítá. Nakonec testujeme  $p$ -hodnotu  $p_{(3)}$  a protože  $p_{(3)} = 0,015 < 0,05 = \alpha$ , tak i hypotézu  $H_{(3)}^0 = H_2^0$  zamítáme. Obdobně, pokud spočítáme upravené  $p$ -hodnoty, dostaneme  $q_{(1)} = 0,015, q_{(2)} = 0,02$  a  $q_{(3)} = 0,02$ , které jsou všechny menší než  $\alpha = 0,05$ , takže dojdeme ke stejnému závěru, tedy že všechny tři hypotézy  $H_1^0, H_2^0$  i  $H_3^0$  zamítáme.

Holmova metoda je vícezkroková sestupná metoda, která zamítá všechny hypotézy zamítnuté Bonferroniho metodou a ještě může zamítnou i další. Nyní se podíváme na Holmovu metodu trochu z jiného úhlu, a to tak, že budeme brát v úvahu princip uzávěru. Pro nulové hypotézy  $H_1^0, \dots, H_m^0$ , splňující podmínku volné kombinace a při aplikaci Bonferroniho testu na každou hypotézu průniku  $H_I = \bigcap_{i \in I} H_i^0, I \subseteq \{1, \dots, m\}$ , je Holmova metoda zkrácením principu uzávěru. To znamená, že Holmova metoda vede pro jednoduché hypotézy  $H_1^0, \dots, H_m^0$  ke stejným závěrům jako Bonferroniho metoda, založená na principu uzávěru.

Uvažujme příklad, kde srovnáváme dva nové léčebné postupy s postupem kontrolním. Předpokládejme, že je Bonferroniho metoda využita k testování hypotézy průniku  $H_{12}^0 = H_1^0 \cap H_2^0$ , což znamená, že  $H_{12}^0$  je zamítnuta, jestliže  $\min(p_1, p_2) \leq \alpha/2$ . Jestliže je hypotéza  $H_{12}^0$  zamítnuta, pak je také hned zamítnuta jedna z jednoduchých hypotéz  $H_1^0, H_2^0$  a jen zbylá hypotéza musí být testována na hladině významnosti  $\alpha$ . Jestliže je například  $p_1 < \alpha/2$ , pak i  $p_1 < \alpha$  a hypotéza  $H_1^0$  je zamítnuta a zbývá testovat hypotézu  $H_2^0$  na hladině významnosti  $\alpha$ . Obecněji, jestliže je hypotéza  $H_I = \bigcap_{i \in I} H_i^0, I \subseteq M$ , zamítnuta, pak existuje index  $i^* \in I$  takový, že  $p_{i^*} \leq \alpha/|I|$  a všechny hypotézy  $H_J$ , kde  $i^* \in J \subseteq I$ , jsou také zamítnuty, protože  $|J| \leq |I|$  a proto  $p_{i^*} \leq \alpha/|I| \leq \alpha/|J|$ .

Jestliže nulové hypotézy  $H_1^0, \dots, H_m^0$  splňují podmínku omezené kombinace, Holmovu metodu lze využít, ale je konzervativní. Uvažujme párové srovnávání

třech hypotéz prostřednictvím  $\mu_1, \mu_2$  a  $\mu_3$ , tedy  $m = 3$  a  $H_{ij}^0 : \mu_i = \mu_j$ . Odpovídající princip uzávěru pro tři nulové hypotézy  $H_{12}^0, H_{13}^0$  a  $H_{23}^0$  je znázorněn na Obrázku 3.1, kde  $H_{123}^0 : \mu_1 = \mu_2 = \mu_3$ . Aplikace Bonferroniho testu vede k zamítnutí hypotézy  $H_{123}^0$ , jestliže  $\min\{p_{(1)}, p_{(2)}, p_{(3)}\} < \alpha/3$ . Jakmile je hypotéza  $H_{123}^0$  zamítnuta, je také zamítnuta jedna z jednoduchých hypotéz a zbývají dvě hypotézy, které je nutno testovat na hladině významnosti  $\alpha$ . Poznamenejme, že pro uspořádanou množinu  $p$ -hodnot  $\{p_{(1)}, p_{(2)}, p_{(3)}\}$ , by Holmova metoda aplikovala hladiny významnosti  $\{\alpha/3, \alpha/2, \alpha\}$ , zatímco by stačily hladiny  $\{\alpha/3, \alpha, \alpha\}$ .



Obrázek 3.1: Schéma principu uzávěru pro tři nulové hypotézy  $H_{ij}^0 : \mu_i = \mu_j, 1 \leq i < j \leq 3$ .

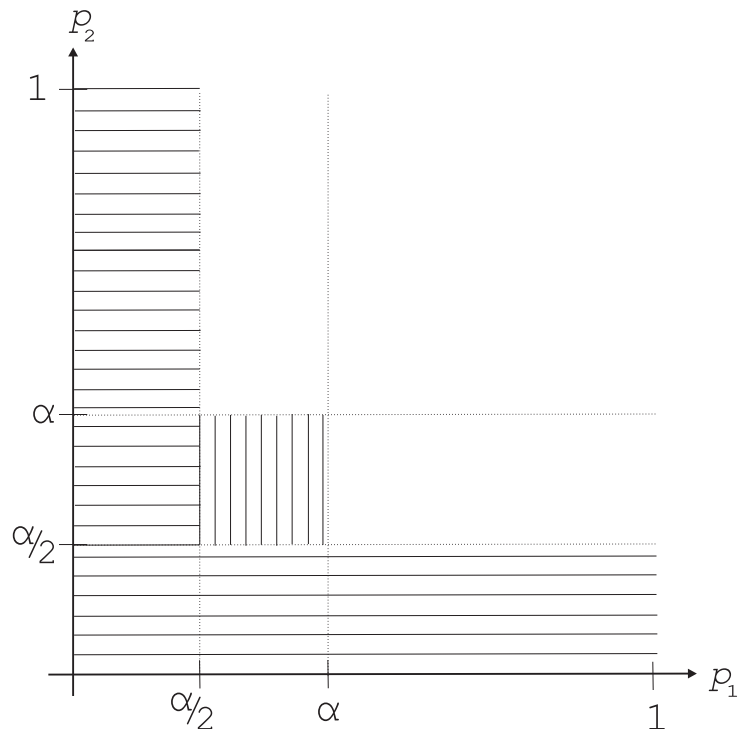
### 3.5. Simesův test

Další modifikace Bonferroniho metody jsou založeny na Simesově testu, který si teď jen krátce představíme. Na Simesově testu jsou obvykle založeny vzestupné metody, které začínají testováním největší  $p$ -hodnoty, tedy nejméně významné  $p$ -hodnoty  $p_{(m)}$ .

Simes navrhl následující modifikaci Bonferroniho metody k testování globální hypotézy průniku  $H = \bigcap_{i \in M} H_i^0$ . Nechť máme znova uspořádané neupravené  $p$ -hodnoty  $p_{(1)} \leq \dots \leq p_{(m)}$ , odpovídající hypotézám  $H_{(1)}^0, \dots, H_{(m)}^0$ . Pomocí Simesova testu zamítneme hypotézu  $H$ , jestliže najdeme index  $i \in M = \{1, \dots, m\}$  takový, že  $p_{(i)} \leq i\alpha/m$ . O jednotlivých hypotézách  $H_{(j)}^0$  však nemůžeme učinit žádný závěr. Nemůžeme zamítnout hypotézu  $H_{(j)}^0$ , když bude  $p_{(j)} \leq j\alpha/m$  pro nějaké  $j \in M$ , protože v tomto případě není kontrolována *FWER*.

Simesův test je silnější než Bonferroniho metoda, protože když zamítneme hypotézu  $H$  Bonferroniho testem, bude také zamítnuta Simesovým testem, ale

ne naopak. Obrázek 3.2 srovnává oblasti zamítnutí Bonferroniho a Simesova testu pro  $m = 2$ . Připomeňme si, že v případě  $m = 2$ , Bonferroniho test zamítá globální hypotézu průniku  $H$ , když je  $p_1 \leq \alpha/2$ , nebo když je  $p_2 \leq \alpha/2$ . Simesův test zamítá hypotézu  $H$ , když je  $p_{(1)} \leq \alpha/2$ , nebo když je  $p_{(2)} \leq \alpha$ . Jak můžeme vidět z Obrázku 3.2, Simesův test přidává k oblasti zamítnutí Bonferroniho testu ještě čtverec  $[\alpha/2, \alpha] \times [\alpha/2, \alpha]$ , a proto je silnější.



Obrázek 3.2: Oblasti zamítnutí pro Bonferroniho test (vodorovné čáry) a pro Simesův test (vodorovné a svislé čáry).

Nutno poznamenat, že Simesův test má určitou výhodu, avšak na úkor přidávaných předpokladů. Simes ukázal, že v případě nezávislosti  $p$ -hodnot  $p_1, \dots, p_m$  je pro jeho test  $FWER = \alpha$ . Pokud není splněn předpoklad nezávislosti, kontrola chyby prvního druhu není vždy jasná.

### 3.6. Hochbergova metoda

Hochberg přišel se vzestupným rozšířením Simesova testu, které umožňuje učinit závěry o jednotlivých hypotézách  $H_i^0$ ,  $i \in M$ . Hochbergova metoda je vlastně obrácená Holmova metoda, protože využívá stejných kritických hodnot, ale testuje v opačném směru. Holmova metoda končí prvním nezamítnutím a Hochbergova metoda končí prvním zamítnutím. Hochbergova metoda je silnější než Holmova metoda.

Hochbergova metoda, stejně jako Holmova metoda, testuje pomocí neupravených  $p$ -hodnot i upravených  $p$ -hodnot  $q_{(j)}$ , které jsou srovnávány s hladinou významnosti  $\alpha$ . Upravené  $p$ -hodnoty pro Hochbergovu metodu jsou dány vztahem:

$$q_{(j)} = \min\{1, \min[(m - j + 1)p_{(j)}, q_{(j+1)}]\},$$

nebo také

$$q_{(j)} = \min_{k=j, \dots, m} \{\min\{(m - k + 1)p_{(k)}, 1\}\}.$$

#### Postup testování pro neupravené $p$ -hodnoty $p_{(j)}$ :

1. Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti  $p_{(1)} \leq \dots \leq p_{(m)}$ .
2. Postupně testujeme od největší  $p$ -hodnoty a metoda končí prvním zamítnutím. Tj. nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když najdeme  $i \in \{j, \dots, m\}$  takové, že  $p_{(i)} \leq \alpha / (m - i + 1)$ . Pokud zamítneme hned první testovanou hypotézu,  $p_{(m)} \leq \alpha$ , metoda končí a zamítají se všechny hypotézy.

#### Postup testování pro upravené $p$ -hodnoty $q_{(j)}$ :

1. Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti  $p_{(1)} \leq \dots \leq p_{(m)}$ .



2. Dále vypočítáme upravené  $p$ -hodnoty  $q_{(j)}$  podle vzorce:

$$q_{(j)} = \min_{k=j, \dots, m} \{ \min \{ (m - k + 1)p_{(k)}, 1 \} \}.$$

3. Nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když  $q_{(j)} \leq \alpha$ . Jinak, pokud  $q_{(j)} > \alpha$ , hypotézu  $H_{(j)}^0$  nelze zamítnout.

Upravené  $p$ -hodnoty pro Hochbergovu metodu můžeme získat pomocí programu R z vektoru neupravených  $p$ -hodnot následovně:

```
p.adjust(p, "hochberg")
```

Poznamenejme, že vzestupné metody jsou často silnější, než jejich sestupné obdoby. Hochbergova metoda je silnější, než metoda Holmova, nicméně za předpokladu, že testovací statistiky jsou nezávislé.

**Příklad 3.6.1** Uvažujme následující početní příklad s  $m = 4$  nulovými hypotézami a hladinou významnosti  $\alpha = 0,05$ . Nechtě  $p_1 = 0,022$ ,  $p_2 = 0,02$ ,  $p_3 = 0,01$  a  $p_4 = 0,09$  jsou neupravené  $p$ -hodnoty.

Nejprve uvažujme Holmovu metodu. Dané  $p$ -hodnoty uspořádáme od nejmenší po největší a začneme testováním té nejmenší, v tomto případě nejprve testujeme  $p_3$ . Protože  $p_{(1)} = p_3 = 0,01 < 0,0125 = \alpha/4$ , hypotéza  $H_{(1)}^0 = H_3^0$  je zamítnuta. Pokračujeme testováním  $p_{(2)} = p_2$  a protože  $p_{(2)} = p_2 = 0,02 > 0,0167 = \alpha/3$ , žádná další hypotéza se nezamítá a Holmova metoda končí.

Hochbergova metoda naopak začíná testováním největší  $p$ -hodnoty a končí prvním zamítnutím. Protože  $p_{(4)} = p_4 = 0,09 > 0,05 = \alpha$ , hypotézu  $H_{(4)}^0 = H_4^0$  nezamítáme a pokračujeme testováním  $p$ -hodnoty  $p_{(3)} = p_1$ . Protože  $p_{(3)} = p_1 = 0,022 < 0,025 = \alpha/2$ , hypotézu  $H_1^0$  zamítáme a stejně tak i hypotézy  $H_2^0$  a  $H_3^0$ . Tabulka 3.1 shrnuje tento příklad a jsou v ní uvedeny i upravené  $p$ -hodnoty.

j	Práh		Rozhodnutí		Upravená $p$ -hodnota	
	$p_{(j)}$	$\alpha/(4 - j + 1)$	Holm	Hochberg	Holm	Hochberg
1	0,01	0,0125	Z	Z	0,04	0,04
2	0,02	0,0167	NZ	Z	0,06	0,044
3	0,022	0,025	NZ	Z	0,06	0,044
4	0,09	0,05	NZ	NZ	0,09	0,09

Tabulka 3.1: Srovnání Holmovy a Hochbergovy metody pro  $m = 4$  hypotézy a  $\alpha = 0,05$ . Z = zamítáme, NZ = nelze zamítnout.

# Kapitola 4

## Metody kontrolující FDR

Další přístup k mnohonásobnému porovnávání byl navržen v roce 1995 Benjaminiem a Hochbergem. Tito autoři tvrdili, že v mnoha situacích může kontrola  $FWER$  vést k přehnaně konzervativním metodám a že malé množství chyb prvního druhu může být tolerováno za předpokladu, že počet chyb prvního druhu je ve srovnání s celkovým počtem zamítnutých hypotéz opravdu malý. Tohle kritérium vede k méně konzervativnímu přístupu, který kontroluje poměr chybně zamítnutých nulových hypotéz a celkově zamítnutých nulových hypotéz, tedy  $FDR$ .

Mezi metody, které kontrolují  $FDR$  na hladině  $\alpha$ , patří *Benjamini-Hochbergova* metoda (dále jen BH metoda) a *Benjamini-Yekutieliho* metoda (dále jen BY metoda).

V dalším textu opět předpokládejme, že máme k dispozici neupravené  $p$ -hodnoty  $p_1, p_2, \dots, p_m$ , odpovídající nulovým hypotézám  $H_1^0, \dots, H_m^0$ , kde  $m$  je počet testovaných hypotéz. Nechť máme také stanovenou hladinu významnosti  $\alpha$ .

Tato kapitola byla sepsána pomocí zdrojů [3, 4, 7, 10, 11, 14].

### 4.1. Benjamini-Hochbergova metoda

BH metoda je vzestupná metoda, která pro nezávislé testovací statistiky kontroluje  $FDR$ .

### Postup testování pro neupravené $p$ -hodnoty $p_{(j)}$ :

1. Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti  
 $p_{(1)} \leq \dots \leq p_{(m)}$ .
2. Postupně testujeme od největší  $p$ -hodnoty  $p_{(m)}$ . Metoda končí prvním zamítnutím a všechny ostatní hypotézy v nižším pořadí se potom také zamítají. Tj. nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když najdeme  $i \in \{j, \dots, m\}$ , takové, že  $p_{(i)} \leq i\alpha/m$ . Pokud zamítneme hned první testovanou hypotézu ( $p_{(m)} \leq \alpha$ ), metoda končí a zamítají se všechny hypotézy.

### Postup testování pro upravené $p$ -hodnoty $q_{(j)}$ :

1. Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti  
 $p_{(1)} \leq \dots \leq p_{(m)}$ .
2. Dále vypočítáme upravené  $p$ -hodnoty  $q_{(j)}$  podle vzorce:

$$q_{(j)} = \min_{k=j, \dots, m} \{ \min\{ (m/k)p_{(k)}, 1 \} \}.$$

3. Nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když  $q_{(j)} \leq \alpha$ . Jinak, pokud  $q_{(j)} > \alpha$ , hypotézu  $H_{(j)}^0$  nelze zamítnout.

