



## **Bakalářská práce**

# **Statistické zpracování dat zabývající se odhalením dyspraxie studentů.**

*Studijní program:*

B0613A140005 Informační technologie

*Studijní obor:*

Aplikovaná informatika

*Autor práce:*

**David Bálik**

*Vedoucí práce:*

Ing. Josef Chudoba, Ph.D.

Ústav nových technologií a aplikované informatiky

Liberec 2023



## Zadání bakalářské práce

# Statistické zpracování dat zabývající se odhalením dyspraxie studentů.

<i>Jméno a příjmení:</i>	<b>David Bálik</b>
<i>Osobní číslo:</i>	M20000032
<i>Studijní program:</i>	B0613A140005 Informační technologie
<i>Specializace:</i>	Aplikovaná informatika
<i>Zadávací katedra:</i>	Ústav nových technologií a aplikované informatiky
<i>Akademický rok:</i>	2022/2023

### Zásady pro vypracování:

V loňském roce proběhlo šetření více než pětiset studentů z některých základních škol v Libereckém kraji. Cílem zkoumání bylo odhalit dyspraxii studentů. Studenti provedli praktické úlohy, které byly zaznamenány do pracovního listu a následně byl proveden ručně přenos dat do tabulky.

Student v rámci bakalářské práce provede:

1. programové zjištění chyb a neúplnosti dat vzniklé přepisem,
2. určení, zda jednotlivé ročníky studentů nevykazují mezi sebou významné statistické rozdíly,
3. určení, zda jednotlivá pohlaví nevykazují mezi sebou významné statistické rozdíly,
4. možné zjednodušení seznamu úkolů (odstraněním úkolů korelujících s jinými).

*Rozsah grafických prací:* dle potřeby dokumentace  
*Rozsah pracovní zprávy:* 30 – 40 stran  
*Forma zpracování práce:* tištěná/elektronická  
*Jazyk práce:* Čeština

### **Seznam odborné literatury:**

- [1] HENDL, Jan. *Přehled statistických metod zpracování dat: analýza a metaanalýza dat*. Praha: Portál, 2004. ISBN 80-7178-820-1.
- [2] MELOUN, Milan a Jiří MILITKÝ. *Kompendium statistického zpracování dat: metody a řešené úlohy*. Vyd. 2., přeprac. a rozš. Praha: Academia, 2006. ISBN 80-200-1396-2.
- [3] ANDĚL, Jiří. *Statistické metody*. 4., upr. vyd. Praha: Matfyzpress, 2007. ISBN 80-7378-003-8.
- [4] VANDERPLAS, Jacob T. *Python data science handbook: essential tools for working with data*. Sebastopol, CA: O'Reilly Media, 2016. ISBN 978-1-491-91205-8.

*Vedoucí práce:* Ing. Josef Chudoba, Ph.D.  
Ústav nových technologií a aplikované informatiky

*Datum zadání práce:* 12. října 2022  
*Předpokládaný termín odevzdání:* 22. května 2023

prof. Ing. Zdeněk Plíva, Ph.D.  
děkan

L.S.

Ing. Josef Novák, Ph.D.  
vedoucí ústavu

V Liberci dne 19. října 2022

## Prohlášení

Prohlašuji, že svou bakalářskou práci jsem vypracoval samostatně jako původní dílo s použitím uvedené literatury a na základě konzultací s vedoucím mé bakalářské práce a konzultantem.

Jsem si vědom toho, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu Technické univerzity v Liberci.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti Technickou univerzitu v Liberci; v tomto případě má Technická univerzita v Liberci právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Současně čestně prohlašuji, že text elektronické podoby práce vložený do IS/STAG se shoduje s textem tištěné podoby práce.

Beru na vědomí, že má bakalářská práce bude zveřejněna Technickou univerzitou v Liberci v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů.

Jsem si vědom následků, které podle zákona o vysokých školách mohou vyplývat z porušení tohoto prohlášení.

# Statistické zpracování dat zabývající se odhalením dyspraxie studentů

## Abstrakt

Tato bakalářská práce řeší statistické zpracování dat zabývající se odhalením dyspraxie studentů pomocí Bruininks-Oseretsky testu motorické zdatnosti. Sběr dat proběhl v některých základních školách a školkách Libereckého kraje na 665 respondentech ve věku mezi 4 roky a 12 lety. Ze získaných dat respondentů byly odstraněny chyby, a tato data byla rozdělena do 11 věkových skupin.

V praktické části se práce zabývá statistickou analýzou získaných dat a vyhodnocuje statistické rozdíly na základě věku, a poté na základě pohlaví a věku. V posledním kroku jsou otestovány jednotlivé prováděné úkoly Bruininks-Oseretsky testu a je navrženo možné zjednodušení tohoto seznamu úkolů a to zjištěním korelace.

**Klíčová slova:** dyspraxie, Bruininks-Oseretsky test, statistické zpracování dat, ANOVA, Kruskal-Wallisův test, korelace, Liberecký kraj

# Statistical processing of data dealing with the detection of students dyspraxia

## Abstract

This bachelor thesis deals with statistical data processing dealing with the detection of dyspraxia in students using the Bruininks-Oseretsky Test of Motor Proficiency. Data collection was conducted in some primary schools and kindergartens of the Liberec region on 665 respondents aged between 4 and 12 years. Errors were removed from the data collected from the respondents, and the data were divided into 11 age groups.

The practical part of the thesis deals with the statistical analysis of the obtained data and evaluates the statistical differences based on age, and then based on gender and age. In the last step, the different tasks performed in the Bruininks-Oseretsky test are tested and a possible simplification of this list of tasks is proposed by establishing correlations.

**Keywords:** dyspraxia, Bruininks-Oseretsky test, statistical data processing, ANOVA, Kruskal-Wallis test, correlation, Liberec region

## Poděkování

Rád bych poděkoval panu Ing. Josefu Chudobovi, Ph.D. za odborné rady a vedení práce. Především mu ale patří dík za ochotu a trpělivost, kterou mi v průběhu práce věnoval.

Práce byla podpořena projektem TA ČR Éta TL03000221 s názvem Diagnostika motorické kompetence v pedagogické praxi: identifikace dysprakticky ohrožených dětí mladšího školního věku.

# Obsah

<b>1</b>	<b>Úvod</b>	<b>12</b>
<b>2</b>	<b>Vstupní data</b>	<b>13</b>
2.1	Charakteristika dat . . . . .	13
2.2	Výpočet výsledku testu . . . . .	14
<b>3</b>	<b>Filtrace a třídění dat</b>	<b>16</b>
3.1	Kvalita dat . . . . .	16
3.2	Filtry . . . . .	17
3.3	Třídění dat . . . . .	19
<b>4</b>	<b>Statistická analýza dat</b>	<b>21</b>
4.1	Analýza dat na základě věku . . . . .	21
	Teorie . . . . .	23
	Statistické vyhodnocení . . . . .	25
4.2	Analýza dat na základě pohlaví a věku . . . . .	36
	Teorie . . . . .	37
	Statistické vyhodnocení . . . . .	38
4.3	Korelace mezi jednotlivými úkoly Bruininks-Oseretsky testu . . . . .	48
	Teorie . . . . .	49
	Statistické vyhodnocení . . . . .	50
<b>5</b>	<b>Závěr</b>	<b>56</b>
	<b>Seznam použité literatury</b>	<b>58</b>
<b>A</b>	<b>Přílohy</b>	<b>59</b>



## Seznam obrázků

2.1	Vstupní tabulka . . . . .	14
4.1	Schéma prováděných jednofaktorových testů pro analýzu dat na základě věku . . . . .	22
4.2	Porovnání distribučních funkcí vstupních dat a normálního rozdělení . . . . .	24
4.3	Instalace matlab engine pro Python . . . . .	24
4.4	Mapy významných rozdílů středních hodnot pro sloupce BI a BJ vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	27
4.5	Mapa významných rozdílů středních hodnot pro sloupec BK vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	28
4.6	Mapa významných rozdílů středních hodnot pro sloupec BM vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	30
4.7	Mapa významných rozdílů středních hodnot pro sloupec BO vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	31
4.8	Mapy významných rozdílů středních hodnot pro sloupce BP a BQ vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	32
4.9	Mapy významných rozdílů středních hodnot pro sloupce BR a BS vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	33
4.10	Mapa významných rozdílů středních hodnot pro sloupec BT vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	35
4.11	Schéma prováděných dvoufaktorových testů pro analýzu dat na základě pohlaví a věku . . . . .	36
4.12	Grafy středních hodnot a 95% intervalových odhadů pro sloupce BP a BQ vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	44
4.13	Grafy středních hodnot a 95% intervalových odhadů pro sloupce BR a BS vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	46
4.14	Graf středních hodnot a 95% intervalových odhadů pro sloupec BT vstupního datového souboru <code>Student_groups_by_age_valid_2.xlsx</code> . . . . .	47
4.15	Schéma prováděných korelačních testů pro zjištění úkolů korelujících s jinými . . . . .	49
4.16	Graf Spearmanova korelačního koeficientu pro podskupinu 1 a 2 . . . . .	51
4.17	Graf Spearmanova korelačního koeficientu pro podskupinu 3 a 4 . . . . .	52
4.18	Graf Spearmanova korelačního koeficientu pro podskupinu 5 a 6 . . . . .	53
4.19	Graf Spearmanova korelačního koeficientu pro podskupinu 7 a 8 . . . . .	54
A.1	Graf rozdílů statistických ukazatelů z Matlabu . . . . .	62

A.2	Graf rozdílů statistických ukazatelů z Matlabu . . . . .	62
A.3	Graf rozdílů statistických ukazatelů z Matlabu . . . . .	63
A.4	Graf rozdílů statistických ukazatelů z Pythonu . . . . .	63

## Seznam tabulek

3.1	Počet záznamů v jednotlivých věkových skupinách . . . . .	20
3.2	Počet záznamů ve vyřazených věkových skupinách . . . . .	20

## Seznam zkratk a použitého značení

<b>DCD</b>	Developmental Coordination Disorder
<b>FP TUL</b>	Fakulta přírodovědně–humanitní a pedagogická Technické univerzity v Liberci
<b>GDPR</b>	General Data Protection Regulation
$\mu$	střední hodnota
$\sigma$	rozptyl
$x_{0.5}$	medián
$H_0$	nulová hypotéza
$H_A$	alternativní hypotéza
<b>ANOVA</b>	analýza rozptylu

# 1 Úvod

Dyspraxie, známá také jako vývojová koordinační porucha (DCD), je neurovývojová porucha, která ovlivňuje schopnost jedince plánovat a koordinovat pohyby, což vede k obtížím při vykonávání každodenních činností. Tato porucha je často poddiagnostikována a může mít významný dopad na studijní a sociální život jedince. [5]

V Libereckém kraji bylo provedeno šetření, kterého se zúčastnilo 665 žáků z několika základních škol a školek. Cílem tohoto šetření bylo odhalit dyspraxii u žáků tím, že žáci plnili praktické úkoly Bruininks-Oseretsky testu motorické zdatnosti. [2] Po konci šetření byla získaná data anonymizována a poskytnuta pro tuto práci ve formě .xlsx souborů, které jsou typické pro program Microsoft Excel.

Tato bakalářská práce si klade za cíl analyzovat data získaná z tohoto šetření a zjistit případné chyby či nesrovnalosti v datech, jenž jsou způsobeny ručním přepisem z testového archu do poskytnuté tabulky.

Dále jsou v kapitole 4.1 zkoumány významné statistické rozdíly v prevalenci dyspraxie mezi žáky 11 věkových skupin, které jsou v této práci diskutovány a je na ně kladen důraz v manuálu Bruininks-Oseretsky testu. [2].

V kapitole 4.2 se práce zabývá vlivem pohlaví a věku, jakožto hlavními faktory ovlivňující dosažená skóre jednotlivých žáků. V obou uvedených kapitolách se nachází schémata, která slouží k vyobrazení algoritmizace výběru statistických testů používaných v této práci. Poté jsou v těchto kapitolách uvedeny testované statistické hypotézy, vztahující se k daným statistickým testům.

V kapitole 4.3 je opět uvedeno schéma, dle kterého bylo postupováno, a také jsou zde popsány testované hypotézy týkající se statistických testů pro vyhodnocení korelace. Těmito testy je zkoumán seznam úkolů Bruininks-Oseretsky testu a je zde navrženo zúžení tohoto seznamu na základě zjištění korelačních koeficientů mezi jednotlivými úkoly.

## 2 Vstupní data

Tato práce se zabývá daty, která byla poskytnuta FP TUL v rámci šetření dyspraxie studentů na některých základních školách v Libereckém kraji. Konkrétně byl na jednotlivých školách prováděn Bruininks-Oseretsky test motorické zdatnosti.

Data poskytnutá pro tuto práci byla anonymizována. Některé informace, jako například jména žáků či školy, které jsou danými studenty navštěvovány, byly změněny, nebo naprosto vyřazeny. Jedná se pouze o ochranu GDPR jednotlivých studentů a informace tohoto typu nemají žádný dopad na výsledky, kterými se bude tato práce zabývat. Na druhou stranu byla ponechána informace o pohlaví daného studenta, datum narození, datum provedení celkového testu či výsledky jednotlivých testů, které byly vykonávány. Pro tuto práci byly celkem poskytnuty 2 datové soubory:

- `data_matrix1.xlsx`
- `data_matrix2.xlsx`

Druhý uvedený soubor byl dodán až později při psaní práce. V podstatě se jedná uspořádáním o 2 totožné soubory. Co se však velmi liší je kvalita a úplnost dat. První soubor obsahuje mnoho neúplných údajů, které byly zjištěny na základě analýzy, a budou diskutovány v kapitole 3. Druhý soubor tedy vznikl upravením, doplněním a zpřesněním prvotního souboru. V souboru `data_matrix1.xlsx` je celkem obsaženo 639 záznamů (záznamem je myšlen řádek v souboru s informacemi o daném studentovi) o 71 sloupcích. Některé z těchto záznamů obsahují nevyplněné hodnoty v příslušných buňkách. Později poskytnutý soubor `data_matrix2.xlsx` je definován 665 záznamy o 71 sloupcích. Pro všechny tyto záznamy platí, že mají plně obsazeny všechny jejich sloupce a jsou většinou bez chyb.

### 2.1 Charakteristika dat

Data byla v rámci testu zaznamenávána do protokolu [2] a poté ručně přepisována do `.xlsx` souboru, používaného tabulkovým programem Microsoft Excel. Dané `.xlsx` soubory disponují níže popsanými sloupci, přičemž řádek představuje právě jeden záznam s informacemi o daném studentovi.

- Škola
- Třída
- Datum narození
- Datum testu

- Pohlaví
- Úkoly 1-53 Bruininks-Oseretsky testu (rozepsány v příloze A.0.1)
- Subtest1 - Subtest8 - jednotlivé podskupiny úkolů 1-53
- Test1 - Test4 - jednotlivé podskupiny subtestu1 - subtestu8
- TMC - celkové výsledné skóre

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	škola	třída	datum narození	datum testu	gender	1 ...		53 subtest1	...		subtest8	test1	...	test4	TMC

Obrázek 2.1: Vstupní tabulka

## 2.2 Výpočet výsledku testu

Výše uvedené sloupce .xlsx souborů budou nyní rozebrány více do detailu. Tím dojde k definování pojmů, které poté slouží k přehlednějšímu popisu struktury daných souborů. Tato kapitola vychází z manuálu Bruininks-Oseretsky testu. [2]

**Raw skóre** – představuje přímo naměřené výsledky 53 úkolů v 8 subtestových kategoriích, absolvovaných žáky, v jednotkách bodového hodnocení jako například počet bodů za vybarvení hvězdy, počet chyb v bludišti, počet úspěšných pokusů při třídění karet atd. Tyto výsledky byly přímo zaznamenávány do protokolů.

- Výsledky nelze vzájemně porovnávat v rámci věku a pohlaví dítěte.
- Výsledky jsou uvedeny ve sloupcích F-BF vstupního .xlsx souboru.

**Point skóre** - Raw skóre jsou přímo v protokolech převáděny na point skóre.

- Výsledky nelze vzájemně porovnávat v rámci věku a pohlaví dítěte.

**Celkové point skóre** - prostý součet jednotlivých point skóre u testových úkolů sdružených do 8 subtestových kategorií:

- |                                |                               |
|--------------------------------|-------------------------------|
| 1. Přesnost v jemné motorice   | 2. Integrace v jemné motorice |
| 3. Manuální zručnost           | 4. Bilaterální koordinace     |
| 5. Rovnováha                   | 6. Rychlost a agilita         |
| 7. Koordinace horních končetin | 8. Síla                       |

- Výsledky nelze vzájemně porovnávat v rámci věku a pohlaví dítěte.

**Škálové skóre** – obdržaná celková point skóre subtestových kategorií 1 až 8 jsou převedena, za pomoci tabulek zohledňujících věk a pohlaví dítěte, na škálové skóre podskupiny.

- Předpoklad Bruininks-Oseretsky testu je, že střední hodnota Škálového skóre je 15 a směrodatná odchylka 5.

(b) Výsledky lze částečně vzájemně porovnávat v rámci věku a pohlaví dítěte.

(c) Výsledky jsou uvedeny ve sloupcích BG-BN vstupního .xlsx souboru

**Standardní skóre** – je získáno tabulkovým převodem prostých součtů škálových skóre příslušných dvojic subtestových kategorií (1+2; 3+7; 4+5; 6+8). Je ukryto pod sloupci Test(1 - 4), kterými jsou určeny výsledky 4 kompozitních kategorií:

- Kontrola v jemné motorice
- Manuální koordinace
- Celková tělesná koordinace
- Síla a agility

(a) Předpoklad je, že střední hodnota standardního skóre je 50 a směrodatná odchylka 10.

(b) Výsledky lze částečně vzájemně porovnávat v rámci věku a pohlaví dítěte.

(c) Výsledky jsou uvedeny ve sloupcích BO-BR vstupního .xlsx souboru

**Celkový motorický koeficient** – je získán tabulkovým převodem standardních skóre 1+2+3+4. Tabulkový převod je závislý na věku a pohlaví dítěte.

(a) Předpoklad je, že střední hodnota celkového motorického koeficientu je 50, směrodatná odchylka 10.

(b) Výsledky lze vzájemně porovnávat v rámci věku a pohlaví dítěte.

(c) Výsledky jsou uvedeny ve sloupci BS vstupního .xlsx souboru

## 3 Filtrace a třídění dat

Tato kapitola je zaměřena na odfiltrování nevalidních, neúplných či nepřesných dat ze vstupních souborů `data_matrix1.xlsx` a `data_matrix2.xlsx`, které jsou popsány v kapitole 2. Tento proces probíhal tak, že byly napsány filtry, nad vstupními daty, díky kterým byly odhaleny nevalidní záznamy. Tyto zjištěné záznamy byly v rámci filtru vždy zaznamenány do .csv souborů, které byly poté poskytnuty k opravě dat. Důvodem použití .csv souborů místo .xlsx, ve kterých jsou původní data je to, že .csv soubory jsou rychlejší, zabírají méně místa a k jejich otevření není zapotřebí tabulkového softwaru. [3] Po odhalení nevalidních záznamů, byla data roztržena do celkem 11 skupin, které jsou definovány věkem.

Před začátkem filtrace a třídění dat bylo bráno v potaz několik nástrojů, za pomoci kterých mohly být tyto kroky programovány. Konkrétně se jedná o programovací jazyky Python, Matlab a R. Všechny uvedené jazyky disponují různými knihovnamí, které se dají velmi dobře použít pro zpracování dat a statistickou analýzu.

