

Danka Haruštiaková, Lucie Pehalová, Michal Uher

Biostatistika

Učební texty pro studenty zdravotnických oborů

MUNI
PRESS

Danka Haruštiaková, Lucie Pehalová, Michal Uher

Biostatistika

Učební texty pro studenty zdravotnických oborů

Masarykova univerzita

Brno, 2020

PŘEDMLUVA

Tyto učební texty jsou určeny především studentům předmětů MIKAM021 Analýza a management dat pro zdravotnické obory a BLKBS051 Biostatistika kombinované formy studia. Jsou ovšem také vhodným doplňujícím výukovým materiálem pro všechny studenty Masarykovy univerzity, kteří studují jakýkoli předmět týkající se biostatistiky či analýzy a zpracování dat. Učební texty mohou být užitečné také pro studenty, kteří využijí statistické zpracování dat pro svoji bakalářskou či diplomovou práci.

Cílem těchto textů je stručně a jasně představit základní biostatistické metody. Předložené učební texty se skládají ze dvou ucelených částí, které jsou koncipovány tak, aby se vzájemně doplňovaly a umožnili čtenáři pochopení základních biostatistických principů jak z teoretického, tak i praktického hlediska.

První část těchto učebních textů má za cíl uvést čtenáře do problematiky jednotlivých oblastí analýzy a zpracování dat z teoretického pohledu. V rámci této části jsou podrobně představeny principy práce s daty, datové typy, jejich základní popis a zpracování. Stěžejní část je věnována testování statistických hypotéz zaměřených na různé typy dat. Největší důraz je kladen na pochopení základních principů jednotlivých statistických testů a jejich korektní využití. Všechny teoretické kapitoly jsou doplněny o ukázkou praktického řešení adekvátního úkolu na cvičném datovém souboru.

Druhou část učebních textů tvoří sbírka příkladů k samostatnému procvičování nabytých vědomostí. V její úvodní kapitole je popsán datový soubor, se kterým se v rámci celé sbírky pracuje. Následující kapitoly jsou věnovány jednotlivým tematickým celkům odpovídajícím výukovým blokům v první části učebních textů. V rámci sbírky je komplexně zahrnuta celá problematika analýzy dat od počáteční přípravy datového souboru až po testování statistických hypotéz. V každé kapitole jsou uvedeny vzorové příklady s detailním popisem řešení a interpretací získaných výsledků. Součástí sbírky jsou i příklady k samostatnému procvičení včetně správného řešení. Ke všem příkladům je uveden návod na zpracování ve statistickém programu Statistica, verze 13, jež mají studenti Masarykovy univerzity k dispozici.

Pro zpracování učebních textů byl využit tabulkový procesor Microsoft Excel a statistický software Statistica. Příkladová data užívaná v těchto textech jsou uložena v informačním systému MU.

Učební texty byly vytvořeny s podporou projektu MUNI/FR/1344/2018.

V Brně 16. 3. 2020

Danka Haruštiaková
Lucie Pehalová
Michal Uher

OBSAH

ČÁST PRVNÍ.....	6
1. Úvod do biostatistiky	7
1.1. Základní pojmy a příprava dat.....	7
2. Představení datového souboru.....	9
3. Typy proměnných	13
4. Popisná statistika	14
4.1. Popis kvalitativních proměnných.....	14
4.2. Popis kvantitavních proměnných.....	16
5. Normální rozdělení.....	20
6. Testování hypotéz	26
6.1. Interpretace výsledku statistického testu.....	27
6.2. Přehled statistických testů.....	27
6.3. Test normality rozdělení.....	30
7. Parametrické testy pro kvantitativní proměnné.....	33
7.1. Jednovýběrový t-test.....	33
7.2. Dvouvýběrový párový t-test	35
7.3. Dvouvýběrový nepárový t-test	38
7.4. Analýza rozptylu (ANOVA).....	41
8. Neparametrické testy pro kvantitativní proměnné	46
8.1. Jednovýběrový Wilcoxonův a znaménkový test.....	46
8.2. Párový Wilcoxonův a znaménkový test.....	50
8.3. Mannův-Whitneyův test.....	52
8.4. Kruskalův-Wallisův test.....	56
9. Testy pro kvalitativní proměnné	60
9.1. Kontingenční tabulka.....	60
9.2. Pearsonův chí-kvadrát test	62
9.3. Fisherův exaktní test.....	65
9.4. McNemarův test.....	68
10. Závislost dvou kvantitativních proměnných: korelace.....	71
10.1. Pearsonův korelační koeficient	72
10.2. Spearmanův korelační koeficient.....	75
Seznam použité literatury	78
ČÁST DRUHÁ	79
Cvičení 1. Práce s datovým souborem	80
1.1. Data z dotazníkového šetření.....	80
1.2. Popis proměnných.....	81

1.3. Práce s daty	82
Cvičení 2. Popisná statistika.....	85
2.1. Typy proměnných	85
2.2. Popisné statistiky kategoriálních proměnných	86
2.3. Popisné statistiky spojitých proměnných.....	88
Cvičení 3. Rozložení dat	92
Cvičení 4. Parametrické testy pro kvantitativní proměnné	100
4.1. Jednovýběrový t-test.....	100
4.2. Dvouvýběrový t-test	102
4.3. Párový t-test	105
4.4. Analýza rozptylu (ANOVA).....	107
Cvičení 5. Neparametrické testy pro kvantitativní proměnné	112
5.1. Jednovýběrový Wilcoxonův a znaménkový test.....	112
5.2. Mannův-Whitneyův test.....	115
5.3. Párový Wilcoxonův a znaménkový test.....	118
5.4. Kruskalův-Wallisův test.....	120
Cvičení 6. Testy pro kategoriální proměnné	125
6.1. Pearsonův chí-kvadrát test	125
6.2. Fisherův exaktní test.....	127
6.3. McNemarův test.....	129
Cvičení 7. Základy korelační analýzy	132
7.1. Pearsonův korelační koeficient	132
7.2. Spearmanův korelační koeficient.....	134

ČÁST PRVNÍ

1. Úvod do biostatistiky

Biostatistika je vědní obor na rozhraní matematické statistiky a věd o živých systémech. Jde v podstatě o aplikaci statistických metod pro řešení biologických a klinických problémů. Biostatistika primárně vychází ze statistiky, jejich hranice však nejsou ostré.

Termín *statistika* má více významů. V laickém chápání je to uspořádaný soubor dat (statistika výsledků hokejových zápasů, statistika hlasování poslanců, statistika dopravních nehod). Statistika jako věda – přesněji nazývána matematická statistika – je nástrojem, jak z výše uvedených informací získat hodnotnou informaci. Je to samostatná vědní disciplína zabývající se prací s daty obsahujícími náhodnou složku. Vyhodnocuje zákonitosti v datech a odlišuje je od náhodné variability. Termín statistika se používá také ve významu veličiny spočítané z dat. Např. obecně známé minimum, maximum, průměr jsou statistikami, charakterizujícími soubor dat. Na závěr: lidová moudrost praví, že statistika je nauka o přesném počítání s nepřesnými čísly.

Díky přístupné výpočetní technice je statistika dostupná pro většinu biologů a kliniků. Již plánování pokusů a pozorování je ovlivněno znalostí statistiky. Pouze data, jež jsou sesbírána správným způsobem, lze statisticky vyhodnotit. Znalost základů statistiky je tedy nezbytná k úspěšné práci v biologických i klinických oborech.

Biostatistika není samostatným vědním oborem, jde o aplikaci statistiky na biologické problémy, která klade důraz na pochopení principů metod a zásad jejich použití. Hlavním cílem těchto učebních textů je představení základních principů práce s daty a zpracování medicínských dat pomocí statistických metod. Teoretické základy principů a statistických metod, které budou představeny, jsou doplněny praktickými příklady řešenými ve statistickém programu Statistica for Windows, verze 13. Jde o program, který je k uživateli přátelský, poskytuje příjemný systém nabídek v menu a dialogových oknech a umožňuje snadnou úpravu grafických prezentací výsledků. Program je k dispozici v anglickém i českém vydání, v této učebnici se odkazujeme na anglickou verzi programu, tudíž jsou v návodech uvedeny anglické termíny.

1.1. Základní pojmy a příprava dat

Hlavním úkolem předcházejícím zpracování dat je jejich sběr, tj. získání informace. Většinou není v našich silách získat informace o celé tzv. *cílové populaci* (*základním souboru*), ale pouze o části cílové populace, tzv. *výběru z cílové populace* (*experimentální vzorek*). Skutečně, kdybychom chtěli měřit např. hmotnost a výšku u všech pacientů s diabetem II. typu v ČR (tj. základní soubor), bylo by to časově a finančně velice náročné. V případě, že pozorování pacientů s diabetem II. typu bude správně navrženo (tj. tak, aby byla zajištěna reprezentativnost experimentálního vzorku), výsledky získané vyhodnocením výběru z cílové populace budou platné pro cílovou populaci.

Experimentální vzorek tedy představuje část cílové populace, který sledujeme. Je to skupina subjektů (např. pacientů, nebo vyšetření), kterou máme k dispozici a která představuje *pozorování* cílové populace. Sledované vlastnosti označujeme jako *znak*, nebo *proměnná*.

Předpokladem úspěšného zpracování dat je jejich správné uložení. Správné a přehledné uložení dat je základem jejich pozdější analýzy. Je vhodné rozmyslet si předem, jak budou data ukládána. Nejvhodnějším způsobem je uložení dat ve formě *databázové tabulky*, kde v sloupcích jsou *proměnné* (např. věk, hmotnost, výška, tepová frekvence atd.) a v řádcích *pozorování* (např. jednotlivý pacienti, nebo jednotlivá vyšetření).

Pro databázovou tabulku platí následovné:

- každý sloupec obsahuje pouze jediný typ dat, identifikovaný hlavičkou sloupce;
- každý řádek obsahuje minimální jednotku dat (např. pacient, vyšetření pacienta apod.);
- není přípustné kombinovat v jednom sloupci číselné a textové hodnoty;
- komentáře jsou uloženy v samostatných sloupcích;
- u textových dat je nezbytné kontrolovat překlapy v názvech kategorií;
- specifickým typem dat jsou datумы, u nichž je nezbytné kontrolovat, zda jsou uloženy ve správném formátu.

Takto uspořádaná data můžeme v tabulkových nebo databázových programech převést na libovolnou výstupní tabulku. Pro uložení a čištění dat je možné využít tabulkový procesor Microsoft Excel.

2. Představení datového souboru

V celém textu budeme pracovat s cvičným datovým souborem Rehabilitace po mozkovém infarktu. Cvičný datový soubor obsahuje záznamy o celkem 407 pacientech hospitalizovaných pro mozkový infarkt na neurologickém oddělení akutní péče, kde jim byla poskytnuta terapie pro obnovu krevního oběhu v postižené části mozku. Po zvládnutí akutní fáze byl u pacientů vyhodnocen stupeň soběstačnosti v základních denních aktivitách pomocí tzv. indexu Barthelové (BI) a byli přeloženi na rehabilitační oddělení. Po dvou týdnech byl opět dle BI vyhodnocen stupeň soběstačnosti a pacienti byli buď propuštěni do ambulantní péče, nebo přeloženi na oddělení následné péče.

U každého pacienta byly sbírané tyto informace:

- základní demografické údaje (pohlaví a věk),
- informace o samotné diagnóze mozkové příhody (etiologie a lokalizace uzávěru cévy),
- informace o léčbě (typ indikované terapie a výskyt komplikací),
- informace o způsobu ukončení rehabilitace,
- stupeň soběstačnosti před rehabilitací a na konci rehabilitace určen na základě dotazníku.

Náhled části datového souboru představuje Tabulka 2.1. Barthelové index byl určen na základě dotazníku uvedeného v Tabulce 2.2.

Tabulka 2.1. Ukázka části datového souboru pacientů po mozkovém infarktu.

ID	M/Ž	Věk	Etiologie	Lokali- zace	Terapie	BI před	Závislost před reh.	Ukončení rehabilitace	BI po	Závislost po reh.
1	M	82	okl/st	mozk t	jiná	25	závislý	propuštěn	85	soběstačný
2	Ž	81	emb	mozk t	jiná	20	závislý	přeložen	40	závislý
3	M	55	okl/st	mozk t	jiná	35	závislý	propuštěn	75	soběstačný
4	Ž	46	emb	mozk t	trombolýza	20	závislý	propuštěn	60	soběstačný
5	M	76	okl/st	mozk t	jiná	45	soběstačný	propuštěn	80	soběstačný
6	M	72	okl/st	mozk t	jiná	25	závislý	přeložen	50	soběstačný
7	M	62	tromb	mozk t	jiná	40	závislý	propuštěn	70	soběstačný
8	M	64	tromb	přív t	jiná	15	závislý	propuštěn	85	soběstačný
9	Ž	82	okl/st	mozk t	jiná	10	závislý	přeložen	30	závislý
10	M	58	tromb	mozk t	jiná	25	závislý	propuštěn	95	soběstačný
11	M	84	okl/st	mozk t	jiná	40	závislý	propuštěn	85	soběstačný
12	Ž	92	okl/st	mozk t	jiná	30	závislý	propuštěn	60	soběstačný
13	Ž	79	emb	mozk t	jiná	40	závislý	propuštěn	60	soběstačný
14	M	69	tromb	mozk t	jiná	45	soběstačný	propuštěn	80	soběstačný
15	M	67	okl/st	mozk t	trombektomie	25	závislý	přeložen	35	závislý
16	Ž	70	tromb	přív t	trombektomie	40	závislý	propuštěn	70	soběstačný
17	Ž	59	tromb	mozk t	jiná	25	závislý	přeložen	55	soběstačný
...
...
...
...
...

ID – číselný identifikátor pacienta; M/Ž – pohlaví pacienta: M – muž, Ž – žena; Věk – věk v době přijetí k hospitalizaci; Etiologie – charakter vzniku uzávěru cévy a infarktu: okl/st – okluze nebo stenóza, emb – embolie, tromb – tromboza; Lokalizace – lokalizace postižené cévy: moz k t – mozkové tepny, přív t – přívodní tepny; Terapie – Způsob použité terapie na akutním lůžku: trombektomie – mechanická trombektomie, trombolýza – intravenózní trombolýza rt-PA, jiná – jiná farmakologická terapie; BI před – Index Barthelové před zahájením rehabilitace; Závislost před reh. – Kategorie míry závislosti před zahájením rehabilitace: závislý – vysoce závislý, soběstačný – částečně soběstačný; Ukončení rehabilitace – Způsob ukončení rehabilitace (další péče o pacienta): propuštěn – propuštěn do ambulantní péče, přeložen – přeložen na lůžko následné péče; BI po – Index Barthelové po ukončení rehabilitace; Závislost po reh. – Kategorie míry závislosti po ukončení rehabilitace: závislý – vysoce závislý, soběstačný – částečně soběstačný.