Upravené  $p$ -hodnoty pro BH metodu můžeme získat pomocí programu R z vektoru neupravených  $p$ -hodnot následovně:

```
p.adjust(p, "BH")
```

## 4.2. Benjamini-Yekutieliho metoda

Benjamini a Yekutieli přišli s BY metodou, což je jednoduchá konzervativní modifikace metod, které kontrolují  $FDR$ , avšak pro libovolné závislé struktury. Jak můžeme vidět dále, BY metoda se od BH metody příliš neliší.

### Postup testování pro neupravené $p$ -hodnoty $p_{(j)}$ :

1. Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti  $p_{(1)} \leq \dots \leq p_{(m)}$ .
2. Postupně testujeme od největší  $p$ -hodnoty  $p_{(m)}$ . Metoda končí prvním zamítnutím a všechny ostatní hypotézy v nižším pořadí jsou rovněž zamítnuty. Tj. nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když najdeme  $i \in \{j, \dots, m\}$ , takové, že  $p_{(i)} \leq \frac{i\alpha}{m \sum_{l=1}^m (1/l)}$ . Pokud zamítneme hned první testovanou hypotézu,  $p_{(m)} \leq \frac{\alpha}{\sum_{l=1}^m (1/l)}$ , metoda končí a zamítají se všechny hypotézy.

### Postup testování pro upravené $p$ -hodnoty $q_{(j)}$ :

1. Nejprve uspořádáme neupravené  $p$ -hodnoty do neklesající posloupnosti  $p_{(1)} \leq \dots \leq p_{(m)}$ .
2. Dále vypočítáme upravené  $p$ -hodnoty  $q_{(j)}$  podle vzorce:

$$q_{(j)} = \min_{k=j, \dots, m} \left\{ \min \left\{ \frac{m \sum_{l=1}^m (1/l)}{k} p_{(k)}, 1 \right\} \right\}.$$

3. Nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když  $q_{(j)} \leq \alpha$ . Jinak, pokud  $q_{(j)} > \alpha$ , hypotézu  $H_{(j)}^0$  nelze zamítnout.

Upravené  $p$ -hodnoty pro BY metodu potom získáme pomocí programu R z vektoru neupravených  $p$ -hodnot následovně:

```
p.adjust(p, "BY")
```

# Kapitola 5

## Přehled metod

V předchozích kapitolách jsme se seznámili s obecnými metodami mnohonásobného porovnávání a v této kapitole je uvedený přehled všech zmíněných metod.

Pro větší přehled a snadnější porozumění metodám si nyní uvedeme stručný přehled metod a vzorců, které jednotlivé metody využívají. Z předchozí kapitoly je jasné, že všechny metody můžeme testovat dvěma způsoby. První možností je testování pomocí neupravených  $p$ -hodnot a upravených hladin významnosti  $\alpha_j$ . Druhým způsobem je potom testování pomocí upravených  $p$ -hodnot  $q_{(j)}$ , které jsou srovnávány s neupravenou hladinou významnosti  $\alpha$ . Nejprve si uvedeme přehled vzorců, sloužících k výpočtu upravených hladin významnosti  $\alpha_j$  a následně vzorce pro výpočet upravených  $p$ -hodnot  $q_{(j)}$ .

### Upravené hladiny významnosti $\alpha_j$ :

- Bonferroniho metoda:  $\alpha_j = \alpha/m$ ,
- Jednokroková Šidákova metoda:  $\alpha_j = 1 - (1 - \alpha)^{1/m}$ ,
- Vícekroková Šidákova metoda:  $\alpha_j = (1 - (1 - \alpha)^{1/(m-j+1)})$ ,
- Holm-Bonferroniho metoda:  $\alpha_j = \alpha/(m - j + 1)$ ,
- Hochbergova metoda:  $\alpha_j = \alpha/(m - j + 1)$ ,
- BH metoda:  $\alpha_j = j\alpha/m$ ,
- BY metoda:  $\alpha_j = j\alpha/(m \sum_{l=1}^m (1/l))$ .

Testování u Bonferroniho a jednokrokové Šidákovy metody potom probíhá tak, že neupravené  $p$ -hodnoty jsou postupně srovnávány s upravenou hladinou významnosti  $\alpha_j$ , která je u těchto dvou metod pro všechny  $p$ -hodnoty stejná. Nulová hypotéza  $H_j^0$  se potom zamítá právě tehdy, když  $p_j \leq \alpha_j$ , kde  $j = 1, \dots, m$ .

Testování u sestupné Šidákovy vícekové metody a sestupné Holm-Bonferroniho metody probíhá tak, že nejprve uspořádáme neupravené  $p$ -hodnoty podle velikosti, následně testujeme od nejmenší  $p$ -hodnoty a metody končí prvním nezamítnutím.

Testování u vzestupné Hochbergovy metody, BH metody a BY metody probíhá tak, že nejprve uspořádáme neupravené  $p$ -hodnoty podle velikosti, následně testujeme od největší  $p$ -hodnoty a metody končí prvním zamítnutím.

#### Upravené $p$ -hodnoty $q_j$ , respektive $q_{(j)}$ :

- Bonferroniho metoda:  $q_j = \min\{mp_j, 1\}$ ,
- Jednokroková Šidáková metoda:  $q_j = 1 - (1 - p_j)^m$ ,
- Víceková Šidáková metoda:  $q_{(j)} = \max_{k=1, \dots, j} \{1 - (1 - p_{(k)})^{(m-k+1)}\}$ ,
- Holm-Bonferroniho metoda:  $q_{(j)} = \max_{k=1, \dots, j} \{\min\{(m - k + 1)p_{(k)}, 1\}\}$ ,
- Hochbergova metoda:  $q_{(j)} = \min_{k=j, \dots, m} \{\min\{(m - k + 1)p_{(k)}, 1\}\}$ ,
- BH metoda:  $q_{(j)} = \min_{k=j, \dots, m} \{\min\{(m/k)p_{(k)}, 1\}\}$ ,
- BY metoda:  $q_{(j)} = \min_{k=j, \dots, m} \left\{ \min \left\{ \frac{m \sum_{l=1}^m (1/l)}{k} p_{(k)}, 1 \right\} \right\}$ .

Testování v případě, že uvažujeme upravené  $p$ -hodnoty probíhá tak, že každá upravená  $p$ -hodnota je srovnána s neupravenou hladinou významnosti  $\alpha$  a nulovou hypotézu  $H_j^0$  zamítneme právě tehdy, když  $q_j \leq \alpha$ , respektive nulovou hypotézu  $H_{(j)}^0$  zamítneme právě tehdy, když  $q_{(j)} \leq \alpha$ .

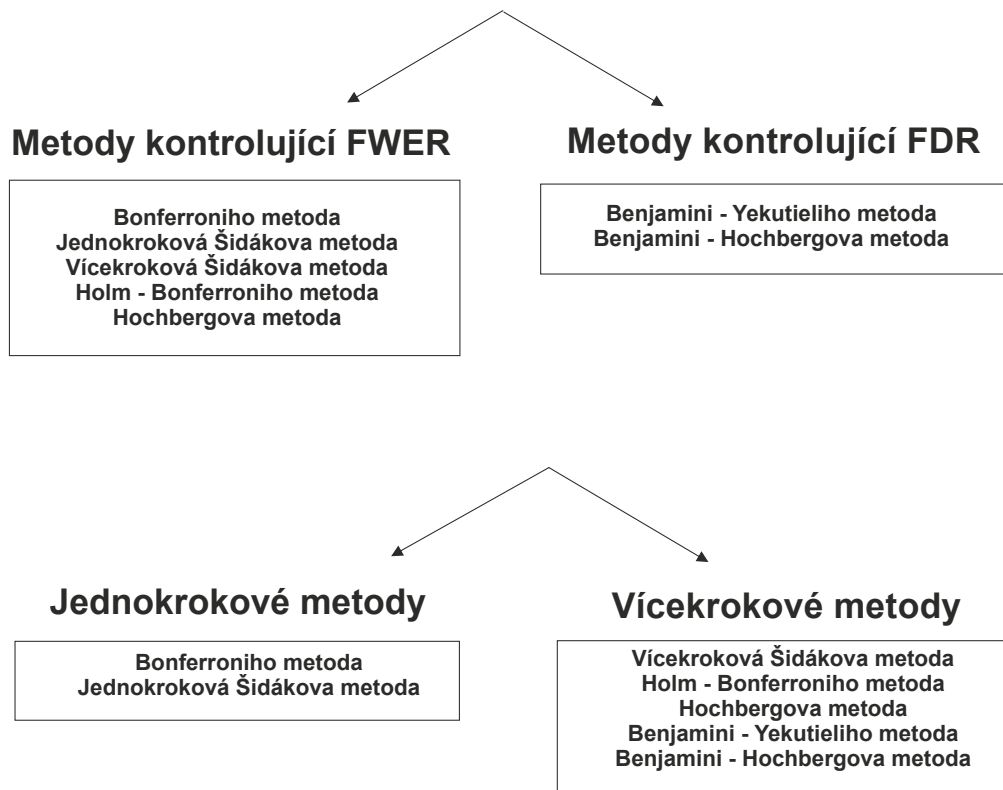
Již víme, že neupravené  $p$ -hodnoty lze snadno vypočítat pomocí programu R. Když zadáme příkaz `p.adjust.methods`, program nám nabídne možnosti,

kterými upravené  $p$ -hodnoty lze přepočítat. Najdeme tam všechny doposud zmíněné metody, kromě Šidákovy jednokrokové a Šidákovy vícekové metody. Pznamenejme, že metoda s příkazem `none`, vrací původní neupravené  $p$ -hodnoty a příkaz `fdr` je ekvivalent pro BH metodu.

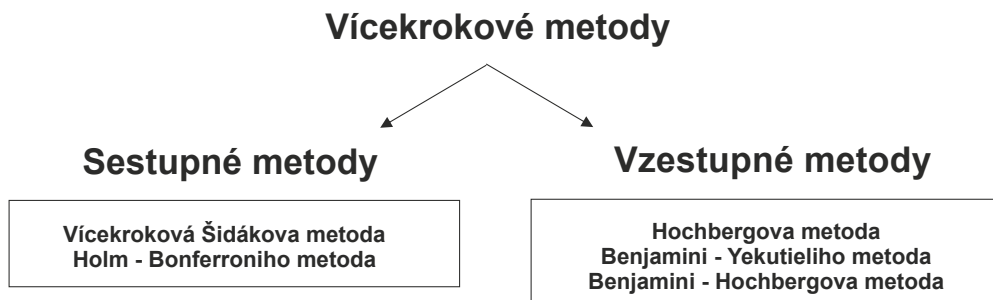
Na závěr této kapitoly je na Obrázku 5.1 a Obrázku 5.2 uvedené přehledné dělení metod mnohonásobného porovnávání.

Tato kapitola byla sepsána pomocí zdrojů [2, 6, 9, 11, 12, 13].





Obrázek 5.1: Dělení metod mnohonásobného porovnávání.



Obrázek 5.2: Dělení vícekových metod mnohonásobného porovnávání.

# Kapitola 6

## Reálné příklady

Nyní si uvedeme několik praktických příkladů, které jsou řešeny pomocí programu R. První příklad pochází z datové knihovny programu R. Po zadání příkazu `chickwts`, se dostaneme k datasetu, který obsahuje 71 pozorování (váž v gramech). Tato pozorování byla získána následovně: čerstvě vylíhnutá kuřata byla rozdělena do šesti skupin a každá skupina byla po dobu šesti týdnů krmena jiným krmivem. Cílem experimentu je zjistit, zda je mezi váhami kuřat z jednotlivých skupin významný rozdíl.

Druhý příklad obsahuje data, která jsem získala vlastním průzkumem. U mladých mužů ve věku 20 - 35 let jsem zjišťovala výšku, váhu a informaci o tom, kolik hodin týdně sportují. Získala jsem tedy skupinu náhodně vybraných mužů v určité věkové a výškové kategorii. Podle informace o počtu hodin v týdnu kdy se věnují sportu, pak spadají do určité (mnou vytvořené) skupiny. Cílem studie je ověřit, že mezi tělesnými váhami mužů z jednotlivých skupin není významný rozdíl.

Data pro třetí, a zároveň poslední příklad, opět pochází z datové knihovny programu R. Tento příklad je založený na testu o regresních parametrech a použítá data pochází z průzkumu mezi pracovníky velké finanční organizace.

Tato kapitola byla sepsána pomocí zdrojů [2, 6, 9, 11, 12, 13].

## 6.1. Příklad 1 - váhy kuřat

Čerstvě vylíhnutá kuřata byla rozdělena do šesti skupin o rozsazích 10, 12, 14, 12, 11 a 12. Každá skupina byla po dobu šesti týdnů krmena jiným typem krmiva a poté bylo každé kuře zváženo. První skupina byla krmena fazolemi (horsebean), druhá lněným semínkem (linseed), třetí sojou (soybean), čtvrtá slunečnicí (sunflower), pátá masovou moučkou (meatmeal) a šestá kaseinem (casein).

Na hladině testu  $\alpha = 0,05$  budeme testovat hypotézy, že mezi jednotlivými typy krmiv není významný rozdíl (co se váhy kuřat týče). Označme fazole F, lněné semínko L, soju SO, slunečnici SL, masovou moučku M a kasein K.

Jak již bylo řečeno, data použitá v tomto příkladu jsou převzata z datové knihovny programu R. Data získáme zadáním příkazu `chickwts` do programu R, který nám dataset následně zobrazí. Data jsou uvedena v Tabulce 6.1. V Příloze A jsou potom uvedeny boxploty pro váhy kuřat (v gramech) rozdělených do skupin a v Příloze B jsou znázorněny histogramy k tomuto příkladu.

Kuře	Fazole	Lněné sem.	Soja	Slunečnice	Masová m.	Kasein
1	179	309	243	423	325	368
2	160	229	230	340	257	390
3	136	181	248	392	303	379
4	227	141	327	339	315	260
5	217	260	329	341	380	404
6	168	203	250	226	153	318
7	108	148	193	320	263	352
8	124	169	271	295	242	359
9	143	213	316	334	206	216
10	140	257	267	322	344	222
11		244	199	297	258	283
12		271	171	318		332
13			158			
14			248			

Tabulka 6.1: Váhy kuřat v gramech po šesti týdnech výkrmu fazolemi, lněným semínkem, sojou, slunečnicí, masovou moučkou nebo kaseinem.

Z důvodu malého počtu pozorování zvolíme neparametrický přístup. Máme šest nezávislých statistických znaků  $F, L, SO, SL, M, K$ , se spojitým rozdělením

a s distribučními funkcemi  $F_F, F_L, F_{SO}, F_{SL}, F_M, F_K$ . Pomocí dvouvýběrového Wilcoxonova testu otestujeme hypotézu, že výběry pochází ze stejného rozdělení. Získáme tak neupravené  $p$ -hodnoty, které nás zajímají. Testujeme nulovou hypotézu  $H_{i,j}^0 : F_i = F_j$ , oproti alternativě  $H_{i,j}^A : F_i \neq F_j$ , kde  $i, j = F, L, SO, SL, M, K$  a hladina významnosti  $\alpha = 0,05$ .

**Poznámka 6.1.1** Dvouvýběrový Wilcoxonův test, který si nyní jen krátce připomeneme, je náhradou dvouvýběrového  $t$ -testu v případě, že není splněn předpoklad normality. Nechť máme dva nezávislé náhodné výběry  $X_1, \dots, X_m$  a  $Y_1, \dots, Y_n$ , které pochází ze spojitého rozdělení a mají distribuční funkce  $F_X$  a  $F_Y$ . Testuje se hypotéza  $H^0 : F_X = F_Y$  proti alternativě  $H^A : F_X \neq F_Y$ . Nulová hypotéza tvrdí, že oba výběry pochází ze stejného rozdělení.