Pro účely této práce byl však dále využíván programovací jazyk Python a to hned z několika důvodů. Z těchto nástrojů disponuje právě Python velmi širokou komunitou, což se odráží na rychlejším vyhledávání problémů a jejich řešení. Další jeho předností může být velká podpora různých knihoven, které budou potřeba ke zpracování poskytnutých dat a jejich následné statistické analýze<sup>1</sup>. Výhodou také může být, že tyto knihovny jsou Open-source a tím pádem nejsou placené. Na druhou stranu nevýhodou by mohlo být, že Matlab, jakožto placený software, bude více optimalizován, tím pádem bude rychlejší a přesnější.

### 3.1 Kvalita dat

Na obou souborech, které jsou uvedeny v kapitole 2, poskytnutých pro tuto práci, byly spuštěny obdobné filtry, jelikož v obou těchto souborech se nachází stejně uspořádaná data. To, co je ve vstupních souborech rozdílné, je však úplnost dat. V prvním datovém souboru bylo odhaleno mnohem více nepřesných či chybějících hodnot. Druhý soubor vznikl právě opravením či doplněním těchto nevalidních hodnot, na které bylo poukázáno, a rozšířením prvotního souboru o další množinu dat stejného typu.

---

<sup>1</sup>Konkrétně se jedná o knihovny: Pandas, SciPy, Statsmodel, Matplotlib a některé další



## 3.2 Filtry

Pro vstupní data bylo naprogramováno, za pomoci Python knihovny Pandas, hned několik filtrů. Filtry jsou zaměřeny na stěžejní informace, které by mohly mít vliv na konečný výsledek a jeho přesnost.

Konkrétně se jedná o filtry na:

- Datum narození
- Datum testu
- Prázdné záznamy
- Celkové point skóre a standardní skóre
- Celkový motorický koeficient
- Sloupcové extrémy
- Nenumerické hodnoty

### 3.2.1 Filtr na datum narození

Pro následné zpracování dat, jenž bude následovat v dalších kapitolách, je datum narození stěžejní hodnotou. Na základě data narození a data vykonání testu je později určen věk žáka v den testu a tato informace je poté použita k rozřazení žáků do jednotlivých věkových skupin.

Tímto filtrem bylo kontrolováno, zda je v záznamu datum narození uvedeno. Pokud není uvedeno nebo nabývá nějaké jiné hodnoty, je převzat index příslušné buňky a poté celý řádek výskytu automaticky zapsán do souboru `birth_date_err.csv`. Filtr byl spuštěn na obou poskytnutých souborech, přičemž v souboru:

- `data_matrix1.xlsx` bylo nalezeno **209** chybných záznamů
- `data_matrix2.xlsx` bylo nalezeno **0** chybných záznamů

### 3.2.2 Filtr na datum testu

Tento filtr slouží ke stejnému účelu jako filtr v kapitole [Filtr na datum narození](#). Pokud datum testu není v záznamu uvedeno, nebo nabývá nějaké jiné hodnoty, je podobně jako v kapitole výše převzat celý řádek a automaticky zapsán do souboru `test_date_err.csv`.

Filtr byl spuštěn na obou poskytnutých souborech, přičemž v souboru:

- `data_matrix1.xlsx` bylo nalezeno **177** chybných záznamů
- `data_matrix2.xlsx` bylo nalezeno **0** chybných záznamů

### 3.2.3 Filtr na prázdné záznamy

V původním poskytnutém souboru bylo vynecháno značné množství záznamů. Tímto filtrem mělo být docíleno spíše toho, že poukáže na nedostatečnost dat. Tento filtr již bere v potaz dva předchozí filtry, a tedy začíná od sloupce F. Vstupem do

tohoto filtru jsou sloupce F-BS vstupních souborů a pokud suma všech odpovídajících indexů je 0, tedy všechny sloupce daného řádku jsou prázdné, jsou odpovídající řádky automaticky zapsány do souboru `missing_tests.csv`.

Filtr byl spuštěn na obou poskytnutých souborech, přičemž v souboru:

- `data_matrix1.xlsx` bylo nalezeno **10** chybných záznamů
- `data_matrix2.xlsx` bylo nalezeno **0** chybných záznamů

### 3.2.4 Filtr na škálové skóre a standardní skóre

Pro výpočet celkového motorického koeficientu je nutné znát hodnoty z předešlých výpočtů, jak je uvedeno v kapitole [Výpočet výsledku testu](#). Tímto filtrem jsou kontrolovány sloupce BG-BR vstupních souborů, jelikož jsou v nich hodnoty důležité k dalšímu výpočtu celkového motorického koeficientu. Filtr načítá tyto sloupce, a pokud odpovídající indexy mají nulovou hodnotu, jsou automaticky zapsány do souboru `missing_subtests.csv`.

Filtr byl spuštěn na obou poskytnutých souborech, přičemž v souboru:

- `data_matrix1.xlsx` bylo nalezeno **52** chybných záznamů
- `data_matrix2.xlsx` bylo nalezeno **0** chybných záznamů

### 3.2.5 Filtr na celkový motorický koeficient

Z kapitoly [Výpočet výsledku testu](#) plyne, že pokud je celý řádek tvořen nenulovými hodnotami, pak celkový motorický koeficient nemůže nabývat nulové hodnoty. Pokud tato situace nastane, jsou příslušné řádky automaticky zapsány do souboru `tmc_is_zero.csv`.

Filtr byl spuštěn na obou poskytnutých souborech, přičemž v souboru:

- `data_matrix1.xlsx` bylo nalezeno **32** chybných záznamů
- `data_matrix2.xlsx` bylo nalezeno **0** chybných záznamů

### 3.2.6 Filtr na extrémny sloupců

Většina sloupců v tabulce má svoji maximální hodnotu. Není tomu tak vždy, nachází se zde jednotlivé testy, které nemají maximální hodnotu a zapisuje se maximum dosažené v testu. Například pod testem, jehož výsledky jsou zapsány ve sloupci BB, se skrývá skok daleký a dosažené Raw skóre, které je poté přepočítáno, nabývá maximální hodnoty skoku v cm, tedy nemá jasně danou maximální hodnotu. Naopak minimální hodnota v každém sloupci nesmí být menší než 0, to je dáno pravidly testu. [2]

Díky manuálním přepisům do tabulky se dá předpokládat, že zde budou chyby, které nebudou dodržovat sloupcová maxima. Pakliže daný sloupec má maximální skóre, je kontrolováno, stejně tak minimální skóre. Pokud filtr odhalí takovéto

hodnoty, nejsou zapsány jako celý řádek, jak tomu bylo doposud, ale pouze jako souřadnice ve formátu „sloupec,řádek“ do souboru `max_value_exceeded.csv`.

Filtr byl použit pouze na souboru `data_matrix2.xlsx`, kde bylo nalezeno **46 chybných záznamů** na celkem **41** různých řádcích. Vzhledem k tomu, že by tyto chybné hodnoty mohly mít vliv na konečné výsledky, byly z původní datové matice, `data_matrix2.xlsx`, odebrány. Důvodem použití pouze na druhém vstupním souboru je to, že filtr byl psán později, kdy už byla k dispozici nová datová matice, která byla upravena na základě předešlých filtrů.

### 3.2.7 Filtr na nenumernické hodnoty

Od sloupce F vstupních souborů se nachází ve všech dalších sloupcích tabulky pouze numerické hodnoty. Tento filtr byl psán z toho důvodu, že pokud je uživatelem Excelu špatně zvolen datový typ sloupce, může dojít ke konverzi datového typu. To může mít za následek například přepsání reálného čísla na datum. Pokud tedy v tabulce, od sloupce F, bude obsažena nějaká nenumernická hodnota, měla by být tímto filtrem odhalena a příslušný řádek zapsán do souboru `non_numeric_values.csv`.

Tento filtr byl dopsán až později při psaní práce, kdy už byla poskytnuta nová datová matice. Filtr byl tedy spuštěn pouze na souboru `data_matrix2.xlsx`, kde byly nalezeny **2 chybné záznamy**.

V prvním případě obsahoval celý sloupec hodnoty typu `object`, což je v knihovně Pandas datový typ používaný pro textové nebo smíšené (nenumernické a numerické) hodnoty. [10] Po zhlédnutí záznamů je však zřejmé, že se zapisovatel snažil o zapsání desetinného čísla s desetinnou čárkou, což se díky typografické chybě nepodařilo. Byla zapsána hodnota s desetinnou tečkou, k tomu bylo přidáno špatné nastavení tabulky a odhalená chybná hodnota se tedy tváří jako datum. Vzhledem ke správnosti ostatních dat v řádku byla hodnota přepsána s desetinnou čárkou a v práci byla dále uvažována.

V druhém případě není možné jednoduše určit, co by se v dané chybné buňce mělo nacházet, a proto je tento záznam odstraněn a nebude s ním dále počítáno.

## 3.3 Třídění dat

Od této kapitoly po zbytek práce se uvažuje jako vstup pouze soubor `data_matrix2.xlsx`, či jeho modifikace, ze kterého bylo odstraněno **42 záznamů** na základě předešlých filtrů. To znamená, že ve zmíněném souboru je nyní obsaženo **622 záznamů**.

Na požadavek FP TUL měla být data rozřazena do 11 skupin na základě věku jednotlivých žáků. Data, která nespádají ani do jedné z věkových skupin, měla být vyřazena na základě nedostatku záznamů v těchto kategoriích.

Obdobně jako v kapitolách výše byla k tomuto úkonu zvolena knihovna Pandas programovacího jazyka Python. Byl tedy vytvořen skript, kterým je převzata vyfiltrovaná datová matice `data_matrix2.xlsx`, a následně pro každého studenta určen jeho věk v den vykonání Bruininks-Oseretsky testu. Dle zjištěného věku byli poté

záznamy studentů řazeny do následujících věkových skupin, které byly přesně takto definovány FP TUL, a je u nich uveden přesný počet studentů v dané skupině, který byl zjištěn.

Skupina	Věkový rozsah	Počet záznamů
1	6 let 4 měsíce - 6 let 7 měsíců	31
2	6 let 8 měsíců - 6 let 11 měsíců	45
3	7 let 0 měsíců - 7 let 3 měsíce	60
4	7 let 4 měsíce - 7 let 7 měsíců	52
5	7 let 8 měsíců - 7 let 11 měsíců	58
6	8 let 0 měsíců - 8 let 5 měsíců	64
7	8 let 6 měsíců - 8 let 11 měsíců	63
8	9 let 0 měsíců - 9 let 5 měsíců	73
9	9 let 6 měsíců - 9 let 11 měsíců	51
10	10 let 0 měsíců - 10 let 5 měsíců	54
11	10 let 6 měsíců - 10 let 11 měsíců	45

Tabulka 3.1: Počet záznamů v jednotlivých věkových skupinách

Všechna data mladších či starších studentů byla odstraněna. Pro zajímavost je zde uvedeno, o kolik dat se jedná.

Věkový rozsah	Počet záznamů
4 roky - 5 let	4
5 let - 6 let	8
6 let - 6 let 3 měsíce	6
11 let - 12 let	8

Tabulka 3.2: Počet záznamů ve vyřazených věkových skupinách

Záznamy jednotlivých věkových skupin byly uloženy do nového .xlsx souboru `Student_groups_by_age_valid_2`, který má stejné sloupce jako vstupní soubor. Pro každou skupinu byl v tomto souboru vytvořen oddělený list, ve kterém se nachází příslušný počet záznamů každé skupiny viz tabulka 3.1. Jednotlivé listy jsou systematicky pojmenovány Group1-Group11. Celkově soubor `Student_groups_by_age_valid_2` čítá 596 záznamů, rozdělených do 11 skupin dle věku a pohlaví.

## 4 Statistická analýza dat

Cílem této kapitoly bylo provést statistické testy na poskytnutých datech a odhalit možné chyby vstupního datového souboru. Každý ze statistických testů použitý v této kapitole byl v minulosti podrobně zkoumán (např. v monografii [1] a [4]), a proto není účelem je jakkoli zdokonalovat či implementovat. Cílem práce bylo vybrat tu nejvhodnější posloupnost testů tak, aby bylo dosaženo co nejuvěrnějších statistických výsledků.

V kapitole 3.3 bylo diskutováno získání souboru `Student_groups_by_age_valid_2.xlsx`, který je pouze modifikovanou obdobou vstupního souboru `data_matrix2.xlsx`. Pro datovou analýzu tedy byl použit právě tento modifikovaný soubor, jelikož byl díky své struktuře vhodnější pro provedení statistických testů.

Stejně jako v předchozích kapitolách bude k těmto účelům použit převážně Python a konkrétně knihovna Scipy, která disponuje potřebnými statistickými testy.[11] Některé statistické výpočty byly však kontrolovány za pomoci softwaru Matlab, jelikož se jedná o placený software, který může být přesnější.

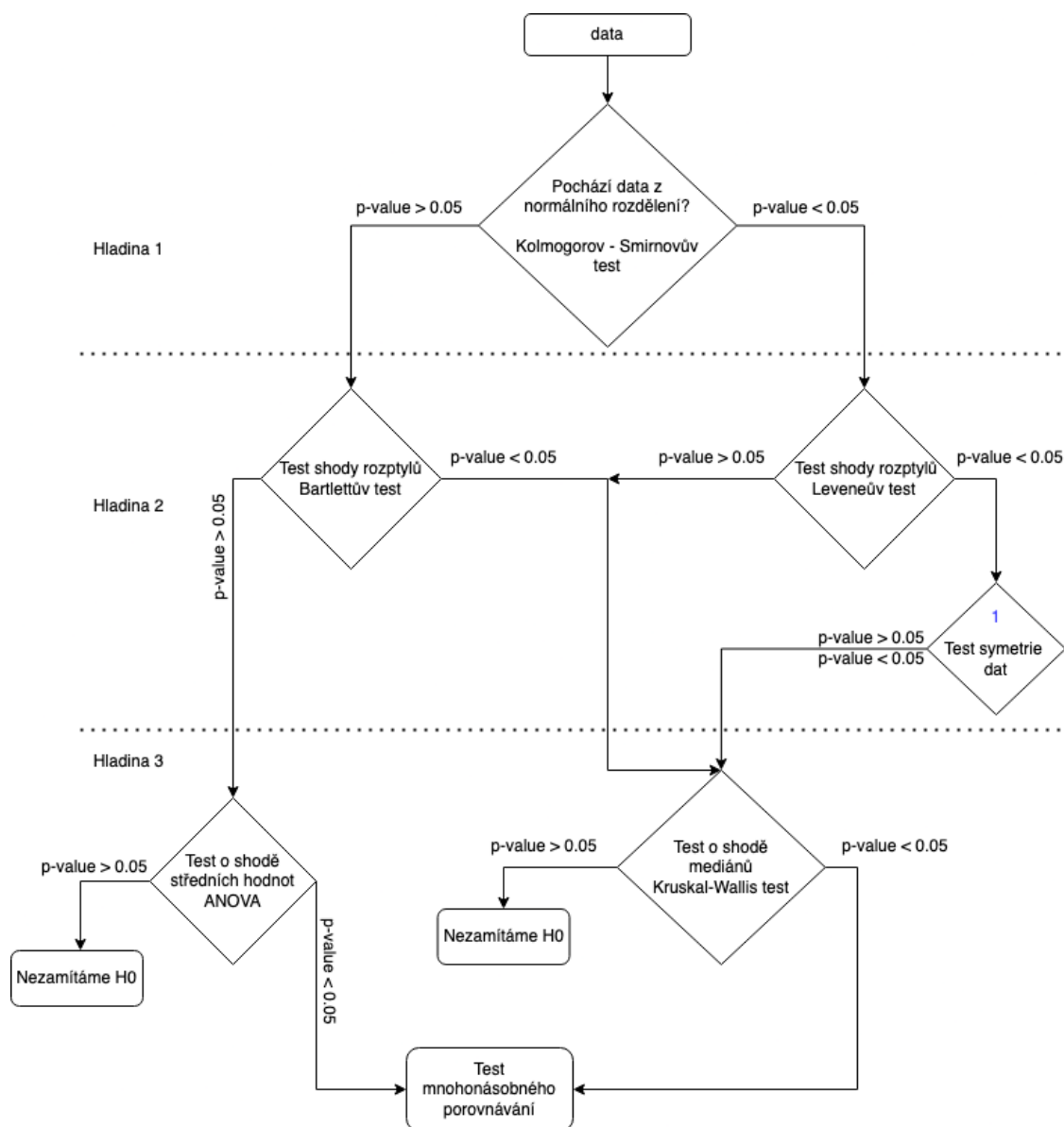
### 4.1 Analýza dat na základě věku

Pro jednotlivé věkové skupiny ve vstupním souboru `Student_groups_by_age_valid_2.xlsx` byly kontrolovány sloupce BH až BT. Pod těmito sloupci se ukrývá:

- Škálové skóre
- Standardní skóre
- Celkový motorický koeficient

Hlavním faktorem vstupujícím do jednotlivých statistických testů byl tedy věk studenta v den vykonání Bruininks-Oseretsky testu. Statistické testy proto berou v potaz věkové skupiny vyobrazené v tabulce 3.1.

Posloupnost prováděných statistických testů se řídila následujícím schématem:



Obrázek 4.1: Schéma prováděných jednofaktorových testů pro analýzu dat na základě věku<sup>1</sup>

Toto schéma je pro přehlednost rozděleno do 3 hladin. Tyto hladiny jsou v následujících kapitolách podrobněji rozebrány a jsou zmíněny jednotlivé testované hypotézy. Poté přijdou na řadu jednotlivé testované sloupce vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`, na kterých jsou demonstrovány výsledky těchto provedených statistických testů, jež se v daných hladinách nacházejí.

<sup>1</sup>V této práci je předpokladem, že jsou data symetrická. Pokud nějaká data nejsou symetrická, nebude to v této práci uvažováno, jelikož by to nemělo mít vliv na celkové výsledky.

## Teorie

Statistické testy a hypotézy zmíněné v následujících kapitolách jsou obecně popsány v této literatuře<sup>2</sup>:

- Kolmogorov-Smirnovův test (resp. Lillieforsův test)
  - str. 104 [1], str. 233 [4]
- Bartlettův a Leveneův test
  - str. 121 [1], str. 205 [6]
- ANOVA a Kruskal-Wallisův test
  - str. 339 (resp. str. 347) [4] a str. 479 (platí pro obojí) [9]

### 4.1.1 Hladina 1

Na této hladině bylo testováno, zda právě zkoumaná data ze souboru `Student_groups_by_age_valid_2.xlsx` pochází z normálního rozdělení.

Byla tedy testována hypotéza:

- $H_0 : X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ , kde  $X_1, X_2, \dots, X_n$  jsou pozorování, která tvoří vzorek z normálního rozdělení s neznámými parametry  $\mu$  a  $\sigma^2$ . Symbol  $\sim$  znamená "je distribuováno jako" a  $N(\mu, \sigma^2)$  značí normální rozdělení s parametry  $\mu$  a  $\sigma^2$ .
- $H_A : X_1, X_2, \dots, X_n \not\sim N(\mu, \sigma^2)$ , kde  $X_1, X_2, \dots, X_n$  jsou pozorování, která tvoří vzorek z normálního rozdělení s parametry  $\mu$  a  $\sigma^2$ . Symbol  $\not\sim$  znamená "není distribuováno jako" a  $N(\mu, \sigma^2)$  značí normální rozdělení s parametry  $\mu$  a  $\sigma^2$ .