Tabulka 2.2. Dotazník k výpočtu indexu Barthelové.

Barthelové index základních všedních činností (BI)	
Identifikace případu: Jméno pacienta _____	
Jméno hodnotitele _____	
Datum hodnocení _____	
Činnost	Skóre
Jedení 10 = samostatně 5 = s pomocí (např. krájení, roztírání másla) nebo s potřebou speciální diety 0 = neprovede	
Přesun z invalidního vozíku na lůžko a zpět 15 = samostatně bez pomoci 10 = s menší pomocí (verbální nebo fyzickou) 5 = s větší pomocí (fyzickou, jednoho nebo dvou lidí), může se posadit 0 = neprovede, neudrží rovnováhu vsedě nebo není schopen používat invalidní vozík	
Provádění osobní hygieny 5 = samostatně umytí rukou, obličeje, čištění zubů, holení 0 = nutná pomoc s osobní hygienou	
Posazení na toaletu a vstání z ní 10 = samostatně bez pomoci (usednutí, otření, oblečení, zvednutí) 5 = potřebuje pomoc, ale zvládá některé úkony samostatně 0 = závisle na pomoci	
Koupání nebo sprchování 5 = samostatné koupání nebo sprchování 0 = závisle na pomoci	
Chůze (pohyb na vozíku) na rovném povrchu 15 = chůze samostatně (případně s oporou, např. holí) nad 50 metrů 10 = chůze s malou pomocí nad 50 metrů 5 = samostatný pohyb na vozíku, včetně zatáčení, nad 50 metrů 0 = imobilní, nebo mobilní do 50 metrů	
Chůze do schodů a ze schodů 10 = samostatně bez pomoci 5 = s pomocí (verbální, fyzickou, s podporou) 0 = nezvládne	
Oblékání a svlékání (včetně zavazování tkaniček, zapínání zipů) 10 = samostatně 5 = potřebuje pomoc, ale zvládá z poloviny samostatně 0 = závisle na pomoci	
Ovládání stolice 10 = kontinentní 5 = příležitostné nehody nebo potřeba pomoci s aplikací klystýru 0 = inkontinentní	
Ovládání močení 10 = kontinentní 5 = příležitostné nehody nebo potřeba pomoci s externí pomůckou 0 = inkontinentní, nebo katetrizovaný bez možnosti samostatného močení	
Celkový součet (0-100)	
Barthelové index základních všedních činností (BI)	
Vyhodnocení stupně závislosti v základních denních aktivitách	

0-40 bodů	vysoce závislý
45-60 bodů	závislost středního stupně
65-95 bodů	lehká závislost
100 bodů	nezávislý

Pokyny k použití

1. Index by měl být používán jako záznam o tom, jaké aktivity pacient aktuálně zvládá, nikoliv jako záznam toho, co by pacient zvládat mohl.
2. Hlavním cílem je stanovit stupeň nezávislosti na jakékoliv pomoci, fyzické nebo verbální, jakkoliv velké a nezávisle na důvodu poskytnutí.
3. Potřeba kontroly znamená, že pacient není nezávislý.
4. Výkon pacienta by měl být stanoven pomocí nejlepších dostupných informačních podkladů. Pomocí dotazování se pacienta, přátel, příbuzných, zdravotnického personálu, což jsou obvyklé zdroje, ale také pomocí přímého pozorování a zdravého rozumu. Přímé testování však není potřeba.
5. Obvykle je podstatný výkon pacienta za posledních 24 až 48 hodin, v některých případech je relevantní i delší období.
6. Střední kategorie naznačují, že pacient k provedení úkolu vynakládá alespoň poloviční množství celkového úsilí.
7. Použití pomůcek neznamena omezení nezávislosti.

3. Typy proměnných

Realitu můžeme popisovat různými typy dat, každý z nich se specifickými vlastnostmi, výhodami, nevýhodami a vlastní sadou využitelných statistických metod. Data reprezentují sledované vlastnosti, znaky, resp. proměnné.

Rozlišujeme tyto typy proměnných:

- *Kvalitativní (kategorální) proměnná*. Lze ji řadit do kategorií, ale nelze ji kvantifikovat, resp. jednotlivým kategoriím lze přiřadit číselné kódy, které však nemají logickou souvislost s úrovní sledovaného znaku. Jako příklad můžeme uvést pohlaví, HIV status apod.

Kvalitativní proměnná může být binární, nominální nebo ordinální.

- *Binární proměnná* nabývá pouze dvou hodnot – kategorií. Tyto kategorie většinou kódujeme pomocí číslic 1 (přítomnost sledovaného znaku) a 0 (nepřítomnost sledovaného znaku). Příkladem binární proměnné je např. přítomnost onemocnění (1 – ano, 0 – ne), pohlaví (1 – muž, 0 – žena).
- *Nominální proměnná* obsahuje několik kategorií, které nelze vzájemně seřadit a u nichž nemá smysl ptát se na relaci větší-menší. Příkladem nominální proměnné je např. krevní skupina (A, B, AB, 0).
- *Ordinální proměnná* také obsahuje více kategorií, tyto lze ovšem vzájemně seřadit, tedy můžeme se ptát, která je větší nebo menší. Příkladem ordinální proměnné je např. stupeň bolesti (mírná, střední, velká), stadium maligního onemocnění (I, II, III, IV).
- *Kvantitativní (numerická) proměnná*. Kvantitativní proměnné můžeme přiřadit číselnou hodnotu, např. výška pacienta, počet hospitalizací apod.

Kvantitativní proměnná může být spojitá nebo diskrétní.

- *Spojité proměnná* nabývá jakýchkoliv hodnot v určitém rozmezí. Příkladem spojitých dat je výška a hmotnost osob, délka časového období od narození po výskyt nemoci, teplota apod.
- *Diskrétní proměnná* může nabývat pouze spočetně mnoha hodnot. Příkladem diskrétních dat je počet krvácivých epizod, počet dětí v rodině apod.

Kvantitativní proměnnou můžeme rozlišovat také dle toho, jestli ji měříme na intervalové nebo poměrové stupnici. V případě *intervalové stupnice* se můžeme ptát na otázku, o kolik jednotek se dvě hodnoty liší (např. teplota měřená ve stupních Celsia). *Poměrová stupnice* nám umožňuje ptát se i na otázku, kolikrát je jedna hodnota větší než druhá. Příkladem poměrových dat jsou např. výška a hmotnost osob, počet krevních buněk v 1 ml krve apod.

4. Popisná statistika

Prvním krokem jakékoliv analýzy bez ohledu na typ proměnné by měla být sumarizace a vizualizace pozorovaných hodnot. K sumarizaci pozorovaných hodnot slouží metody *popisné statistiky*. Tyto metody jsou potřebné ke shrnutí základních informací o proměnných. Objektivně není možné pouhým pohledem na data odečíst potřebné informace ze souboru např. 1 000 pacientů. Potřebujeme popisné statistiky, abychom mohli o datech vůbec přemýšlet.

K základním popisným statistikám patří např. obecně známý průměr, minimum, maximum a další statistiky. V následujícím textu si je představíme podrobněji.

4.1. Popis kvalitativních proměnných

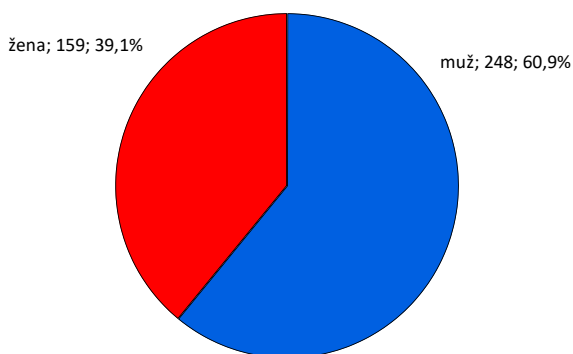
U kvalitativních proměnných předpokládáme, že se jednotlivé hodnoty opakují, počet možných kategorií je totiž známý. Proto je logické je sumarizovat pomocí tabulky s četnostmi možných hodnot (*tabulka četností, frekvenční tabulka*). V takové tabulce uvádíme pozorovanou četnost a relativní četnost každé kategorie. Tabulku četností proměnné "Pohlaví" z našeho datového souboru uvádíme v Tabulce 4.1.

Tabulka 4.1. Sumarizace kvalitativní proměnné "Pohlaví" v tabulce četností.

Kategorie	Absolutní četnost (počet)	Relativní četnost (%)
Muž	248	60,9
Žena	159	39,1
Celkem	407	100

Vzhledem k tomu, že hodnoty kvalitativní proměnné nelze seřadit dle velikosti, používá se jako těžiště této proměnné tzv. *mód*, nebo *modus*. Je to kategorie kvalitativní proměnné s nejvyšší četností. V našem příkladu z Tabulky 2 je modus 'muž' (tj. kategorie s nejvyšší četností).

Pro zobrazení zastoupení kategorií v souboru je vhodný výsečový – *koláčový graf* (Obrázek 4.1).



Obrázek 4.1. Ukázka zobrazení kvalitativní proměnné "Pohlaví" ve výsečovém – koláčovém grafu.

Příklad: Popis kvalitativních dat

Zadání: Proved'te základní popis zastoupení pohlaví u pacientů s mozkovým infarktem. Následně srovnajte zastoupení pohlaví mezi třemi skupinami pacientů dle etiologie mozkové příhody.

Postup v programu Statistica

1. Pro popis dat je vhodné určit absolutní a relativní četnosti. V menu *Statistics* zvolíme *Basic statistics* a vybereme *Frequency tables*. Vybereme proměnnou (*Variables*), kterou chceme analyzovat a na záložce *Options* zaškrtneme možnost *Percentages (relative frequencies)*. Chceme-li získat výsledky zvlášť pro podskupiny jiné proměnné, aktivujeme funkci *By Group* (vybereme třídící proměnnou a zaškrtneme *Accumulate tabular results in a single spreadsheet*). Kliknutím na tlačítko *Summary* získáme výsledky.
2. Grafické znázornění je ideální pomocí koláčového grafu. V menu *Graphs* zvolíme nabídku *2D* a vybereme *Pie Charts*. Vybereme proměnnou (*Variables*), kterou chceme analyzovat a na záložce *Advanced* vybereme v nastavení legendy možnost *Text and Percent*. Chceme-li získat výsledky zvlášť pro podskupiny jiné proměnné, aktivujeme *By Group* (opět vybrat třídící proměnnou). Kliknutím na tlačítko *OK* získáme výstupy.

Výsledky v programu Statistica

Frekvenční tabulka proměnné pohlaví u pacientů s CMP.

Category	Frequency table: Pohlaví	
	Count	Percent
muž	248	60,93366
žena	159	39,06634

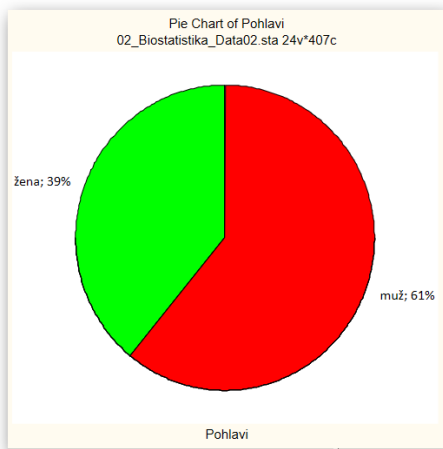
Frekvenční tabulka proměnné pohlaví u pacientů s CMP dle etiologie centrální mozkové příhody.

Category	Aggregate Results Frequency table: Pohlaví (02_Biostatistika_Dat)		
	Etiologie	Count	Percent
muž	okluze nebo stenóza	125	62,18905
žena	okluze nebo stenóza	76	37,81095
žena	embolie	36	46,15385
muž	embolie	42	53,84615
muž	trombóza	81	63,28125
žena	trombóza	47	36,71875

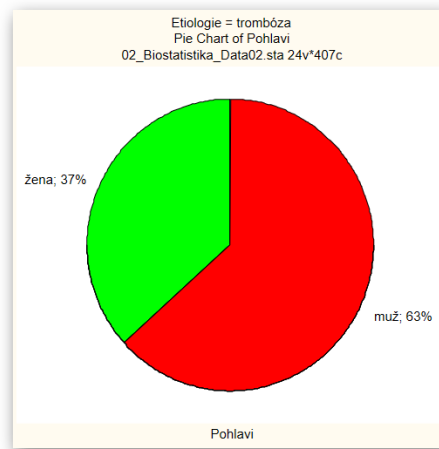
Zastoupení mužů a žen v celém souboru je 61 % oproti 39 %. Při srovnání pacientů dle etiologie mozkového infarktu se nejvíce liší pacienti s embolií, u kterých je podíl mužů jen 54 % oproti 46 % žen.

Koláčový graf proměnné pohlaví u pacientů s CMP:

všichni pacienti

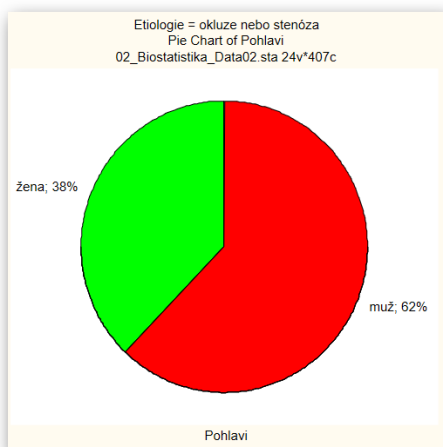


pacienti s trombózou



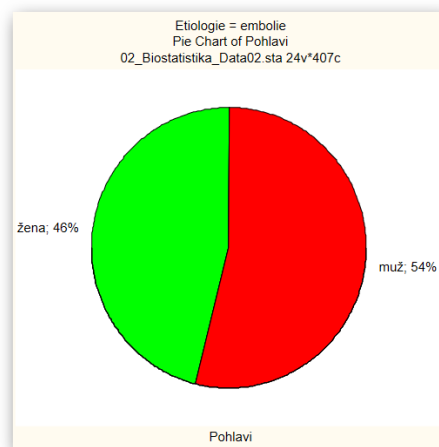
Koláčový graf proměnné pohlaví:

pacienti s okluzí nebo stenózou



Koláčový graf proměnné pohlaví:

pacienti s embolií



4.2. Popis kvantitavních proměnných

Nejznámější statistikou, která charakterizuje proměnnou jedním číslem a představuje typickou hodnotu, kolem které mají ostatní hodnoty tendenci kolísat, je *průměr*. Průměr lze jednoduše spočítat jako součet pozorovaných hodnot dělený jejich počtem. Je ukazatelem střední hodnoty.