Nejprve označíme výběry tak, aby platilo, že  $m \leq n$ . Poté všech  $m+n$  hodnot  $(x_1, \dots, x_m, y_1, \dots, y_n)$  uspořádáme vzestupně podle velikosti. Označme  $T_1$  součet pořadí hodnot  $x_1, \dots, x_m$ ,  $T_2$  součet pořadí hodnot  $y_1, \dots, y_n$  a vypočítáme

$$U_1 = mn + \frac{m(m+1)}{2} - T_1, U_2 = mn + \frac{n(n+1)}{2} - T_2.$$

Pokud je hodnota  $\min(U_1, U_2)$  menší nebo rovna kritické hodnotě pro tento test (kritické hodnoty jsou uvedeny ve speciálních tabulkách pro tento test), tak nulovou hypotézu na hladině testu nejvýše  $\alpha$  zamítáme.

Pokud jsou hodnoty  $m, n$  velké (větší než 10), užívá se testové kritérium

$$U_0 = \frac{U_1 - \frac{1}{2}mn}{\sqrt{\frac{mn}{12}(m+n+1)}},$$

které má za platnosti  $H_0$  asymptoticky rozdělení  $N(0,1)$ . Nulovou hypotézu zamítáme ve prospěch oboustranné alternativy, jestliže  $u_0$  leží v kritickém oboru  $W = (-\infty, -u_{\frac{\alpha}{2}}) \cup (u_{\frac{\alpha}{2}}, +\infty)$ , kde  $u_{\frac{\alpha}{2}}$  je příslušná kritická hodnota.

Neupravené  $p$ -hodnoty získané pomocí dvouvýběrového Wilcoxonova testu uložíme do matice výsledky, která je symetrická a na diagonále jsou jedničky. V programu R to provedeme následovně:

```

>vysledky=matrix(0, nrow=pocet_krmiva,ncol=pocet_krmiva)
>colnames(vysledky)=krmiva
>rownames(vysledky)=krmiva
>for (i in 1:nrow(vysledky))
{for (j in 1:ncol(vysledky))
{vysledky[i,j]=wilcox.test(data$weight[which(data$feed==krmiva[i])],
data$weight[which(data$feed==krmiva[j])])$p.val}}
>vysledky

```

Z matice výsledky vypíšeme vektor neupravených  $p$ -hodnot a uložíme ho do proměnné  $p$  (proměnná dvojice přiřadí k  $p$ -hodnotám dvojici krmiv). Vektor  $p$ -hodnot uspořádáme podle velikosti a uložíme ho do proměnné  $p\_usp$ .

```

>p <- c(vysledky[1,2:6], vysledky[2,3:6], vysledky[3,4:6],
vysledky[4,5:6], vysledky[5,6])
>p <- as.array(p)
>dvojice <- c("kasein-fazole", "kasein-lnenesem", "kasein-masovam",
"kasein-soja", "kasein-slunecnice", "fazole-lnenesem",
"fazole-masovam", "fazole-soja", "fazole-slunecnice",
"lnenesem-masovam", "lnenesem-soja", "lnenesem-slunecnice",
"masovam-soja", "masovam-slunecnice", "soja-slunecnice")
>rownames(p)=dvojice
>p_usp <- sort(p)

```

Získali jsme tedy vektor neupravených uspořádaných  $p$ -hodnot  $p\_usp$ , který odpovídá po řadě nulovým hypotézám:  $H_{F,SL}^0, H_{K,F}^0, H_{L,SL}^0, H_{F,M}^0, H_{F,SO}^0, H_{K,L}^0, H_{SO,SL}^0, H_{K,SO}^0, H_{F,L}^0, H_{L,M}^0, H_{M,SL}^0, H_{K,M}^0, H_{L,SO}^0, H_{M,SO}^0, H_{K,SL}^0$ . Hypotéza  $H_{i,j}^0$  potom říká, že mezi krmivy  $i$  a  $j$ , kde  $i, j = F, L, SO, SL, M, K$ , není významný rozdíl. Neupravené uspořádané  $p$ -hodnoty jsou uvedeny v Tabulce 6.2.

Upravené  $p$ -hodnoty pro Bonferroniho, Holm-Bonferroniho, Hochbergovu, BH a BY metodu získáme pomocí programu R, zadáním následujících příkazů:

```

>metody=c("bonferroni", "holm", "hochberg", "BH", "BY")
>pupr=matrix(0, nrow=length(p_usp), ncol=length(metody))

```

```

>for(i in 1:length(metody))pupr[,i]=p.adjust(p_usp,method=metody[i])
>rownames(pupr) <- rownames(p_usp)
>colnames(pupr) <- metody

```

Upravené  $p$ -hodnoty pro jednokrokovou Šidákovu metodu získáme pomocí programu R zadáním příkazu:

```

>m=dim(p)
>sidak=1-(1-p_usp)^(m)

```

Upravené  $p$ -hodnoty pro víceřadkovou Šidákovu metodu získáme pomocí programu R zadáním příkazu:

```

>r <- rep(0,dim(p))
>for (i in 1:dim(p) ){r[i]=1-(1-p_usp[i]) ^ (m-i+1)}
>sidvic <- rep(0,dim(p))
>for(i in 1:dim(p)){sidvic[i] <- max(r[1:i])}

```

Nyní je potřeba zobrazit všechny upravené  $p$ -hodnoty v jedné matici. Vytvoříme matici `vysledek`, která obsahuje upravené  $p$ -hodnoty pro všechny metody. Výsledné upravené  $p$ -hodnoty jsou uvedeny v Tabulce 6.2 a zamítnuté hypotézy jsou zde označeny hvězdičkou.

```

>vysledek <- cbind(pupr,sidak,sidvic)
>vysledek <- vysledek[,c(1,6,7,2:5)]

```

Stejně závěry získáme i pomocí neupravených  $p$ -hodnot a upravených hladin významnosti. Výsledné upravené hladiny významnosti jsou v Tabulce 6.3.

U hypotéz, které byly zamítnuty, nás zajímá, které krmivo má lepší výsledky. To zjistíme tak, že vypočítáme průměrné váhy kuřat v jednotlivých skupinách a následně pro jednotlivé hypotézy vypočítáme rozdíly těchto průměrných vah (průměrná váha kuřat ze skupiny v řádku - průměrná váha kuřat ze skupiny ve sloupci). Kladné číslo potom značí, že kuřata krmená krmivem v řádku jsou těžší, tedy krmivo uvedené v řádku je lepší než krmivo uvedené ve sloupci. Tyto rozdíly jsou uvedeny v Tabulce 6.4.

```

>rozdily <- matrix(0,nrow=pocet_krmiva, ncol=pocet_krmiva)
>rownames(rozdily) <- krmiva

```

```

>colnames(rozdily) <- krmiva
>for(i in 1:nrow(rozdily)){for(j in 1:ncol(rozdily))
rozily[i,j] <- mean(data$weight[data$feed==krmiva[i]])-
mean(data$weight[data$feed==krmiva[j]])}

```

Závěr: Z tabulek je jasně vidět, že ať testujeme pomocí neupravených  $p$ -hodnot a upravených hladin významnosti nebo upravených  $p$ -hodnot a neupravených hladin významnosti, závěr je stejný. Bonferroniho a Šidákova metoda zamítají pouze 7 hypotéz. Holmova metoda a Hochbergova metoda zamítají 8 hypotéz a Šidákova vícekroková metoda, BH metoda a metoda BY zamítají 9 hypotéz.

Když se podíváme na Tabulku 6.2, vidíme, že prvních 7 hypotéz zamítají všechny metody. Z Tabulky 6.3 potom vyčteme, že kasein, masová moučka, soja a slunečnice jsou lepší než fazole. Kasein a slunečnice jsou lepší než lněné semínko a slunečnice je lepší než soja. Šidákova vícekroková metoda, Holmova metoda, Hochbergova metoda, BH metoda a BY metoda zamítají navíc ještě hypotézu  $H_{K,SO}^0$  a podle těchto metod je ještě navíc lepší kasein než soja. Šidákova vícekroková metoda, metoda BH a metoda BY navíc ještě zamítají hypotézu  $H_{F,L}^0$  a podle těchto metod je ještě navíc lepší lněné semínko než fazole. Fazole tedy vyšly jako nejhorší krmivo.

$H_{i,j}^0$	neupravené	Bonferroni	Šidák J.	Šidák V.
$H_{F,SL}^0$	0,000006185765	9,278647e-05*	9,278245e-05*	9,278245e-05*
$H_{K,F}^0$	0,00002165018	3,247526e-04*	3,247034e-04*	3,030598e-04*
$H_{L,SL}^0$	0,00004955335	7,433003e-04*	7,430425e-04*	6,440021e-04*
$H_{F,M}^0$	0,0002551628	3,827442e-03*	3,820613e-03*	3,057660e-03*
$H_{F,SO}^0$	0,0007582898	1,137435e-02*	1,131417e-02*	8,309635e-03*
$H_{K,L}^0$	0,001221273	1,831910e-02*	1,816332e-02*	1,214583e-02*
$H_{SO,SL}^0$	0,001557454	2,336181e-02*	2,310883e-02*	1,393008e-02*
$H_{K,SO}^0$	0,00591939	8,879085e-02	8,520446e-02	4,638555e-02*
$H_{F,L}^0$	0,007144558	1,071684e-01	1,019711e-01	4,895264e-02*
$H_{L,M}^0$	0,03634269	5,451403e-01	4,260959e-01	1,991785e-01
$H_{M,S}^0$	0,06881704	1	6,568173e-01	2,998758e-01
$H_{K,M}^0$	0,09084239	1	7,603444e-01	3,167861e-01
$H_{L,SO}^0$	0,2366518	1	9,825884e-01	5,551967e-01
$H_{M,SO}^0$	0,2616558	1	9,894347e-01	5,551967e-01
$H_{K,SL}^0$	1	1	1	1
$H_{i,j}^0$	Holm	Hochberg	BH	BY
$H_{F,SL}^0$	9,278647e-05*	9,278647e-05*	9,278647e-05*	0,0003078868*
$H_{K,F}^0$	3,031025e-04*	3,031025e-04*	1,623763e-04*	0,0005388018*
$H_{L,SL}^0$	6,441936e-04*	6,441936e-04*	2,477668e-04*	0,0008221469*
$H_{F,M}^0$	3,061954e-03*	3,061954e-03*	9,568605e-04*	0,0031750822*
$H_{F,SO}^0$	8,341188e-03*	8,341188e-03*	2,274869e-03*	0,0075485379*
$H_{K,L}^0$	1,221273e-02*	1,221273e-02*	3,053184e-03*	0,0101311621*
$H_{SO,SL}^0$	1,401709e-02*	1,401709e-02*	3,337402e-03*	0,0110742635*
$H_{K,SO}^0$	4,735512e-02*	4,735512e-02*	1,109886e-02*	0,0368285476*
$H_{F,L}^0$	5,001191e-02	5,001191e-02	1,190760e-02*	0,0395121338*
$H_{L,M}^0$	2,180561e-01	2,180561e-01	5,451403e-02	0,1808900289
$H_{M,S}^0$	3,440852e-01	3,440852e-01	9,384141e-02	0,3113872951
$H_{K,M}^0$	3,633696e-01	3,633696e-01	1,135530e-01	0,3767948246
$H_{L,SO}^0$	7,099555e-01	5,233116e-01	2,730598e-01	0,9060750196
$H_{M,SO}^0$	7,099555e-01	5,233116e-01	2,803455e-01	0,9302505697
$H_{K,SL}^0$	1	1	1	1

Tabulka 6.2: Neupravené a upravené  $p$ -hodnoty (data chickwts).



$H_{i,j}^0$	Uspoř. neupr. $p$ -hodnoty	Bonferroni	Šidák J.	Šidák V.
$H_{F,SL}^0$	0,000006185765	0,003333*	0,003414*	0,003414*
$H_{K,F}^0$	0,00002165018	0,003333*	0,003414*	0,003657*
$H_{L,SL}^0$	0,00004955335	0,003333*	0,003414*	0,003938*
$H_{F,M}^0$	0,0002551628	0,003333*	0,003414*	0,004265*
$H_{F,SO}^0$	0,0007582898	0,003333*	0,003414*	0,004652*
$H_{K,L}^0$	0,001221273	0,003333*	0,003414*	0,005116*
$H_{SO,SL}^0$	0,001557454	0,003333*	0,003414*	0,005683*
$H_{K,SO}^0$	0,00591939	0,003333	0,003414	0,006391*
$H_{F,L}^0$	0,007144558	0,003333	0,003414	0,007301*
$H_{L,M}^0$	0,03634269	0,003333	0,003414	0,008512
$H_{M,S}^0$	0,06881704	0,003333	0,003414	0,010206
$H_{K,M}^0$	0,09084239	0,003333	0,003414	0,012741
$H_{L,SO}^0$	0,2366518	0,003333	0,003414	0,016952
$H_{M,SO}^0$	0,2616558	0,003333	0,003414	0,025321
$H_{K,SL}^0$	1	0,003333	0,003414	0,050000
$H_{i,j}^0$	Holm	Hochberg	BH	BY
$H_{F,SL}^0$	0,003333*	0,003333*	0,003333*	0,001005*
$H_{K,F}^0$	0,003571*	0,003571*	0,006667*	0,002009*
$H_{L,SL}^0$	0,003846*	0,003846*	0,010000*	0,003014*
$H_{F,M}^0$	0,004167*	0,004167*	0,013333*	0,004018*
$H_{F,SO}^0$	0,004545*	0,004545*	0,016667*	0,005023*
$H_{K,L}^0$	0,005000*	0,005000*	0,020000*	0,006027*
$H_{SO,SL}^0$	0,005556*	0,005556*	0,023333*	0,007032*
$H_{K,SO}^0$	0,006250*	0,006250*	0,026667*	0,008036*
$H_{F,L}^0$	0,007143	0,007143	0,030000*	0,009041*
$H_{L,M}^0$	0,008333	0,008333	0,033333	0,010046
$H_{M,S}^0$	0,010000	0,010000	0,036667	0,011050
$H_{K,M}^0$	0,012500	0,012500	0,040000	0,012055
$H_{L,SO}^0$	0,016667	0,016667	0,043333	0,013059
$H_{M,SO}^0$	0,025000	0,025000	0,046667	0,014064
$H_{K,SL}^0$	0,050000	0,050000	0,050000	0,015068

Tabulka 6.3: Upravené hladiny významnosti (data chickwts).

Krmiva	Prům.	Kasein	Fazole	Lněné s.	Mas. m.	Soja	Sluneč.
Kasein	323,58	0	163,38	104,83	46,67	77,15	-5,33
Fazole	160,2	-163,38	0	-58,55	-116,71	-86,23	-168,72
Lněné s.	218,75	-104,83	58,55	0	-58,16	-27,68	-110,17
Masová m.	276,91	-46,67	116,71	58,16	0	30,48	-52,01
Soja	246,43	-77,15	86,23	27,68	-30,48	0	-82,49
Slunečnice	328,92	5,33	168,72	110,17	52,01	82,49	0

Tabulka 6.4: Průměrné váhy ve skupinách a rozdíly průměrných vah kuřat v gra-  
mech: průměrná váha kuřat ze skupiny v řádku - průměrná váha kuřat ze skupiny  
ve sloupci (data chickwts).

## 6.2. Příklad 2 - mladí muži

Neméně zajímavé bude aplikovat metody mnohonásobného porovnávání na data získaná vlastním průzkumem. Náhodně vybraných mužů jsem se ptala na jejich věk, výšku, váhu a v neposlední řadě na to, kolik hodin týdně věnují sportu. Ze všech získaných dat jsem potom vytvořila skupinu mužů ve věku 20 až 35 let, kteří měří 180-195 centimetrů. Obecně je známo, že nezáleží na tom kolik člověk váží, ale na tom jak se cítí, jak se sám sobě líbí a hlavně na tom, jestli je zdravý. Také je známo, že každý z nás má v těle jiný podíl tuku a jiný podíl svalové hmoty. Zajímá mě tedy, jestli má počet hodin týdně strávených sportem vliv na váhu mladého muže. Ze získaných údajů o jednotlivých mužích jsem zjistila, že někdo nesportuje vůbec a někdo naopak tráví sportem několik hodin denně. Muže jsem rozdělila do pěti skupin podle toho, kolik hodin týdně sportují:  $A = 0$ ,  $B = (0, 3)$ ,  $C = (3, 6)$ ,  $D = (6, 9)$ ,  $E = (9, 168)$ . V Tabulce 6.5 jsou uvedené váhy mužů rozděleny do výše uvedených skupin.