Bylo k tomu použito **Kolmogorova-Smirnovového jednovýběrového**, resp. Lillieforsova, **testu** rozdělení.[4] Tento test nebyl v počátku brán v potaz. Vzhledem k tomu, že data poskytnutá pro tuto práci nejsou spojitá, a tím pádem nemají spojitou distribuční funkci, bylo by přesnější testovat normalitu dat za pomoci  $\chi^2$  testu dobré shody s očekávaným rozdělením. [1]

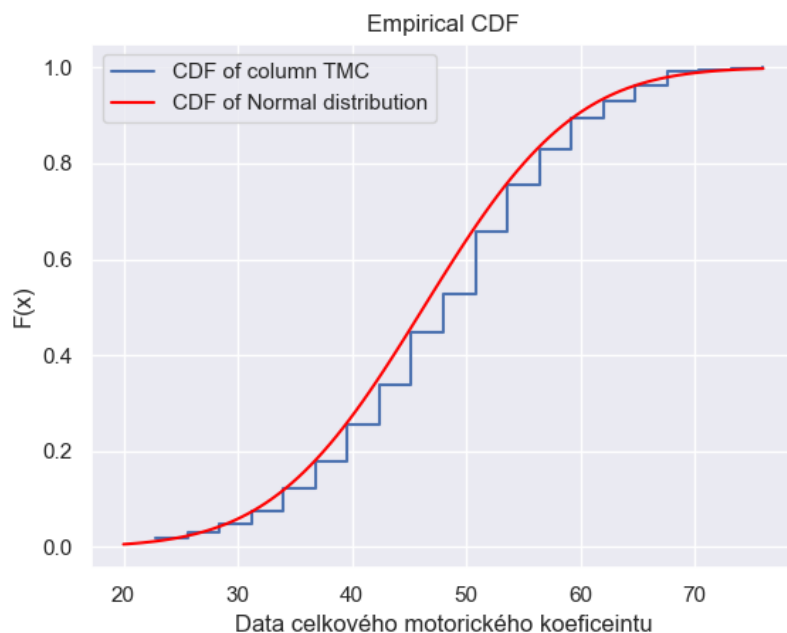
Provedení  $\chi^2$  testu dobré shody se v Pythonu ukázalo jako velmi problematické. Vstupem této funkce jsou očekávané a pozorované hodnoty a je zde předpoklad, že tyto hodnoty mají stejné sumy, pokud tomu tak není, je zde nastavena tolerance této odlišnosti na  $1 \times 10^{-8}$ . Tento předpoklad o stejné hodnotě sum nebyl nikdy dosažen, přestože byly vyzkoušeny různé způsoby, které byly nalezeny na Githubu Scipy<sup>4</sup> a tento problém zde byl diskutován.

---

<sup>2</sup>Zmíněna je vždy počáteční strana kapitoly

<sup>3</sup>Vyobrazená distribuční funkce je spočítána ze sloupce BT vstupního souboru `Student_groups_by_age_valid_2`.

<sup>4</sup>Odkaz na diskuze, kde je tento problém zmiňován lze nalézt [zde](#) a [zde](#).  
(<https://github.com/scipy/scipy/issues/14298>,  
<https://github.com/UDST/synthpop/issues/75>)



Obrázek 4.2: Porovnání distribučních funkcí vstupních dat a normálního rozdělení <sup>3</sup>

Pro dosažení přesnosti při testování normality dat bylo vyzkoušeno volání Matlabové funkce *chi2gof()* za pomoci pythonového skriptu, kterým byla daná problematika řešena. Pro volání této funkce bylo nutné připojit Matlabové API pro Python.

Zde muselo být vytvořeno nové prostředí pro Python za pomoci programu Anaconda kvůli dosavadní jiné použité verzi Pythonu. Jelikož verze Matlabu byla používána ve verzi *R2021b*, bylo nutné použít kompatibilní verzi Pythonu <sup>5</sup>. Poté bylo toto prostředí aktivováno a za pomoci terminálu byl vyhledán, spuštěn a nainstalován soubor *setup.py* v kořenovém adresáři, kam byl nainstalován Matlab. [7]

```

→ python conda activate matlab_test
(matlab_test) → python pwd
/Applications/MATLAB_R2021b.app/extern/engines/python
(matlab_test) → python python3 setup.py install

```

Obrázek 4.3: Instalace matlab engine pro Python <sup>6</sup>

Připojení Matlabu k pythonovému skriptu bylo úspěšné. Avšak volání potřebné statistické funkce *chi2gof()* již úspěšné nebylo. Vyskytla se zde syntaktická chyba, kterou nebylo možné odstranit. Volání jiných jednodušších funkcí, jako například *magic()* či *sum()* bylo správné, čímž byla ověřena funkčnost tohoto volání.

Některá data byla poté převedena do Matlabu, za účelem zjistit, jakých hodnot dosahuje *p*-value při  $\chi^2$  testu dobré shody. Hodnoty poskytnuté Matlabem dosahovaly však tak nízkých hodnot (například: pro sloupec BG je *p*-value =  $8.6 \times 10^{-6}$ , pro

<sup>5</sup>Byl použit Python 3.9

<sup>6</sup>Demonstrace příkazů použitých v terminálu k instalaci matlab engine pro python



sloupec F je  $p$ -value =  $3.2 \times 10^{-121}$ ), že je možné, bez markantního rozdílu na výsledek testu, použít právě funkci **Kolmogorova-Smirnovového jednovýběrového testu**.

### 4.1.2 Hladina 2

Zde byl proveden test o shodě rozptylů. Ze schématu 4.1 plyne, že pokud jsou data z normálního rozdělení, byl proveden Bartlettův test. Pokud data naopak nejsou z normálního rozdělení, byl proveden Leveneův test, jenž je méně citlivý na porušení normality dat.

Byly tedy testovány hypotézy:

- $H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$
- $H_A : \exists i, j \in 1, 2, \dots, k, i \neq j : \sigma_i^2 \neq \sigma_j^2$ , kde  $k$  je počet srovnávaných skupin,  $\sigma_i^2$  a  $\sigma_j^2$  jsou rozptyly  $i$ -té a  $j$ -té skupiny, které se porovnávají.

### 4.1.3 Hladina 3

Stejně jako tomu je v kapitole 4.1.2, tak i v této kapitole bylo přihlíženo k tomu, zda data pochází z normálního rozdělení. Dále do této hladiny vstupuje i výsledek testu shodnosti rozptylů. Pokud byla v obou případech  $p$ -value  $> 0.05$ , tak byla vybrána jednofaktorová ANOVA a testuje se hypotéza o shodnosti středních hodnot:

- $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$
- $H_A : \exists i, j \in 1, 2, \dots, k, i \neq j$  tak, že  $\mu_i \neq \mu_j$

Na druhou stranu, pakliže není splněn předpoklad o normálním rozdělení, nebo data nevykazují shodnost rozptylů či platí oboje zároveň, byl vybrán Kruskal-Wallisův test o rovnosti mediánů a byla testována hypotéza:

- $H_0 : x_{0.5_1} = x_{0.5_2} = \dots = x_{0.5_k}$
- $H_A : \exists i, j \in 1, 2, \dots, k, i \neq j : x_{0.5_i} \neq x_{0.5_j}$

Pokud byla zamítnuta nulová hypotéza  $H_0$ , byl proveden test mnohonásobného porovnávání. Zamítnutí nulové hypotézy znamená odchylky ve středních hodnotách či mediánech ve zkoumaných datech a bylo tedy žádoucí provést test, kterým bylo zjištěno mezi kterými skupinami jsou významné statistické rozdíly. Pro tuto práci byl jako test mnohonásobného porovnávání zvolena Tukey-Kramerova metoda.

## Statistické vyhodnocení

### 4.1.4 Škálové skóre - Subtest 1

V sloupci BH datového souboru Student\_groups\_by\_age\_valid\_2.xlsx, ve všech jeho listech, jsou výsledky zaměřeny na **přesnost v jemné motorice** viz [Výpočet](#)

výsledku testu. Bylo zde zkoumáno, zda má věk studentů vliv na jejich dosažené celkové point skóre v této kategorii.

**Hladina 1:** Prvním krokem bylo testování hypotéz zmíněných v kapitole 4.1.1 na hladině významnosti 5 %. Po provedení testu byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení ( $p$ -value  $\approx 0.04$ ).

**Hladina 2:** Na základě zamítnutí nulové hypotézy  $H_0$  uvedené v kapitole 4.1.1 byla vstupní data dále testována na hladině významnosti 5 % Leveneovým testem o shodě všech rozptylů z kapitoly 4.1.2. U tohoto testu byla přijata nulová hypotéza  $H_0$  o shodě všech rozptylů ( $p$ -value  $\approx 0.6$ ).

**Hladina 3:** Jelikož data nepodléhají normálnímu rozdělení, byl použit Kruskal-Wallisův test, u kterého byly testovány hypotézy o shodě mediánů z kapitoly 4.1.3. Po provedení testu bylo možné konstatovat, že jednotlivé skupiny mezi sebou nevykazují významné statistické rozdíly, jelikož byla na hladině významnosti 5 % přijata nulová hypotéza  $H_0$  o shodě všech mediánů ( $p$ -value  $\approx 0.4$ ), a nebylo nutné provádět dále test mnohonásobného porovnávání.

### 4.1.5 Škálové skóre - Subtest 2

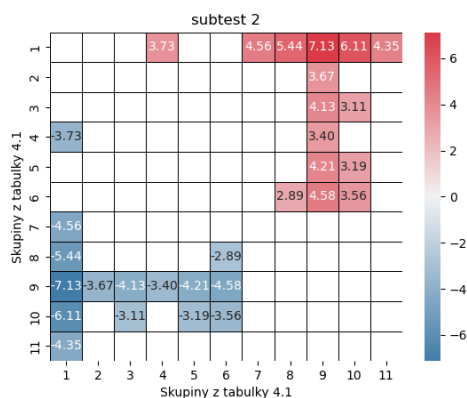
Ve vstupním datovém souboru `Student_groups_by_age_valid_2.xlsx`, ve všech jeho listech, se ve sloupci BI nachází výsledky zaměřené na **integraci v jemné motorice**. Stejně jako v předešlé kapitole je tento fakt zmíněn v kapitole [Výpočet výsledku testu](#).

**Hladina 1:** Nejprve byla testována normalita dat dle příslušných hypotéz v kapitole 4.1.1. Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.04$ ). Jelikož data v sloupci BI prvního listu s názvem Group1 vstupního souboru nepochází z normálního rozdělení, další data nejsou otestována a toto je automaticky převedeno na všechny listy.

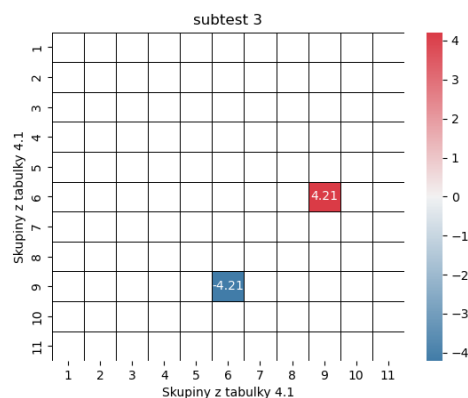
**Hladina 2:** Dále je zvolen Leveneův test na základě výsledku hypotéz popsanych v kapitole 4.1.1. Zde nebyla na hladině významnosti 5 % zamítnuta hypotéza  $H_0$  o shodnosti všech rozptylů ( $p$ -value  $\approx 0.98$ ) z kapitoly 4.1.2. Jelikož data nepochází z normálního rozdělení, ale není zamítnuta nulová hypotéza  $H_0$  o shodnosti rozptylů, dalším testem v pořadí je Kruskal-Wallisův test viz 4.1.3. Na hladině významnosti 5 % je zamítnuta nulová hypotéza  $H_0$  o shodnosti mediánů mezi jednotlivými skupinami.

**Hladina 3:** Nyní bylo zkoumáno mezi kterými skupinami dochází k významným statistickým rozdílům. Byl proveden Tukey-Kramerův test mnohonásobného porovnávání. Z této metody byl pro přehlednost ve výsledcích vytvořen níže uvedený graf. Je zde rovnou uveden i graf pro následující kapitolu.

V grafu 4.4 (a) je vidět, že největší rozdíly jsou mezi odlehlejšími věkovými skupinami. Tedy mezi věkovými skupinami, které mezi sebou mají velký věkový rozdíl. Naopak nikdy není rozdíl mezi věkovými skupinami, které jsou hned vedle sebe, tedy je mezi nimi věkový rozdíl nejmenší.



(a) Subtest 2



(b) Subtest 3

Obrázek 4.4: Mapy významných rozdílů středních hodnot pro sloupce BI a BJ vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`

### 4.1.6 Škálové skóre - Subtest 3

Všechny listy vstupního datového souboru `Student_groups_by_age_valid_2.xlsx` obsahují, ve sloupci BJ, informace o výsledcích zaměřených na **manuální zručnost**.

**Hladina 1:** Jako první krok po načtení těchto dat ze souboru byly na hladině významnosti 5 % testovány hypotézy o normalitě dat z kapitoly 4.1.1. Dle dosažených výsledků u první testované skupiny byla zamítnuta nulová hypotéza  $H_0$  o normalitě dat ( $p$ -value  $\approx 0.04$ ). Stejně jako v předešlých kapitolách byla tedy automaticky všechna data brána, jako by nepocházela z normálního rozdělení.

**Hladina 2:** Dle schématu 4.1 byl dále vybrán Leveneův test shody rozptylů a byly testovány hypotézy, na hladině významnosti 5 %, z kapitoly 4.1.2. Dle dosažených výsledků v tomto testu o shodnosti rozptylů nebyla zamítnuta nulová hypotéza  $H_0$  o shodě rozptylů ( $p$ -value  $\approx 0.28$ ).

**Hladina 3:** Předposledním testem v této kapitole byl Kruskal-Wallisův test z kapitoly 4.1.3. Na hladině významnosti 5 % byly tedy testovány hypotézy o shodě všech mediánů. Jelikož byla zde spočítána hodnota  $p$ -value  $\approx 0.007$ , byla zamítnuta nulová hypotéza  $H_0$  o shodě všech mediánů.

Po provedení těchto testů a zjištění nedostatečných hodnot  $p$ -value bylo nutné provést test mnohonásobného porovnávání, kterým byly zjištěny výsledky zobrazené v grafu 4.4 (b). Stejně jako v předchozí kapitole lze z tohoto grafu konstatovat, že ačkoli se statisticky významně od sebe liší pouze 2 skupiny, tak jsou to věkové skupiny, které nejsou hned vedle sebe.

Tato skutečnost se vyskytla po druhé za sebou a začala být sledována po získání výsledků demonstrováných v této kapitole. V příštích kapitolách bude pozorováno, zda platí tento trend o významnějším rozdílu skupin, které jsou spíše dále od sebe (alespoň o 1 skupinu) nežli v těsné blízkosti (skupina vedle skupiny).

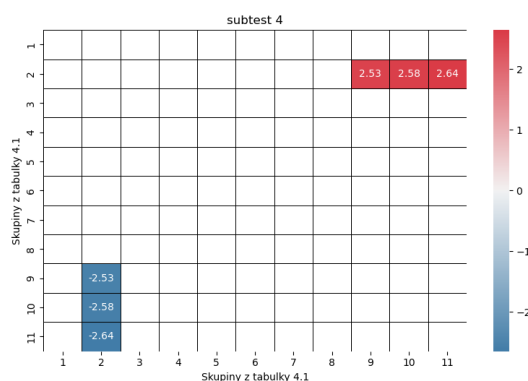
### 4.1.7 Škálové skóre - Subtest 4

Dále byl testován sloupec BK datového souboru. Výsledky v tomto sloupci jsou, ve všech listech souboru `Student_groups_by_age_valid_2.xlsx`, zaměřeny na **bilaterální koordinaci**.

**Hladina 1:** Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.1.1 o tom, zda data pochází z normálního rozdělení ( $p$ -value  $\approx 0.001$ ).

**Hladina 2:** Na základě zamítnutí této nulové hypotézy o normalitě dat bylo dále přistoupeno k Leveneově testu o shodě rozptylů. Na hladině významnosti 5 % byly testovány hypotézy o shodě rozptylů z kapitoly 4.1.2. Po provedení tohoto testu nebyla zamítnuta nulová hypotéza  $H_0$  o shodě rozptylů mezi jednotlivými skupinami ( $p$ -value  $\approx 0.23$ ).

**Hladina 3:** Dalším provedeným testem byl Kruskal-Wallisův test o shodě mediánů a byly testovány hypotézy z kapitoly 4.1.3. Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  o shodě všech mediánů ( $p$ -value  $\approx 0.002$ ).



Obrázek 4.5: Mapa významných rozdílů středních hodnot pro sloupec BK vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`

Vzhledem k zamítnutí nulové hypotézy  $H_0$  z kapitoly 4.1.3, konkrétně z Kruskal-Wallisova testu, byl dále proveden test mnohonásobného porovnávání pro zjištění významných statistických rozdílů mezi jednotlivými skupinami. Výsledky tohoto testu jsou vyobrazeny v grafu 4.5.

Tento graf ukazuje největší odchylky ve středních hodnotách mezi 2. a 9. skupinou, dále mezi 2. a 10. skupinou a také mezi 2. a 11. skupinou<sup>7</sup>. Tímto zjištěním se potvrzuje nastolený trend v předešlých kapitolách - a to takový, že významné statistické rozdíly jsou markantní mezi skupinami žáků, jež jsou od sebe více věkově vzdálené.

### 4.1.8 Škálové skóre - Subtest 5

Ve sloupci BL souboru `Student_groups_by_age_valid_2.xlsx` jsou, ve všech listech, výsledky testů zaměřených na **rovnováhu**. Tento fakt je zmíněn v kapitole

<sup>7</sup>Skupiny a jejich věkové rozmezí jsou popsány v tabulce 3.1

Výpočet výsledku testu, v sekci **Celkové point skóre**.

**Hladina 1:** Po stanovení a následném testování hypotéz z kapitoly 4.1.1 byla na hladině významnosti 5 % zamítnuta nulová hypotéza  $H_0$  o normalitě dat ( $p$ -value  $\approx 0.02$ ). Zmíněná hodnota  $p$ -value byla opět nalezena v 1. testované skupině a další skupiny nebyly testovány, a je stanoveno tvrzení, že data, ve všech skupinách, nepochází z normálního rozdělení.

**Hladina 2:** Vzhledem k tomuto zjištění byl dále proveden Leveneův test o shodnosti všech rozptylů na hladině významnosti 5 %. U tohoto testu nebyla zamítnuta nulová hypotéza  $H_0$  o shodnosti rozptylů, z kapitoly 4.1.2, mezi všemi 11 věkovými skupinami ( $p$ -value  $\approx 0.33$ ).

**Hladina 3:** V dalším kroku byly, na hladině významnosti 5 %, testovány hypotézy z kapitoly 4.1.3 o shodě všech mediánů za pomoci Kruskal-Wallisova testu. Díky nízké hodnotě  $p$ -value  $\approx 0.0001$ , dosažené v tomto testu, byla zamítnuta nulová hypotéza  $H_0$  o shodnosti všech rozptylů.

Jelikož byla zamítnuta nulová hypotéza o shodě všech mediánů, bylo dále testem mnohonásobného porovnávání zkoumáno, které skupiny mezi sebou vykazují významných statistických rozdílů. Tímto testem byla vrácena prázdná výsledná matice, což znamená, že nebyly nalezeny žádné skupiny, mezi kterými by docházelo k významným statistickým rozdílům. Přestože nebyly testem mnohonásobného porovnávání nalezeny žádné významnější statistické rozdíly, mohou zde přesto být očekávány vzhledem k  $p$ -value  $\approx 0.0001$  z Kruskal-Wallisova testu.

### 4.1.9 Škálové skóre - Subtest 6

Soubor `Student_groups_by_age_valid_2.xlsx` ve sloupci BM, v každém listu, obsahuje výsledky testů zaměřených na **rychlost a agility**, jak je uvedeno v kapitole [Výpočet výsledku testu](#).