Dalším ukazatelem střední hodnoty souboru je *medián*, který definujeme jako prostřední pozorovanou hodnotu souboru. Medián rozděluje datový soubor na dvě poloviny, kdy polovina hodnot je menších než medián a polovina hodnot je větších než medián; tj. jde o 50procentní kvantil.

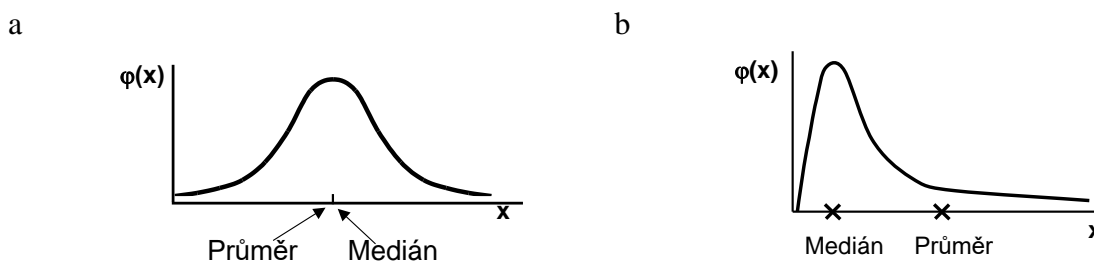
Kvantil je číslo, které dělí uspořádaný datový soubor na části obsahující určitý podíl hodnot. α -kvantil x_α je číslo, které rozděluje uspořádaný datový soubor na dolní úsek, obsahující aspoň podíl α všech dat a na horní úsek obsahující aspoň podíl $1-\alpha$ všech dat. *Dolní kvantil* $x_{0,25}$

(25procentní kvantil) odděluje 25 % hodnot menších nebo rovných než $x_{0,25}$ a 75 % hodnot větších nebo rovných než $x_{0,25}$. Horní kvantil $x_{0,75}$ (75procentní kvantil) odděluje 75 % hodnot menších nebo rovných než $x_{0,75}$ a 25 % hodnot větších nebo rovných než $x_{0,75}$.

Medián je dobrým odhadem frekvenčního středu dat. Průměr jím je pouze v tom případě, když jsou data symetrická a neobsahují odlehlé hodnoty (tzv. odlehlá pozorování). Musíme pamatovat na to, že průměr je silně ovlivněn odlehlými hodnotami, medián jimi ovlivněn není. V případě symetrického rozdělení jsou průměr a medián v podstatě shodné (Obrázek 4.2a), v případě asymetrického rozdělení však nikoliv (Obrázek 4.2b).

Výpočet střední hodnoty je nezbytné doplnit také informací o tom, jak jsou kolem této hodnoty rozložena ostatní pozorování. Nejjednodušší charakteristikou variability pozorovaných dat je *rozsah hodnot (rozpětí)*, který je dán minimální a maximální pozorovanou hodnotou. Toto je ovšem velice citlivé k odlehlým hodnotám. Proto se častěji používá tzv. interkvartilové rozpětí, tj. rozsah hodnot dán dolním a horním kvantilem.

Další mírou variability je *rozptyl*, který hodnotí fluktuaci pozorovaných hodnot kolem průměru, a jeho odmocnina, kterou nazýváme *směrodatná odchylka*.



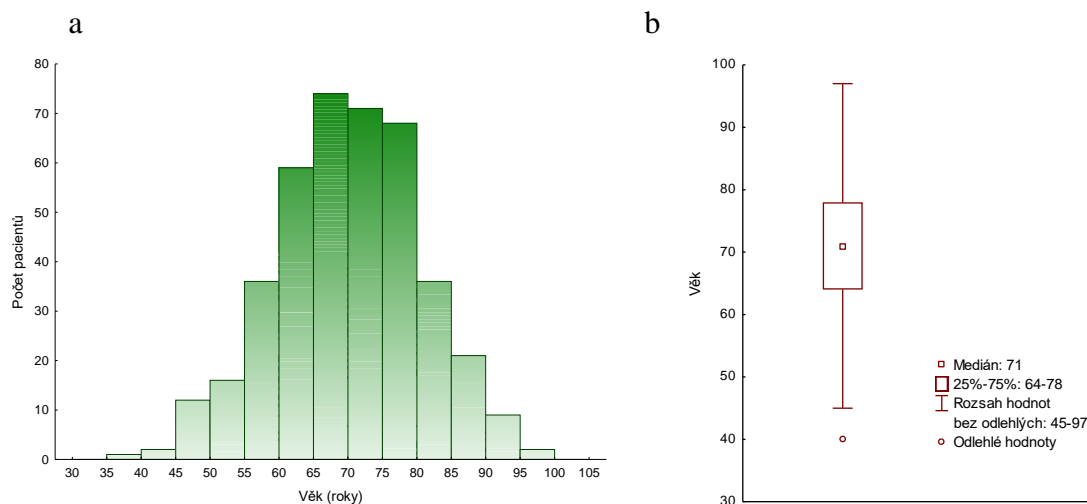
Obrázek 4.2. Ukázka pozice průměru a mediánu u proměnné se symetrickým rozdělením (a) a proměnné s asymetrickým rozdělením (b).

Popisné statistiky proměnné "věk" z našeho datového souboru uvádíme v Tabulce 4.2.

Tabulka 4.2. Popisné statistiky proměnné "Věk" z příkladu v Tabulce 2.1.

Popisná statistika	Hodnota statistiky
Počet hodnot (N)	407
Průměr	70,6
Medián	71,0
Minimum–maximum	40,0–97,0
Dolní kvartil – horní kvartil	64,0–78,0
Směrodatná odchylka	10,3

Pro zobrazení zastoupení hodnot v souboru je vhodný *histogram* (Obrázek 4.3a) nebo *krabicový graf* (Obrázek 4.3b).



Obrázek 4.3. Ukázka zobrazení kvantitativní proměnné "Věk" v histogramu (a) a krabicovém grafu (b).

Příklad: Popis kvantitativních dat

Zadání: Proveďte základní popis soběstačnosti dle indexu Barthelové na konci rehabilitace po mozkovém infarktu. Následně také tento ukazatel srovnajte podle míry komplikací během léčby.

Postup v programu Statistica

1. Pro popis dat je vhodné určit průměr, medián, směrodatnou odchylku, případně minimum a maximum. V menu *Statistics* zvolíme *Basic statistics* a vybereme *Descriptive statistics*. Vybereme proměnnou (*Variables*), kterou chceme analyzovat a na záložce *Advanced* zaškrtneme možnosti výpočtu *Mean, Median, Std. Dev, Min. & Max*. Chceme-li získat výsledky zvlášť pro podskupiny jiné proměnné, použijeme *By Group* (vybereme třídící proměnnou a zaškrtneme *Accumulate tabular results in a single spreadsheet*). Kliknutím na tlačítko *Summary* získáme výsledky.
2. Grafické znázornění je ideální pomocí histogramu. V případě srovnávání různých skupin je vhodný krabicový graf. V menu *Graphs* zvolíme rovnou *Histogram* nebo *Box* (krabicový graf). U histogramu pouze vybereme proměnnou (*Variables*), kterou chceme analyzovat. U krabicového grafu vybereme proměnnou (*Variables*), kterou chceme analyzovat (*dependent*), a proměnnou obsahující skupiny, které srovnáváme (*grouping*). Kliknutím na tlačítko *OK* získáme výstupy.

Výsledky v programu Statistica

Popisná statistika indexu Barthelové na konci rehabilitace u pacientů s CMP

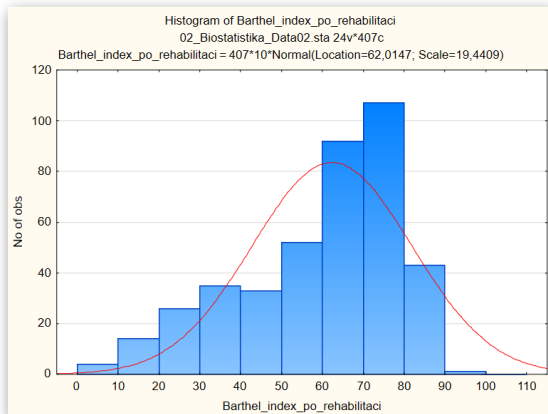
All Groups						
Descriptive Statistics (02_Biostatistika_Data02.sta)						
Variable	Valid N	Mean	Median	Minimum	Maximum	Std.Dev.
Barthel_index_po_rehabilitaci	407	62,01474	70,00000	10,00000	95,00000	19,44095

Popisná statistika indexu Barthelové na konci rehabilitace u pacientů s CMP dle stupně komplikací

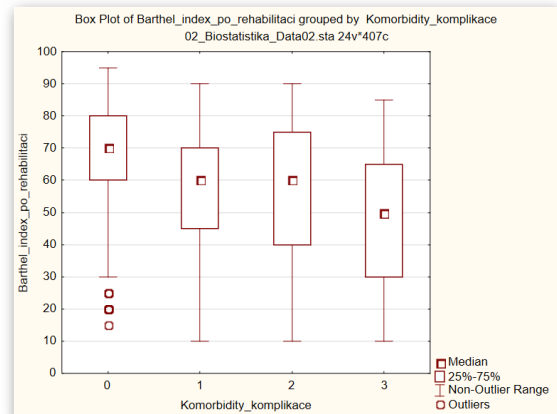
Aggregate Results							
Descriptive Statistics (02_Biostatistika_Data02.sta)							
Variable	Komorbidity_komplikace	Valid N	Mean	Median	Minimum	Maximum	Std.Dev.
Barthel_index_po_rehabilitaci	0	239	66,84100	70,00000	15,00000	95,00000	17,45221
Barthel_index_po_rehabilitaci	1	79	57,15190	60,00000	10,00000	90,00000	19,15998
Barthel_index_po_rehabilitaci	2	50	56,70000	60,00000	10,00000	90,00000	20,01556
Barthel_index_po_rehabilitaci	3	39	49,10256	50,00000	10,00000	85,00000	21,36376

Celková průměrná hodnota indexu Barthelové je 62 bodů s mediánem 70 bodů. V závislosti na skóre komorbidit a komplikací během léčby je zřetelný pokles výsledné soběstačnosti s průměrem 67 bodů u nekomplikovaných případů až k 49 bodům u pacientů se stupněm komplikací 3. Tvar rozdělení indexu Barthelové je asymetrický s hodnotami vyskytujícími se hlavně v rozmezí cca 50 až 90 bodů. V závislosti na skóre komorbidit a komplikací během léčby je zřetelný pokles výsledné soběstačnosti.

Histogram indexu Barthelové na konci rehabilitace u pacientů s CMP

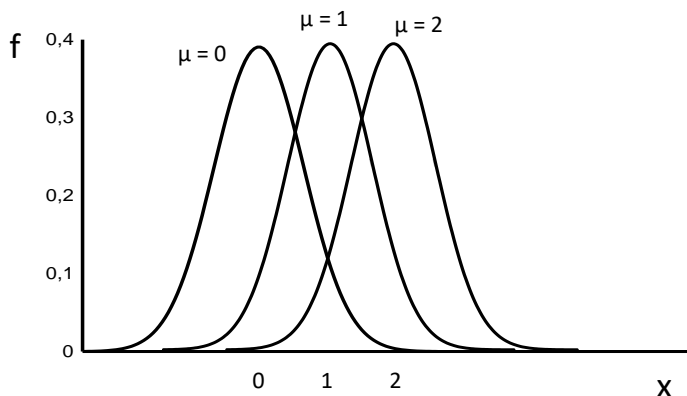


Krabicový graf indexu Barthelové na konci rehabilitace dle stupně komplikací



5. Normální rozdělení

Mnoho kvantitativních proměnných se chová podle jistých pravidel. Toto chování proměnných lze popsat rozdělením pravděpodobnosti. Nejklasičtějším modelovým rozdělením je tzv. *normální rozdělení*, známé též jako Gaussovo rozdělení. Normální rozdělení je rozdělením spojité proměnné na intervalové škále. Běžně se jím ale charakterizují i spojitá data na poměrové stupnici, pokud je průměr alespoň o několik směrodatných odchylek větší než nula a také v některých případech diskrétní data, pokud mohou nabývat dostatečného počtu diskrétních hodnot. Normální rozdělení se vyznačuje symetrií a nepřítomností odlehlých hodnot. Je kompletně popsáno dvěma parametry – průměrem a rozptylem (Obrázek 5.1.).



Obrázek 5.1. Ukázka normálního rozdělení s hodnotami průměru (μ) 0, 1 nebo 2. Hodnota rozptylu (σ) je stejná, vždy 1.