Muž č.	A = 0	B = (0, 3)	C = (3, 6)	D = (6, 9)	E = (9, 168)
1	83	90	90	95	72
2	83	83	90	87	92
3	100	83	85	100	78
4	75	100	83	94	81
5	92	85	72	86	89
6	83	100	76	81	110
7	83	90	108	78	97
8	92	85	95	87	80
9	76	68	87	94	84
10	86	82	93	82	77
11	64	96	87	80	90
12	83	94	84	98	85
13	75	81	95	81	95
14	83	74	100	82	84
15	85	96	86	84	95

Tabulka 6.5: Váhy mužů v kg, ve věku 20 až 35 let, vysokých 180 až 195 cm rozdělené do 5 skupin podle počtu hodin strávených sportem za týden.

Příloha C obsahuje kompletní netříděná data z průzkumu, kde navíc můžeme vidět věk a výšku jednotlivých mužů. V Příloze D pak najdeme souhrn charakteristik pro tyto data, v Příloze E boxploty a v Příloze F histogramy k těmto datům.

Abychom mohli s daty pracovat v programu R, musíme data nejprve importovat. Data uložená v Excelu naimportujeme pomocí příkazu `read.csv(...)`. Následně musíme ověřit normalitu testovaných dat, abychom mohli dále vybrat vhodný test. K tomu použijeme Shapirův-Wilkův test normality a budeme testovat nulovou hypotézu  $H_j^0$  : Data  $j$  pochází z normálního rozdělení, oproti alternativě  $H_j^A$  : Data  $j$  nepochází z normálního rozdělení, kde  $j = A, B, C, D, E$ . Hladina významnosti je  $\alpha = 0,05$ .

```
>data <- read.csv("C:/Documents and Settings/Home/Plocha/vahy_muzi/
vahy_muzi.csv", sep=";")
>urovne <- levels(data$skupina) #skupiny
>pocet_urovni <- length(as.factor(urovne)) #pocet skupin
>shapiro = tapply(data$vaha,data$skupina, FUN="shapiro.test")
#vektor p-hodnot ze shapirova testu
>p.val <- rep(0, dim(shapiro))
>p.val <- cbind(levels(data$skupina),p.val)
>for(i in 1:dim(shapiro)){ p.val[i,2] <- shapiro[[i]]$p.val}
```

Skupina	$p$ -hodnoty
A	0,286902980493173
B	0,561213310908769
C	0,93271801042817
D	0,116445647458986
E	0,742511579678453

Tabulka 6.6:  $P$ -hodnoty ze Shapirova-Wilkova testu.

Výsledné  $p$ -hodnoty jsou uvedeny v Tabulce 6.6. Všechny  $p$ -hodnoty získané pomocí Shapirova-Wilkova testu normality jsou větší než hladina významnosti

$\alpha = 0,05$ , tudíž nulové hypotézy nezamítáme a budeme dále předpokládat normalitu dat. Nyní je potřeba ověřit rovnost rozptylů. To provedeme pomocí Bartlettova testu a budeme testovat nulovou hypotézu  $H^0$  : všechny výběry mají shodné rozptyly, oproti alternativě  $H^A$  : všechny výběry nemají shodné rozptyly, kde výběry jsou  $A, B, C, D, E$ .

```
>bartlett.test(vaha ~ skupina,data=data)
```

```
Bartlett test of homogeneity of variances
```

```
data: vaha by skupina
```

```
Bartlett's K-squared = 1.4106, df = 4, p-value = 0.8423
```

Bartlettův test nezamítl rovnost rozptylů, tudíž budeme dále předpokládat, že rozptyly jsou si rovny.

Máme pět nezávislých statistických znaků  $A, B, C, D, E$ , které mají normální rozdělení se stejným (neznámým) rozptylem, tj.  $A \sim N(\mu_A, \sigma^2)$ ,  $B \sim N(\mu_B, \sigma^2)$ ,  $C \sim N(\mu_C, \sigma^2)$ ,  $D \sim N(\mu_D, \sigma^2)$ ,  $E \sim N(\mu_E, \sigma^2)$  a chceme na hladině významnosti  $\alpha = 0,05$  otestovat nulové hypotézy o rovnosti středních hodnot, tj.  $H_{i,j}^0 : \mu_i = \mu_j$ , oproti alternativě  $H_{i,j}^A : \mu_i \neq \mu_j$ , kde  $i, j = A, B, C, D, E$ . K tomu využijeme dvouvýběrový t-test, díky kterému získáme neupravené  $p$ -hodnoty.

Vypočítané neupravené  $p$ -hodnoty uložíme do matice výsledky, která je symetrická a na diagonále jsou jedničky. Z této matice vypíšeme vektor neupravených  $p$ -hodnot a uložíme ho do proměnné  $p$ . Vektor  $p$ -hodnot uspořádáme dle velikosti a uložíme ho do proměnné  $p\_usp$ . Uspořádané neupravené  $p$ -hodnoty jsou uvedeny v Tabulce 6.7. V programu R to provedeme stejně jako v předchozím příkladu. Změníme pouze následující část (t.test):

```
>for (i in 1:nrow(vysledky)){for (j in 1:ncol(vysledky))
{vysledky[i,j]=t.test(data$vaha[which(data$skupina==urovne[i])],
data$vaha[which(data$skupina==urovne[j])])$p.val}}
```

Máme vektor neupravených  $p$ -hodnot a nyní je potřeba vypočítat upravené  $p$ -hodnoty. To provedeme opět stejně jako v předchozím příkladu. Výsledné upravené  $p$ -hodnoty jsou uvedeny v Tabulce 6.7.

Pokud budeme testovat dané hypotézy pomocí neupravených  $p$ -hodnot a upra-

vených hladin významnosti, dojdeme ke stejnému závěru. Vypočítané upravené hladiny významnosti jsou uvedeny v Tabulce 6.8.

Závěr: Z tabulek je jasné vidět, že ať testujeme pomocí neupravených  $p$ -hodnot a upravených hladin významnosti nebo upravených  $p$ -hodnot a neupravené hladiny významnosti, závěr bude opět stejný a to takový, že žádná z metod nezamítla žádnou z hypotéz. Počet hodin, které muži věnují týdně sportu, tedy významně neovlivňuje jejich váhu. Poměr svalů a tuků v jejich těle bude zřejmě jiný, ale na váhu to významný vliv nemá.

**Poznámka 6.2.1** Klasickým způsobem bychom nepoužili dvouvýběrový t-test, ale metodu zvanou ANOVA (analysis of variance neboli analýza rozptylu), která je jeho zobecněním pro případ více výběrů. V případě testování více skupin není dvouvýběrový t-test vhodný, protože nezaručuje splnění podmínky, že pravděpodobnost chyby prvního druhu je nejvýše  $\alpha$ . Proto jsme museli následně použít metody mnohonásobného porovnávání. V případě ANOVY bychom nejprve testovali nulovou hypotézu  $H^0 : \mu_A = \dots = \mu_E$ , která tvrdí, že všechny střední hodnoty jsou stejné oproti alternativě, že alespoň jedna dvojice středních hodnot se liší, tj.  $H^A : \exists i \in \{A, \dots, E\} : \mu_i \neq \mu_j, j = A, \dots, E$ . Testuje se pomocí statistiky

$$F_A = \frac{S_A}{S_e} \frac{n - k}{k - 1},$$

kde  $S_A$  je skupinový součet čtverců,  $S_e$  je reziduální součet čtverců,  $n$  je celkový počet pozorování (75) a  $k$  je počet skupin (5). Statistika  $F_A$  má za platnosti  $H^0$  rozdělení  $F_{k-1, n-k}$ . Nulovou hypotézu  $H^0$  potom zamítáme na hladině  $\alpha$ , jestliže  $f_A \geq F_{k-1, n-k; \alpha}$ , kde  $f_A$  je hodnota testového kritéria a  $F_{k-1, n-k; \alpha}$  je příslušná kritická hodnota.

V našem případě vyšla hodnota statistiky  $F_A$  přibližně 0,965433, a když tuto hodnotu porovnáme s kvantilem  $F_{4, 70; 0, 95}$ , který je přibližně roven 2,525, tak dojdeme k závěru, že hypotézu  $H_0$  nelze zamítnout.

V případě, že bychom nulovou hypotézu o rovnosti středních hodnot na hladině  $\alpha$  zamítli, na téže hladině je následně potřeba rozhodnout, které výběry se

od sebe liší, tj. otestovat platnost hypotézy  $H_{ij}^0 : \mu_i = \mu_j$  proti  $H_{ij}^A : \mu_i \neq \mu_j$ , kde  $i, j = A, \dots, E$ . K tomu lze využít například Scheffeho nebo Tukeyho metodu mnohonásobného porovnávání. Vzorce pro tyto dvě metody a přesnější postupy výpočtů lze najít v [9]. Výpočtům se zde nebudu podrobněji věnovat, nicméně Tabulka 6.9 dokazuje, že bychom došli ke stejným závěrům a nezamítli bychom žádnou z dílčích hypotéz, protože pravá strana jak pro Scheffého, tak pro Tukeyho metodu je vždy větší než absolutní hodnota rozdílů průměrů jednotlivých skupin.

$H_{i,j}^0$	neupravené	Bonferroni	Šidák J.	Šidák V.
$H_{A,C}^0$	0,07576388	0,7576388	0,5451907	0,5451907
$H_{A,D}^0$	0,13514453	1	0,7658817	0,7292978
$H_{A,E}^0$	0,19540386	1	0,8862950	0,8243600
$H_{A,B}^0$	0,19814481	1	0,8901097	0,8243600
$H_{C,D}^0$	0,62392656	1	0,9999434	0,9971710
$H_{B,C}^0$	0,63404154	1	0,9999569	0,9971710
$H_{C,E}^0$	0,66966414	1	0,9999845	0,9971710
$H_{B,D}^0$	0,96507609	1	1	0,9999574
$H_{B,E}^0$	0,96947768	1	1	0,9999574
$H_{D,E}^0$	1	1	1	1
$H_{i,j}^0$	Holm	Hochberg	BH	BY
$H_{A,C}^0$	0,7576388	0,7576388	0,4953620	1
$H_{A,D}^0$	1	1	0,4953620	1
$H_{A,E}^0$	1	1	0,4953620	1
$H_{A,B}^0$	1	1	0,4953620	1
$H_{C,D}^0$	1	1	0,9566631	1
$H_{B,C}^0$	1	1	0,9566631	1
$H_{C,E}^0$	1	1	0,9566631	1
$H_{B,D}^0$	1	1	1	1
$H_{B,E}^0$	1	1	1	1
$H_{D,E}^0$	1	1	1	1

Tabulka 6.7: Neupravené a upravené  $p$ -hodnoty (data mladí muži).

$H_{i,j}^0$	Uspoř. neupr. $p$ -hodnoty	Bonferroni	Šidák J.	Šidák V.
$H_{A,C}^0$	0,07576388	0,005	0,005116197	0,005116197
$H_{A,D}^0$	0,13514453	0,005	0,005116197	0,005683045
$H_{A,E}^0$	0,19540386	0,005	0,005116197	0,006391151
$H_{A,B}^0$	0,19814481	0,005	0,005116197	0,007300832
$H_{C,D}^0$	0,62392656	0,005	0,005116197	0,008512445
$H_{B,C}^0$	0,63404154	0,005	0,005116197	0,010206218
$H_{C,E}^0$	0,66966414	0,005	0,005116197	0,012741455
$H_{B,D}^0$	0,96507609	0,005	0,005116197	0,016952428
$H_{B,E}^0$	0,96947768	0,005	0,005116197	0,025320566
$H_{D,E}^0$	1	0,005	0,005116197	0,05
$H_{i,j}^0$	Holm	Hochberg	BH	BY
$H_{A,C}^0$	0,005	0,005	0,005	0,014644841
$H_{A,D}^0$	0,005555556	0,005555556	0,01	0,029289683
$H_{A,E}^0$	0,00625	0,00625	0,015	0,043934524
$H_{A,B}^0$	0,007142857	0,007142857	0,02	0,058579365
$H_{C,D}^0$	0,008333333	0,008333333	0,025	0,073224206
$H_{B,C}^0$	0,01	0,01	0,03	0,087869048
$H_{C,E}^0$	0,0125	0,0125	0,035	0,102513889
$H_{B,D}^0$	0,016666667	0,016666667	0,04	0,11715873
$H_{B,E}^0$	0,025	0,025	0,045	0,131803571
$H_{D,E}^0$	0,05	0,05	0,05	0,146448413

Tabulka 6.8: Upravené hladiny významnosti (data mladí muži).

dvojice	absolutní hodnota rozdílů průměrů	Scheffe: pravá strana nerovnosti	Tukey: pravá strana nerovnosti
A, B	4,267	40,703	8,977
A, C	5,867	40,703	8,977
A, D	4,4	40,703	8,977
A, E	4,4	40,703	8,977
B, C	1,6	40,703	8,977
B, D	0,133	40,703	8,977
B, E	0,133	40,703	8,977
C, D	1,467	40,703	8,977
C, E	1,467	40,703	8,977
D, E	0	40,703	8,977

Tabulka 6.9: Hodnoty pro Scheffého a Tukeyho metodu (data mladí muži).



### 6.3. Příklad 3 - regrese

Byl proveden průzkum spokojenosti mezi pracovníky velké finanční organizace. Z 30 oddělení bylo náhodně vybráno 35 zaměstnanců a zjišťovalo se, zda jsou spokojeni se 7 oblastmi firemní kultury. Daná data potom vyjadřují procentuální podíl příznivých odpovědí v každém odvětví pro každou oblast. Data pro třetí a zároveň poslední příklad opět pochází z datové knihovny softwaru R a zobrazí se po zadání příkazu `attitude`. Cílem úlohy je testovat významnost jednotlivých parametrů.

Data `attitude` obsahují jednu závislou proměnnou  $Y = \text{rating}$  (celkové hodnocení) a šest nezávislých proměnných  $x_1 = \text{complaints}$  (vyřizování stížností zaměstnanců),  $x_2 = \text{privileges}$  ((ne)poskytování zvláštních privilegií),  $x_3 = \text{learning}$  (možnost vzdělávání se),  $x_4 = \text{raises}$  (zvyšování platu v závislosti na výkonu),  $x_5 = \text{critical}$  (kritizování) a  $x_6 = \text{advance}$  (možnost povýšení). Cílem je otestovat významnost všech parametrů zároveň. Hodnoty jsou shrnuty v Tabulce 6.10.

Jak již bylo naznačeno, budeme pracovat s regresní funkcí. Nejprve si tedy nadefinujeme regresní model a pomocí příkazu `summary(...)`, jehož výsledky jsou uvedeny na Obrázku 6.1, získáme  $p$ -hodnoty z dílčích t-testů. Dílčí t-testy jsou testy, které nám umožňují testovat oprávněnost setrvání dané vysvětlující proměnné v regresním modelu, tj. testují významnost jednotlivých parametrů. Pokud testujeme na hladině významnosti  $\alpha = 0,05$ , na Obrázku 6.1 můžeme vidět, že za významný byl označen pouze parametr `complaints`. Získané  $p$ -hodnoty uspořádáme podle velikosti a uložíme je do proměnné `p_ustp`.

```
> data <- attitude
> model <- lm(data$rating ~ data$complaints+data$privileges+
data$learning+data$raises+data$critical+data$advance)
> summary(model)
> p <- summary(model)$coef[,4]
> p_ustp <- sort(p)
```

Nyní je potřeba vypočítat upravené  $p$ -hodnoty a zjistit, které nulové hypotézy budou dle jednotlivých metod zamítnuty, resp. které parametry jsou dle jednotlivých

metod významné. To provedeme pomocí stejných příkazů jako v předchozích příkladech. V Tabulce 6.11 jsou uvedeny výsledné upravené  $p$ -hodnoty a můžeme vidět, že všechny metody zamítly pouze hypotézu týkající se parametru complaints a můžeme tedy říct, že podle všech metod je jediným významným parametrem právě tento. Abychom mohli ještě hypotézy otestovat pomocí neupravených  $p$ -hodnot a upravených hladin významnosti, opět musíme tyto upravené hladiny významnosti vypočítat. Upravené hladiny významnosti jsou uvedeny v Tabulce 6.12.