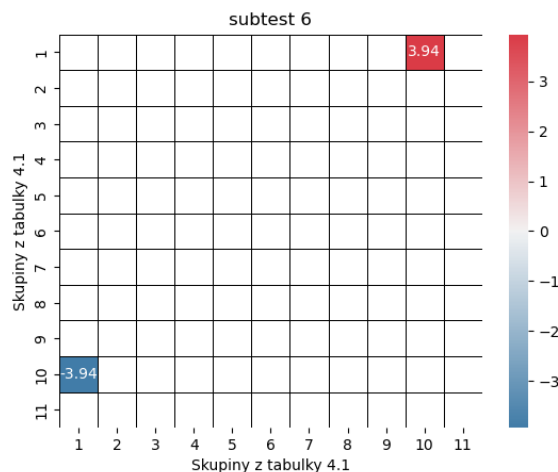
**Hladina 1:** Na hladině 5 % byly testovány hypotézy z kapitoly 4.1.1. V první testované věkové skupině byl odhalen výsledek  $p$ -value  $\approx 0.022$ . Tímto byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení.

**Hladina 2:** Dalším krokem bylo zjistit, zda jednotlivé skupiny nevykazují statisticky významné rozdíly mezi jejich příslušnými rozptyly. Proto byl na základě tvrzení o tom, že data nepochází z normálního rozdělení, vybrán Leveneův test a na hladině významnosti 5 % byly testovány hypotézy z kapitoly 4.1.2 o shodě všech rozptylů. Zde byla zamítnuta nulová hypotéza  $H_0$  o shodnosti rozptylů mezi jednotlivými věkovými skupinami ( $p$ -value  $\approx 0.015$ ).

**Hladina 3:** Dále byl dodržen postup uvedený ve schématu 4.1 a byl proveden Kruskal-Wallisův test. Na hladině významnosti 5 % byly testovány hypotézy o shodě mediánů z kapitoly 4.1.3. S  $p$ -value  $\approx 0.001$  byla zamítnuta nulová hypotéza  $H_0$  o shodnosti všech mediánů.

Dále bylo testem mnohonásobného porovnávání zkoumáno, mezi kterými skupinami dochází ke statisticky významným rozdílům.

Z grafu 4.6 je patrné, že statisticky významný rozdíl středních hodnot se opět vyskytuje mezi dvěma skupinami, které jsou od sebe věkově téměř nejvzdálenější.



Obrázek 4.6: Mapa významných rozdílů středních hodnot pro sloupec BL vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`

#### 4.1.10 Škálové skóre - Subtest 7

Další data, která byla testována, se nachází ve všech listech vstupního souboru `Student_groups_by_age_valid_2.xlsx`, ve sloupci BN. Data z tohoto sloupce jsou zaměřena na **koordinaci horních končetin** viz kapitola [Výpočet výsledku testu](#).

**Hladina 1:** Na hladině významnosti 5 % bylo testováno, zda výsledky testů jednotlivých skupin pochází z normálního rozdělení, jak je uvedeno v kapitole 4.1.1. Byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení, přestože  $p$ -value = 0.0498. U takovéto hodnoty  $p$ -value už by se dalo říci, že data téměř pochází z normálního rozdělení, ale byla striktně dodržena hladina významnosti 5 %. Uvedená  $p$ -value byla nalezena při testování dat 1. věkové kategorie.

**Hladina 2:** Dále byly na hladině 5 % testovány hypotézy z kapitoly 4.1.2. Na základě předchozího výsledku o rozdělení dat byl k těmto účelům použit Leveneův test. Po provedení tohoto testu nebyla zamítnuta nulová hypotéza  $H_0$  o shodnosti rozptylů ( $p$ -value  $\approx$  0.35).

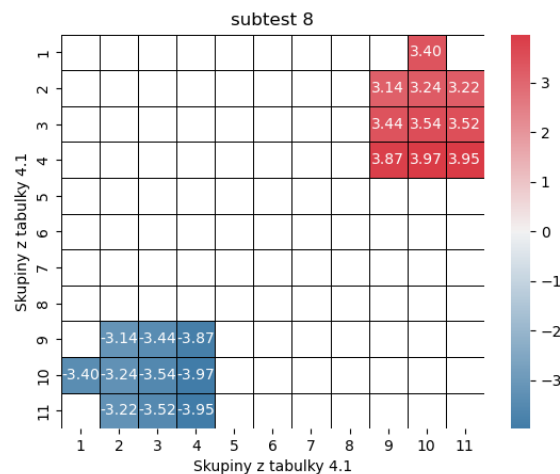
**Hladina 3:** Ke zjištění, zda jsou mezi daty významné statistické rozdíly mezi mediány, byl zvolen Kruskal-Wallisův test a na hladině významnosti 5 % byly testovány hypotézy o shodnosti mediánů z kapitoly 4.1.3. Nebyla zde zamítnuta nulová hypotéza  $H_0$  o shodnosti všech mediánů ( $p$ -value  $\approx$  0.33). Tím pádem nebyly výsledky dále zkoumány.

#### 4.1.11 Škálové skóre - Subtest 8

Výsledky testů zaměřených na **sílu** jsou uvedeny ve sloupci BO souboru `Student_groups_by_age_valid_2.xlsx`.

**Hladina 1:** Prvním krokem bylo testování hypotéz, na hladině významnosti 5 %, o tom, zda data pochází z normálního rozdělení. Tyto testované hypotézy jsou uvedeny v kapitole 4.1.1. U výsledků 1. věkové skupiny byla zamítnuta nulová

hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení ( $p$ -value  $\approx 0.02$ ) a tato skutečnost byla rozšířena na data ze všech skupin.



Obrázek 4.7: Mapa významných rozdílů středních hodnot pro sloupec B0 vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`

**Hladina 2:** Na hladině významnosti 5 % byly dále testovány hypotézy z kapitoly 4.1.2 o shodnosti všech rozptylů. Jelikož data nepochází z normálního rozdělení, byla data testována pomocí Leveneova testu. Po tomto výpočtu nezamítáme nulovou hypotézu  $H_0$  o shodnosti rozptylů ( $p$ -value  $\approx 0.43$ ).

**Hladina 3:** Dále byl použit Kruskal-Wallisův test o shodnosti mediánů. Na hladině významnosti 5 % byly testovány hypotézy z kapitoly 4.1.3. Nulová hypotéza  $H_0$  byla zamítnuta s ohledem na to, že  $p$ -value =  $2.03 \times 10^{-7}$ .

Testem mnohonásobného porovnávání bylo dále zkoumáno, mezi kterými skupinami dochází ke statisticky významným rozdílům.

V grafu 4.7 je viditelný rozdíl mezi středními hodnotami jednotlivých skupin. Opět se jedná výhradně o významné statistické rozdíly mezi skupinami, jež jsou od sebe věkově značně vzdálené.

#### 4.1.12 Standardní skóre - Test 1

V této a dalších 3 kapitolách byly testovány výsledky **Standardních skóre** mezi jednotlivými skupinami. Výpočet **Standardního skóre** je podrobněji rozebrán v kapitole [Výpočet výsledku testu](#).

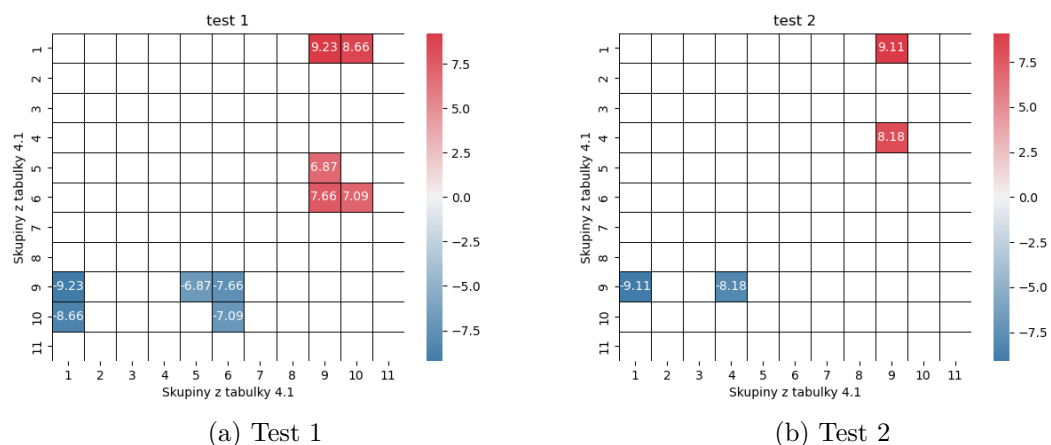
V této kapitole byly testovány výsledky ve sloupci BP vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`, kterými je určena **kontrola v jemné motorice**.

**Hladina 1:** Stejně jako tomu bylo doposud, byly prvním testem zkoumány hypotézy o tom, zda data pochází z normálního rozdělení z kapitoly 4.1.1. Po provedení testu nebyla, na hladině významnosti 5 %, zamítnuta hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení, a tedy po otestování všech 11 skupin platí, že data pochází z normálního rozdělení.

**Hladina 2:** Dále byl na hladině významnosti 5 % proveden Bartlettův test o shodě rozptylů. Byly testovány hypotézy z kapitoly 4.1.2. Zde nebyla zamítnuta nulová hypotéza  $H_0$  o shodě všech rozptylů ( $p$ -value  $\approx 0.28$ ).

**Hladina 3:** Dalším krokem, uvedeným v diagramu 4.1, byl na základě předchozích výsledků test o shodě středních hodnot ANOVA. Díky spočtené  $p$ -value  $\approx 0.0001$  byla, na hladině významnosti 5 %, zamítnuta nulová hypotéza  $H_0$  o shodě středních hodnot.

Dále bylo testem mnohonásobného porovnávání zkoumáno, mezi kterými skupinami dochází ke statisticky významným rozdílům.



Obrázek 4.8: Mapy významných rozdílů středních hodnot pro sloupce BP a BQ vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`

Z grafu 4.8 (a) je patrné, že rozdíly ve středních hodnotách věkových skupin, uvedených na osách, se chovají stejně jako v předchozích kapitolách a to tak, že jsou zde pozorovatelné významné statistické rozdíly mezi věkově vzdálenějšími skupinami studentů.

### 4.1.13 Standardní skóre - Test 2

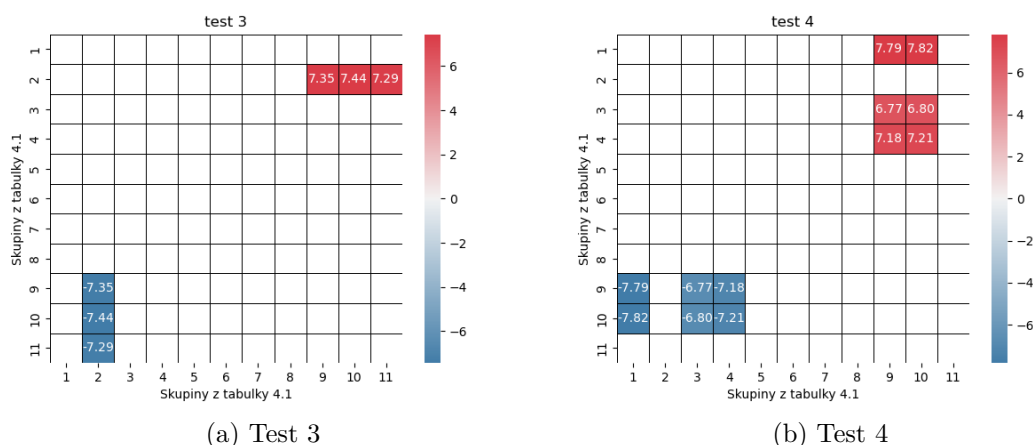
Ve sloupci BQ vstupního datového souboru `Student_groups_by_age_valid_2.xlsx` jsou výsledky **Standardního skóre** zaměřeny na **manuální koordinaci**, viz [Výpočet výsledku testu](#).

**Hladina 1:** Na hladině významnosti 5 % byly testovány hypotézy z kapitoly 4.1.1 o tom, zda data pochází z normálního rozdělení. Na základě výsledku testu  $p$ -value  $\approx 0.036$ , který byl nalezen již v první testované skupině, byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení.

**Hladina 2:** Dalším krokem bylo provedení Leveneova testu o shodnosti rozptylů a otestování hypotéz v kapitole 4.1.2. Zde byla, na hladině 5 %, zamítnuta nulová hypotéza  $H_0$  o shodnosti všech rozptylů ( $p$ -value  $\approx 0.03$ ).

**Hladina 3:** Poté byl proveden Kruskal-Wallisův test. Byly testovány hypotézy o shodě všech mediánů z kapitoly 4.1.3. Na hladině významnosti 5 % byla zamítnuta





Obrázek 4.9: Mapy významných rozdílů středních hodnot pro sloupce BR a BS vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`

nulová hypotéza  $H_0$  o shodě všech mediánů.

Vzhledem k předchozím výsledkům bylo nutné prozkoumat, mezi kterými skupinami dochází k významným statistickým rozdílům. K tomuto účelu byl proveden test mnohonásobného porovnávání. Tímto testem byly zjištěny rozdíly mezi středními hodnotami zobrazené v grafu 4.8 (b). Nadále bylo pozorováno, že významné rozdíly jsou odhaleny mezi věkově vzdálenými skupinami.

#### 4.1.14 Standardní skóre - Test 3

Dalším testovaným sloupcem je sloupec BR, ve kterém se, ve všech listech vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`, nachází výsledky zaměřené na **celkovou tělesnou koordinaci**, viz [Výpočet výsledku testu](#).

**Hladina 1:** Stejně jak tomu bylo doposud, bylo prvním krokem testování hypotéz z kapitoly 4.1.1 o tom, zda data pochází z normálního rozdělení. Zde byla po kontrole první skupiny zjištěna  $p$ -value  $\approx 0.004$ . Jelikož byl test proveden na hladině významnosti 5 %, je zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení.

**Hladina 2:** Vzhledem k výsledku předchozího testu o normalitě dat bylo dále přistoupeno k Leveneově testu o shodnosti všech rozptylů. Zde byla zjištěna  $p$ -value  $\approx 0.96$  a nebyla zamítnuta nulová hypotéza  $H_0$  o shodnosti rozptylů mezi všemi 11 testovanými skupinami.

**Hladina 3:** Na základě předchozích výsledků byl pro další testování dat zvolen Kruskal-Wallisův test. Tímto testem byly, na hladině významnosti 5 %, otestovány hypotézy z kapitoly 4.1.3. Byla zamítnuta nulová hypotéza  $H_0$  o shodnosti všech mediánů.

Data byla dále analyzována testem mnohonásobného porovnávání pro zjištění, mezi kterými skupinami dochází k statisticky významným rozdílům.

Opět byl pro účely lepšího zobrazení výsledků vytvořen graf, na němž jsou zobra-

zeny rozdílů ve středních hodnotách mezi jednotlivými skupinami, které mezi sebou vykazují významnějších statistických rozdílů.

Pro tuto kapitolu je to graf 4.9 (a), ve kterém je viditelné, že statisticky významnějších rozdílů mezi skupinami je dosaženo mezi věkově vzdálenými skupinami.

#### 4.1.15 Standardní skóre - Test 4

Další výsledky, které je možné porovnávat v rámci věku studenta, se nachází ve sloupci BS vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`. Tyto výsledky jsou zaměřeny na **sílu a agilitu**.

**Hladina 1:** Na hladině významnosti 5 % byl proveden test z kapitoly 4.1.1 o tom, zda data pochází z normálního rozdělení. První skupina však tuto skutečnost nevykazuje ( $p$ -value  $\approx 0.02$ ), proto byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení.

**Hladina 2:** Jelikož data nepochází z normálního rozdělení, bylo potřebné využít Leveneova testu a otestovat hypotézy popsané v kapitole 4.1.2. Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  o shodnosti všech rozptylů ( $p$ -value  $\approx 0.001$ ).

**Hladina 3:** Díky zamítnutí nulové hypotézy z Leveneova testu bylo dále nutné použít Kruskal-Wallisův test a otestovat, na hladině významnosti 5 %, hypotézy popsané v kapitole 4.1.3. S  $p$ -value =  $4.6 \times 10^{-5}$  byla zamítnuta nulová hypotéza  $H_0$  o shodě všech mediánů.

Dále byl proveden test mnohonásobného porovnávání, aby bylo zjištěno, mezi kterými skupinami dochází ke statisticky významným rozdílům. Z tohoto testu byl vytvořen graf 4.9 (b), na kterém jsou vyobrazeny rozdílů mezi středními hodnotami jednotlivých věkových skupin a bylo dále kontrolováno tvrzení, že statisticky významné skupiny se nacházejí věkově dále od sebe.

#### 4.1.16 Celkový motorický koeficient - TMC

V této kapitole byla provedena analýza **Celkového motorického koeficientu**, jakožto celkového dosaženého skóre v testu. V kapitole [Výpočet výsledku testu](#) bylo pojednáváno o tom, jak je tento celkový výsledek dopočítáván.

Výsledky **Celkového motorického koeficientu** jsou uvedeny ve sloupci BT vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`.

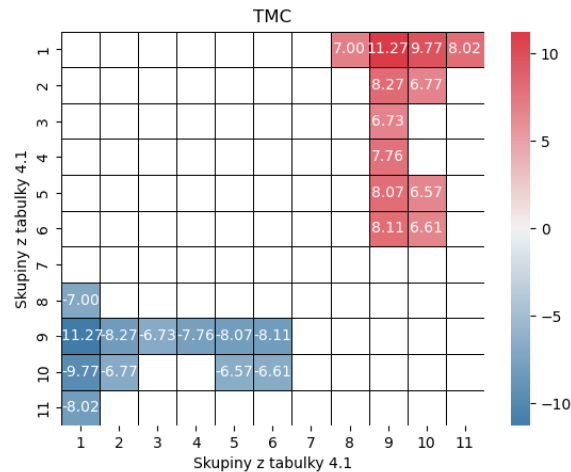
**Hladina 1:** Nejdříve byly na hladině významnosti 5 % testovány hypotézy z kapitoly 4.1.1 o tom, zda data pochází z normálního rozdělení. U 1. testované skupiny byla zjištěna  $p$ -value  $\approx 0.013$ , a tím pádem byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení.

**Hladina 2:** Dalším krokem v analýze **Celkového motorického koeficientu** bylo, na hladině 5 %, provést Leveneův test o shodě rozptylů a otestovat tak hypotézy z kapitoly 4.1.2. Po výpočtu Leveneova testu nebyla zamítnuta nulová hypotéza  $H_0$  o shodě všech rozptylů ( $p$ -value  $\approx 0.6$ ).

**Hladina 3:** Poté byly testovány, na hladině významnosti 5 %, hypotézy z kapitoly 4.1.3 o shodě všech mediánů. Pro tento účel byl vzhledem k předchozím výsledkům

vybrán Kruskal-Wallisův test. Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  o shodnosti všech mediánů ( $p$ -value =  $6.3 \times 10^{-8}$ ).

Testem mnohonásobného porovnávání bylo poté dále zkoumáno, mezi kterými věkovými skupinami dochází ke statisticky významným rozdílům.



Obrázek 4.10: Mapa významných rozdílů středních hodnot pro sloupec BT vstupního datového souboru Student\_groups\_by\_age\_valid\_2.xlsx

Graf 4.10 zobrazuje, mezi kterými skupinami dochází ke statisticky významným rozdílům u celkového dosaženého skóre v Bruininks-Oseretsky testu. Dále jsou zde vidět rozdíly jednotlivých středních hodnot mezi věkovými skupinami. Z grafu se dá tedy vyčíst, že i u celkového dosaženého skóre dochází ke statisticky významným rozdílům mezi věkovými skupinami, mezi kterými je významnější věkový rozdíl.

#### 4.1.17 Závěr analýzy dat na základě věku

V kapitolách 4.1.4 - 4.1.16, které se věnují analýze dat na základě věku byly nalezeny významnější statistické rozdíly, mezi 11 věkovými skupinami, v 11 případech. Ve zbylých dvou případech a to v kapitolách 4.1.4 a 4.1.10 nebyly nalezeny významnější statistické rozdíly a na základě výsledků nebylo provedeno další zkoumání.