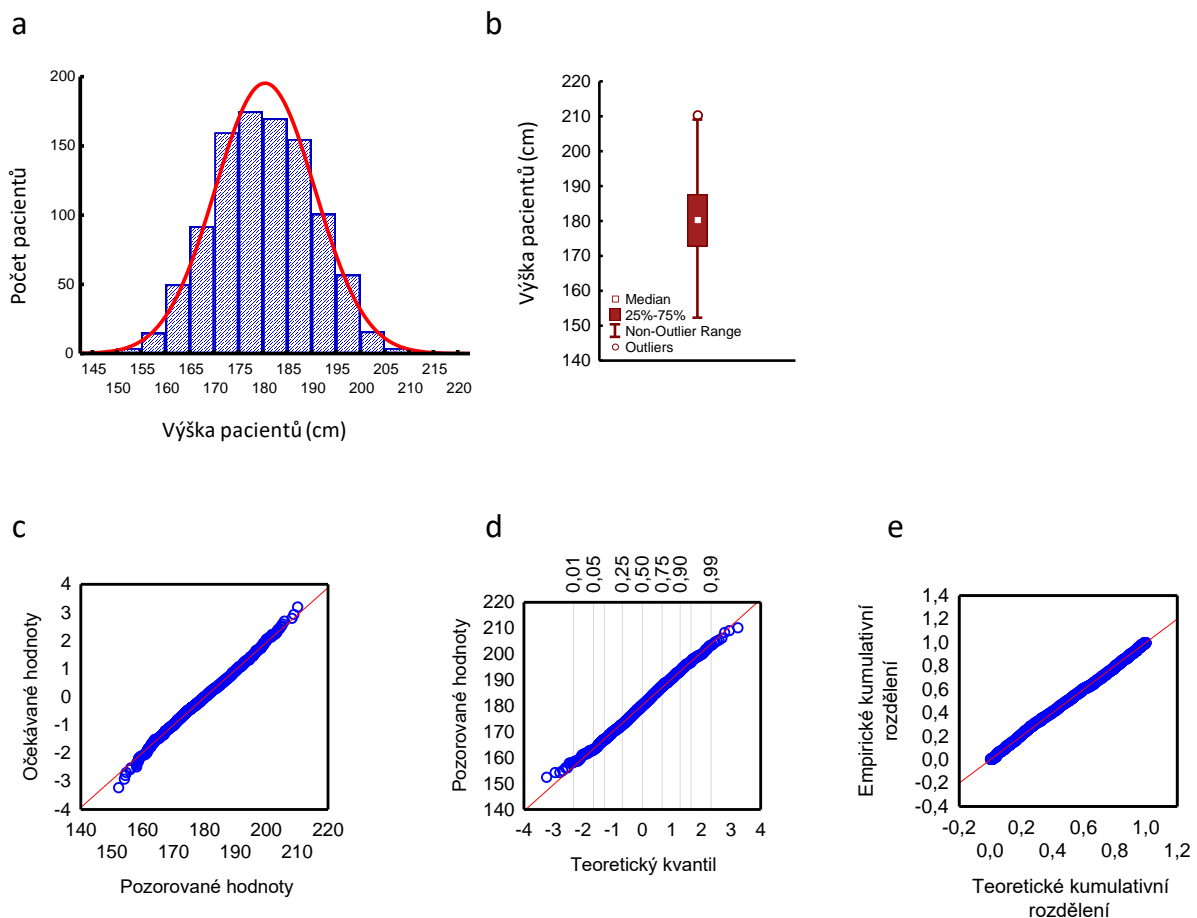
Velké množství statistických metod předpokládá, že data, která zpracováváme, mají normální rozdělení. Proto je ve statistice vždy potřebné zjistit, zda sledovaná proměnná má nebo nemá normální rozdělení. Pro ověření normality existuje řada grafických popisných metod i statistických testů.

Nejjednodušším způsobem vizuálního ověření normality je posouzení vzhledu histogramu četností sledované proměnné. U histogramu je nevyhnutné si všimnout, zda je symetrický a zda je srovnatelný s teoretickou křivkou hustoty pravděpodobnosti normálního rozdělení, která má stejnou střední hodnotu a směrodatnou odchylku (Obrázek 5.2a). Symetrii můžeme ověřit i u krabicového grafu zobrazujícího medián jako střední hodnotu a zachycujícího variabilitu pomocí dolního a horního kvartilu (Obrázek 5.2b).

Dalším způsobem vizuálního ověření normality je použití diagnostických grafů: *normálně-pravděpodobnostní graf*, *kvantil-kvantilový graf* a *pravděpodobnostně-pravděpodobnostní graf*. Normálně-pravděpodobnostní graf (N-P graf) vykresluje pozorované hodnoty vůči hodnotám teoretického normálního rozdělení. Pozorované hodnoty jsou vykresleny vůči očekávaným hodnotám pocházejícím z normálního rozdělení takovým způsobem, aby tvořili přibližně přímku. Odchyly od přímky znamenají odchylky od normálního rozdělení (Obrázek 5.2c). Kvantil-kvantilový graf (Q-Q graf) zobrazuje kvantily

dvou sad proti sobě. V případě použití teoretického normálního rozdělení jako druhé sady dat lze zobrazit pozorovaný kvantil vůči teoretickému kvantilu z normálního rozdělení (Obrázek 5.2d). Pravděpodobnostně-pravděpodobnostný graf (P-P graf) podobně jako Q-Q graf zobrazuje dvě sady dat proti sobě. V případě použití teoretického normálního rozdělení jako druhé sady dat lze porovnat pozorovaná data s normálním rozdělením vynesením pozorovaných a očekávaných hodnot kumulativního rozdělení (Obrázek 5.2e). V obou grafech platí podobně jako u N-P grafu, že když zobrazené body tvoří přibližně přímku, data pocházejí z normálního rozdělení. Odchyšky od přímky znamenají odchylky od normálního rozdělení.

Statistickým testům pro ověření normality rozdělení se budeme věnovat v další kapitole.



Obrázek 5.2. Vizuální ověření normality pomocí histogramu (a), krabicového grafu (b), normálně-pravděpodobnostního grafu (c), kvantil-kvantilového grafu (d) a pravděpodobnostně-pravděpodobnostního grafu (e).

Příklad: Vizualní ověření normality věku

Zadání: Ověřte vizuálními metodami normalitu věku při mozkovém infarktu.

Postup v programu Statistica

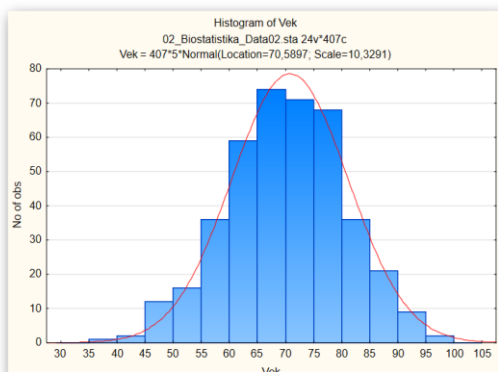
1. Srovnání průměru a mediánu. V menu *Statistics* zvolíme *Basic Statistics* a vybereme *Descriptive Statistics*. Zvolíme proměnnou (*Variables*), kterou chceme analyzovat a na záložce *Advanced* zaškrtneme možnosti *Mean*, *Median*. Kliknutím na tlačítko *Summary* získáme výsledky.
2. Histogram. V menu *Graphs* zvolíme *Histogram*. Vybereme proměnnou (*Variables*), kterou chceme analyzovat. Kliknutím na tlačítko *OK* získáme výstupy.
3. Krabicový graf. V menu *Graphs* zvolíme *2D* a vybereme *Box Plots*. Vybereme proměnnou (*Variables*), kterou chceme analyzovat (*dependent*). Kliknutím na tlačítko *OK* získáme výstupy.
4. Diagnostický N-P graf. V menu *Graphs* zvolíme *2D* a vybereme *Normal Probability Plots*. Vybereme proměnnou (*Variables*), kterou chceme analyzovat. Kliknutím na tlačítko *OK* získáme výstupy.

Výsledky v programu Statistica

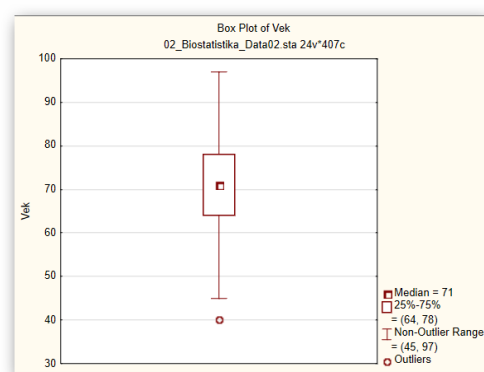
Srovnání průměru a mediánu věku pacientů s CMP.

Descriptive Statistics (02_Biostatistik			
Variable	Valid N	Mean	Median
Vek	407	70,58968	71,00000

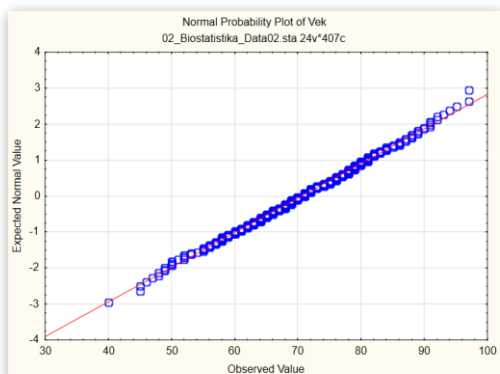
Histogram věku u pacientů s CMP



Krabicový graf věku u pacientů s CMP



N-P graf věku u pacientů s CMP



Průměr a medián jsou téměř shodné (cca 71 let) a data jsou tedy nejspíš alespoň symetrická. Symetrie je patrná i z krabicového grafu. Navíc histogram naprosto jasně odpovídá průběhu normálního rozdělení. Z N-P grafu také nejsou patrné odchylky od normality.

Příklad: Vizuální ověření normality věku s odlehlou hodnotou

Zadání: Ověřte vizuálními metodami normalitu věku při mozkovém infarktu obsahující odlehlou hodnotu.

Postup v programu Statistica

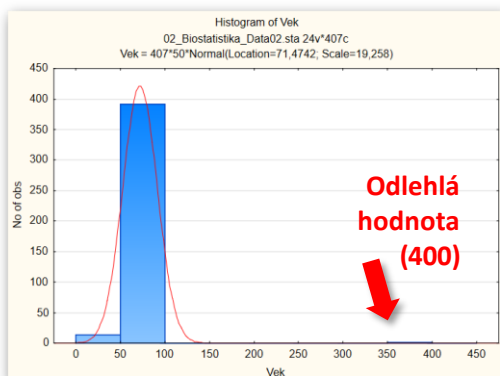
1. Vytvoříme odlehlou hodnotu přepsáním hodnoty 40 na řádku 124 na hodnotu 400.
2. Postupujeme všemi kroky stejně jako v příkladu 5.1, tj. srovnáme průměr a medián, vykreslíme histogram, krabicový graf a diagnostický N-P graf.

Výsledky v programu Statistica

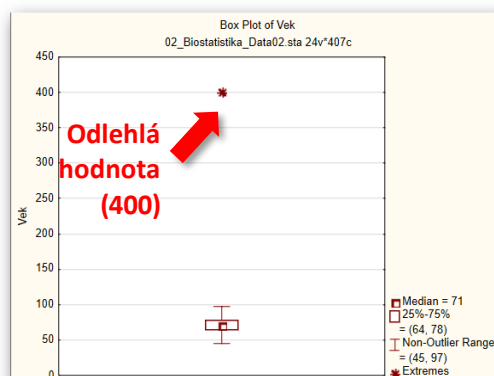
Srovnání průměru a mediánu věku pacientů s CMP.

Variable	Descriptive Statistics (02_Biostatistik		
	Valid N	Mean	Median
Vek	407	71,47420	71,00000

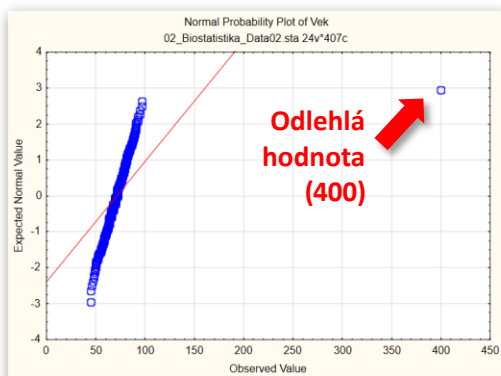
Histogram věku s odlehlou hodnotou u pacientů s CMP



Krabicový graf věku s odlehlou hodnotou u pacientů s CMP



N-P graf věku s odlehlou hodnotou u pacientů s CMP



Průměr a medián jsou stále podobné (cca 71 let) a data by tedy mohla být alespoň symetrická. Ze všech tří grafických nástrojů lze identifikovat výskyt odlehlé/chybné hodnoty, jejíž přítomnost zkresluje pohled na zbytek souboru.

Příklad: Vizuelní ověření normality indexu Barthelové

Zadání: Ověřte vizuálními metodami normalitu indexu Barthelové (vyjadřuje stupeň soběstačnosti v základních denních aktivitách) na konci akutní hospitalizační péče o pacienty s CMP.

Postup v programu Statistica

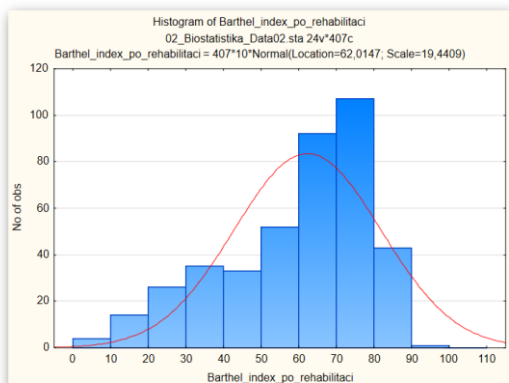
1. Postupujeme všemi kroky stejně jako v příkladu 5.1, tj. srovnáme průměr a medián, vykreslíme histogram, krabicový graf a diagnostický N-P graf.

Výsledky v programu Statistica

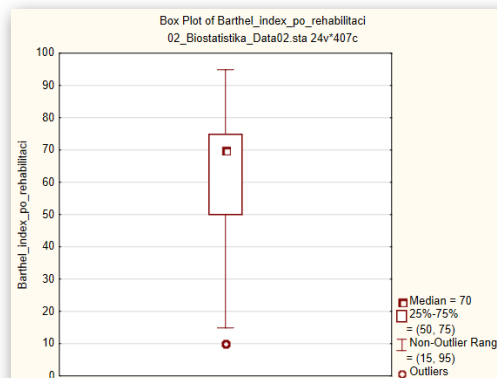
Srovnání průměru a mediánu věku pacientů s CMP.

Variable	Descriptive Statistics (02_Biostatistik)		
	Valid N	Mean	Median
Barthel_index_po_rehabilitaci	407	62,01474	70,00000

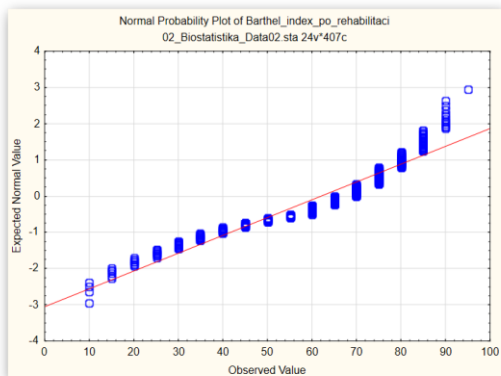
Histogram indexu Barthelové u pacientů s CMP



Krabicový graf indexu Barthelové u pacientů s CMP



N-P graf indexu Barthelové u pacientů s CMP



Průměr a medián se výrazně liší (průměr 62 bodů, medián 70 bodů), což znamená, že data jsou nejspíše asymetrická. Asymetrie je patrná i z krabicového grafu a histogramu. Z histogramu je navíc zřetelně vidět odlišnost od normálního rozdělení. Odchytky od normality jsou patrné i z N-P grafu.