Závěr: Z tabulek je jasné vidět, že ať testujeme pomocí neupravených  $p$ -hodnot a upravených hladin významnosti nebo upravených  $p$ -hodnot a neupravené hladiny významnosti, závěr bude zase stejný a to takový, že podle všech metod je jediným významným parametrem parametr complaints. Výsledný model pak bude zahrnovat pouze tento parametr:

```
> model2 <- lm(data$rating ~ data$complaints)
> summary(model2)
```

a `summary(model2)` pro výsledný model je uvedeno v dolní části Obrázku 6.1.

	rating	complaints	privileges	learning	raises	critical	advance
1	43	51	30	39	61	92	45
2	63	64	51	54	63	73	47
3	71	70	68	69	76	86	48
4	61	63	45	47	54	84	35
5	81	78	56	66	71	83	47
6	43	55	49	44	54	49	34
7	58	67	42	56	66	68	35
8	71	75	50	55	70	66	41
9	72	82	72	67	71	83	31
10	67	61	45	47	62	80	41
11	64	53	53	58	58	67	34
12	67	60	47	39	59	74	41
13	69	62	57	42	55	63	25
14	68	83	83	45	59	77	35
15	77	77	54	72	79	77	46
16	81	90	50	72	60	54	36
17	74	85	64	69	79	79	63
18	65	60	65	75	55	80	60
19	65	70	46	57	75	85	46
20	50	58	68	54	64	78	52
21	50	40	33	34	43	64	33
22	64	61	52	62	66	80	41
23	53	66	52	50	63	80	37
24	40	37	42	58	50	57	49
25	63	54	42	48	66	75	33
26	66	77	66	63	88	76	72
27	78	75	58	74	80	78	49
28	48	57	44	45	51	83	38
29	85	85	71	71	77	74	55
30	82	82	39	59	64	78	39

Tabulka 6.10: Data attitude.

```

##summary pro celkovy model

> summary(model)

Call:
lm(formula = data$rating ~ data$complaints + data$privileges +
    data$learning + data$raises + data$critical + data$advance)

Residuals:
    Min       1Q   Median       3Q      Max
-10.9418  -4.3555   0.3158   5.5425  11.5990

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   10.78708    11.58926   0.931 0.361634
data$complaints  0.61319     0.16098   3.809 0.000903 ***
data$privileges -0.07305     0.13572  -0.538 0.595594
data$learning   0.32033     0.16852   1.901 0.069925 .
data$raises     0.08173     0.22148   0.369 0.715480
data$critical   0.03838     0.14700   0.261 0.796334
data$advance   -0.21706     0.17821  -1.218 0.235577
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.068 on 23 degrees of freedom
Multiple R-squared:  0.7326, Adjusted R-squared:  0.6628
F-statistic: 10.5 on 6 and 23 DF,  p-value: 1.24e-05

#####
##summary pro model2, obsahujici pouze parametr complaints

> summary(model2)

Call:
lm(formula = data$rating ~ data$complaints)

Residuals:
    Min       1Q   Median       3Q      Max
-12.8799  -5.9905   0.1783   6.2978   9.6294

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   14.37632     6.61999   2.172  0.0385 *
data$complaints  0.75461     0.09753   7.737 1.99e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.993 on 28 degrees of freedom
Multiple R-squared:  0.6813, Adjusted R-squared:  0.6699
F-statistic: 59.86 on 1 and 28 DF,  p-value: 1.988e-08

```

Obrázek 6.1: Shrnutí celkového modelu a modelu, který obsahuje pouze významný parametr complaints.

Parametr	neupravené	Bonferroni	Šidák J.	Šidák V.
complaints	0,000903	0,006320075*	0,006302982*	0,006302982*
learning	0,069925	0,489477422	0,397960945	0,352698138
advance	0,235577	1	0,847476759	0,738983122
privileges	0,595594	1	0,998231017	0,9338617
raises	0,71548	1	0,999849065	0,9338617
critical	0,796334	1	0,999985465	0,9338617
Parametr	Holm	Hochberg	BH	BY
complaints	0,006320075*	0,006320075*	0,006320075*	0,01638705*
learning	0,419552076	0,419552076	0,244738711	0,63457251
advance	1	0,796334264	0,54967978	1
privileges	1	0,796334264	0,796334264	1
raises	1	0,796334264	0,796334264	1
critical	1	0,796334264	0,796334264	1

Tabulka 6.11: Neupravené a upravené  $p$ -hodnoty (data attitude).

Parametr	Usp. neupr. $p$ -hodnoty	Bonferroni	Šidák J.	Šidák V.
complaints	0,000903	0,008333*	0,008512*	0,008512*
learning	0,069925	0,008333	0,008512	0,010206
advance	0,235577	0,008333	0,008512	0,012741
privileges	0,595594	0,008333	0,008512	0,016952
raises	0,71548	0,008333	0,008512	0,025321
critical	0,796334	0,008333	0,008512	0,05
Parametr	Holm	Hochberg	BH	BY
complaints	0,008333*	0,008333*	0,008333*	0,003401*
learning	0,01	0,01	0,016667	0,006803
advance	0,0125	0,0125	0,025	0,010204
privileges	0,016667	0,016667	0,033333	0,013605
raises	0,025	0,025	0,041667	0,017007
critical	0,05	0,05	0,05	0,020408

Tabulka 6.12: Upravené hladiny významnosti (data attitude).

# Kapitola 7

## Simulace

Na závěr celé práce se podíváme na simulační studii, která je věnovaná srovnávání metod mnohonásobného porovnávání. Soustředít se budeme zejména na otázku, jak jednotlivé metody dodržují hladinu významnosti a jakou mají sílu testu. Pro jednoduchost budeme uvažovat pouze nezávislé testové statistiky s normálním rozdělením a bude nás zajímat, jak se výsledky mění v závislosti na počtu porovnávaných skupin a počtu pozorování v každé skupině.

Pro všechny simulace bude platit, že budeme testovat na hladině významnosti  $\alpha = 0,05$  a budeme provádět 10 000 iterací. Všechna data budeme generovat z normálního rozdělení pomocí příkazu `rnorm`. Příkaz `set.seed(...)` nám zajistí, aby vygenerované hodnoty byly pro každou novou simulaci stejné. Čtenář si tak může celou simulační studii sám zopakovat a dostane stejné výsledky.

Označme nyní počet skupin  $m$  a počet pozorování v každé skupině  $n$ . To jsou parametry, které kromě počtu iterací nebo hladiny významnosti můžeme měnit. Celková nulová hypotéza, kterou chceme otestovat říká, že střední hodnoty pro všechny výběry jsou rovny nule, tj.  $H^0 : \mu_1 = \mu_2 = \dots = \mu_m = 0$ . K zamítnutí celkové hypotézy však dochází vlivem zamítnutí dílčích testů. Pro  $m$  skupin máme  $\binom{m}{2}$  dílčích testů, které testují nulovost středních hodnot jednotlivých výběrů, tj.  $H_{ij}^0 : \mu_i = \mu_j$ , pro  $i, j = 1, \dots, m$ . Z důvodu grafického zobrazení silofunkcí budeme dále předpokládat, že pro prvních  $m - 1$  výběrů je střední hodnota rovna 0, čímž budeme mít zajištěnou platnost hypotéz  $H_{ij}^0$ , pro  $i, j = 1, 2, \dots, m - 1$ . U zbývajících  $m$ -tého náhodného výběru budeme měnit hodnotu střední hodnoty

(parametru  $\mu_m$ ) v rozsahu od 0 do 3. Hodnoty měnícího se parametru  $\mu_m$  potom udávají míru porušení nulové hypotézy  $H_{im}^0$ , pro  $i = 1, 2, \dots, m - 1$ .

Z důvodu dosažení „pěkných“ grafů byl zvolen krok 0,01, což vzhledem k faktu, že bylo prováděno 10 000 iterací, ve výsledku vedlo k několika hodinovému výpočtu a provádění simulací se značně časově protáhlo. Výsledkem jsou vykreslené funkce, které představují řezy silofunkcí pro měnící se  $\mu_m$ , při pevných  $\mu_1 = 0, \mu_2 = 0, \dots, \mu_{m-1} = 0$  (vždy pro  $m$  skupin a  $n$  pozorování v každé skupině). V detailech obrázků můžeme pozorovat chování silofunkcí v okolí nulové hypotézy ( $\mu_m$  je blízko 0 a hypotéza  $H_{im}^0$  je zde porušena jen velmi málo).

Na Obrázku 7.1 je vykreslený řez silofunkce pro 3 výběry s 10 pozorováními v každé skupině. V simulaci jsou střední hodnoty pro první dva výběry nulové a třetí se mění. Pro  $m=3$  pracujeme s třemi dílčími testy. K zamítnutí celkové hypotézy  $H^0 : \mu_1 = \mu_2 = \mu_3 = 0$ , dochází i vlivem některých z dílčích hypotéz  $H_{12}^0 : \mu_1 = \mu_2, H_{13}^0 : \mu_1 = \mu_3, H_{23}^0 : \mu_2 = \mu_3$ . Při výpočtu hodnot silofunkce je tedy nutno brát v úvahu všechny dílčí testy. Na Obrázku 7.1 můžeme sledovat celkový průběh silofunkcí. Je to řez silofunkce pro měnící se  $\mu_3$  při  $\mu_1 = 0$  a  $\mu_2 = 0$ . Na Obrázku 7.2 je jeho detail, zobrazující chování funkcí v okolí nulové hypotézy a na Obrázku 7.3 je detail grafu, kde můžeme najít takové hodnoty  $\mu_3$ , které testy zamítají s jistotou. Funkce pro Bonferroniho metodu je skrytá pod funkcí Holmovy metody. Na obrázcích vidíme, že všechny metody, s výjimkou metody BY, jsou srovnatelné.

Na Obrázku 7.4 a Obrázku 7.5 jsou vykresleny silofunkce také pro případ 3 výběrů, ale tentokrát je v každé skupině 40 pozorování. Když se podíváme na průběh funkcí, hned vidíme, že zvýšením počtu pozorování v každé skupině došlo k mnohem rychlejšímu stoupání směrem k jedničce. Jak můžeme vidět na Obrázku 7.6, tak zatímco v případě 10 pozorování testy zamítají s jistotou vychýlení  $\mu_3$  až okolo 2,5, pro 40 pozorování testy zamítají s jistotou vychýlení  $\mu_3$  už okolo 1,2. Silofunkce pro Bonferroniho metodu je opět skrytá pod funkcí Holmovy metody a silofunkce pro Hochbergovu a BH metodu jsou v jejich těsné blízkosti. Empirické hladiny významnosti jednotlivých metod v případě  $n = 10$

a  $n = 40$  jsou v okolí nulové hypotézy srovnatelné. Přesné hodnoty jsou uvedeny v Tabulce 7.1 a v Tabulce 7.2.

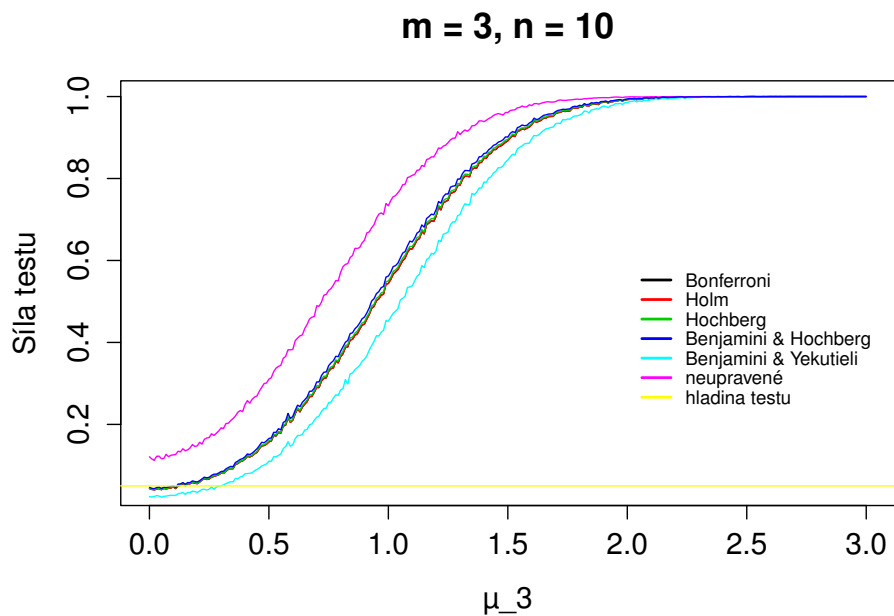
Na dalších obrázcích v pořadí jsou vykresleny silofunkce pro 5 výběrů, nejdříve s 20 pozorováními v každé skupině (Obrázek 7.7, Obrázek 7.8 a Obrázek 7.9) a následně s 50 pozorováními ve skupině (Obrázek 7.10, Obrázek 7.11 a Obrázek 7.12). Na Obrázku 7.13, Obrázku 7.14 a Obrázku 7.15 jsou vykresleny silofunkce ještě pro větší počet skupin a pozorování. Konkrétně jsou to silofunkce v případě 10 výběrů, kde v každém výběru je 50 pozorování. V těchto třech případech spolu metody Bonferroniho, Holmova a Hochbergova (skoro všude) splývají. V Příloze G je pro tento případ uveden kód, který lze snadno upravovat pro libovolné kombinace počtu skupin a počtu pozorování.

Metody jsou navrženy tak, aby snížily hladinu významnosti. Na obrázcích tak můžeme vidět, že empirická hladina významnosti pro nepravené  $p$ -hodnoty (a velká  $m$ ) je výrazně nad hodnotou 0,05 a pro upravené  $p$ -hodnoty (s výjimkou metody BY) se pak k hodnotě 0,05 přibližuje. Kromě celkového průběhu silofunkcí je pak také zajímavé, jak se tyto funkce chovají v okolí nulové hypotézy (v okolí nuly) a kdy metody zamítají nulovou hypotézu  $H_{im}^0$  již s jistotou. To můžeme vidět v detailech jednotlivých obrázků.

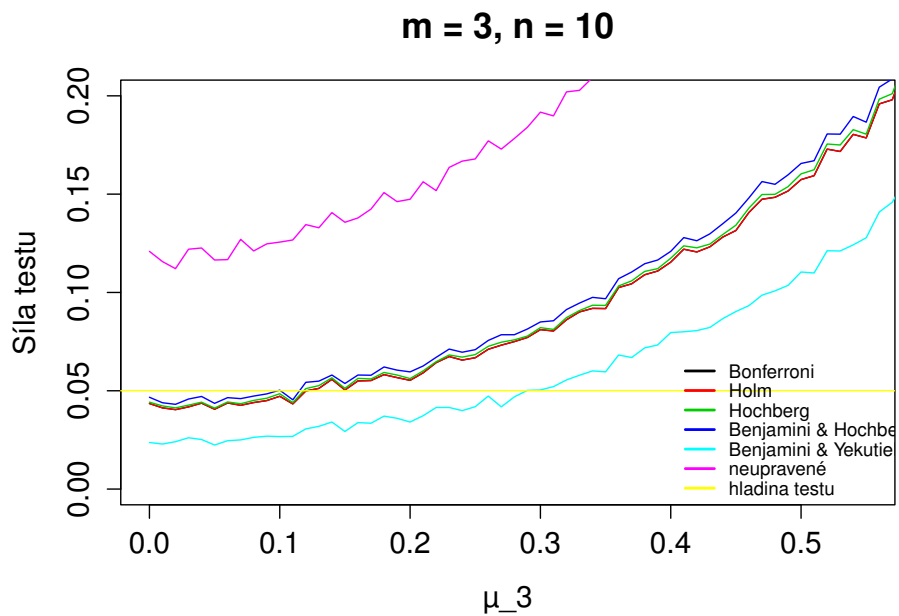
Na obrázcích je jasně vidět průběh silofunkcí a můžeme tak zodpovědět otázky, které jsme zmínili na začátku kapitoly. S rostoucím počtem skupin se zvětšují vzdálenosti mezi křivkami a křivky tak méně splývají. Obecně tak můžeme říct, že s rostoucím počtem skupin se zvětšují rozdíly mezi metodami. Při pohledu na obrázky je také patrné, že pro větší počty pozorování silofunkce rychleji stoupá k 1.

Na obrázcích je občas vidět drobný pokles síly testu. Tento jev rozhodně není správný, ale lze ho přičíst malému počtu iterací. S větším počtem replikací by toto kolísání bylo odstraněno. Nicméně vzhledem k časové náročnosti jsou grafy ponechány v současné podobě, tak jak byly získány z 10 tisíce opakování.