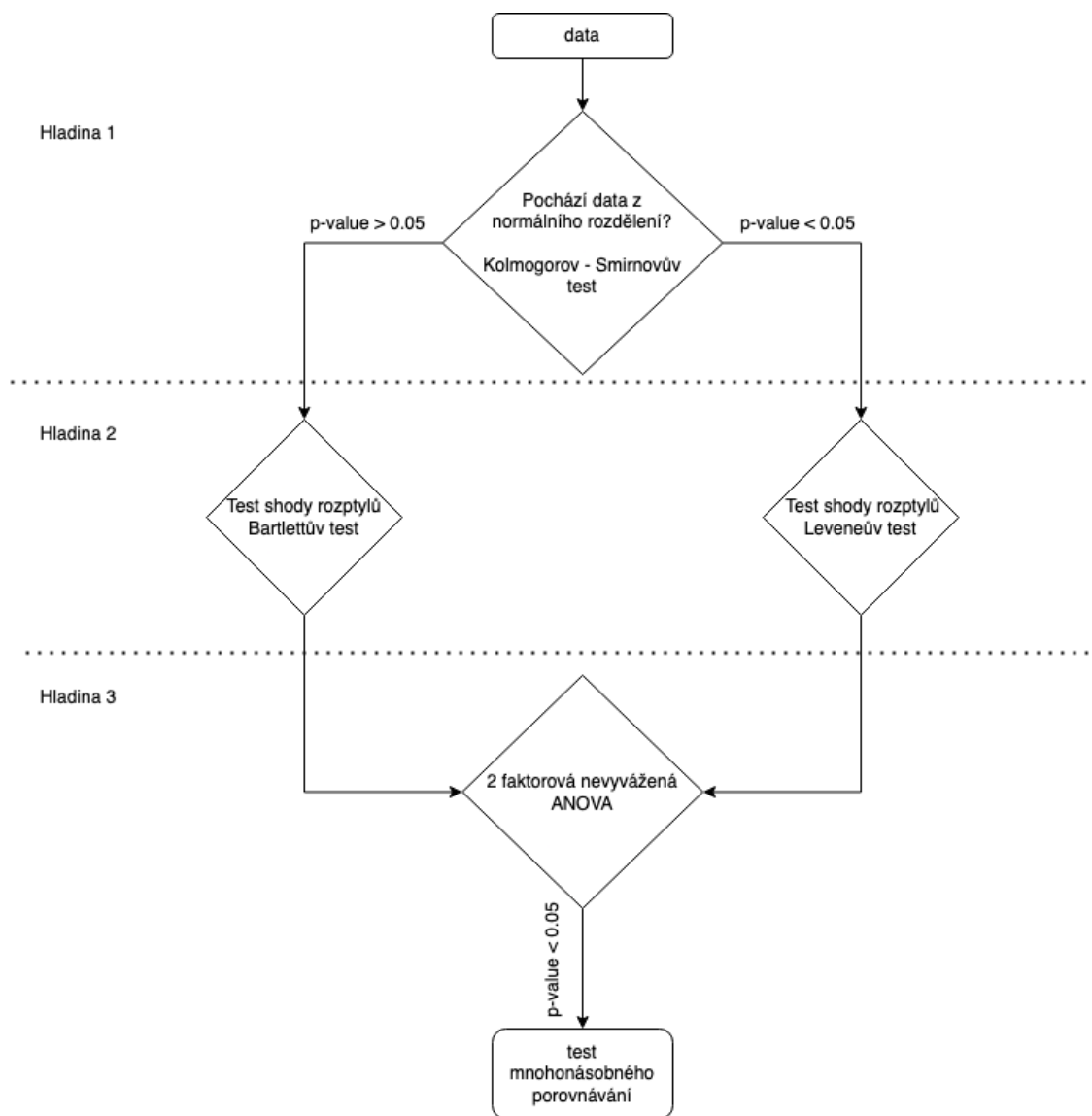
Ve všech ostatních kapitolách byly nalezeny významnější statistické rozdíly. Z uvedených grafů, rozdílů středních hodnot, je viditelné, že data vykazují rostoucí či klesající trendy vzhledem k věku. Bloky skupin 8, 9, 10, 11 vykazují statisticky odlišnější výsledků, než-li bloky skupin 1, 2, 3, 4, 5, 6, 7.

Z grafů se dá také sledovat, že významnějších statistických rozdílů vždy nabývají věkové skupiny, které jsou od sebe relativně vzdálené a nikdy mezi sebou nevykazují statistické rozdíly věkové skupiny, jenž jsou hned vedle sebe a tím pádem je mezi nimi nejmenší věkový rozdíl.

Zřejmě nejčastěji zastoupenou skupinou, která vykazuje statistické rozdíly proti ostatním skupinám je 9. věková skupina, která ku příkladu v kapitole o celkovém motorickém koeficientu 4.1.16 vykazuje rozdíl středních hodnot se 6 jinými věkovými skupinami.

## 4.2 Analýza dat na základě pohlaví a věku

V kapitole 4.1 bylo prozkoumáno, zda jednotlivé věkové skupiny demonstrované v tabulce 3.1 mezi sebou nevykazují významné statistické rozdíly středních hodnot. V této kapitole byl přidán vliv pohlaví, jakožto další faktor, kterým by mohly být výsledky ovlivněny.



Obrázek 4.11: Schéma prováděných dvoufaktorových testů pro analýzu dat na základě pohlaví a věku

V souboru `Student_groups_by_age_valid_2.xlsx` byly kontrolovány sloupce BH až BT. A byl zjišťován vliv pohlaví (1. faktor) a věku (2. faktor) na celkové dosažené výsledky v uvedených sloupcích. Data v těchto sloupcích jsou zaměřena na:

- Škálové skóre
- Standardní skóre
- Celkový motorický koeficient

Výběr statistických testů v této kapitole probíhal dle schématu 4.11.

## Teorie

Statistické testy a hypotézy zmíněné v následujících kapitolách jsou obecně popsány v této literatuře<sup>8</sup>:

- Kolmogorov-Smirnovův test (resp. Lillieforsův test)
  - str. 104 [1], str. 233 [4]
- Bartlettův a Leveneův test
  - str. 121 [1], str. 205 [6]
- dvoufaktorová nevyvážená ANOVA
  - str. 509 [9]

### 4.2.1 Hladina 1

Dle schématu 4.11 bylo na této hladině testováno, zda data ze souboru `Student_groups_by_age_valid_2.xlsx` pochází z normálního rozdělení. Bylo k tomu použito stejného testu jako v kapitole 4.1.1, a tím byly testovány i stejné hypotézy:

- $H_0 : X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ , kde  $X_1, X_2, \dots, X_n$  jsou pozorování, která tvoří vzorek z normálního rozdělení s neznámými parametry  $\mu$  a  $\sigma^2$ . Symbol  $\sim$  znamená "je distribuováno jako" a  $N(\mu, \sigma^2)$  značí normální rozdělení s parametry  $\mu$  a  $\sigma^2$ .
- $H_A : X_1, X_2, \dots, X_n \not\sim N(\mu, \sigma^2)$ , kde  $X_1, X_2, \dots, X_n$  jsou pozorování, která netvoří vzorek z normálního rozdělení s parametry  $\mu$  a  $\sigma^2$ . Symbol  $\not\sim$  znamená "není distribuováno jako" a  $N(\mu, \sigma^2)$  značí normální rozdělení s parametry  $\mu$  a  $\sigma^2$ .

### 4.2.2 Hladina 2

Dalším krokem ve schématu 4.11 bylo otestovat za pomoci Bartlettova či Leveneova testu shodnost rozptylů. Postup i výsledky testu jsou shodné s kapitolou 4.1.2, proto zde byly testovány následující hypotézy:

- $H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$
- $H_A : \exists i, j \in 1, 2, \dots, k, i \neq j : \sigma_i^2 \neq \sigma_j^2$ , kde  $k$  je počet srovnávaných skupin,  $\sigma_i^2$  a  $\sigma_j^2$  jsou rozptyly  $i$ -té a  $j$ -té skupiny, které se porovnávají.

---

<sup>8</sup>Zmíněna je opět pouze počáteční strana kapitoly

### 4.2.3 Hladina 3

V této hladině byla na vstupních datech provedena **dvoufaktorová nevyvážená ANOVA**. Tento test byl vybrán přesto, že některé výsledky z předešlých hladin nemusí sedět se vstupními požadavky tohoto testu.

V "pravé" větvi grafu 4.11, kde je proveden Leveneův test, byl jako následující nejdříve uvažován Friedmanův test, který nepředpokládá normalitu dat. Tento test ale předpokládá vyváženost jednotlivých výběrů<sup>9</sup>. To znamená, že ve všech 11 věkových skupinách by musel být stejný počet záznamů, což splněno není viz tabulka 3.1. Vhodněji se jevílo použít dvoufaktorovou nevyváženou ANOVU, která by mohla docílit větší přesnosti.

Pro dvoufaktorovou nevyváženou ANOVU byly testovány tyto hypotézy:

Hypotéza faktoru A se zabývá vlivem tohoto (prvního) faktoru na závislou proměnnou.

- $H_0$  : Neexistuje žádný vliv faktoru A na závislou proměnnou
- $H_A$  : Faktor A má vliv na závislou proměnnou

Hypotéza faktoru B se zabývá vlivem tohoto (druhého) faktoru na závislou proměnnou.

- $H_0$  : Neexistuje žádný vliv faktoru B na závislou proměnnou
- $H_A$  : Faktor B má vliv na závislou proměnnou

Pokud to dále bylo zapotřebí a za pomoci dvoufaktorové nevyvážené ANOVY byly nalezeny statisticky významnější rozdíly, bylo dále testem mnohonásobného porovnávání zjišťováno, která data mezi sebou vykazují odchylky.

## Statistické vyhodnocení

### 4.2.4 Škálové skóre - Subtest 1

V sloupci BH datového souboru `Student_groups_by_age_valid_2.xlsx`, ve všech jeho listech<sup>10</sup>, jsou výsledky zaměřeny na **přesnost v jemné motorice** viz [Výpočet výsledku testu](#). Bylo zde zkoumáno, zda má pohlaví a věk studentů vliv na jejich dosažené celkové point skóre v této kategorii.

**Hladina 1:** Nejprve byla zkoumána normalita dat a na základě  $p$ -value  $\approx 0.04$  byla, na hladině významnosti 5 %, zamítnuta nulová hypotéza  $H_0$  o tom, zda data pocházejí z normálního rozdělení. Tato hypotéza je popsána v kapitole 4.2.1.

<sup>9</sup>V knize [4] je uveden příklad Friedmanova testu na vyvážených datech. Obdobně je tento test popisován v dokumentaci Scipy [12] a Matlabu [8]. Kniha [9] tento test zmiňuje u nevyvážené ANOVY, čímž by se dalo chápat, že je možné využít Friedmanův test i na nevyvážených datech. Jelikož byly výpočty testů dosaženy za pomoci Pythonu či Matlabu, neuvažuje se Friedmanův test jako možný pro nevyvážená data, se kterými je v této práci počítáno.

<sup>10</sup>V každém listu je uložena příslušná věková skupina z tabulky 3.1.

**Hladina 2:** Poté byl proveden Leveneův test o shodě všech rozptylů a, na hladině významnosti 5 %, byly testovány hypotézy popsané v kapitole 4.2.2. Byla přijata nulová hypotéza  $H_0$  o shodě všech rozptylů vzhledem k  $p$ -value  $\approx 0.6$ .

**Hladina 3:** Dalším krokem vyobrazeným v grafu 4.11 bylo provedení dvoufaktorové nevyvážené ANOVY. Byly testovány hypotézy, na hladině významnosti 5 %, z kapitoly 4.2.3, a to pro oba vstupní faktory (pohlaví, věk):

- Pohlaví: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.008$ )
- Věk: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.56$ )

Po otestování výše uvedených hypotéz a zamítnutí nulové hypotézy u faktoru pohlaví, byl proveden test mnohonásobného porovnávání, kterým nebyly zjištěny žádné statisticky významnější odchylky, které se přesto dají očekávat. Pro nenalezení těchto odchylek, byl obdobný test proveden i v softwaru Matlab, aby došlo ke kontrole výsledků. V Matlabu byly zjištěny stejné hodnoty u dvoufaktorové nevyvážené ANOVY i u testu mnohonásobného porovnávání.

## 4.2.5 Škálové skóre - Subtest 2

Ve vstupním datovém souboru `Student_groups_by_age_valid_2.xlsx`, ve všech jeho listech, se ve sloupci BI nachází výsledky zaměřené na **integraci v jemné motorice**. Stejně jako v předešlé kapitole je tento fakt zmíněn v kapitole [Výpočet výsledku testu](#). Bylo zde zkoumáno, zda má pohlaví a věk studentů vliv na jejich dosažené celkové point skóre v této kategorii.

**Hladina 1:** Prvním krokem bylo otestovat, zda data pocházejí z normálního rozdělení. Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.1 vzhledem k  $p$ -value  $\approx 0.04$ .

**Hladina 2:** Po provedení Leveneova testu nebyla, na hladině významnosti 5 %, zamítnuta nulová hypotéza  $H_0$  o shodnosti všech rozptylů ( $p$ -value  $\approx 0.98$ ) z kapitoly 4.2.2.

**Hladina 3:** Provedením dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3:

- Pohlaví: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.65$ )
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 3.7 \times 10^{-10}$ )

Poté byl proveden test mnohonásobného porovnávání a bylo zjištěno, mezi kterými skupinami dochází k významnějším statistickým rozdílům. Opět byl tento krok proveden i softwarem Matlab, kvůli kontrole výsledků, které mohou být v Matlabu přesnější. Pro přehlednost textu je tento graf v příloze A.0.2, jelikož bylo nalezeno celkem 15 skupin, které vykazují rozdíly, a grafy by zde byly nečitelné.

### 4.2.6 Škálové skóre - Subtest 3

Všechny listy vstupního datového souboru `Student_groups_by_age_valid_2.xlsx` obsahují, ve sloupci BJ, informace o výsledcích zaměřených na **manuální zručnost** a byl zkoumán vliv pohlaví a věku na tyto výsledky.

**Hladina 1:** Na hladině významnosti 5 % byly testovány hypotézy z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení, a s  $p$ -value  $\approx 0.04$  byla zamítnuta nulová hypotéza  $H_0$ .

**Hladina 2:** Po provedení Leveneova testu, vzhledem ke schématu 4.11, nebyla, na hladině významnosti 5 %, zamítnuta nulová hypotéza  $H_0$  o shodnosti rozptylů ( $p$ -value  $\approx 0.28$ ).

**Hladina 3:** Provedením dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3:

- Pohlaví: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.26$ )
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.0011$ )

Dále byl proveden test mnohonásobného porovnávání, kterým byl za pomoci Pythonu nalezen 1 záznam, který vykazuje statisticky významnější rozdíly (rozdíl střední hodnoty je 6.3) mezi děvčaty 7. věkové skupiny a chlapci 9. věkové skupiny<sup>11</sup>.

### 4.2.7 Škálové skóre - Subtest 4

V této kapitole byl testován sloupec BK datového souboru. Výsledky v tomto sloupci jsou, ve všech listech souboru `Student_groups_by_age_valid_2.xlsx`, zaměřeny na **bilaterální koordinaci**.

**Hladina 1:** Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení  $p$ -value  $\approx 0.001$ .

**Hladina 2:** Dále byl proveden Leveneův test a na hladině významnosti 5 % byly testovány hypotézy z kapitoly 4.2.2, kde nebyla zamítnuta nulová hypotéza  $H_0$  o shodě rozptylů mezi skupinami.

**Hladina 3:** Provedením dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

- Pohlaví: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.5$ )
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.002$ )

Jelikož výsledky hypotéz nejsou shodné a vliv na výsledky má věk, a nikoli pohlaví, byl dále proveden test mnohonásobného porovnávání pro bližší prozkoumání. Tímto testem, za pomoci Pythonu, byly objeveny významnější statistické rozdíly mezi děvčaty 2. věkové skupiny a děvčaty 9. věkové skupiny<sup>12</sup>.

<sup>11</sup>Vzhledem k tabulce 3.1 se jedná o děvčata mezi 8 lety a 6 měsíci až 8 lety a 11 měsíci, a o chlapce mezi 9 lety a 6 měsíci až 9 lety a 11 měsíci.

<sup>12</sup>Vzhledem k tabulce 3.1 se jedná o děvčata mezi 6 lety a 8 měsíci až 6 lety a 11 měsíci, a o děvčata mezi 9 lety a 6 měsíci až 9 lety a 11 měsíci.



## 4.2.8 Škálové skóre - Subtest 5

Ve sloupci BL souboru `Student_groups_by_age_valid_2.xlsx` jsou, ve všech listech, výsledky testů zaměřených na **rovnováhu**. Tento fakt je zmíněn v kapitole [Výpočet výsledku testu](#), v sekci **Celkové point skóre**. Zde bylo zkoumáno, jaký vliv má pohlaví a věk na dosažené škálové skóre v této disciplíně.

**Hladina 1:** Na této hladině byly testovány hypotézy z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení. Na hladině významnosti 5 % zde byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení ( $p$ -value  $\approx 0.02$ ).

**Hladina 2:** Jelikož byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení, byl dále proveden Leveneův test, a tím pádem byly, na hladině významnosti 5 %, testovány hypotézy z kapitoly 4.2.2. Vzhledem k  $p$ -value  $\approx 0.33$  nebyla zamítnuta nulová hypotéza  $H_0$  o shodě všech rozptylů.

**Hladina 3:** Po provedení dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

- Pohlaví: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.24$ )
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.007$ )

Jak lze vidět z uvedených  $p$ -value, tak výsledky hypotéz nejsou shodné, to znamená, že výsledky této podskupiny úkolů jsou závislé především na věku, a nikoli na pohlaví. Proto bylo přistoupeno k testu mnohonásobného porovnávání, kterým nebyly nalezeny žádné významnější statistické rozdíly, přesto zde mohou být očekávány. Stejně výsledky byly dosaženy kontrolou pomocí softwaru Matlab.

## 4.2.9 Škálové skóre - Subtest 6

Soubor `Student_groups_by_age_valid_2.xlsx` ve sloupci BM, v každém listu, obsahuje výsledky testů zaměřených na **rychlost a agility**, jak je uvedeno v kapitole [Výpočet výsledku testu](#). Bylo kontrolováno, jaký má vliv na výsledky v této kategorii úkolů pohlaví a věk.

**Hladina 1:** Na hladině významnosti 5 % bylo zkoumáno, zda data pochází z normálního rozdělení. Byla zamítnuta nulová hypotéza  $H_0$ , z kapitoly 4.2.1, s  $p$ -value  $\approx 0.022$  o tom, zda data pochází z normálního rozdělení.

**Hladina 2:** Dále byl díky předchozímu výsledku zvolen Leveneův test a na hladině významnosti 5 % byly testovány hypotézy z kapitoly 4.2.2. Byla zamítnuta nulová hypotéza  $H_0$  o shodnosti rozptylů ( $p$ -value  $\approx 0.015$ ).

**Hladina 3:** Po provedení dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

- Pohlaví: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.0003$ )
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.0021$ )

Zde jsou výsledky hypotéz shodné, což znamená, že na hladině významnosti 5 % jsou výsledky škálového skóre v disciplíně rychlost a agility závislé nejen na pohlaví, ale i na věku žáků.

Dále bylo zkoumáno, mezi kterými skupinami nastávají významnější statistické rozdíly, k čemuž byl zvolen test mnohonásobného porovnávání. Tímto testem byly zjištěny významnější statistické rozdíly mezi<sup>13</sup>:

- chlapci 1. věkové skupiny a dívkami 10. věkové skupiny (rozdíl středních hodnot: 5.84)
- chlapci 3. věkové skupiny a dívkami 10. věkové skupiny (rozdíl středních hodnot: 4.6)
- chlapci 4. věkové skupiny a dívkami 10. věkové skupiny (rozdíl středních hodnot: 4.82)

V příloze A.0.2 je vyobrazen graf, který demonstruje zjištěné výsledky uvedené v této kapitole.