6. Testování hypotéz

V předešlé části jsme představili popisné statistiky a normální rozdělení pravděpodobnosti. Pokud se chceme posunout od pouhého popisu ke srovnávacím analýzám, musíme představit princip *testování hypotéz*. Pomocí statistických testů budeme pak schopni:

- rozhodnout o charakteru rozdělení pravděpodobnosti proměnné; ověřit normalitu rozdělení
- porovnat hodnoty proměnné s referenční hodnotou, tj. porovnat výběrovou charakteristiku s předpokládanou hodnotou (pod pojmem *výběr* rozumíme soubor hodnot dané proměnné u sledované skupiny pacientů apod.),
- porovnat výběrovou charakteristiku dvou náhodných výběrů mezi sebou,
- porovnat výběrovou charakteristiku tří nebo více náhodných výběrů mezi sebou,
- zhodnotit změnu v hodnotách sledované proměnné,
- rozhodnout o nezávislosti dvou kvalitativních proměnných,
- rozhodnout o významnosti korelace dvou kvantitativních proměnných.

Statistické hypotézy nejsou nic jiného než tvrzení, které lze na základě pozorovaných hodnot pomocí statistických metod ohodnotit. Rozlišujeme tzv. nulovou a alternativní hypotézu. *Nulová hypotéza* (H_0) je tvrzení, které je vždy postaveno jako nepřítomnost efektu. Jinak řečeno, nulová hypotéza odráží fakt, že se něco nestalo nebo neprojevilo, a je tedy stanovena jako opak toho, co chceme experimentem prokázat. *Alternativní hypotéza* (H_A) je tvrzení, které popírá nulovou hypotézu. Vymezuje, jaká situace nastává, když nulová hypotéza neplatí. Platnost hypotéz ověřujeme pomocí *statistického testu*. Výsledné rozhodnutí nabývá pouze dvě možnosti: nulovou hypotézu H_0 nezamítáme nebo naopak, nulovou hypotézu H_0 zamítáme.

Rozhodování o zamítnutí nebo nezamítnutí nulové hypotézy je spojeno s dvěma typy chyb. Ty jsou standardně označovány jako chyba I. druhu (její pravděpodobnost značíme jako α) a chyba II. druhu (její pravděpodobnost značíme jako β). Pravděpodobnost chyby I. druhu souvisí s falešně pozitivním závěrem testu, kdy na základě výsledku testu zamítneme nulovou hypotézu, která ale ve skutečnosti platí. Podobně, pravděpodobnost chyby II. druhu souvisí s falešně negativním závěrem testu, kdy na základě výsledku testu nezamítneme nulovou hypotézu, která ale ve skutečnosti neplatí. V biostatistice je za důležitější považována chyba I. druhu, kterou se snažíme omezit na přijatelné minimum. Jako standardní hranice je přijímána hranice 5 %, tj. $\alpha = 0,05$. Je to vlastně hladina významnosti, na které statistický test provádíme. Na tomto místě je ještě potřebné představit pojem *síla testu*. Jde o pravděpodobnostně vyjádřenou schopnost rozpoznat neplatnost nulové hypotézy (označujeme ji $1 - \beta$).

Rozhodnutí o zamítnutí nebo nezamítnutí nulové hypotézy činíme dle hodnoty *testové statistiky* jejím porovnáním s kritickou hodnotou nebo dle tzv. *p-hodnoty*, která vyjadřuje pravděpodobnost za platnosti H_0 s níž bychom získali stejnou nebo extrémnější (ještě méně pravděpodobnou) hodnotu testové statistiky. Hodnotu testové statistiky i p-hodnotu získáme při statistickém testování ve statistickém programu. Pro zjednodušení se můžeme orientovat pouze

podle p-hodnoty. Stručně řečeno, vyjde-li nám p-hodnota nižší než stanovená hladina významnosti α (standardně je přijímána hranice 5 %), znamená to, že nulovou hypotézu můžeme zamítnout. Je-li p-hodnota větší než stanovená hladina významnosti α , pak nulovou hypotézu nezamítáme.

Statistické testování probíhá následovně:

1. Formulujeme nulovou hypotézu H_0 , která tvrdí, že sledovaný efekt je nulový. Např. v našem případě H_0 : věk u mužů a žen se neliší. Zároveň zvolíme hladinu významnosti testu α (nejčastěji $\alpha = 0,05$), která představuje pravděpodobnost získání falešně pozitivního výsledku.
2. Formulujeme alternativní hypotézu H_A . V našem případě H_A : věk u mužů a žen se liší.
3. Zvolíme adekvátní statistický test, vypočítáme hodnotu testové statistiky a p-hodnotu výsledku, kterou srovnáme se zvolenou hladinou významnosti testu. Pokud je p-hodnota menší než stanovená hladina významnosti α , zamítáme nulovou hypotézu H_0 . Pokud není p-hodnota menší než α , pak H_0 nezamítáme. V našem příkladu o věku mužů a žen byla výsledná p-hodnota menší než 0,05, tudíž zamítáme nulovou hypotézu, tj. věk mužů a žen se liší.

6.1. Interpretace výsledku statistického testu

Při interpretaci výsledku statistického testu je potřebné pamatovat na následující situace.

1. Nezamítnutí nulové hypotézy neznamená automaticky její přijetí. Může se jednat o situaci, kdy pro zamítnutí nulové hypotézy nemáme dostatečné množství informace.
2. Dosažená hladina významnosti testu (ať už 5 %, 1 % nebo 10 %) nesmí být slepě brána jako hranice pro (ne)existenci testovaného efektu.
3. Malá p-hodnota nemusí znamenat velký efekt. Hodnota testové statistiky a p-hodnota mohou být ovlivněny velkou velikostí vzorku a malou variabilitou pozorovaných dat.
4. Na výsledky testování musí být nahlíženo kriticky. Jedná se o závěr založený „pouze“ na jednom výběrovém souboru.
5. Statistická významnost indikuje, že pozorovaný rozdíl není náhodný, ale nemusí znamenat, že je významný i ve skutečnosti. Důležitá je i praktická (klinická) významnost.

6.2. Přehled statistických testů

V těchto učebních textech se budeme nejdříve věnovat testování hypotéz o spojitých (kvantitativních) proměnných, tedy těch, které mohou nabývat jakýchkoliv hodnot v určitém intervalu. Klasickými příklady jsou výška postavy, hmotnost jedince, nebo časové a teplotní měření. Tyto testy můžeme použít i v případě diskrétních proměnných, ale pouze v případě, že nabývají velký počet možných hodnot. Příkladem může být počet červených krvinek v 1 ml krve, což není spojitá náhodná veličina (počet krvinek je jenom celé číslo, nejsou možné

desetinná čísla), ale počet možných hodnot je natolik velký, že nás opravňuje použít pro testování hypotéz testy pro spojité proměnné. Podobně můžeme za spojitou proměnnou považovat i věk z našeho příkladu.

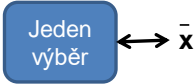
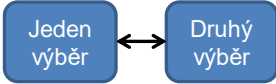
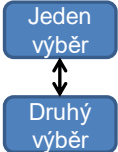

Statistické testy můžeme rozdělit na dvě velké skupiny: *parametrické testy* a *neparametrické testy* (Obr. 6.1). Zásadním rozdílem mezi nimi je zejména nutnost předpokladu o rozdělení dat u parametrických testů. Neparametrické testy nemají předpoklady o rozdělení vstupujících dat. Velkou výhodou neparametrických testů je fakt, že pracují s pořadími hodnot, a tím nejsou citlivá vůči odlehlým hodnotám. Nevýhodou neparametrických testů je ovšem jejich snížená síla testu, tj. snížená schopnost zamítnout nulovou hypotézu, což je následně nutné kompenzovat větší velikostí vzorku. To může být problém hlavně v případě menších souborů. Parametrické testy předpokládají normální rozdělení dat.

Dále rozlišujeme testy jednovýběrové, dvouvýběrové a testy pro více výběrů. Testy jednovýběrové srovnávají jeden výběr s referenční hodnotou. Dvouvýběrové testy srovnávají dva výběry a testy pro více výběrů srovnávají více výběrů navzájem.



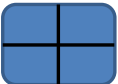
Nakonec musíme zmínit párově a nepárově uspořádaná pozorování. Párově uspořádaná data jsou vzájemně závislá, resp. vázaná nějakým společným prvkem. Klasickým příkladem párových pozorování jsou hodnoty dvou po sobě jdoucích měření na stejném pacientovi, které samozřejmě nelze považovat za nezávislé, neboť jsou vázány osobou pacienta. U nepárově uspořádaných pozorování jsou data na sobě zcela nezávislá, např. dvě skupiny pacientů s odlišnou léčbou a pod. (Obr. 6.1).

V další části uvedeme metody pro hodnocení závislosti kvalitativních proměnných (např. pohlaví a výskyt atopického ekzému). Stejně jako v případě kvantitativních proměnných, tak i u kvalitativních proměnných můžeme hodnotit vztahy mezi nimi pomocí parametrických i neparametrických testů (Obrázek 6.2).

V poslední části uvedeme základy regresní analýzy a představíme parametrický i neparametrický korelační koeficient.

TEST		Parametrický test	Neparametrický test
Jednovýběrový		Jednovýběrový t-test	Wilcoxonův test Znaménkový test
Dvouvýběrový pro závislé výběry		Dvouvýběrový t-test pro závislé výběry (párový t-test)	Wilcoxonův test Znaménkový test
Dvouvýběrový pro nezávislé výběry		Dvouvýběrový t-test pro nezávislé výběry (nepárový t-test)	Mannův-Whitneyův U test
Vícevýběrový pro nezávislé výběry		ANOVA	Kruskalův-Wallisův test

Obrázek 6.1. Přehled statistických testů pro hodnocení kvantitativních proměnných.

TEST		Parametrický test	Neparametrický test
Jeden výběr		Jednovýběrový binomický test	
Více výběrů párová data			McNemarův test
Více výběrů nepárová data		Pearsonův chí-kvadrát test	Fisherův exaktní test

Obrázek 6.2. Přehled statistických testů pro hodnocení kvalitativních proměnných.

6.3. Test normality rozdělení

V této části si představíme statistické testy pro ověření normality rozdělení. V předešlé kapitole jsme představili normální rozdělení a zdůraznili jeho důležitost pro mnoho statistických testů. Je proto nevyhnutné, abychom dokázali normalitu rozdělení dat vyhodnotit. Kromě vizuálního ověření normality bude užitečné i její statistické testování.

Nejčastěji se k testování normality rozdělení používají *Kolmogorov-Smirnov test* a *Shapiro-Wilk test*. Kolmogorov-Smirnov test se zaměřuje zejména na distribuční funkci. Častěji se používá v jeho modifikaci – Lilieforsův test. Shapiro-Wilk test je neparametrický test použitelný i při velmi malých počtech pozorování s dobrou silou testu. Je zaměřen na testování symetrie.

Všechny testy normality pracují s nulovou hypotézou, že není rozdíl mezi pozorovaným rozdělením a normálním rozdělením. Přesnější formulace zní, že data jsou realizací náhodné proměnné s normálním rozdělením.

Příklad: Ověření normality věku

Zadání: Ověřte statistickým testem normalitu věku při mozkovém infarktu.

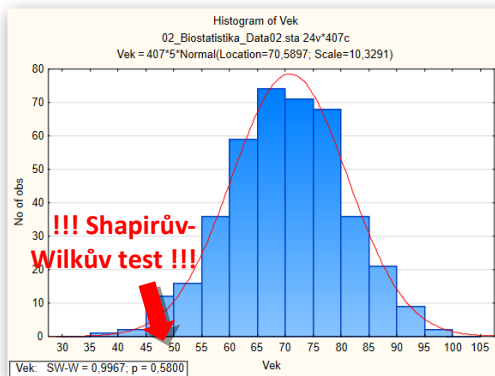
Postup v programu Statistica

1. Při nastavení histogramu zvolíme provedení Shapirova-Wilkova testu nebo Lilieforsovy modifikace Kolmogorovova-Smirnovova testu. V menu **Graphs** zvolíme **Histogram**. Vybereme proměnnou (**Variables**), kterou chceme analyzovat. Na záložce **Advanced** v části **Statistics** zaškrtneme příslušný test. Kliknutím na tlačítko **OK** získáme výstupy.
2. Při nastavení diagnostického N-P grafu zvolíme provedení statistického testu. V menu **Graphs** zvolíme **2D** a vybereme **Normal Probability Plots**. Vybereme proměnnou (**Variables**), kterou chceme analyzovat. Na záložce **Quick** v části **Statistics** zaškrtneme test. Kliknutím na tlačítko **OK** získáme výstupy.

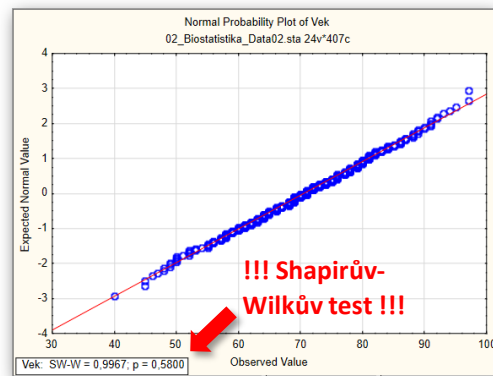
Výsledky v programu Statistica

Připomeňme si, že na hladině významnosti $\alpha = 0,05$ testujeme nulovou hypotézu H_0 : není rozdíl mezi pozorovanými daty a teoretickým normálním rozdělením. Když je vypočítaná p-hodnota menší než hladina významnosti, zamítáme nulovou hypotézu.

Histogram věku u pacientů s CMP



N-P graf věku u pacientů s CMP



Na základě p-hodnoty 0,580 nezamítáme nulovou hypotézu o normalitě (tj. nezamítáme, že není rozdíl mezi pozorovanými daty a teoretickým normálním rozdělením, tj. data jsou normálně rozdělená).