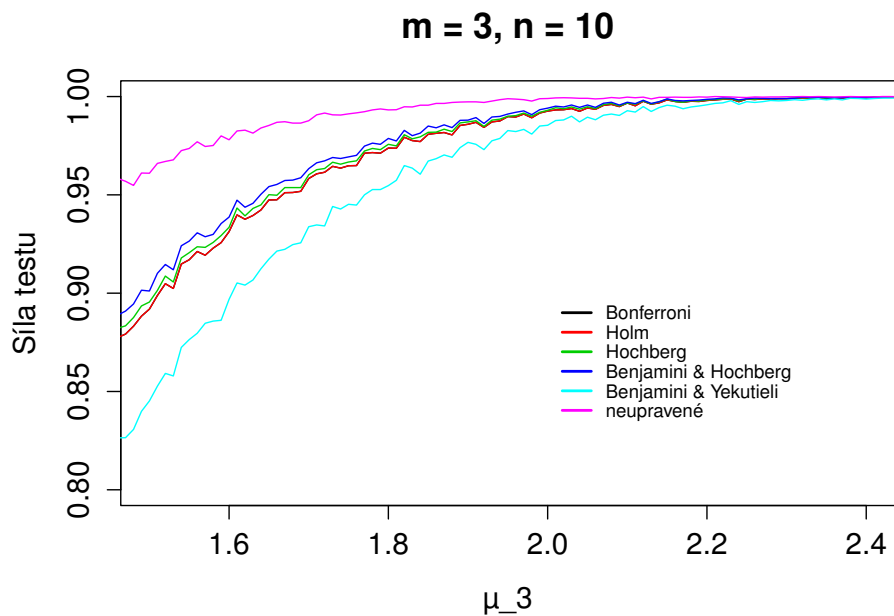




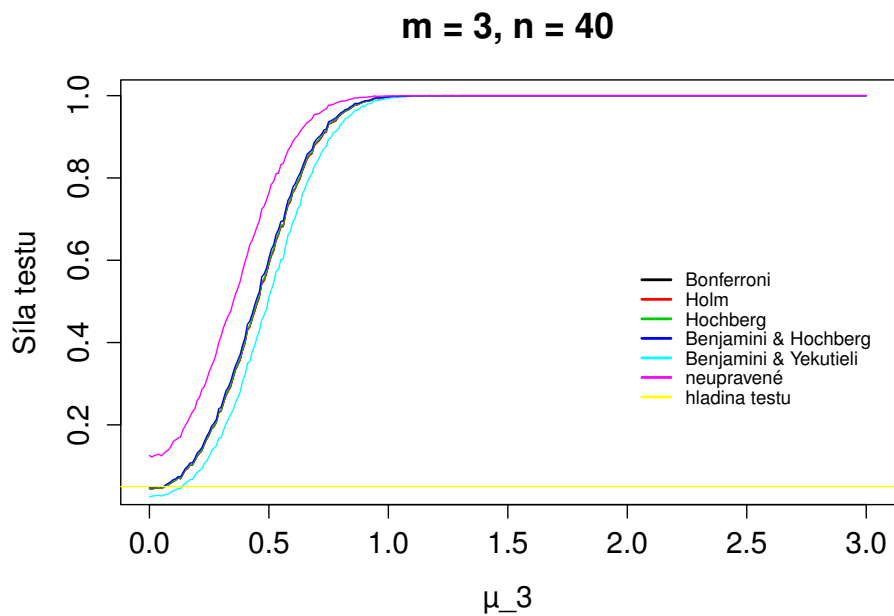
Obrázek 7.1: Řez silofunkce pro měnící se  $\mu_3$  při  $\mu_1 = 0$  a  $\mu_2 = 0$  (3 výběry, 10 pozorování ve skupině).



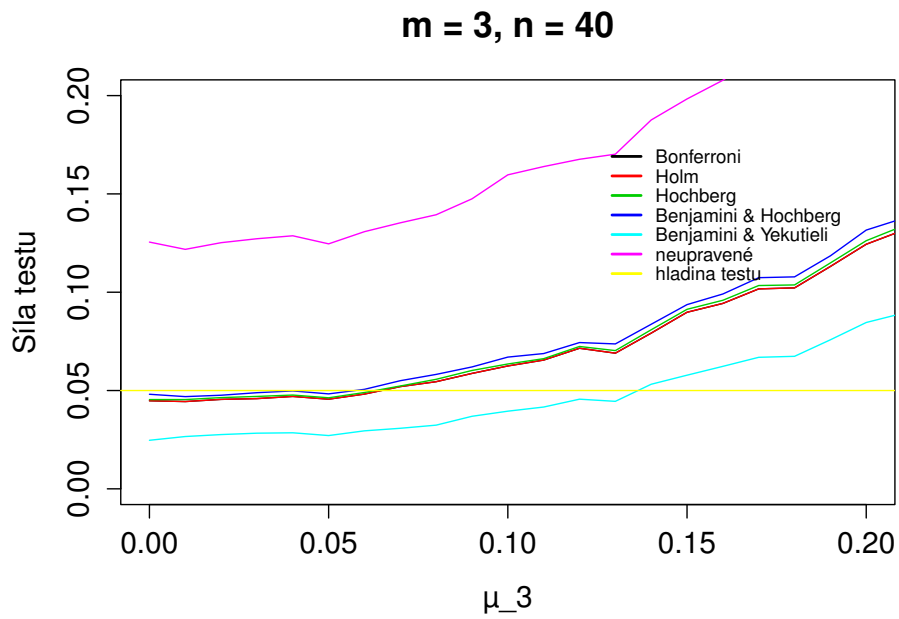
Obrázek 7.2: Detail řezu silofunkce pro měnící se  $\mu_3$  při  $\mu_1 = 0$  a  $\mu_2 = 0$  (3 výběry, 10 pozorování ve skupině).



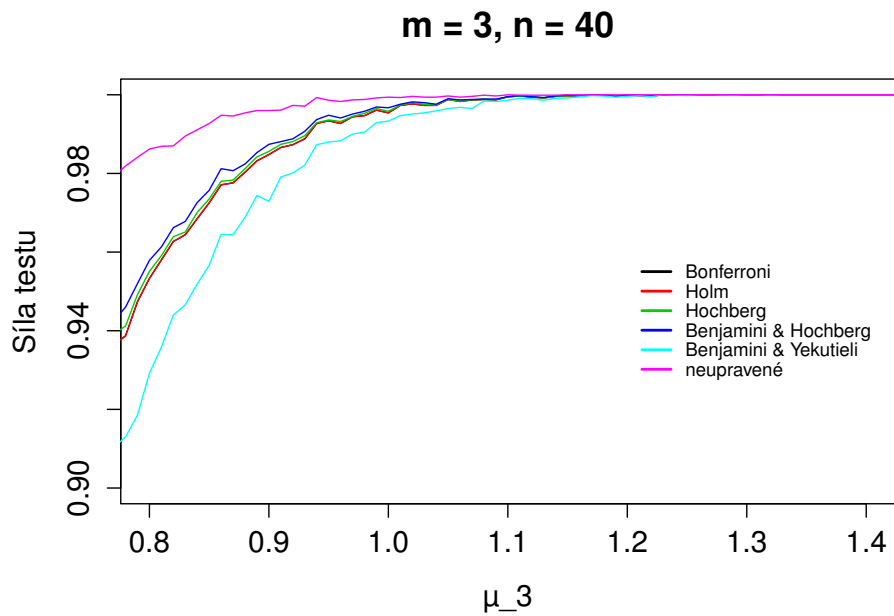
Obrázek 7.3: Detail řezu silofunkce pro měnící se  $\mu_3$  při  $\mu_1 = 0$  a  $\mu_2 = 0$  (3 výběry, 10 pozorování ve skupině).



Obrázek 7.4: Řez silofunkce pro měnící se  $\mu_3$  při  $\mu_1 = 0$  a  $\mu_2 = 0$  (3 výběry, 40 pozorování ve skupině).



Obrázek 7.5: Detail řezu silofunkce pro měnící se  $\mu_3$  při  $\mu_1 = 0$  a  $\mu_2 = 0$  (3 výběry, 40 pozorování ve skupině).



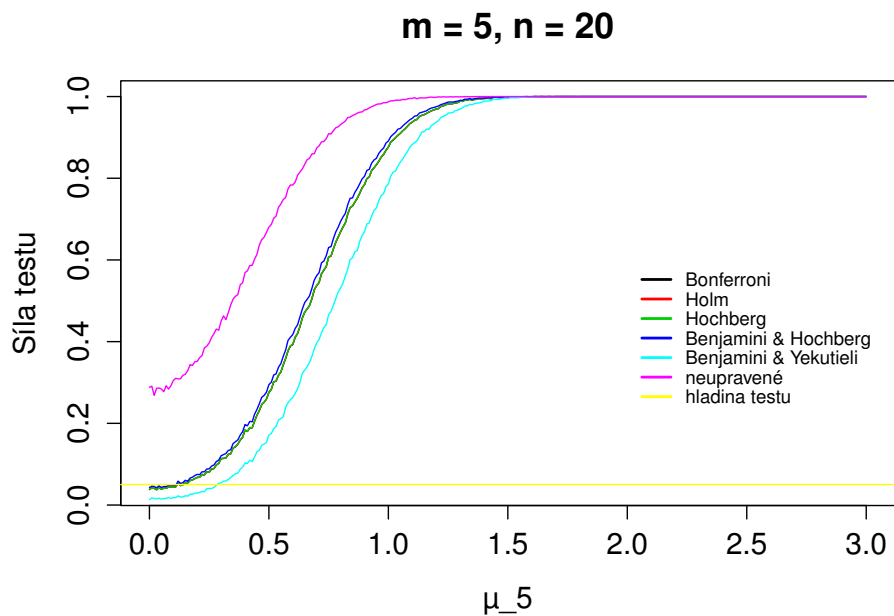
Obrázek 7.6: Detail řezu silofunkce pro měnící se  $\mu_3$  při  $\mu_1 = 0$  a  $\mu_2 = 0$  (3 výběry, 40 pozorování ve skupině).

$\mu_3$	Bonferroni	Holm	Hochberg	BH	BY	neupravené
0,00	0,0435	0,0435	0,0442	0,0467	0,0237	0,1209
0,01	0,0413	0,0413	0,0423	0,0439	0,0229	0,1157
0,02	0,0404	0,0404	0,0413	0,0431	0,0241	0,1121
0,03	0,0418	0,0418	0,0427	0,0458	0,0261	0,1220
0,04	0,0437	0,0437	0,0442	0,0471	0,0252	0,1226
0,05	0,0406	0,0406	0,0412	0,0436	0,0224	0,1165
0,06	0,0438	0,0438	0,0443	0,0465	0,0246	0,1167
0,07	0,0426	0,0426	0,0434	0,0460	0,0250	0,1270
0,08	0,0440	0,0440	0,0451	0,0473	0,0263	0,1211
0,09	0,0450	0,0450	0,0462	0,0484	0,0269	0,1247
0,10	0,0472	0,0472	0,0485	0,0503	0,0267	0,1256

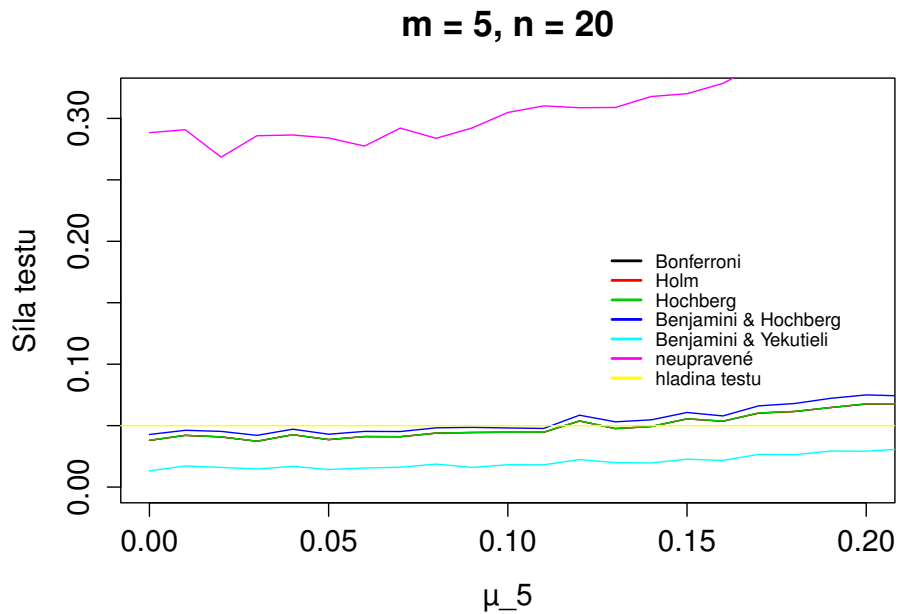
Tabulka 7.1: Hodnoty silofunkcí testu ( $\mu_1 = \mu_2 = \mu_3$ ) ve vybraných bodech  $\mu_3$  při  $\mu_1 = 0$  a  $\mu_2 = 0$  a pro  $m = 3$  a  $n = 10$ .

$\mu_3$	Bonferroni	Holm	Hochberg	BH	BY	neupravené
0,00	0,0448	0,0448	0,0454	0,0481	0,0247	0,1255
0,01	0,0444	0,0444	0,0454	0,0469	0,0266	0,1218
0,02	0,0455	0,0455	0,0464	0,0476	0,0276	0,1252
0,03	0,0459	0,0459	0,0470	0,0489	0,0283	0,1272
0,04	0,0470	0,0470	0,0477	0,0498	0,0285	0,1287
0,05	0,0456	0,0456	0,0463	0,0483	0,0271	0,1246
0,06	0,0482	0,0482	0,0490	0,0506	0,0295	0,1308
0,07	0,0520	0,0520	0,0523	0,0550	0,0308	0,1353
0,08	0,0545	0,0545	0,0557	0,0582	0,0324	0,1394
0,09	0,0587	0,0587	0,0602	0,0620	0,0369	0,1475
0,10	0,0625	0,0625	0,0635	0,0670	0,0395	0,1597

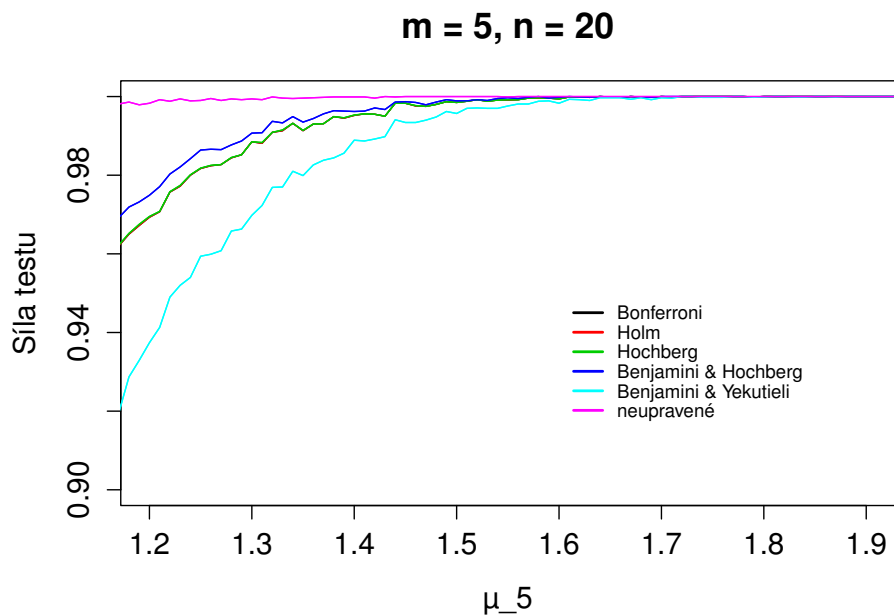
Tabulka 7.2: Hodnoty silofunkcí testu ( $\mu_1 = \mu_2 = \mu_3$ ) ve vybraných bodech  $\mu_3$  při  $\mu_1 = 0$  a  $\mu_2 = 0$  a pro  $m = 3$  a  $n = 40$ .