#### 4.2.10 Škálové skóre - Subtest 7

Další data, která byla testována, se nachází ve všech listech vstupního souboru `Student_groups_by_age_valid_2.xlsx`, ve sloupci BN. Data z tohoto sloupce jsou zaměřena na **koordinaci horních končetin** viz kapitola [Výpočet výsledku testu](#). Zde bylo zkoumáno, jaký dopad má pohlaví a věk žáka na tyto výsledky.

**Hladina 1:** Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení ( $p$ -value = 0.0498).

**Hladina 2:** Dalším testem v pořadí byl Leveneův test. Tímto testem byla zjištěna  $p$ -value  $\approx 0.33$  a tím nebyla, na hladině významnosti 5 %, zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.2 o shodnosti rozptylů.

**Hladina 3:** Po provedení dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

- Pohlaví: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.0006$ )
- Věk: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.23$ )

Z uvedených  $p$ -value je patrné, že výsledky v této kategorii jsou závislé na pohlaví, a nikoli na věku žáků. Proto bylo dále přistoupeno k testu mnohonásobného porovnávání, kterým nebyly v Pythonu nalezeny žádné statisticky významnější rozdíly ve středních hodnotách. Toto tvrzení bylo zkoumáno softwarem Matlab, kterým naopak byly nalezeny významnější statistické rozdíly mezi pohlavími, a to například mezi chlapci z 9. věkové skupiny a dívkami ze 4. a 1. věkové skupiny. Graf z Matlabu lze nalézt v příloze A.0.2.

---

<sup>13</sup>Tyto skupiny jsou vyobrazeny v tabulce 3.1. Jelikož by se jednalo o rozsáhlejší text, nebudou zde již znovu uvedeny.

### 4.2.11 Škálové skóre - Subtest 8

Výsledky testů zaměřených na **sílu** jsou uvedeny ve sloupci BO souboru `Student_groups_by_age_valid_2.xlsx`. Tyto výsledky zde byly zkoumány v rámci pohlaví a věku žáka.

**Hladina 1:** Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení ( $p$ -value = 0.002).

**Hladina 2:** Po provedení Leveneova testu, na hladině významnosti 5 %, nebyla zamítnuta nulová hypotéza  $H_0$  o shodě rozptylů z kapitoly 4.2.2 ( $p$ -value  $\approx$  0.43).

**Hladina 3:** Po provedení dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

- Pohlaví: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx$  0.21)
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx$   $2.5 \times 10^{-7}$ )

Jelikož výsledky hypotéz nejsou shodné a vliv na výsledky má věk, a nikoli pohlaví, byl dále proveden test mnohonásobného porovnávání pro zjištění významnějších statistických rozdílů. Tímto testem bylo zjištěno, že k významnějším rozdílům dochází mezi<sup>14</sup>:

- chlapci 4. věkové skupiny a chlapci 10. věkové skupiny (rozdíl středních hodnot: 4.5)
- chlapci 4. věkové skupiny a chlapci 11. věkové skupiny (rozdíl středních hodnot: 4.4)

### 4.2.12 Standardní skóre - Test 1

V této a dalších 3 kapitolách byly testovány výsledky **Standardních skóre**. Výpočet **Standardního skóre** je podrobněji rozebrán v kapitole [Výpočet výsledku testu](#).

V této kapitole byly testovány výsledky ve sloupci BP vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`, kterými je určena **kontrola v jemné motorice**, a také bylo zkoumáno, jaký dopad na ně má pohlaví a věk žáka.

**Hladina 1:** Na hladině významnosti 5 % nebyla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení vzhledem k tomu, že data ze všech 11 věkových skupin dosahují v tomto testu  $p$ -value  $>$  0.05.

**Hladina 2:** Dále byl na hladině významnosti 5 % proveden Bartlettův test o shodě rozptylů. Nebyla zde zamítnuta nulová hypotéza  $H_0$  o shodě rozptylů z kapitoly 4.2.2 ( $p$ -value  $\approx$  0.28).

**Hladina 3:** Po provedení dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

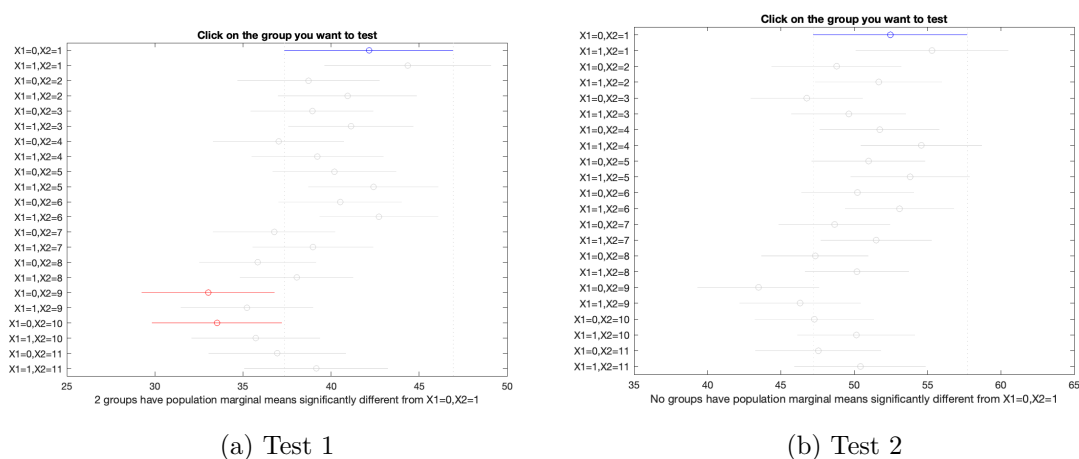
- Pohlaví: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx$  0.011)

---

<sup>14</sup>Tyto skupiny jsou vyobrazeny v tabulce 3.1

- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.0001$ )

Výsledky hypotéz jsou shodné, tj. výsledky zkoumané v této kapitole závisí nejen na pohlaví, ale i na věku žáků. Déle bylo testem mnohonásobného porovnávání zkoumáno, mezi kterými výsledky dochází ke statisticky významnějším rozdílům. Bylo nalezeno celkem 13 různých statisticky významnějších výsledků, které jsou závislé na pohlaví a věku. Rozdíly středních hodnot jsou vyobrazeny v grafu A.4.



Obrázek 4.12: Grafy středních hodnot a 95% intervalových odhadů pro sloupce BP a BQ vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`

### 4.2.13 Standardní skóre - Test 2

Ve sloupci BQ vstupního datového souboru `Student_groups_by_age_valid_2.xlsx` jsou výsledky **Standardního skóre** zaměřeny na **manuální koordinaci**, viz [Výpočet výsledku testu](#). Zde bylo zkoumáno jaký vliv má faktor pohlaví a věku na tyto výsledky.

**Hladina 1:** Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení ( $p$ -value  $\approx 0.036$ ).

**Hladina 2:** Po provedení Leveneova testu, na hladině významnosti 5 %, byla zamítnuta nulová hypotéza  $H_0$  o shodě rozptylů z kapitoly 4.2.2 ( $p$ -value  $\approx 0.03$ ).

**Hladina 3:** Po provedení dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

- Pohlaví: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.0028$ )
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.0063$ )

Z výše uvedených  $p$ -value plyne, že výsledky v této kapitole jsou závislé jak na pohlaví, tak na věku. Vzhledem k tomuto faktu byl dále proveden test mnohonásobného porovnávání. Tímto testem byly nalezeny významnější statistické rozdíly mezi:

- dívkami 5. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 13.68)
- dívkami 6. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 10.61)
- dívkami 7. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 12.71)
- chlapci 4. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 12.31)

Je patrné, že skupina chlapců v 9. věkové skupině vykazuje statistickou odlišnost, která by mohla být dále zkoumána.

### 4.2.14 Standardní skóre - Test 3

Dalším testovaným sloupcem je sloupec BR, ve kterém se, ve všech listech vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`, nachází výsledky zaměřené na **celkovou tělesnou koordinaci**, viz [Výpočet výsledku testu](#). Byly zde zkoumány faktory pohlaví a věku, jimiž by mohly být tyto výsledky ovlivněny.

**Hladina 1:** Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení ( $p$ -value  $\approx 0.004$ ).

**Hladina 2:** Po provedení Leveneova testu, na hladině významnosti 5 %, nebyla zamítnuta nulová hypotéza  $H_0$  o shodě rozptylů z kapitoly 4.2.2 ( $p$ -value  $\approx 0.96$ ).

**Hladina 3:** Po provedení dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

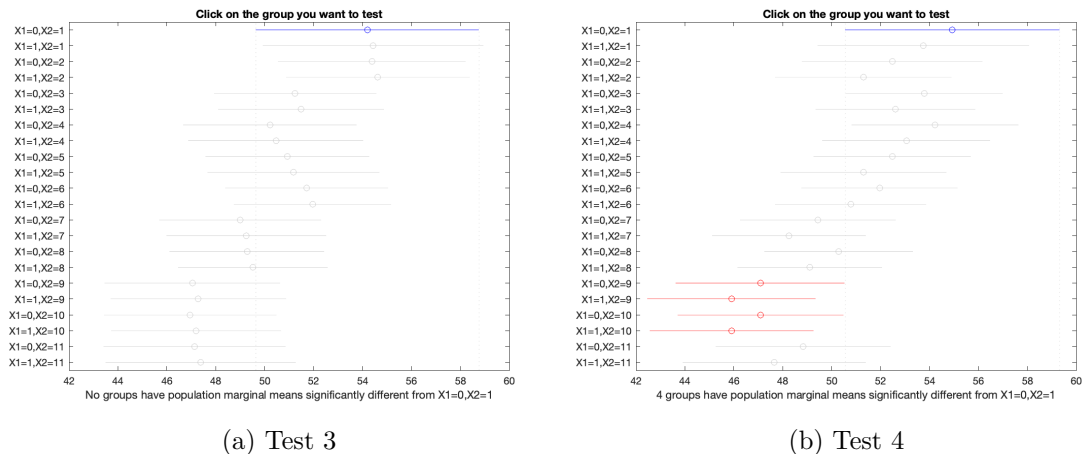
- Pohlaví: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.77$ )
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.0006$ )

Výsledky hypotéz nejsou shodné a byl formulován závěr, že výsledky v této kapitole nejsou závislé na pohlaví, ale jsou závislé na věku. Proto byly testem mnohonásobného porovnávání výsledky dále zkoumány. Tímto testem byl zjištěn rozdíl mezi dívkami 2. věkové skupiny a dívkami 9. věkové skupiny (rozdíl středních hodnot: 10.6).

### 4.2.15 Standardní skóre - Test 4

Další výsledky, které je možné porovnávat v rámci věku a pohlaví studenta, se nachází ve sloupci BS vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`. Tyto výsledky jsou zaměřeny na **sílu a agilitu**. Byly zde zkoumány faktory pohlaví a věku, jimiž by mohly být tyto výsledky ovlivněny.

**Hladina 1:** Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení ( $p$ -value  $\approx 0.002$ ).



Obrázek 4.13: Grafy středních hodnot a 95% intervalových odhadů pro sloupce BR a BS vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`

**Hladina 2:** Po provedení Leveneova testu, na hladině významnosti 5 %, byla zamítnuta nulová hypotéza  $H_0$  o shodě rozptylů z kapitoly 4.2.2 ( $p$ -value  $\approx 0.001$ ).

**Hladina 3:** Po provedení dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

- Pohlaví: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.14$ )
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.000017$ )

Po spočtení výše uvedených  $p$ -value bylo možné konstatovat, že výsledky nejsou závislé na pohlaví, ale jsou závislé na věku. Dále byly tyto výsledky zkoumány testem mnohonásobného porovnávání, kterým bylo zjištěno významnějších statistických rozdílů mezi:

- chlapci 4. věkové skupiny a dívkami 9. věkové skupiny (rozdíl středních hodnot: 10.12)
- chlapci 4. věkové skupiny a dívkami 10. věkové skupiny (rozdíl středních hodnot: 10.56)

### 4.2.16 Celkový motorický koeficient - TMC

V této kapitole byla provedena analýza **Celkového motorického koeficientu**, jakožto celkového dosaženého skóre v testu. V kapitole [Výpočet výsledku testu](#) je pojednáváno o tom, jak je tento celkový výsledek dopočítáván.

Výsledky **Celkového motorického koeficientu** jsou uvedeny ve sloupci BT vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`. Zde byly zkoumány faktory pohlaví a věku, jimiž by mohly být tyto celkové výsledky ovlivněny.

**Hladina 1:** Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.2.1 o tom, zda data pochází z normálního rozdělení ( $p$ -value  $\approx 0.013$ ).

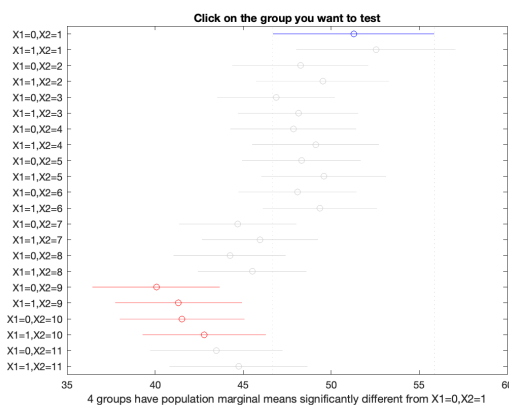
**Hladina 2:** Po provedení Leveneova testu, na hladině významnosti 5 %, nebyla zamítnuta nulová hypotéza  $H_0$  o shodě rozptylů z kapitoly 4.2.2 ( $p$ -value  $\approx 0.6$ ).

**Hladina 3:** Po provedení dvoufaktorové nevyvážené ANOVY, na hladině významnosti 5 %, a otestováním hypotéz z kapitoly 4.2.3 pro faktory pohlaví a věku:

- Pohlaví: nebyla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0.13$ )
- Věk: byla zamítnuta nulová hypotéza  $H_0$  ( $p$ -value  $\approx 0$ )

Výsledky hypotéz nejsou shodné a to znamená, že celkový motorický koeficient není závislý na pohlaví, ale je závislý na věku. Testem mnohonásobného porovnávání bylo dále zjištěno, že k významnějším statistickým rozdílům dochází mezi <sup>15</sup>:

- dívkami 1. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 12.38)
- dívkami 5. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 11.39)
- dívkami 6. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 9.73)
- dívkami 7. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 9.54)
- chlapci 1. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 12.38)
- chlapci 4. věkové skupiny a chlapci 9. věkové skupiny (rozdíl středních hodnot: 11.2)



Obrázek 4.14: Graf středních hodnot a 95% intervalových odhadů pro sloupec BT vstupního datového souboru `Student_groups_by_age_valid_2.xlsx`

<sup>15</sup>Věkové skupiny jsou demonstrovány v tabulce 3.1

### 4.2.17 Závěr analýzy dat na základě pohlaví a věku

V kapitolách 4.2.4, 4.2.9, 4.2.10, 4.2.12 a 4.2.13 byl zjištěn vliv pohlaví na závislou proměnnou. Naopak u faktoru věku nebyla zamítnuta nulová hypotéza  $H_0$  pouze ve 2 případech, a to v kapitolách 4.2.4 a 4.2.10, čímž jsou zároveň potvrzeny výsledky z kapitoly 4.1, kde byl v odpovídajících kapitolách konstatován stejný výsledek.

Byly nalezeny statistické rozdíly především mezi chlapci v 9. věkové skupině a ostatními (toto tvrzení je patrné z výsledků v kapitole 4.2.13 a 4.2.16). Tyto výsledky by mohly být dále zkoumány. V této skupině se nachází žáci ze 2 různých základních škol, což je zjistitelné po nahlédnutí do souboru `Student_groups_by_age_valid_2.xlsx`. Ku příkladu by tyto školy mohly být sportovněji zaměřené, a tudíž by žáci dosahovali lepších výsledků v motorickém testu, nebo by se zde mohly projevovat různé sociální faktory, které v této práci nejsou uvažovány.

## 4.3 Korelace mezi jednotlivými úkoly Bruininks-Oseretsky testu

Statistická korelace je míra vztahu mezi dvěma nebo více kvantitativními proměnnými. Korelace tedy ukazuje, do jaké míry se jedna proměnná mění společně s druhou proměnnou. Je vyjádřena korelačním koeficientem, který se pohybuje v intervalu  $[-1, 1]$ . Hodnota  $-1$  znamená dokonalou negativní korelaci, kdy se u jedné proměnné zvyšuje hodnota, zatímco u závislé proměnné se hodnota snižuje. Hodnota  $+1$  znamená dokonalou pozitivní korelaci, kdy obě proměnné rostou nebo klesají společně. Hodnota  $0$  znamená, že mezi proměnnými žádná korelace neexistuje.

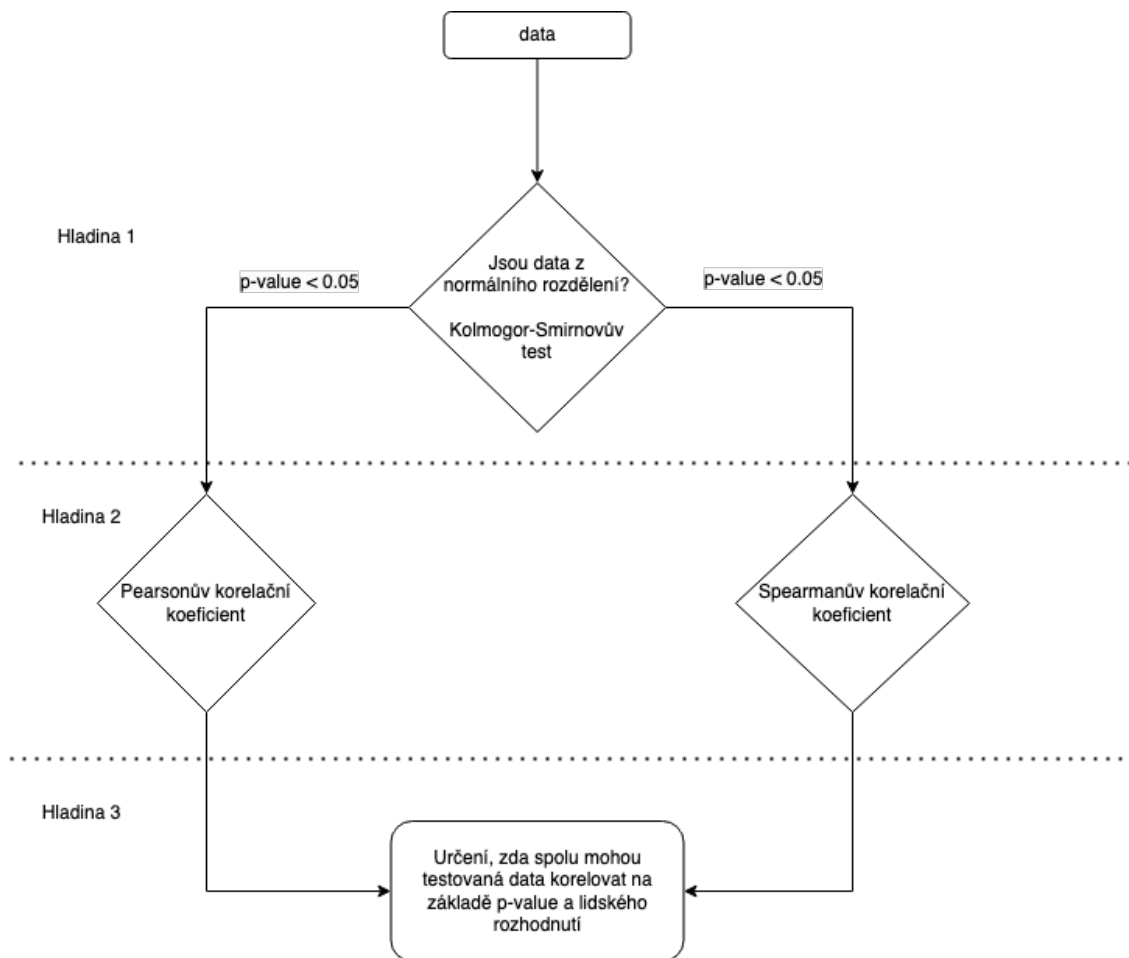
Bruininks-Oseretsky test motorické zdatnosti obsahuje celkem 53 jednotlivých úkolů, které jsou zaměřeny na různé oblasti fyzické zdatnosti viz příloha A.0.1. Jelikož je tento seznam úkolů značně obsáhlý, je také cílem této práce zjistit, které úkoly by mezi sebou mohly do značné míry korelovat, čímž by mohlo dojít ke zkrácení tohoto seznamu. Zároveň tím dojde i k nemalému ušetření času stráveného vykonáváním Bruininks-Oseretsky testu motorické zdatnosti.