Příklad: Ověření normality věku s odlehlou hodnotou

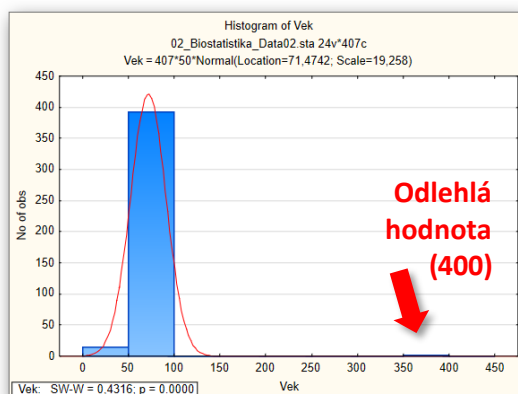
Zadání: Ověřte statistickým testem normalitu věku při mozkovém infarktu obsahující odlehlou hodnotu.

Postup v programu Statistica

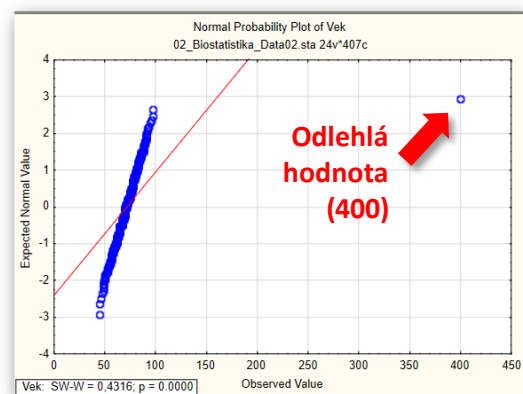
1. Vytvoříme odlehlou hodnotu přepsáním hodnoty 40 na řádku 124 na hodnotu 400.
2. Postupujeme všemi kroky stejně jako v příkladu 6.1, tj. nastavíme provedení Shapirova-Wilkova testu u histogramu i diagnostického N-P grafu.

Výsledky v programu Statistica

Histogram věku s odlehlou hodnotou u pacientů s CMP



N-P graf věku s odlehlou hodnotou u pacientů s CMP



Na základě p-hodnoty $< 0,001$ zamítáme nulovou hypotézu o normalitě (tj. zamítáme, že není rozdíl mezi pozorovanými daty a teoretickým normálním rozdělením, tj. data nejsou normálně rozdělená).

Příklad: Ověření normality indexu Barthelové

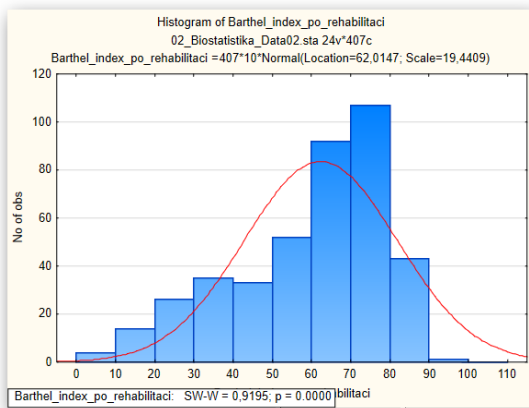
Zadání: Ověřte statistickým testem normalitu indexu Barthelové (vyjadřuje stupeň soběstačnosti v základních denních aktivitách) na konci akutní hospitalizační péče o pacienty s mozkovým infarktem.

Postup v programu Statistica

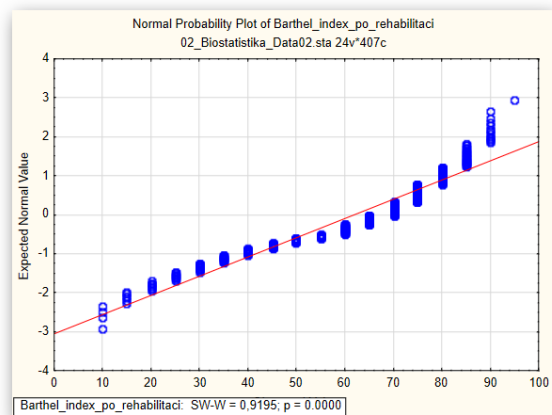
1. Postupujeme všemi kroky stejně jako v příkladu 6.1, tj. nastavíme provedení Shapiro-Wilkova testu u histogramu i diagnostického N-P grafu.

Výsledky v programu Statistica

Histogram indexu Barthelové u pacientů s CMP



N-P graf indexu Barthelové u pacientů s CMP



Na základě p-hodnoty $< 0,001$ zamítáme nulovou hypotézu o normalitě (tj. zamítáme, že není rozdíl mezi pozorovanými daty a teoretickým normálním rozdělením, tj. data nejsou normálně rozdělená).

7. Parametrické testy pro kvantitativní proměnné

V této části uvedeme základní parametrické testy pro testování kvantitativních proměnných. Jde o tzv. t-testy, které jsou v praktické statistice velice často používány k testování hypotéz o jednom a dvou výběrových souborech, a analýzu rozptylu (ANOVA), která testuje hypotézu o více výběrech. Tyto parametrické testy se vyznačují silným předpokladem normality dat, jež je třeba vždy ověřit, např. pomocí grafických metod a pomocí Shapiro-Wilkova testu.

7.1. Jednovýběrový t-test

Cílem jednovýběrového t-testu je srovnání výběrového průměru s hypotetickým průměrem základního souboru. Z hlediska statistické teorie jde o ověření, zda daný vzorek pochází z rozdělení se stejnou střední hodnotou jako je předpokládaná konstanta μ_0 . Předpokladem je, že data pochází z normálního rozdělení. Tento předpoklad je vhodné ověřit vizuálně i statistickým testem. U jednovýběrového t-testu nepředpokládáme znalost rozptylu σ .

Postup výpočtu jednovýběrového t-testu je následovný (poté, co ověříme předpoklady).

1. Stanovíme nulovou a alternativní hypotézu. $H_0: \bar{x} = \mu_0$, $H_A: \bar{x} \neq \mu_0$.
2. Vypočítáme aritmetický průměr a rozptyl výběru a určíme počet pozorování.
3. Vypočítáme hodnotu testové statistiky $t = \frac{\bar{x} - \mu_0}{s} \sqrt{n}$ a odpovídající p-hodnotu.
4. Vypočítané t porovnáme s kritickou hodnotou nebo porovnáme p-hodnotu s hladinou významnosti α .
5. Když je $|t| >$ kritická hodnota nebo p-hodnota $\leq \alpha$, zamítáme nulovou hypotézu.

Výpočet t-testu je k narušení předpokladu normality rozložení dat robustní a to tím víc, čím větší výběr zpracováváme. Z výše uvedeného postupu testování je zřejmé, že do vzorce pro t-test nevstupují původní data, ale průměr a výběrová směrodatná odchylka. Předpoklad normality původních dat je nutný kvůli zajištění normálního rozdělení výběrového průměru. Pro průměr spočtený z velkého množství dat to obvykle platí i v případě, když se původní data od normálního rozdělení odchylují.

Příklad: Jednovýběrový t-test

Zadání: Ústav zdravotnických informací a statistiky (ÚZIS) v rámci celorepublikové zdravotnické statistiky publikoval průměrný věk pacientů s mozkovým infarktem 71,6 let. Ověřte, zda váš datový soubor věkově odpovídá celorepublikové hodnotě, anebo zda se vámi hodnocení pacienti věkově vymykají obecnému průměru.

Postup v programu Statistica

1. Ověříme předpoklady testu – normalita rozložení věku pacientů (ověříme vizuálně i statistickým testem – Shapiro-Wilkův test). Postupujeme všemi kroky stejně jako v příkladu 6.1, tj. nastavíme provedení Shapiro-Wilkova testu u histogramu i diagnostického N-P grafu.

Postup výpočtu t-testu (po ověření předpokladu normality)

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu $H_0: \bar{x} = 71,6$, $H_A: \bar{x} \neq 71,6$.
2. Vypočítáme **aritmetický průměr** a **rozptyl** výběrového souboru a určíme počet pozorování.
3. Vypočítáme testovou statistiku **t** a odpovídající **p-hodnotu**. V menu *Statistics* zvolíme *Basic statistics*, vybereme *t-test, single sample*. Vybereme proměnnou, kterou chceme testovat (*Variables*). Na kartě *Quick* napíšeme do pole *Test all means against* velikost střední hodnoty populace (lze také na kartě *Advanced, Options*). Kliknutím na *Summary t-test* nebo na *Summary* získáme výstupy.

$$t = \frac{\bar{x} - \mu}{s} \sqrt{n} = \frac{70,6 - 71,6}{10,3} \sqrt{407} = -1,973, \text{ odpovídající p-hodnota} = 0,049.$$

4. Vypočítané t porovnáme s kritickou hodnotou, nebo porovnáme p-hodnotu s hladinou významnosti $\alpha = 0,05$.

Je-li **p-hodnota** $\leq \alpha$, **zamítáme H_0** . Věk pacientů je odlišný od celorepublikového průměru.

Výsledky v programu Statistica

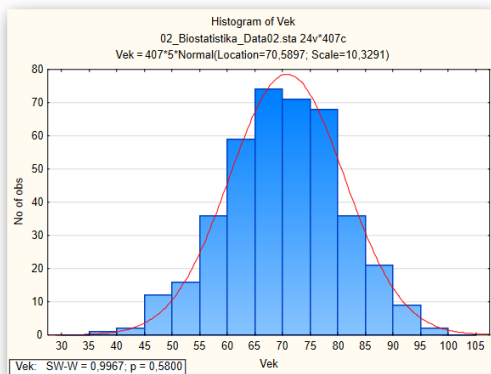
Srovnání průměru a mediánu

Variable	Descriptive Statistics (02_Biostatistik)		
	Valid N	Mean	Median
Věk	407	70,58968	71,00000

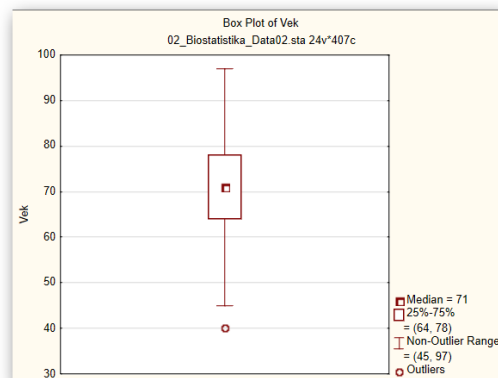
Průměr a medián jsou téměř shodné (cca 71 let) a data jsou tedy nejspíš alespoň symetrická.

Ověření normality rozložení dat

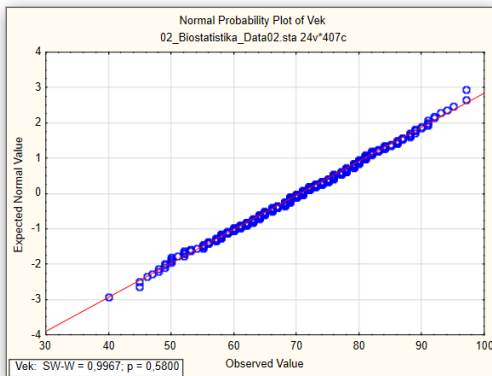
Histogram věku pacientů s CMP



Krabicový graf věku pacientů s CMP



N-P graf věku pacientů s CMP



Symetrie je patrná i z krabicového grafu. Navíc histogram naprosto jasně odpovídá průběhu normálního rozdělení. Z N-P grafu také nejsou patrné odchylky od normality. Na základě p-hodnoty 0,580 nezamítáme nulovou hypotézu o normalitě (tj. nezamítáme, že není rozdíl mezi pozorovanými daty a teoretickým normálním rozdělením – data jsou normálně rozdělená).

Variable	Mean	Std.Dev.	N	Std.Err.	Reference Constant	t-value	df	p
Vek	70,58968	10,32913	407	0,511996	71,60000	-1,97330	406	0,049140

Výběrový průměr (pozorovaných dat) → Mean
Rozsah výběru → N
Hodnota testové statistiky → t-value
Výběrová směrodatná odchylka (pozorovaných dat) → Std.Dev.
Referenční konstanta (předpokládaná velikost střední hodnoty) → Reference Constant
p-hodnota t-testu → p

Pozorovaný průměrný věk je 70,6 let, což je **o rok méně** než reference 71,6 let. P-hodnota statistické významnosti této pozorované odchylky je $p = 0,049$, což na hladině významnosti 0,05 značí hraničně **významný rozdíl**, a lze tedy usuzovat, že **naši pacienti jsou v průměru mírně mladší ve srovnání s celou populací mozkových infarktů v ČR**.

7.2. Dvouvýběrový párový t-test

Ke srovnání dvou závislých výběrů, kdy jsou skupiny propojeny přes objekt měření, se používá dvouvýběrový párový t-test. Oba výběry musí mít shodný počet pozorování, protože všechna měření v jednom souboru musí být spárována s měřením v druhém souboru. Při vlastním výpočtu se potom počítá se změnou hodnot (diferencí) subjektů. Párový t-test je parametrický test. Nemá žádné předpoklady o rozdělení vstupních dat, protože je počítán až na základě jejich diferencí. Předpokladem pro použití párového t-testu je tedy normalita rozdělení těchto diferencí. Otázkou v párovém t-testu je, zda se průměrná hodnota diferencí rovná nějaké teoretické hodnotě. Typicky jde o srovnání s nulou jako důkaz neexistence změny mezi oběma výběry. Jako příklad použití párového t-testu můžeme uvést srovnání hodnot krevního parametru u pacientů před léčbou a po léčbě apod.

Postup výpočtu párového t-testu je následovný (poté, co ověříme předpoklady).

1. Stanovíme nulovou a alternativní hypotézu. $H_0: \mu_1 - \mu_2 = 0$, $H_A: \mu_1 - \mu_2 \neq 0$.
2. Pro novou proměnnou diferencí prvního a druhého měření vypočítáme aritmetický průměr a rozptyl výběru a určíme počet pozorování.

3. Vypočítáme hodnotu testové statistiky $t = \frac{\bar{x}-0}{s}\sqrt{n}$ a odpovídající p-hodnotu stejně jako u jednovýběrového t-testu oproti nule.
4. Vypočítané t porovnáme s kritickou hodnotou nebo porovnáme p-hodnotu s hladinou významnosti α .
5. Když je $|t| >$ kritická hodnota nebo p-hodnota $\leq \alpha$, zamítáme nulovou hypotézu.

Příklad: Párový t-test

Zadání: Pacientům s mozkovým infarktem byla na lůžku akutní péče poskytnuta terapie pro obnovu krevního oběhu v postižené části mozku. Po zvládnutí akutní fáze byl u pacientů vyhodnocen stupeň soběstačnosti v základních denních aktivitách pomocí indexu Barthelové (BI) a byli přeloženi na rehabilitační oddělení. Po dvou týdnech byl opět vyhodnocen stupeň soběstačnosti dle BI. Zjistěte, zda poskytnutá rehabilitační péče vedla k jeho zlepšení.