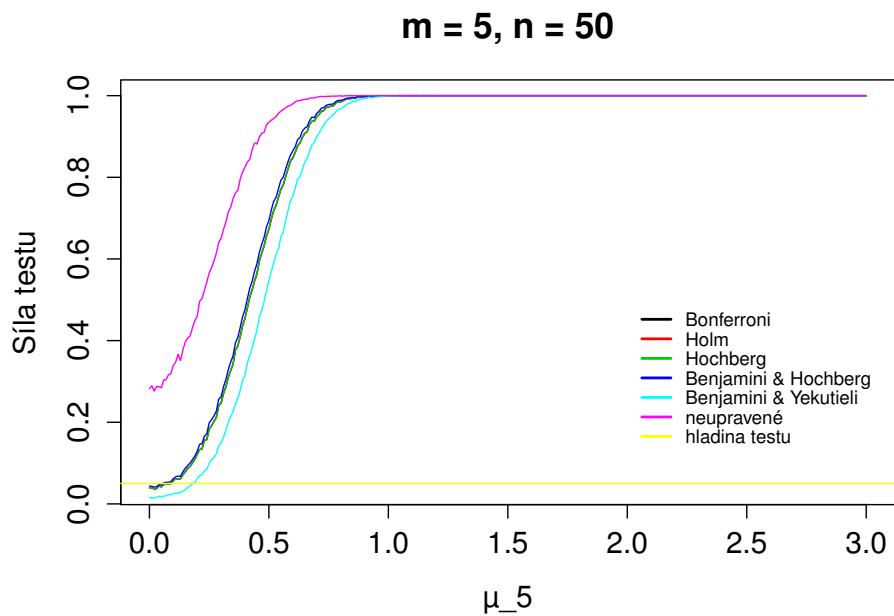
Obrázek 7.7: Řez silofunkce pro měnící se  $\mu_5$  při  $\mu_1 = 0, \dots, \mu_4 = 0$  (5 výběrů, 20 pozorování ve skupině).



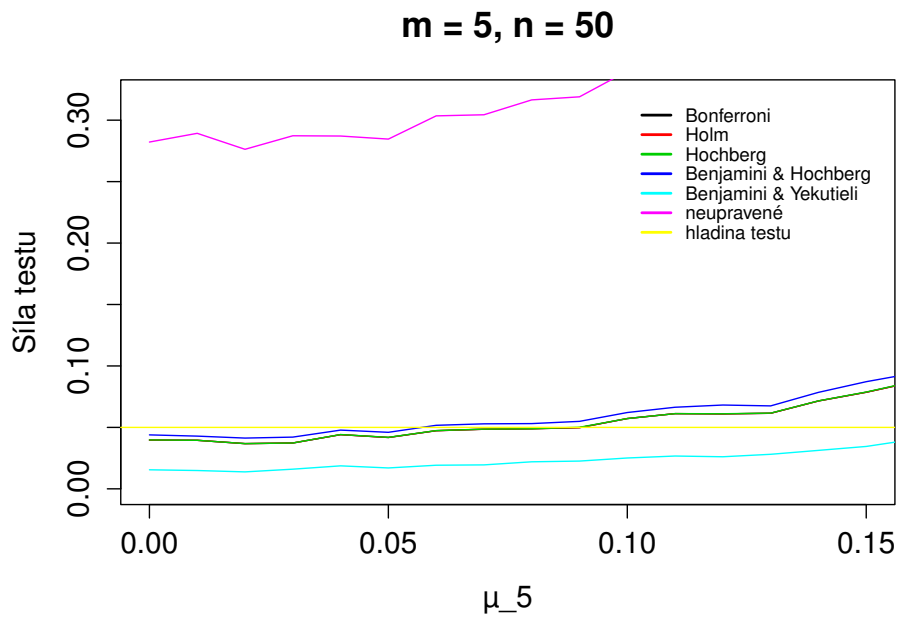
Obrázek 7.8: Detail řezu silofunkce pro měnící se  $\mu_5$  při  $\mu_1 = 0, \dots, \mu_4 = 0$  (5 výběrů, 20 pozorování ve skupině).



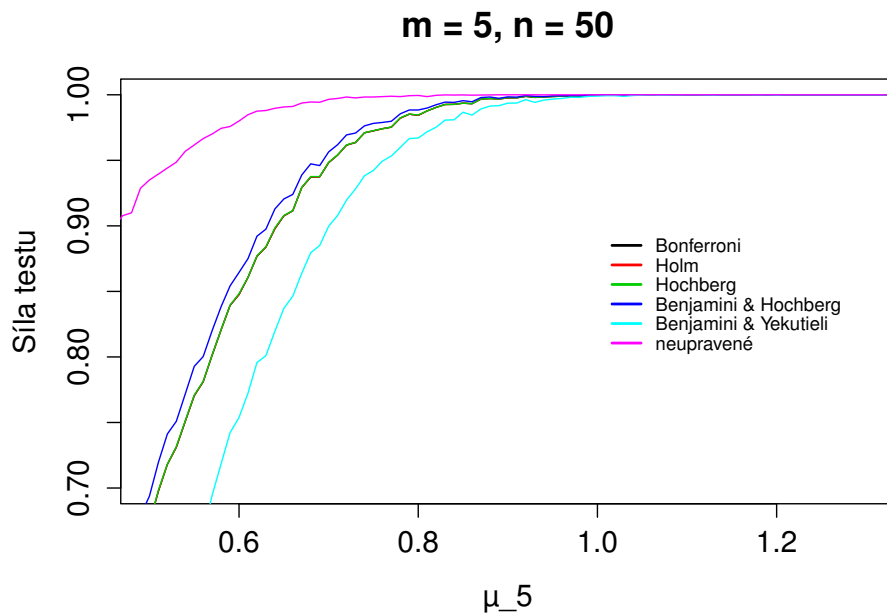
Obrázek 7.9: Detail řezu silofunkce pro měnící se  $\mu_5$  při  $\mu_1 = 0, \dots, \mu_4 = 0$  (5 výběrů, 20 pozorování ve skupině).



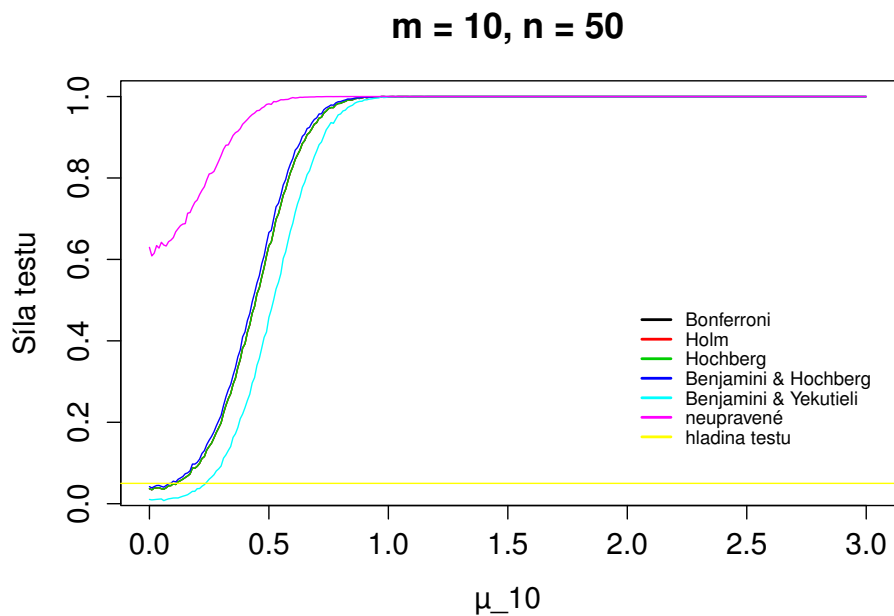
Obrázek 7.10: Řez silofunkce pro měnící se  $\mu_5$  při  $\mu_1 = 0, \dots, \mu_4 = 0$  (5 výběrů, 50 pozorování ve skupině).



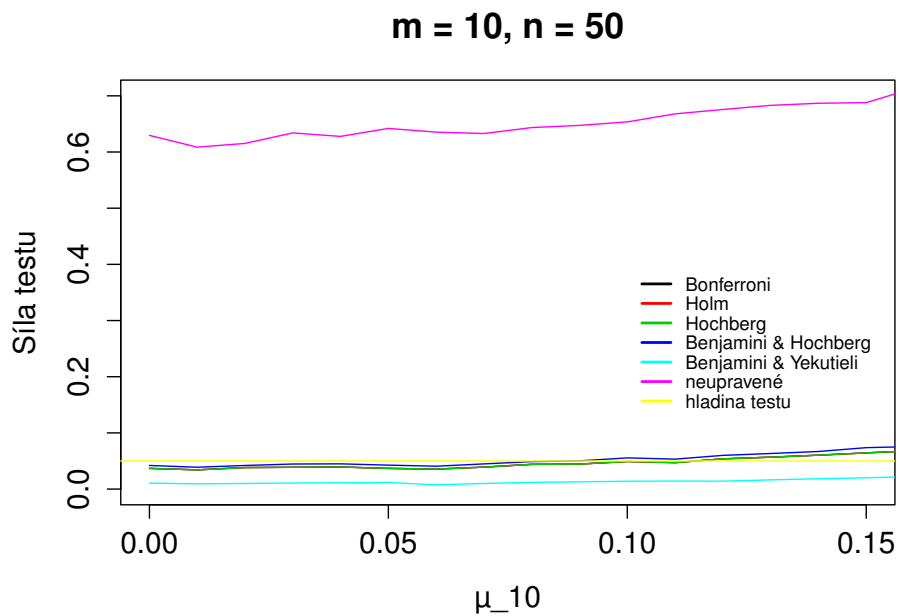
Obrázek 7.11: Detail řezu silofunkce pro měnící se  $\mu_5$  při  $\mu_1 = 0, \dots, \mu_4 = 0$  (5 výběrů, 50 pozorování ve skupině).



Obrázek 7.12: Detail řezu silofunkce pro měnící se  $\mu_5$  při  $\mu_1 = 0, \dots, \mu_4 = 0$  (5 výběrů, 50 pozorování ve skupině).

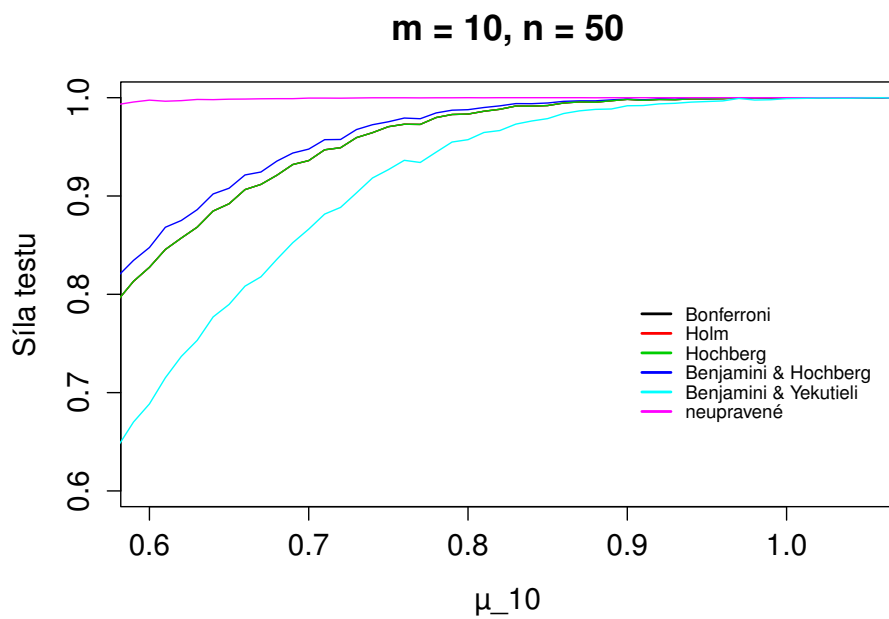


Obrázek 7.13: Řez silofunkce pro měnící se  $\mu_{10}$  při  $\mu_1 = 0, \dots, \mu_9 = 0$  (10 výběrů, 50 pozorování ve skupině).



Obrázek 7.14: Detail řezu silofunkce pro měnící se  $\mu_{10}$  při  $\mu_1 = 0, \dots, \mu_9 = 0$  (10 výběrů, 50 pozorování ve skupině).





Obrázek 7.15: Detail řezu silofunkce pro měnící se  $\mu_{10}$  při  $\mu_1 = 0, \dots, \mu_9 = 0$  (10 výběrů, 50 pozorování ve skupině).

# Závěr

Cílem práce bylo zpracovat teoretický základ vybraných metod mnohonásobného porovnávání, aplikovat metody na reálná data a provést simulační studii. Cíl práce byl dle mého úsudku splněn. Myslím, že by tato práce mohla poskytnout čtenáři ucelený přehled o metodách a zároveň ukázat využití metod v praxi. Snažila jsem se práci napsat tak, aby byla i přes poměrně obsáhlou teorii čtivá a snadno pochopitelná.

Dalším (neoficiálním) cílem bylo naučit se lépe pracovat s anglickou literaturou a se softwarem R. Téměř veškerá dostupná literatura byla v anglickém jazyce a občas nebylo snadné si vše přeložit a následně pochopit. Před samotným začátkem psaní práce jsem navíc měla jen základní znalosti pro práci v softwaru R, takže ani to nebylo snadné. Nakonec jsem se ale za pomoci nejen různých překladačů a příruček, ale hlavně díky pomoci paní doc. RNDr. Evy Fišerové, Ph.D., pročetla a propsala k (snad) zdárnému výsledku.

Nezbývá, než na závěr ještě jednou poděkovat všem, kteří mi pomáhali a podporovali mě, poskytovali mi cenné rady a konzultace, čímž také přispěli k dokončení této diplomové práce. V neposlední řadě bych chtěla tímto poděkovat také mé rodině, která mě podporovala po dobu nejen celého studia, ale po celý můj život. Největší dík tedy patří mým rodičům, kteří mi umožnili se vzdělávat a vždy mi byli oporou.

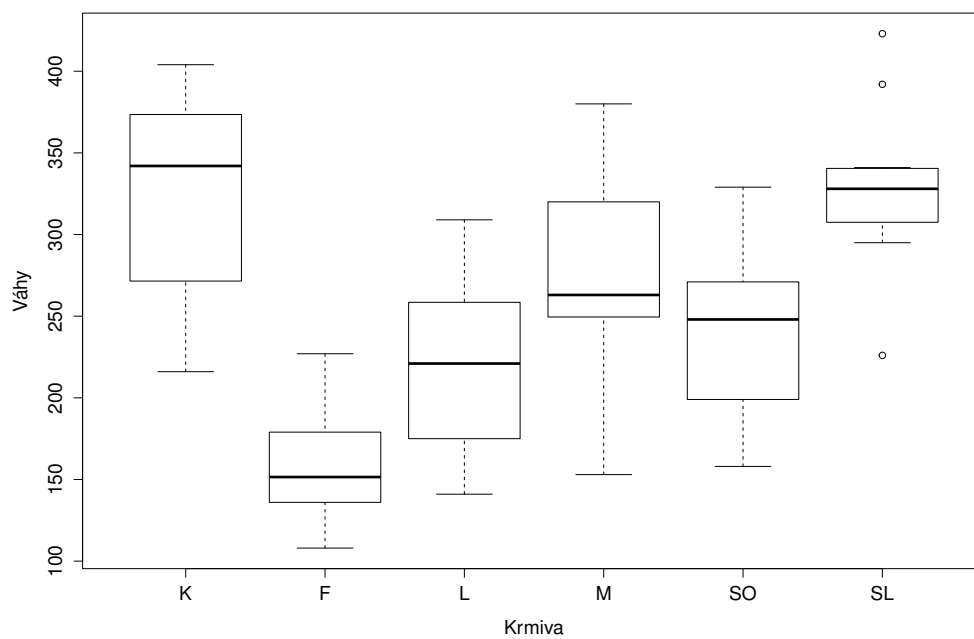
# Literatura

- [1] Aickin, M., Gensler, H.: *Adjusting for Multiple Testing When Reporting Research Results: The Bonferroni vs Holm Methods*. American Journal of Public Health, vol. 86, no. 5, 1996.
- [2] Bedáňová, I., Večerek, V.: *Základy statistiky pro studující veterinární medicíny a farmacie*. Brno: Veterinární a farmaceutická univerzita, 2007.
- [3] Benjamini, Y., Hochberg, Y.: *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing*. Journal of the Royal Statistical Society, series B, vol. 57, no. 1, 1995, 289-300.
- [4] Bretz F., Hothorn T., Westfall P.: *Multiple Comparisons Using R*, CRC Press, 2011.
- [5] Budíková, M.: *Statistika II: Distanční studijní opora*, 1. vydání. Brno: Masarykova univerzita, 2006. ISBN 80-210-4105-6.
- [6] Dubjaková, E., *Diplomová práce: Metody mnohonásobného porovnávání pro jednoduché třídění*, Brno: Masarykova Univerzita, 2009.
- [7] Dudoit, S., Shaffer, J. P., Boldrick J. C.: *Multiple Hypothesis Testing in Microarray Experiments*. Institute of Mathematical Statistics, vol. 18, no. 1, 2003, 71-103.
- [8] Hochberg, Y.: *A sharper Bonferroni procedure for multiple tests of significance*. Biometrika, vol. 75, no. 4, 1988.
- [9] Hron K., Kunderová P.: *Základy počtu pravděpodobnosti a matematické statistiky*, 1. vydání. Olomouc: Vydavatelství Univerzity Palackého, 2013, 330 s. ISBN 978-80-244-3396-7.
- [10] Kadavý, M., *Bakalářská práce: Mnohonásobné testování*, Praha: Univerzita Karlova, 2007.
- [11] Katina, S., Králík, M., Hupková, A.: *Aplikovaná statistická inferencia I*. Brno: MUNI Press, 1. vydání, 2015.