Stejně jako v předchozích kapitolách bude k tomuto účelu využíváno datového souboru `Student_groups_by_age_valid_2.xlsx`. Z toho budou vytaženy příslušné sloupce a poté bude proveden statistický test. Tyto statistické testy se budou řídit uvedeným diagramem 4.15.

Korelace mezi jednotlivými úkoly bude vyhodnocována dle příslušných podskupin - subtestových kategorií - v příloze A.0.1. Pro každou podskupinu bude určena korelační matice, která bude zobrazena ve formě grafu. Následně bude tento graf okomentován a bude navrženo či zamítnuto zjednodušení seznamu úkolů.

Diagram 4.15 je opět rozdělen do 3 hladin, jež budou níže popsány. Po stanovení teorie potřebné k testování budou popsány korelační matice jednotlivých podskupin.





Obrázek 4.15: Schéma prováděných korelačních testů pro zjištění úkolů korelujících s jinými

## Teorie

Opět je zde nejdříve uvedená literatura, ze které bylo vycházeno, a ve které se dají najít obecně popsané statistické testy, které jsou zmíněny v následujících kapitolách:

- Kolmogorov-Smirnovův test (resp. Lillieforsův test)
  - str. 104 [1], str. 233 [4]
- Pearsonův (resp. Spearmanův) korelační koeficient
  - str. 243 (resp. 257) [4], str. 276 (resp. 280) [6]

### 4.3.1 Hladina 1

Na této hladině bude testováno, zda data ze vstupního souboru `Student_groups_by_age_valid_2.xlsx` pochází z normálního rozdělení. Stanovené hypotézy a průběh testu je stejný jako v kapitole 4.1.1. Proto tedy jsou testovány následné hypotézy:

- $H_0 : X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ , kde  $X_1, X_2, \dots, X_n$  jsou pozorování, která tvoří vzorek z normálního rozdělení s neznámými parametry  $\mu$  a  $\sigma^2$ . Symbol  $\sim$  znamená "je distribuováno jako" a  $N(\mu, \sigma^2)$  značí normální rozdělení s parametry  $\mu$  a  $\sigma^2$ .
- $H_A : X_1, X_2, \dots, X_n \not\sim N(\mu, \sigma^2)$ , kde  $X_1, X_2, \dots, X_n$  jsou pozorování, která netvoří vzorek z normálního rozdělení s parametry  $\mu$  a  $\sigma^2$ . Symbol  $\not\sim$  znamená "není distribuováno jako" a  $N(\mu, \sigma^2)$  značí normální rozdělení s parametry  $\mu$  a  $\sigma^2$ .

### 4.3.2 Hladina 2

V grafu 4.15 jsou uvedeny 2 možné testy, kterými lze zjistit korelační koeficienty.

Pearsonův korelační koeficient  $\rho$  se používá, jestliže vstupní data pochází z normálního rozdělení a zároveň mohou nabývat spojitéch hodnot. Testuje se lineární závislost/nezávislost dat:

- $H_0 : \rho = 0$
- $H_A : \rho \neq 0$

Spearmanův korelační koeficient se používá, jestliže vstupní data nepochází z normálního rozdělení, ale mohou nabývat spojitéch hodnot. Obdobně jako Kruskal-Wallisův test pracuje Spearmanův test s pořadím hodnot, a nikoli přímo s naměřenými hodnotami. Toto by mohl být problematický faktor u vstupních dat této práce, jelikož úkoly příslušných podskupin nabývají velmi podobných hodnot.

Testují se hypotézy o závislosti/nezávislosti náhodných veličin:

- $H_0 : X \perp\!\!\!\perp Y$
- $H_A : X \not\perp\!\!\!\perp Y$

### 4.3.3 Hladina 3

Jelikož korelační testy někdy vykazují korelaci mezi úkoly, jež spolu nemají zjevný vztah, musí být výsledné hodnoty jednotlivých korelačních testů dále prozkoumány. Na základě p-value a korelační matice byla použita příloha A.0.1, kde jsou vypsány jednotlivé testy, a poté bylo vyhodnoceno, zda spolu příslušné úkoly mohou korelovat či nikoli.

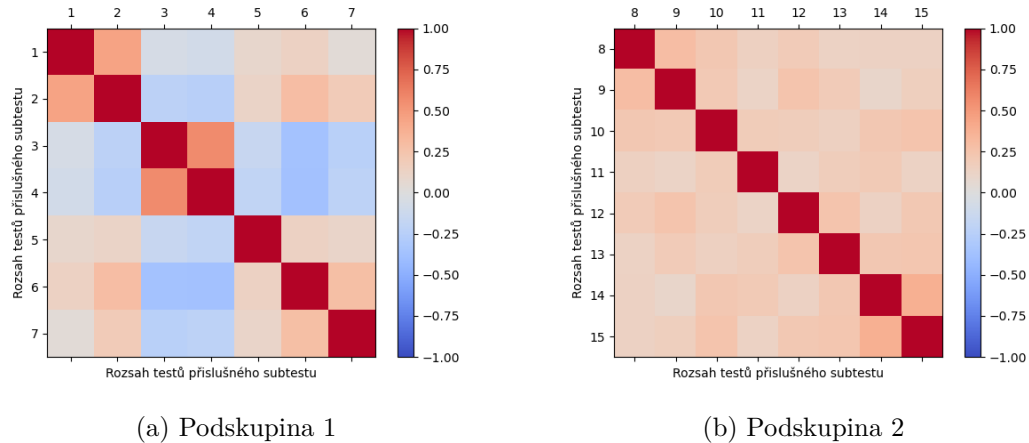
## Statistické vyhodnocení

### 4.3.4 Korelační analýza mezi úkoly podskupiny 1

V této kapitole budou na korelaci testovány jednotlivé úkoly z podskupiny 1, viz příloha A.0.1. Tyto úkoly jsou stejně jako v kapitole 4.1.4 zaměřeny na **přesnost v jemné motorice**.

**Hladina 1:** Nejdříve byl proveden test, na hladině významnosti 5 %, o tom, zda data pochází z normálního rozdělení. Zde byla zamítnuta nulová hypotéza  $H_0$  z kapitoly 4.1.1 o tom, zda data pochází z normálního rozdělení ( $p$ -value =  $1 \times 10^{-3}$ ).

**Hladina 2:** Po zamítnutí nulové hypotézy  $H_0$  o tom, zda data pochází z normálního rozdělení, byl dále zjišťován Spearmanův korelační koeficient. Byly tedy testovány hypotézy z kapitoly 4.3.2 na hladině významnosti 5 %.



Obrázek 4.16: Graf Spearmanova korelačního koeficientu pro podskupinu 1 a 2

**Hladina 3:** V grafu 4.16 (a) lze vidět vyšší pozitivní korelaci mezi úkoly 1 a 2 (hodnota z matice = 0.45), dále také mezi úkoly 3 a 4 (hodnota z matice = 0.57)<sup>16</sup>. Po nahlédnutí do seznamu úkolů bylo vyhodnoceno, že tyto úkoly mezi sebou opravdu mohou korelovat, jelikož se jedná téměř o totožné úkoly, které jsou pouze lehce upravené. Díky této skutečnosti by bylo možné ze seznamu úkolů vyřadit vždy jeden úkol ze zmíněné dvojice.

Zbylé úkoly mezi sebou nevykazují jakkoli silnější pozitivní či negativní korelaci, a tedy není další zhodnocení potřebné.

### 4.3.5 Korelační analýza mezi úkoly podskupiny 2

Obdobně, jak je tomu v kapitole 4.1.5, jsou výsledky této podskupiny (subtestu 2) zaměřeny na **integraci v jemné motorice**.

**Hladina 1:** Jako první krok byl proveden test na to, zda data pochází z normálního rozdělení, a byly testovány hypotézy z kapitoly 4.3.1. Na hladině významnosti 5 % byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení ( $p$ -value =  $1 \times 10^{-3}$ ).

**Hladina 2:** Jelikož byla zamítnuta nulová hypotéza  $H_0$  o normalitě dat, bylo dále přistoupeno ke zjišťování Spearmanova korelačního koeficientu. Na hladině významnosti 5 % byly testovány hypotézy z kapitoly 4.3.2.

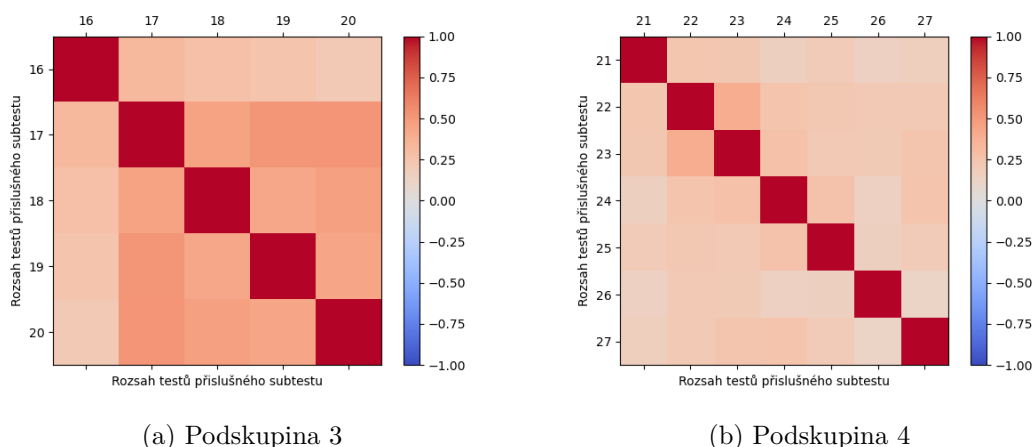
<sup>16</sup>úloha 1 – vyplňování tvaru – kruh, úloha 2 – vyplňování tvaru – hvězda, úloha 3 – kreslení čar ve vyznačené křivé dráze, úloha 4 – kreslení čar ve vyznačené oblé dráze

**Hladina 3:** Zjištěné koeficienty ze Spearmanovy korelační matice byly přeneseny do grafu 4.16 (b), kde je viditelné, že žádné úkoly mezi sebou nevykazují zjevnou korelaci, a tím pádem v této kapitole není navrženo žádné zjednodušení seznamu úkolů.

### 4.3.6 Korelační analýza mezi úkoly podskupiny 3

V této kapitole byla zjišťována korelace mezi úkoly, které jsou zaměřeny na **manuální zručnost**.

**Hladina 1 a 2:** Po testování a následném zamítnutí, na hladině významnosti 5 %, nulové hypotézy  $H_0$  z kapitoly 4.3.1 o tom, zda data pochází z normálního rozdělení ( $p$ -value =  $1 \times 10^{-3}$ ), byl dále zjištěn Spearmanův korelační koeficient viz kapitola 4.3.2.



Obrázek 4.17: Graf Spearmanova korelačního koeficientu pro podskupinu 3 a 4

**Hladina 3:** Matice korelačních koeficientů, pro tuto kapitolu, je předvedena v grafu 4.17 (a). Byly zde opět nalezeny 2 korelační koeficienty, které dosahují hodnoty  $> 0.5$ . Tyto hodnoty byly nalezeny mezi úkoly 17 - 19 a také 17 - 20 <sup>17</sup>.

Jelikož se jedná o úkoly, které jsou zaměřeny na **manuální zručnost**, teoreticky by mohla být smazána úloha 17, jelikož vykazuje korelaci s dvěma dalšími testy.

### 4.3.7 Korelační analýza mezi úkoly podskupiny 4

V další podskupině úkolů se Bruininks-Oseretsky test zaměřuje na **bilaterální koordinaci**.

**Hladina 1:** Nejdříve byly, na hladině významnosti 5 %, testovány hypotézy z kapitoly 4.3.1 o tom, zda data pochází z normálního rozdělení. Zde byla zamítnuta nulová hypotéza  $H_0$  o normalitě dat ( $p$ -value =  $1 \times 10^{-3}$ ).

<sup>17</sup>úloha 17 – přenos mincí, úloha 19 – rovnání karet, úloha 20 – navlékání kostiček na provázek

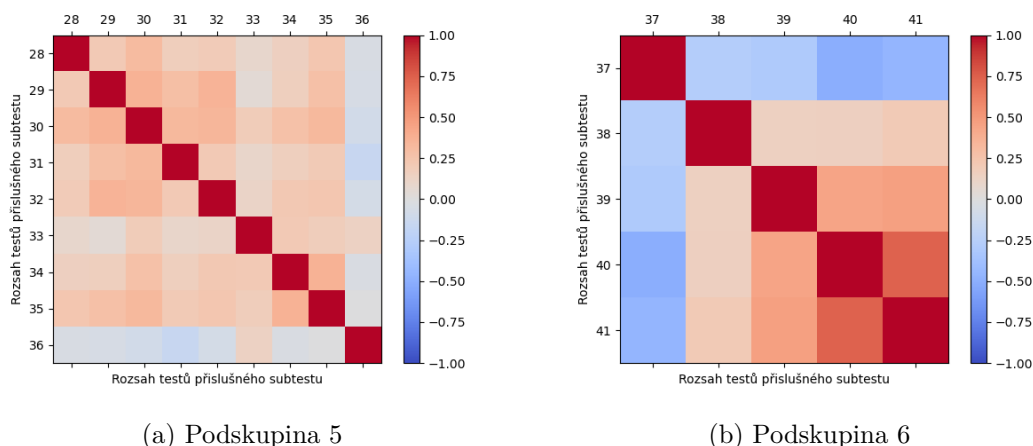
**Hladina 2 a 3:** Poté bylo přistoupeno k testování hypotéz z kapitoly 4.3.2 a zjišťování Spearmanova korelačního koeficientu. Korelační koeficienty jsou zobrazeny v grafu 4.17 (b). Nebyla zjištěna žádná statisticky zajímavější hodnota korelačního koeficientu, a tím pádem zde nebyl seznam úkolů nijak upraven.

### 4.3.8 Korelační analýza mezi úkoly podskupiny 5

Úkoly podskupiny 5 jsou zaměřeny na **rovnováhu**.

**Hladina 1:** V prvním kroku byly, na hladině významnosti 5 %, testovány hypotézy z kapitoly 4.3.1, zda data pochází z normálního rozdělení. Stejně jako v předchozích kapitolách byla zjištěna hodnota  $p\text{-value} = 1 \times 10^{-3}$  a tím zamítnuta nulová hypotéza  $H_0$ . Tato hodnota  $p\text{-value}$  je vrácena vždy, pokud jsou rozdíly mezi distribučními funkcemi normálního rozdělení a zkoumanými daty příliš velké.

**Hladina 2:** Dalším krokem bylo testování hypotéz z kapitoly 4.3.2 a byl zjišťován Spearmanův korelační koeficient.



Obrázek 4.18: Graf Spearmanova korelačního koeficientu pro podskupinu 5 a 6

**Hladina 3:** Z grafu 4.18 (a) je vidět, že se zde nenachází žádné testy, které by spolu mohly korelovat, dle Spearmanova korelačního koeficientu. Nebudou tedy zde navrženy žádné úpravy seznamu úkolů.

### 4.3.9 Korelační analýza mezi úkoly podskupiny 6

Cílem úkolů v podskupině 6 je otestovat žáky na **rychlost**. Zde je zajímavým faktem, že tyto úkoly nemají jasně stanovenou maximální hodnotu, jak je uvedeno v příloze A.0.1, a tedy záleží čistě na žácích, kolik opakování jsou schopni vykonat.

**Hladina 1:** U testování hypotéz z kapitoly 4.3.1 byla zjištěna opět konstantní hodnota  $p\text{-value} = 1 \times 10^{-3}$ , a tím byla, na hladině významnosti 5 %, zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení.

**Hladina 2 a 3:** Dále byl tedy dle grafu 4.15 zjišťován Spearmanův korelační koeficient. Výsledky tohoto výpočtu jsou demonstrovány na grafu 4.18 (b), kde byla

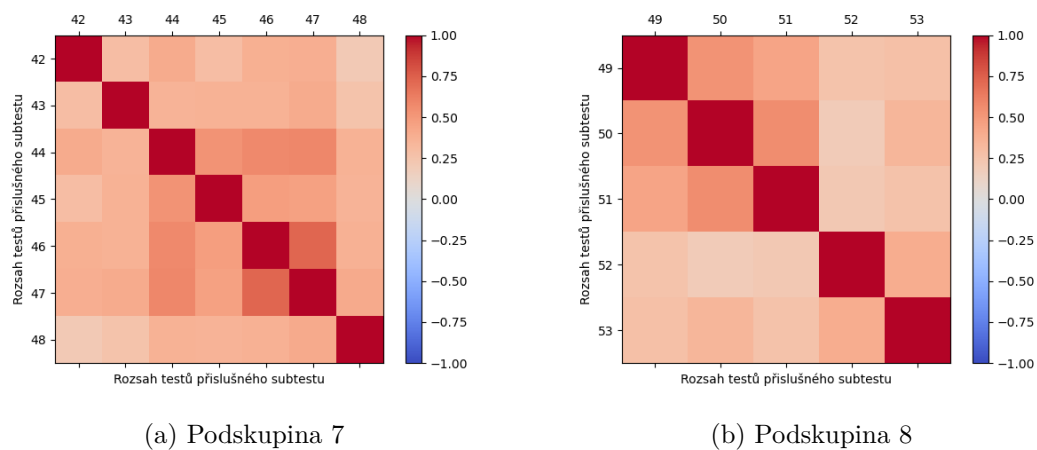
nalezena celkem nejvyšší hodnota korelačního koeficientu vůbec (0.74) mezi 40. a 41. úkolem<sup>18</sup>. Jelikož se jedná o téměř totožné testy, mohl by být jeden z těchto úkolů vyřazen. Konkrétně by mohlo být pro žáky bezpečnější vyřazení úkolu 40, který se zaměřuje pouze na skok po jedné noze.

### 4.3.10 Korelační analýza mezi úkoly podskupiny 7

Úkoly v podskupině 7 jsou cíleny na **koordinaci horních končetin**.

**Hladina 1:** Nejprve byly, na hladině významnosti 5 %, testovány hypotézy z kapitoly 4.3.1 o tom, zda data pochází z normálního rozdělení. Byla zamítnuta nulová hypotéza  $H_0$  o tom, zda data pochází z normálního rozdělení  $p$ -value =  $1 \times 10^{-3}$ .