Postup v programu Statistica

1. Ověříme předpoklady testu: normalita rozložení rozdílů hodnot BI (vizuálně i Shapiro-Wilkovým testem). Postupujeme všemi kroky stejně jako v příkladu 6.1, tj. nastavíme provedení Shapiro-Wilkova testu u histogramu i diagnostického N-P grafu.

Postup výpočtu párového t-testu (po ověření předpokladu normality)

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu $H_0: \mu_1 - \mu_2 = 0$, $H_A: \mu_1 - \mu_2 \neq 0$.
2. Pro novou proměnnou diferencí prvního a druhého měření vypočítáme aritmetický průměr a rozptyl výběru a určíme počet pozorování.
3. Vypočítáme testovou statistiku t a odpovídající **p-hodnotu** stejně jako u jednovýběrového t-testu oproti nule. V menu *Statistics* zvolíme *Basic statistics*, vybereme *t-test, dependent samples*. Zvolíme obě proměnné (*Variables*). Kliknutím na *Summary* získáme výstupy.

$$t = \frac{\bar{x}-0}{s}\sqrt{n} = \frac{-30,2-0}{15,7}\sqrt{407} = -38,8, \text{ odpovídající p-hodnota} < 0,001.$$

4. Vypočítané t porovnáme s kritickou hodnotou, nebo porovnáme p-hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p-hodnota $\leq \alpha$, zamítáme H_0** . Během rehabilitace došlo ke změně soběstačnosti pacientů.

Výsledky v programu Statistica

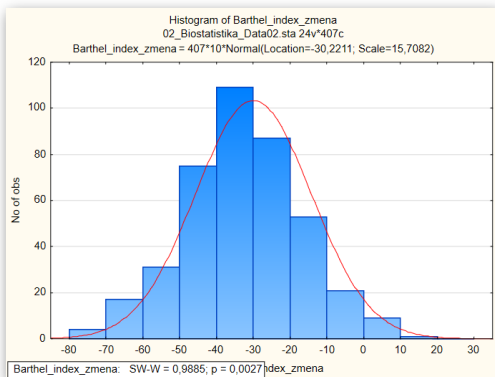
Srovnání průměru a mediánu

Variable	Descriptive Statistics (02_Biostatistik		
	Valid N	Mean	Median
Barthel_index_zmena	407	-30,2211	-30,0000

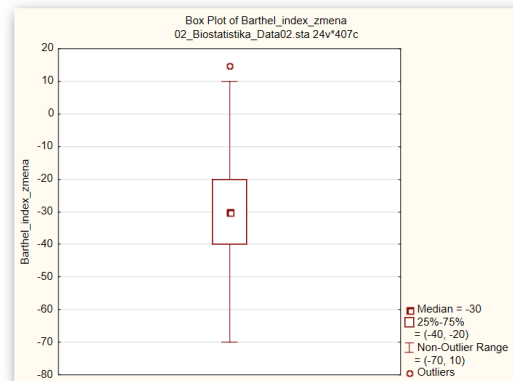
Průměr a medián jsou téměř shodné (cca -30) a data jsou tedy nejspíš symetrická.

Ověření normality rozložení dat

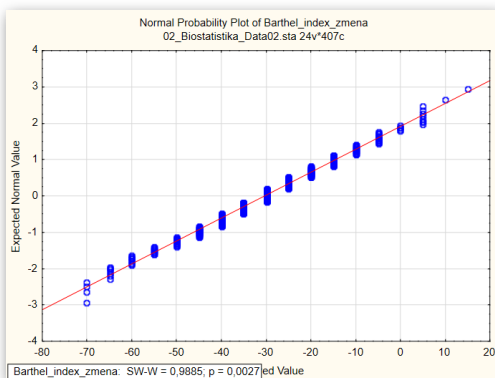
Histogram rozdílů hodnot BI



Krabicový graf rozdílů hodnot BI



N-P graf rozdílů hodnot BI



Symetrie je patrná i z krabicového grafu. Navíc histogram je svým průběhem velmi podobný normálnímu rozdělení. Z N-P grafu také nejsou patrné odchylky od normality. Na základě p-hodnoty 0,003 zamítáme nulovou hypotézu o normalitě (tj. zamítáme, že není rozdíl mezi pozorovanými daty a teoretickým normálním rozdělením, tj. data formálně dle testu nejsou normálně rozdělená). Můžeme si přesto dovolit použít t-test?

	Výběrové průměry obou měření		Rozsah výběru	Hodnota testové statistiky						
Variable	Mean	Std.Dv.	N	Diff.	Std.Dv. Diff.	t	df	p	Confidence -95,000%	Confidence +95,000%
Barthel_index_pred_rehabilitaci	31,79361	9,88115	407	-30,2211	15,70825	-38,8133	406	0,00	-31,7518	-28,6905
Barthel_index_po_rehabilitaci	62,01474	19,44095								

Výběrové směrodatné odchylky obou měření
 Průměr a směrodatná odchylka rozdílu obou měření
 p-hodnota t-testu

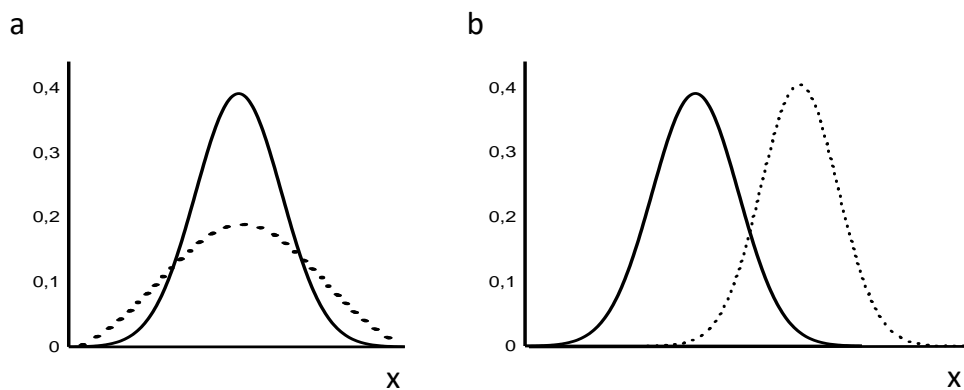
Pozorovaný průměrný Barthelové index na začátku je 31,8 a po rehabilitaci pak 62,0, což je zlepšení o 30,2 bodů. P-hodnota statistické významnosti této pozorované změny je $p < 0,001$, což na hladině významnosti 0,05 značí **významný rozdíl**, a lze tedy prohlásit, že **průměrný stupeň soběstačnosti v základních denních aktivitách se během péče viditelně zlepšil**.

7.3. Dvouvýběrový nepárový t-test

Jedním z nejčastějších úkolů statistické analýzy dat je srovnání spojitých dat ve dvou skupinách pacientů. Základním testem pro srovnání středních hodnot dvou nezávislých výběrů je nepárový dvouvýběrový t-test, který testuje, zda náhodné výběry pochází z rozdělení se středními hodnotami, jejichž rozdíl je daná konstanta c . Umožňuje nám tak posoudit, zda se hodnoty dvou výběrů liší (např. srovnání věku u mužů a žen). Jedná se o parametrický test s těmito předpoklady:

- nezávislost obou srovnávaných výběrů,
- normální rozdělení proměnné v rámci skupin (malé odchylky od normality jsou přípustné, t-test je dostatečně robustní proti drobným odchylkám od tohoto předpokladu),
- shodný rozptyl v obou výběrech (shodnost rozptylu ověřujeme tzv. F-testem).

Přibližná shoda rozptylů obou výběrů je stejně důležitým předpokladem jako normalita jejich rozdělení. Při shodných rozptylech jsou rozdíly středních hodnot lépe identifikovatelné (Obrázek 7.1)



Obrázek 7.1. Hustota pravděpodobnosti dvou normálních rozdělení, lišících se rozptylem (a) a střední hodnotou (b).

Shodu rozptylů testujeme F-testem, jehož nulová hypotéza zní $H_0: \sigma_1 = \sigma_2$ (rozptyly se neliší). V případě shodných rozptylů je vše v pořádku a je možné pokračovat ve výpočtu t-testu. Při zjištění rozdílných rozptylů sledovaných výběrů (tj. když je nulová hypotéza F-testu o shodě rozptylů zamítnuta) není vhodné t-test počítat v jeho původní formě. Existuje ovšem řešení této situace, nejčastěji jde o úpravu výpočtu t-testu pro rozdílné rozptyly tzv. Welchovou korekcí.

Postup výpočtu t-testu pro dva nezávislé výběry je následovný (poté, co ověříme předpoklady normality rozdělení v každém ze sledovaných výběrů Shapirovým-Wilkovým testem a shody rozptylů F-testem).

1. Stanovíme nulovou a alternativní hypotézu. $H_0: \mu_1 = \mu_2$, $H_A: \mu_1 \neq \mu_2$.
2. Pro oba výběry vypočítáme aritmetický průměr a rozptyl výběrového souboru a určíme počet pozorování.
3. Vypočítáme testovou statistiku t a odpovídající p -hodnotu.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

4. Vypočítané t porovnáme s kritickou hodnotou nebo porovnáme p -hodnotu s hladinou významnosti α .
5. Když je $|t| >$ kritická hodnota nebo p -hodnota $\leq \alpha$, zamítáme nulovou hypotézu.

Příklad: Nepárový t-test

Zadání: V literatuře se často uvádí, že mozkový infarkt postihuje ženy v pozdějším věku než muže. Zjistěte na základě příkladových dat, zda je věk pacientů dle pohlaví stejný, anebo zda se věk mužů a žen skutečně liší.

Postup v programu Statistica

1. Ověříme předpoklady testu: normalita rozložení věku mužů a žen samostatně (vizuálně i Shapiro-Wilkovým testem). Postupujeme všemi kroky stejně jako v příkladu 6.1, tj. nastavíme provedení Shapiro-Wilkova testu u histogramu i diagnostického N-P grafu. Shodu rozptylů věku žen a mužů ověříme později F-testem, protože je součástí výpočtu t-testu.

Postup výpočtu nepárového t-testu (po ověření předpokladu normality věku žen a věku mužů)

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu $H_0: \mu_1 = \mu_2$, $H_A: \mu_1 \neq \mu_2$.
2. Pro obě skupiny vypočítáme aritmetický průměr a rozptyl výběrového souboru a určíme počet pozorování.
3. Vypočítáme **testovou statistiku t** a odpovídající **p -hodnotu**. V menu *Statistics* zvolíme *Basic statistics*, vybereme *t-test, independent, by groups*. Vybereme proměnnou, kterou chceme testovat (*dependent*) a proměnnou obsahující skupiny, které srovnáváme (*grouping*). V záložce *Options* zaškrtneme možnost *Test w/separate variance estimates* (umožňuje získat validní výsledek i při nesplnění předpokladu homogenity rozptylů). Kliknutím na *Summary* získáme výstupy.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \dots = -3,42 \quad p = 0,001$$

4. Vypočítané t porovnáme s kritickou hodnotou nebo porovnáme p -hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p -hodnota $\leq \alpha$, zamítáme H_0** . Věk mužů a žen při mozkovém infarktu se liší. U žen se vyskytuje později.

Výsledky v programu Statistica

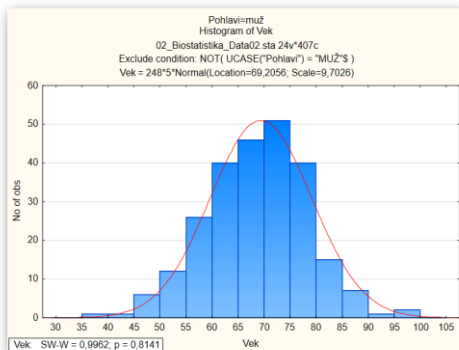
Srovnání průměru a mediánu věku mužů

Pohlavi=muž			
Descriptive Statistics (02_Biostatistik			
Variable	Valid N	Mean	Median
Vek	248	69,20565	69,00000

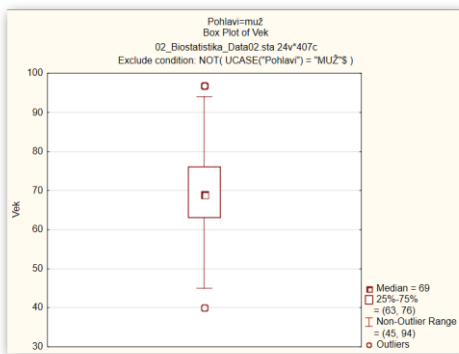
Průměr a medián jsou téměř shodné (cca 69 let) a data jsou tedy nejspíš alespoň symetrická.

Ověření normality rozložení věku mužů

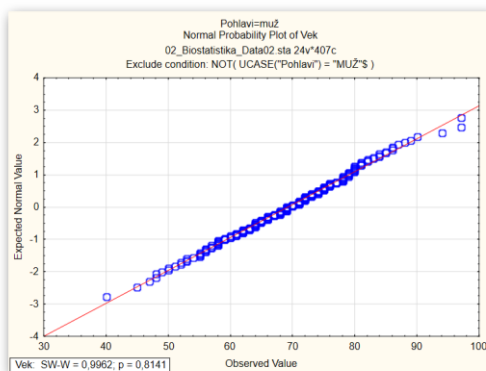
Histogram věku mužů



Krabicový graf věku mužů



N-P graf věku mužů



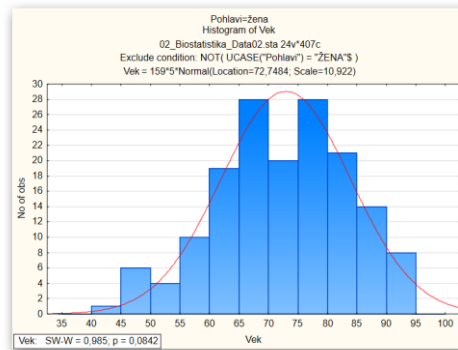
Srovnání průměru a mediánu věku žen

Pohlavi=žena			
Descriptive Statistics (02_Biostatistik			
Variable	Valid N	Mean	Median
Vek	159	72,74843	72,00000

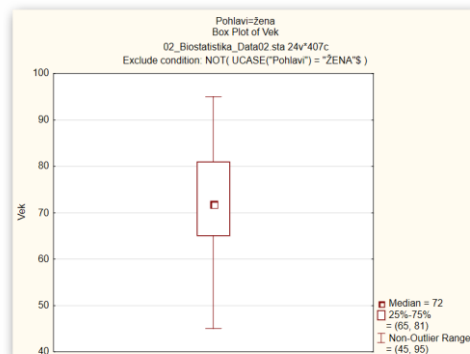
Průměr a medián jsou podobné (cca 72 až 73 let) a data jsou tedy nejspíš alespoň symetrická.