- [12] Konečná, K., Koláček, J.: *Jak pracovat s jazykem R* [online], dostupné z: [http://www.math.muni.cz/xkonecn3/vyuka/MUNI/VMS/navod\\_R.pdf](http://www.math.muni.cz/xkonecn3/vyuka/MUNI/VMS/navod_R.pdf), [citováno 17. 9. 2015].
- [13] Meloun, M., Militký, J.: *Statistická analýza experimentálních dat*. Praha: Academia, 2. vydání, 953 s., 2004. ISBN 80-200-1254-0.
- [14] Shaffer, J. P.: *Multiple hypothesis testing*. Annual Review of Psychology, vol. 46, 1995, 561-584.
- [15] Varmuza, K., Filzmoser, P.: *Introduction to Multivariate Statistical Analysis in Chemometrics*. Boca Raton, Fla.: CRC Press, 13, 321 s., 2009. ISBN 1-4200-5947-5.
- [16] *Bonferroni Inequalities*. *Wofram Math World* [online], dostupné z: <http://mathworld.wolfram.com/BonferroniInequalities.html>, [citováno 9. 6. 2016].

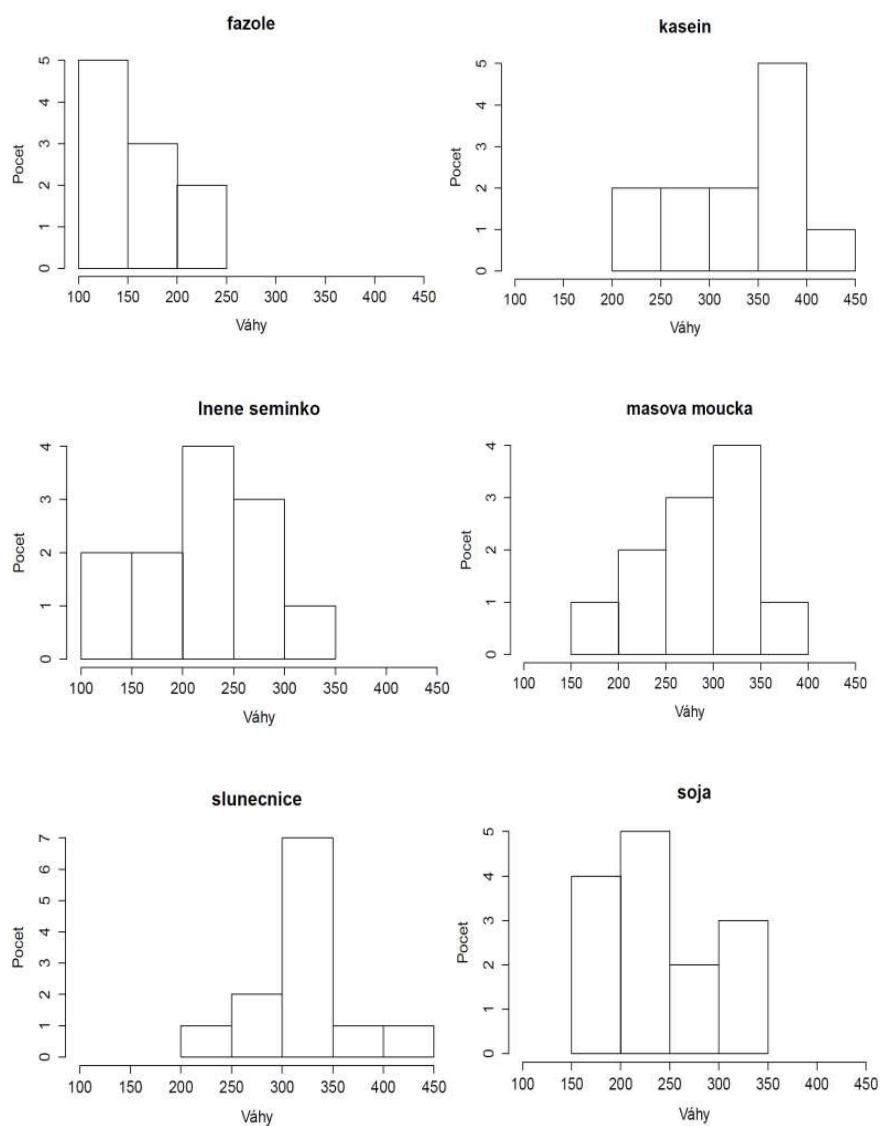
# Příloha A

Boxploty pro váhy kuřat (v gramech) rozdělených do skupin dle krmiva (data chickwts). Symboly značí K - kasein, F - fazole, L - lněné semínko, M - masová moučka, SO - soja, SL - slunečnice.



# Příloha B

Histogramy pro váhy kuřat rozdělených do skupin dle krmiva (data chickwts).



# Příloha C

Data z pŕzkumu - vĕk, vŕška, vĀha (v kilogramech) a hodiny vĕnovanĕ tŕdnĕ sportu (data mladŕ muŕi).

Muŕ	Vĕk	Vŕška	VĀha	Hod.	Muŕ	Vĕk	Vŕška	VĀha	Hod.
1	35	187	83	0	39	31	195	87	6
2	30	182	83	0	40	25	183	93	5
3	28	180	100	0	41	29	182	87	6
4	28	180	75	0	42	23	183	84	6
5	29	185	92	0	43	31	187	95	4
6	31	180	83	0	44	27	183	100	5
7	31	194	83	0	45	34	184	86	6
8	31	193	92	0	46	26	189	95	8
9	28	182	76	0	47	30	180	87	7
10	27	180	86	0	48	30	193	100	8
11	28	188	64	0	49	26	190	94	8
12	35	182	83	0	50	25	192	86	8
13	29	183	75	0	51	29	187	81	7
14	25	180	83	0	52	28	185	78	8
15	28	181	85	0	53	26	187	87	7
16	30	191	90	2	54	35	180	94	7
17	28	181	83	1	55	25	186	82	8
18	29	193	83	3	56	31	180	80	8
19	28	187	100	3	57	28	196	98	7
20	27	193	85	3	58	24	183	81	8
21	24	189	100	2	59	27	185	82	9
22	26	188	90	1	60	27	181	84	9
23	30	180	85	3	61	27	182	72	10
24	26	186	68	3	62	24	195	92	16
25	30	185	82	2	63	26	180	78	10
26	29	180	96	2	64	30	181	81	12
27	28	180	94	3	65	29	188	89	15

Pokračování:

Muž	Věk	Výška	Váha	Hod.	Muž	Věk	Výška	Váha	Hod.
28	27	188	81	2	66	24	195	110	17
29	34	180	74	3	67	23	184	97	12
30	30	185	96	2	68	28	185	80	14
31	27	180	90	4	69	24	184	84	20
32	29	187	90	4	70	33	182	77	10
33	26	188	85	5	71	21	192	90	12
34	26	190	83	6	72	35	186	85	10
35	33	184	72	4	73	28	187	95	12
36	24	181	76	6	74	28	188	84	12
37	31	195	108	6	75	26	192	95	11
38	30	184	95	4					



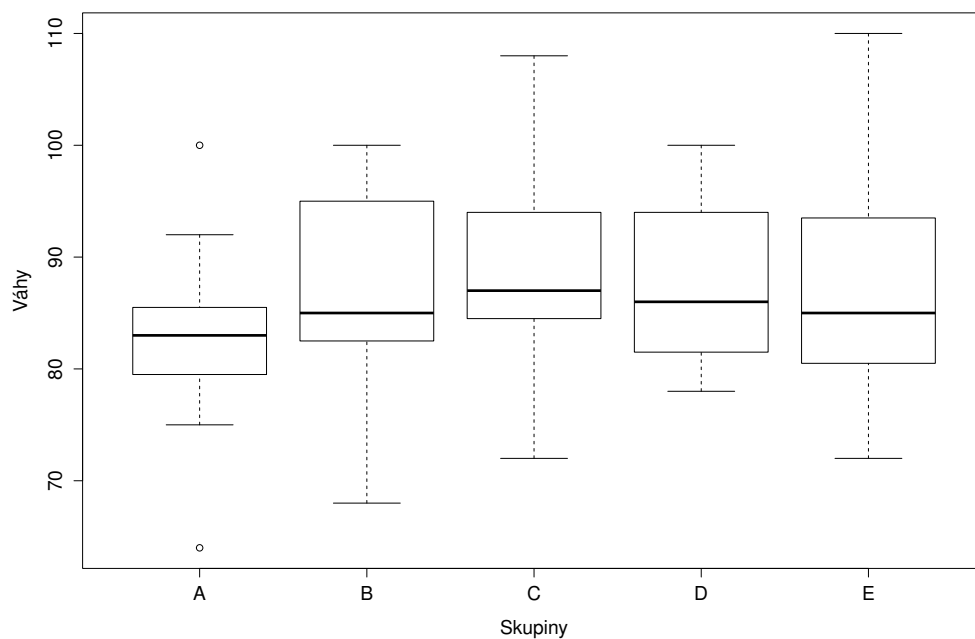
# Příloha D

Charakteristiky pro jednotlivé skupiny (data mladí muži). Muži jsou rozděleni do skupin podle toho, kolik hodin týdně sportují:  $A = 0$ ,  $B = (0, 3)$ ,  $C = (3, 6)$ ,  $D = (6, 9)$ ,  $E = (9, 168)$ .

	A	B	C	D	E
Minimum	64	68	72	78	72
Maximum	100	100	108	100	110
Medián	83	85	87	86	85
Průměr	82,87	87,13	88,73	87,27	87,27
Směrodatná odchylka	8,46	9,25	8,96	7,14	9,66

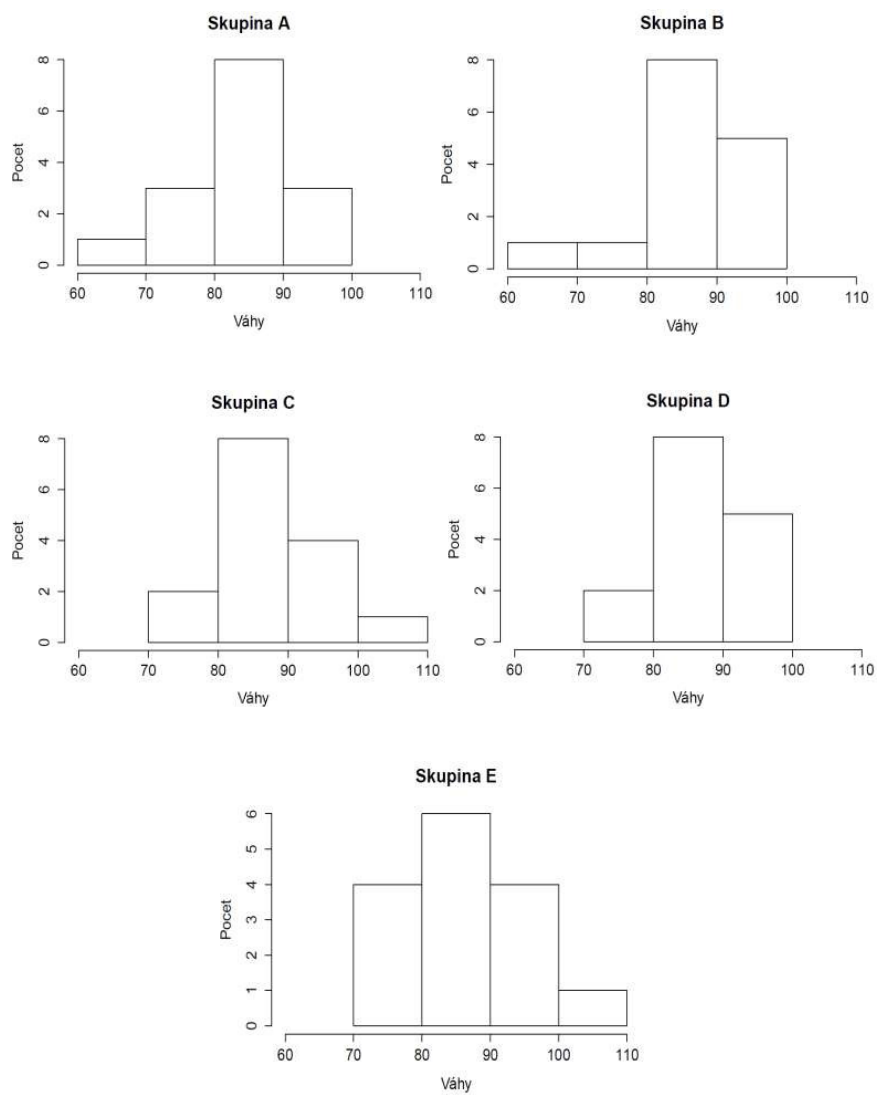
# Příloha E

Boxploty pro váhy mužů rozdělených do skupin (data mladí muži). Data jsou rozdělena do skupin podle hodin týdně strávených sportem: A = 0, B = (0, 3), C = (3, 6), D = (6, 9), E = (9, 168).



# Příloha F

Histogramy pro váhy mužů rozdělených do skupin (data mladí muži).



# Příloha G

Vzorový kód pro simulace v případě 10 skupin a 50 pozorování.

```
#set.seed(...) ZAŘÍDÍ GENEROVÁNÍ STEJNÝCH NÁHODNÝCH ČÍSEL
set.seed(12)
#POČET ITERACÍ
pocet <- 10000
metody=c("bonferroni","holm","hochberg","BH","BY","none")
#POČET SKUPIN
m <- 10
#POČET POZOROVANI VE SKUPINE
n <- 50
sigma <- 1
#ROZSAH A DELKA KROKU
mean <- seq(from=0,to=3,by=0.01)
sila <- matrix(0, nrow=length(mean), ncol=length(metody))
colnames(sila) <- metody

for(mean_i in 1:length(mean)){
  zamitnuto <- matrix(0,nrow=pocet,ncol=length(metody))
  for(iterace in 1:pocet)
  {data <- matrix(0, nrow=n*m, ncol=2)
  data[,1] <- rnorm(n*m, mean=c(rep(0, (m-1)*n),
rep(mean[mean_i],n)),sigma)
  data[,2] <- rep(c(1:m), each=n)
  for(j in 1:length(metody))
  {mm <- pairwise.t.test(data[,1],data[,2],
  p.adjust.method=metody[j],
  pool.sd=T, alternative = "two.sided")$p.value
  if(length(which(mm<0.05))>0)
  zamitnuto[iterace,j] <- 1}}
  pocet_zamitnuti <- rep(0, length(metody))
  for(i in 1:length(metody) )
  pocet_zamitnuti[i] <- sum(zamitnuto[,i])
  #sila testu
  sila[mean_i, ] <- pocet_zamitnuti/pocet}

silofunkce <- cbind(mean,sila)

#GRAF (parametry lze menit)
plot(silofunkce[,2]~silofunkce[,1], type="l",
  main="m = 10, n = 50", xlab="  $\mu_{10}$ ", ylab="Sila testu")
lines(silofunkce[,3]~silofunkce[,1], type="l", col=2)
lines(silofunkce[,4]~silofunkce[,1], type="l", col=3)
lines(silofunkce[,5]~silofunkce[,1], type="l", col=4)
lines(silofunkce[,6]~silofunkce[,1], type="l", col=5)
lines(silofunkce[,7]~silofunkce[,1], type="l", col=6)
abline(b=0, h=0.05, type="l", col=7)
legend(x=2, y=0.4, col=c(1:7), legend=c("Bonferroni",
  "Holm","Hochberg","Benjamini & Hochberg",
  "Benjamini & Yekutieli","neupravené","hladina testu"),
  cex=0.65,lwd=2, bty="n")
```