**Hladina 2:** Dále byl zjišťován Spearmanův korelační koeficient, a tedy testovány hypotézy z kapitoly 4.3.2.



Obrázek 4.19: Graf Spearmanova korelačního koeficientu pro podskupinu 7 a 8

**Hladina 3:** Pokud zde bude dodrženo, stejně jako v předchozích kapitolách, že budou zmíněny úkoly, které mají korelační koeficient větší než 0.5, potom je možné konstatovat, že takových úkolů se zde nachází více a lze to aspoň částečně poznat z uvedeného grafu 4.19 (a). Konkrétně byla sledovaná hodnota korelačního koeficientu větší než 0.5 nalezena celkem mezi čtyřmi dvojicemi. Jedná se o dvojice úkolů 44 - 45, 44 - 46, 44 - 47 a také o 46 - 47<sup>19</sup>. Nejvyšší hodnota korelačního koeficientu byla zjištěna mezi poslední zmíněnou dvojicí úkolů (0.72), u ostatních se hodnota korelačního koeficientu pohybovala v intervalu (0.5, 0.6).

Všechny úkoly v podskupině 7 jsou velmi podobné a zahrnují práci s míčkem. Bylo by tedy možné vyřadit úlohu 44, jelikož má zjevný vztah s více úkoly. Poté by také bylo dobré zvážit vyřazení jednoho úkolu z dvojice 46 - 47, kde jsou vykázány významnější statistické hodnoty.

<sup>18</sup>úloha 40 – skok jednož stranou, úloha 41 – skok snožmo stranou

<sup>19</sup>úloha 44 – pouštění a chytání míčku – jedna ruka, úloha 45 – chytání hozeného míčku – jedna ruka, úloha 46 – driblování míčkem – jedna ruka, úloha 47 – driblování míčkem – střídání rukou

### 4.3.11 Korelační analýza mezi úkoly podskupiny 8

V poslední 8. podskupině jsou úkoly, které prověřují sílu žáků.

**Hladina 1:** Na hladině významnosti 5 % byly testovány hypotézy z kapitoly 4.3.1 o tom, zda data pochází z normálního rozdělení. Byla zamítnuta nulová hypotéza  $H_0$ , a tím pádem nebyl splněn předpoklad, že data pochází z normálního rozdělení ( $p$ -value =  $1 \times 10^{-3}$ ).

**Hladina 2:** Poté byly zjišťovány Spearmanovy korelační koeficienty a testovány hypotézy z kapitoly 4.3.2. Zde byl ve dvou případech nalezen korelační koeficient větší než 0.5, a to konkrétně mezi 49. a 50. úkolem a poté mezi 50. a 51. úkolem<sup>20</sup>. Toto lze pozorovat i na příslušném grafu 4.19 (b).

**Hladina 3:** Po nahlédnutí do seznamu úkolů bylo zjištěno, že 49. úkolem je skok daleký, který vykazuje statistický vztah s 50. úkolem, pod kterým se ukrývají kliky. Jelikož každý z úkolů prověřuje sílu zcela odlišných končetin, nebude zde navržena žádná úprava.

Korelace mezi 50. a 51. úkolem je o něco zjevnější (Spearmanův kor. koef. = 0.56). V tomto případě by bylo vhodné poradit se s odborníkem na dětskou fyzickou přípravu, zda tyto dva testy mohou mít vztah například zapojením středu těla a hlubšího stabilizačního systému, a až poté navrhnout možnou úpravu seznamu úkolů.

### 4.3.12 Závěr korelační analýzy mezi jednotlivými úkoly Bruininks-Oseretsky testu

V kapitolách 4.3.5, 4.3.7 a 4.3.8 (podskupiny 2, 4 a 5) nebyly zjištěny vyšší hodnoty Spearmanova korelačního koeficientu mezi úkoly, a proto v těchto kapitolách nebyly navrženy žádné úpravy seznamu.

V ostatních podskupinách byly nalezeny vyšší hodnoty Spearmanova korelačního koeficientu, mezi úkoly, a bylo zde navrženo možné zjednodušení seznamu:

- podskupina 1 - vyřazení jednoho ze dvojice úkolů 1-2 a 3-4,
- podskupina 3 - vyřazení úkolu 17,
- podskupina 6 - vyřazení úkolu 40,
- podskupina 7 - vyřazení úkolu 44 a poté jednoho z dvojice 46-47,
- podskupina 8 - vyřazení jednoho ze dvojice úkolů 50-51.

---

<sup>20</sup>úloha 49 – skok daleký, úloha 50 – kliky (popř. kliky na kolenou), úloha 51 – sedy lehy

## 5 Závěr

Hlavním cílem této bakalářské práce bylo provedení statistické analýzy dat, jež byla získána šetřením na 665 žácích v Libereckém kraji. Tito žáci prováděli Bruininks-Oseretsky test motorické zdatnosti, který slouží k vyhodnocení prevalence dyspraxie.

V kapitole 2 jsou popsány vstupní datové matice poskytnuté pro tuto práci, způsob výpočtu mezivýsledků a celkového výsledku Bruininks-Oseretsky testu, jakožto i definování pojmů, které jsou v práci dále využívány.

V následující kapitole 3 byly odstraněny chyby či nesrovnalosti, které vznikly ručním přepisem z testových archů do poskytnutých datových souborů. Po opravách těchto záznamů byla data roztržena do celkem 11 věkových skupin, na které je kladen důraz v manuálu Bruininks-Oseretsky testu. Data, která nespadała do žádné ze sledovaných věkových skupin, jsou v této kapitole sice uvedena, ale pro další výpočet bylo těchto 26 záznamů odstraněno.

Po odstranění nevalidních záznamů a roztržení dat, dle věkových skupin, byly v kapitole 4 provedeny celkem 3 druhy statistické analýzy.

V této části, a to nejdříve v kapitole 4.1, byly provedeny statistické testy, kterými byly zjištěny významnější rozdíly mezi 11 věkovými skupinami v celkem 13 různých výsledkových kategoriích Bruininks-Oseretsky testu. Těchto 13 kategorií je možno rozřadit do 3 různých typů - škálové skóre, standardní skóre a celkový motorický koeficient. Výsledky provedených výpočtů jsou v této kapitole znázorněny grafy, z nichž je patrný rozdíl středních hodnot u skupin, jež mezi sebou vykazují statisticky významnější rozdíly. Bylo pozorováno 11 kategorií, s významným vlivem věku na dosažené skóre, a to v kapitolách - subtest 2, 3, 4, 5, 6, 8, test 1, 2, 3, 4 a TMC.

Dalším krokem, o kterém je pojednáváno v kapitole 4.2, bylo otestovat vliv pohlaví a věku na stejných 13 výsledkových kategoriích. Výsledky zjišťované v této kapitole byly počítány dvoufaktorovými statistickými testy, do nichž vstupovalo pohlaví, jakožto první faktor a věk žáka, jakožto druhý faktor. Byly zjištěny vlivy těchto dvou faktorů na závislou proměnnou v podobě testovaného výsledku z 13 výsledkových kategorií. Vliv pohlaví byl pozorován celkem v 5 různých kategoriích, kdežto vliv věku byl pozorován u 11 kategorií.

Poslední částí, v kapitole 4, bylo navrhnout možné zjednodušení seznamu úkolů Bruininks-Oseretsky testu. V kapitole 4.3 bylo zkoumáno celkem 53 úkolů, které se třídí do celkem 8 různých podskupin. Každá z těchto 8 podskupin je zaměřena na jinou fyzickou oblast, a proto byly zjišťovány korelace mezi úkoly výhradně v rámci jednotlivých podskupin. Po zjištění korelačních koeficientů v rámci podskupiny bylo diskutováno možné odstranění úkolů, které vykazovaly zjevný vztah. Celkem bylo



navrženo zjednodušení o 7 úkolů, který vykazují významnější vztah s jinými úkoly.

## Seznam použité literatury

- [1] Jiří Anděl. *Statistické metody*. 4., upr. vyd. Praha: Matfyzpress, 2007. ISBN: 80-7378-003-8.
- [2] Robert H. Bruininks a Brett D. Bruininks. *BOT2. Bruininks-Oseretsky Test of Motor Proficiency*. Bloomington, MN, USA, 2005.
- [3] Susan Gipson. “CSV vs Excel – Difference Between Them”. In: (). URL: <https://www.guru99.com/excel-vs-csv.html> (cit. 07. 05. 2023).
- [4] Jan Hendl. *Přehled statistických metod zpracování dat. Analýza a metaanalýza dat*. Vyd. 1. Praha: Portál, 2004. ISBN: 80-7178-820-1.
- [5] P. Kolář, J. Smržová a A. Kobešová. “Vývojová porucha koordinace - vývojová dyspraxie. Developmental Coordination Disorder - Developmental Dyspraxia”. In: *Česká a slovenská neurologie a neurochirurgie* (2011), s. 533–535.
- [6] Martina Litschmannová. *Úvod do statistiky*. Ostrava, 2011.
- [7] *MathWorks. Python Setup Script to Install MATLAB Engine API*. c1994-2023. URL: [https://www.mathworks.com/help/matlab/matlab\\_external/python-setup-script-to-install-matlab-engine-api.html](https://www.mathworks.com/help/matlab/matlab_external/python-setup-script-to-install-matlab-engine-api.html) (cit. 20. 04. 2023).
- [8] *MathWorks. friedman*. c1994-2023. URL: <https://www.mathworks.com/help/stats/friedman.html> (cit. 20. 04. 2023).
- [9] Milan Meloun a Jiří Militký. *Kompendium statistického zpracování dat*. Vyd. 2. Praha: Academia, 2006. ISBN: 80-200-1396-2.
- [10] *Pandas documentation*. c2023. URL: <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.dtypes.html> (cit. 07. 05. 2023).
- [11] *SciPy documentation. Statistical functions (scipy.stats)*. c2008-2023. URL: <https://docs.scipy.org/doc/scipy/reference/stats.html>.
- [12] *SciPy documentation. scipy.stats.friedmanchisquare*. c2008-2023. URL: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.friedmanchisquare.html>.

## A Přílohy

V přílohách tohoto dokumentu je možné najít některé doplňující informace.

### A.0.1 Jednotlivé testy Bruininks-Oseretsky testu

#### Podskupina 1 – jemná motorika (přesnost)

	<b>Raw skóre</b>
úloha 1 – vyplňování tvaru – kruh	0 až 3 body,
úloha 2 – vyplňování tvaru – hvězda	0 až 3 body,
úloha 3 – kreslení čar ve vyznačené křivé dráze	0 až 21 chyb,
úloha 4 – kreslení čar ve vyznačené oblé dráze	0 až 21 chyb,
úloha 5 – spojování čar	0 až 12 bodů,
úloha 6 – přehýbání papíru	0 až 12 bodů,
úloha 7 – vystřížení kruhu	0 až 12 bodů.

#### Podskupina 2 – jemná motorika (integrace)

	<b>Raw skóre</b>
úloha 8 – kopírování kruhu	0 až 4 body,
úloha 9 – kopírování čtverce	0 až 5 body,
úloha 10 – kopírování překrývajících se kruhů	0 až 6 bodů,
úloha 11 – kopírování vlnovky	0 až 4 body,
úloha 12 – kopírování trojúhelníku	0 až 5 bodů,
úloha 13 – kopírování kosočtverce	0 až 5 bodů,
úloha 14 – kopírování hvězdy	0 až 5 bodů,
úloha 15 – kopírování překrývajících se tužek	0 až 6 bodů.

#### Podskupina 3 – manuální zručnost

	<b>Raw skóre</b>
úloha 16 – tvorba bodů v kruhu	0 až 80 bodů,
úloha 17 – přenos mincí	0 až 20 bodů,
úloha 18 – umístování kolíčků do děrované desky	0 až 30 bodů,
úloha 19 – rovnání karet	0 až 50 bodů,
úloha 20 – navlékání kostiček na provázek	0 až 15 bodů,

#### Podskupina 4 – bilaterální koordinace

	<b>Raw skóre</b>
úloha 21 – dotýkání se nosu ukazováčkem při zavřených očích	0 až 4 body
úloha 22 – jumping jack	0 až 5
úloha 23 – skákání na místě – synchronizace stejné strany	0 až 5 bodů
úloha 24 – skákání na místě – synchronizace opačné strany	0 až 5 bodů
úloha 25 – otáčení palců a ukazováčků	0 až 5 bodů
úloha 26 – dotýkání se chodidel a prstů – synchronizace stejné strany	0 až 10 bodů
úloha 27 – dotýkání se chodidel a prstů – synchronizace opačné strany	0 až 10 bodů

#### Podskupina 5 – rovnováha

	<b>Raw skóre</b>
úloha 28 – stání na čáře s otevřenýma očima a nohama na krok vzdálenýma od sebe	0 až 10 s,
úloha 29 – chůze vpřed po čáře	0 až 6 kroků,
úloha 30 – stání s otevřenýma očima na čáře na jedné noze	0 až 10 s,
úloha 31 – stání na čáře se zavřenýma očima a nohama na krok vzdálenýma od sebe	0 až 10 s,
úloha 32 – chůze po čáře, pata za špičkou	0 až 6 kroků,
úloha 33 – stání se zavřenýma očima na čáře na jedné noze	0 až 10 s,
úloha 34 – stání na kladině s otevřenýma očima na jedné noze	0 až 10 s,
úloha 35 – stání na kladině, pata za špičkou	0 až 10 s,
úloha 36 – stání na kladině se zavřenýma očima na jedné noze	0 až 10 s.

#### Podskupina 6 – rychlost

	<b>Raw skóre</b>
úloha 37 – člunkový běh	neomezený čas,
úloha 38 – překračování kladiny	neomezený počet kroků,
úloha 39 – skok jednož na místě	neomezený počet skoků,
úloha 40 – skok jednož stranou	neomezený počet skoků,
úloha 41 – skok snožmo stranou	neomezený počet skoků.

### **Podskupina 7 – koordinace horní končetiny**

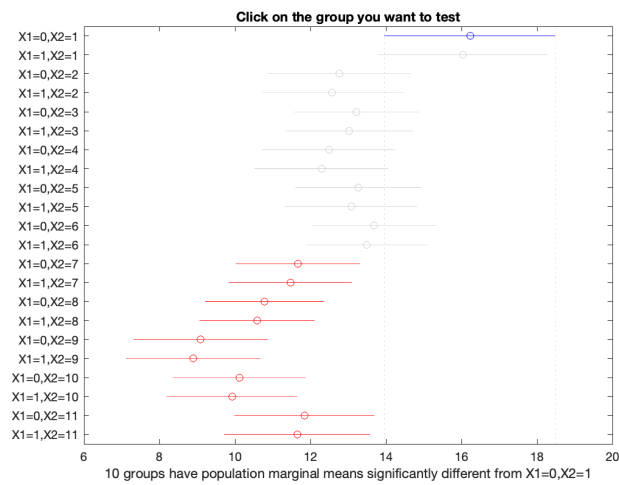
	<b>Raw skóre</b>
úloha 42 – pouštění a chytání míčku – obě ruce	0 až 5 bodů,
úloha 43 – chytání hozeného míčku – obě ruce	0 až 5 bodů,
úloha 44 – pouštění a chytání míčku – jedna ruka	0 až 5 bodů,
úloha 45 – chytání hozeného míčku – jedna ruka	0 až 5 bodů,
úloha 46 – driblování míčkem – jedna ruka	0 až 10 bodů,
úloha 47 – driblování míčkem – střídání rukou	0 až 5 bodů,
úloha 48 – házení míčku na cíl	0 až 5 bodů.

### **Podskupina 8 – síla**

	<b>Raw skóre</b>
úloha 49 – skok daleký	neomezená vzdálenost,
úloha 50 – kliky (popř. kliky na kolenou)	neomezený počet,
úloha 51 – sedy lehy	neomezený počet,
úloha 52 – podřep u stěny	0 až 60 s,
úloha 53 – V-up	0 až 60 s.

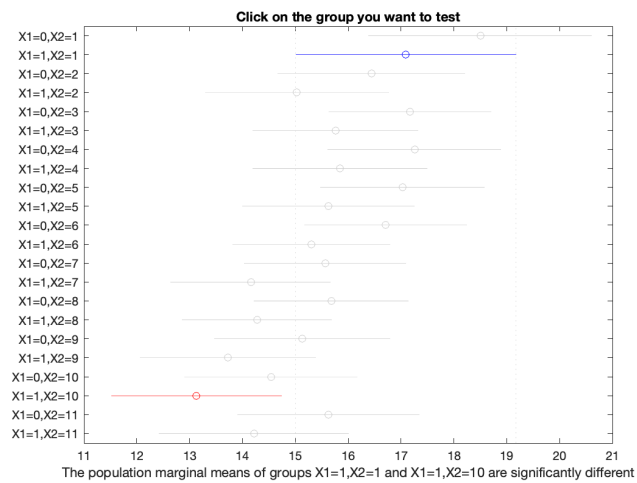
## A.0.2 Grafy ke kapitole 4.2

### Kapitola 4.2.5



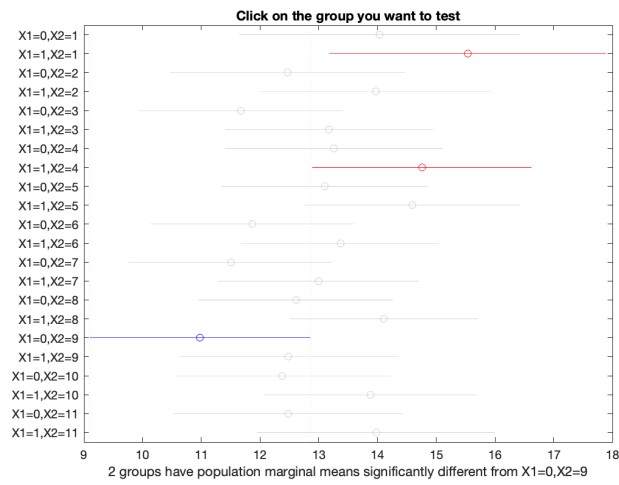
Obrázek A.1: Graf rozdílů statistických ukazatelů z Matlabu

### Kapitola 4.2.9



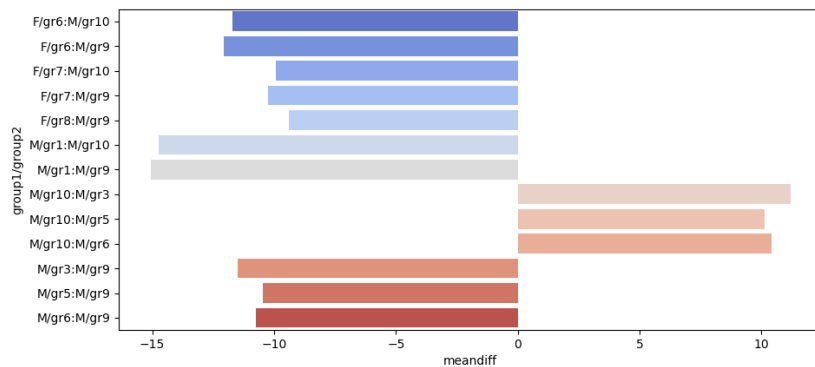
Obrázek A.2: Graf rozdílů statistických ukazatelů z Matlabu

## Kapitola 4.2.10



Obrázek A.3: Graf rozdílů statistických ukazatelů z Matlabu

## Kapitola 4.2.12



Obrázek A.4: Graf rozdílů statistických ukazatelů z Pythonu

### A.0.3 Příložené CD

`code_data`: obsahuje skripty na třídění dat, dále obsahuje filtry, které jsou provedeny na 2 různých vstupních datových souborech a to ve složkách `new_data_filters` a `old_data_filters` -> v každé složce jsou výstupy z těchto filtrů

`new_data_filters`:

`age_groups`:

zde je soubor pro 11 věkových skupin (validní soubor je `Student_groups_by_age_valid_2.xlsx`)

`statistics`:

obsahuje statistické skripty, které byly použity pro výpočet jednotlivých statistických ukazatelů, potřebných pro práci

**matlab**: obsahuje kontroly různých zjišťovaných hodnot, soubory jsou popsány tak, aby z nich bylo viditelné, co se v nich zjišťuje