Ověření normality rozložení věku žen

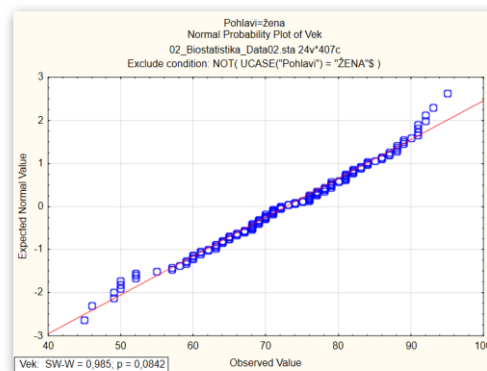
Histogram věku žen



Krabicový graf věku žen



N-P graf věku žen



Symetrie je patrná i z krabicového grafu. Navíc histogram naprosto jasně odpovídá průběhu normálního rozdělení. Z N-P grafu také nejsou patrné odchylky od normality. Na základě p-hodnoty 0,814 nezamítáme nulovou hypotézu o normalitě (tj. nezamítáme, že není rozdíl mezi pozorovanými daty a teoretickým normálním rozdělením, tj. data jsou normálně rozdělená).

Spíše symetrie je patrná i z krabicového grafu. Navíc histogram přibližně odpovídá průběhu normálního rozdělení. Z N-P grafu nejsou patrné výrazné odchylky od normality. Na základě p-hodnoty 0,084 nezamítáme nulovou hypotézu o normalitě (tj. nezamítáme, že není rozdíl mezi pozorovanými daty a teoretickým normálním rozdělením, tj. data jsou normálně rozdělená).

	Výběrové průměry obou skupin				Rozsahy výběru obou skupin				Výběrové směrodatné odchylky obou skupin					
Variable	Mean muž	Mean žena	t-value	df	p	t separ. var. est.	df	p 2-sided	Valid N muž	Valid N žena	Std.Dev. muž	Std.Dev. žena	F-ratio Variances	p Variances
Vek	69,20565	72,74843	-3,42023	403	0,000689	-3,33299	307,908	0,000964	248	159	9,702563	10,92203	1,26716	0,096008

p-hodnota t-testu (při stejných rozptylech)
p-hodnota t-testu (při různých rozptylech)
p-hodnota F-testu pro ověření předpokladu shody rozptylů
 - Pokud je $p \leq 0,05$, pak jsou rozptyly různé.
 - Pokud je $p > 0,05$, pak jsou rozptyly stejné.

Pozorovaný průměrný věk mužů je 69,2 let a u žen 72,7 let. V našich datech jsou tedy **ženy starší o 3,5 roku**. P-hodnota statistické významnosti F-testu je 0,096, což znamená, že na hladině významnosti 0,05 nezamítáme nulovou hypotézu o shodě rozptylů mužů a žen (tj. **rozptyly jsou v obou skupinách stejné**). Na základě p-hodnoty t-testu při stejných rozptylech $p = 0,001$ vyhodnotíme pozorovaný rozdíl 3,5 let jakožto statisticky **významný** výsledek a lze tedy prohlásit, že průměrný věk se u mužů a žen liší (tj. **ženy skutečně postihuje mozkový infarkt později**).

7.4. Analýza rozptylu (ANOVA)

Jak jsme si ukázali v předešlých podkapitolách, ke srovnání dvou nezávislých výběrů používáme t-test. Pokud chceme porovnat více výběrů, nelze t-test použít (tj. nemůžeme t-testem porovnat všechny možné páry výběrů), protože jednotlivé testy nejsou nezávislé a jejich opakovaným použitím k testování jedné hypotézy bychom zvyšovali chybu prvního druhu.

Chceme-li porovnat tři a více nezávislých výběrů používáme analýzu rozptylu (ANOVA). Příkladem může být srovnání krevního tlaku u tří skupin pacientů léčených třemi různými léky, srovnání kognitivních schopností u čtyř skupin pacientů apod. ANOVA testuje hypotézu, že se střední hodnoty všech výběrů mezi sebou neliší.

Analýza rozptylu má několik předpokladů:

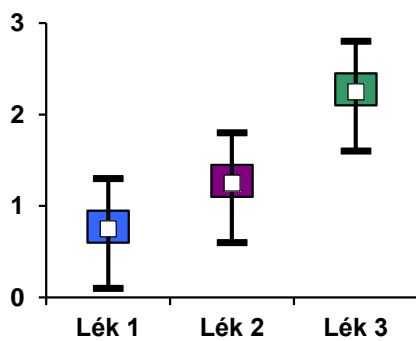
- nezávislost srovnávaných skupin,

- normální rozdělení proměnné ve všech skupinách (malé odchylky od normality jsou přípustné, ANOVA je dostatečně robustní proti drobným odchylkám od tohoto předpokladu),
- shodný rozptyl (homogenita rozptylů) ve všech skupinách (test homogenity rozptylů: Levenův test nebo Bartlettův test).

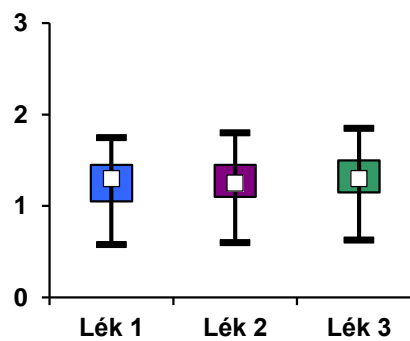
Protože předpokladem analýzy rozptylu je shoda rozptylů, můžeme si představit, že za platnosti nulové hypotézy se jedná o několik výběrů z téhož souboru. Principem ANOVA je srovnání pozorované variability mezi výběry a pozorované variability uvnitř výběrů. Na základě rozptylů uvnitř jednotlivých skupin je odhadnut rozptyl základního souboru. Podle něj je možné předpovědět, jaká je variabilita mezi skupinami a tuto předpokládanou variabilitu porovnat se skutečnou pozorovanou variabilitou mezi skupinami. Pokud je variabilita mezi skupinami nepravděpodobně velká (ANOVA pracuje s F-testem), potom zamítáme nulovou hypotézu o rovnosti průměrů (Obrázek 8.1).

Když označíme počet srovnávaných výběrů jako k , nulová hypotéza analýzy rozptylu zní:
 $H_0: \mu_1 = \mu_2 = \dots = \mu_k$.

a: rozdíl mezi všemi třemi skupinami



b: žádný rozdíl mezi skupinami



Obrázek 8.1. Ukázka tří skupin pacientů léčených třemi různými léky ovlivňujícími hodnoty měřeného parametru (a) a neovlivňující hodnoty měřeného parametru (b).

Postup výpočtu ANOVA je následovný (poté, co ověříme předpoklady normality rozdělení v každém ze sledovaných výběrů Shapirovým-Wilkovým testem a shody rozptylů Levenovým nebo Bartlettovým testem).

1. Stanovíme nulovou a alternativní hypotézu. $H_0: \mu_1 = \mu_2 = \dots = \mu_k$,
 H_A : nejméně jedno μ_i je odlišné od ostatních.
2. Vypočítáme variabilitu v rámci jednotlivých skupin (S_e) a variabilitu mezi skupinami (S_A).
3. Vypočítáme testovou statistiku F a odpovídající p-hodnotu.

$$F = \frac{S_A}{k-1} \cdot \frac{n-k}{S_e}$$
4. Vypočítané F porovnáme s kritickou hodnotou nebo porovnáme p-hodnotu s hladinou významnosti α .
5. Když je $F >$ kritická hodnota nebo p-hodnota $\leq \alpha$, zamítáme nulovou hypotézu.

Případným zamítnutím nulové hypotézy ovšem zjistíme jen to, že se alespoň jedna střední hodnota liší, tj. že všechny nejsou stejné. Které se ovšem liší, z výsledku ANOVA nezjistíme. K tomu musíme použít další metody, označované jako mnohonásobná porovnání nebo post-hoc testy.

Příklad: Analýza rozptylu

Zadání: Porovnejte věk pacientů s mozkovým infarktem dle terapie, která jim byla indikována (mechanická trombektomie, intravenózní trombolýza rt-PA nebo jiná farmakologická léčba), a zjistěte, zda se jedná o statisticky významný rozdíl.

Postup v programu Statistica

1. Ověříme předpoklady testu: normalita rozložení věku ve všech skupinách (vizuálně i Shapiro-Wilkovým testem). Postupujeme všemi kroky stejně jako v příkladu 6.1, tj. nastavíme provedení Shapiro-Wilkova testu u histogramu i diagnostického N-P grafu. Shodu rozptylů ověříme později při výpočtu ANOVA, jejíž je test homogenity součástí.

Postup výpočtu ANOVA (po ověření předpokladu normality věku všech skupin)

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu $H_0: \mu_1 = \mu_2 = \mu_3$, H_A : alespoň jedno μ_i se liší od ostatních.
2. Vypočítáme variabilitu v rámci jednotlivých skupin (S_e) a variabilitu mezi skupinami (S_A).
3. Vypočítáme **testovou statistiku F** a odpovídající **p-hodnotu**. V menu *Statistics* zvolíme *Basic Statistics*, vybereme *Breakdown & one-way ANOVA*. Vybereme proměnnou, kterou chceme testovat (*dependent*) a proměnnou obsahující skupiny (*grouping*) – *OK*. Na záložce *ANOVA & tests* zvolíme *Levene tests*, na záložce *ANOVA & tests* pak *Analysis of Variance*.

$$F = \frac{S_A}{k-1} \cdot \frac{n-k}{S_e} = \dots = 6,41, \quad p = 0,002$$

4. Vypočítané F porovnáme s kritickou hodnotou nebo porovnáme p-hodnotu s hladinou významnosti α .
5. Je-li **p-hodnota $\leq \alpha$** , **zamítáme H_0** . Existuje alespoň jedna dvojice terapie mozkového infarktu, která se liší v průměrném věku pacientů.
6. Provedeme mnohonásobné porovnání. Na záložce *Post-hoc* zvolíme *Tukey HSD*. Získáme tak výsledky mnohonásobného porovnání mezi všemi skupinami.

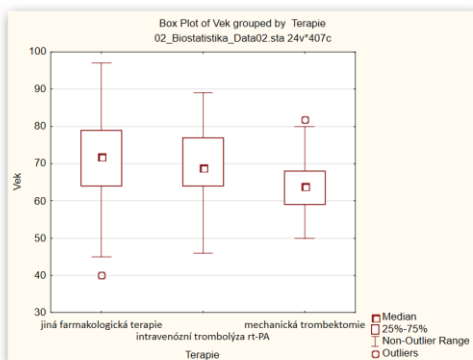
Výsledky v programu Statistica

Srovnání průměru a mediánu

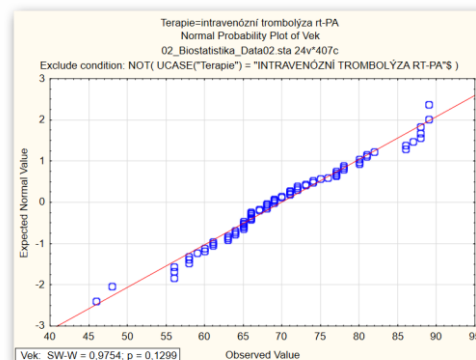
Aggregate Results				
Descriptive Statistics (02_Biostatistika_Data02.sta)				
Variable	Terapie	Valid N	Mean	Median
Vek	jiná farmakologická terapie	299	71,36455	72,00000
Vek	intravenózní trombolýza rt-PA	79	69,93671	69,00000
Vek	mechanická trombektomie	29	64,37931	64,00000

Ověření normality věku v skupinách

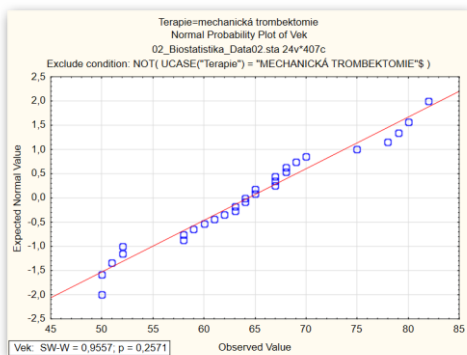
Krabicový graf – všechny skupiny



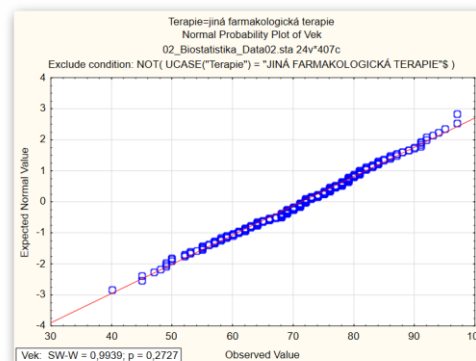
N-P graf, pacienti s intravenózní trombolýzou



N-P graf, pacienti s mechanickou trombektomií



N-P graf, pacienti s jinou terapií



Základní popis i grafické srovnání ukazuje možný rozdíl mezi skupinami, a to především u pacientů s mechanickou trombektomií oproti ostatním pacientům (průměrný věk při mechanické trombektomii je 64 let, při rt-PA trombolýze 70 let a u jiné léčby je průměr 71 let). Normalitu dat nezamítáme u žádné skupiny ($p = 0,273$, $p = 0,130$ a $p = 0,257$) s tím, že ani u jedné skupiny není z N-P grafu patrné výrazné porušení normality.

Levene Test of Homogeneity of Variances (02_Biostatistika_Data02.sta)							
Marked effects are significant at p < .05000							
Variable	SS	df	MS	SS	df	MS	F
	Effect	Effect	Effect	Error	Error	Error	
Vek	94,39532	2	47,19766	15579,57	404	38,56329	1,223901
							0,295169

p-hodnota Levenova testu pro ověření předpokladu shody rozptylů

- Pokud je $p \leq 0,05$, pak jsou rozptyly různé.
- Pokud je $p > 0,05$, pak jsou rozptyly stejné.

P-hodnota statistické významnosti Levenova testu je 0,295, což znamená, že na hladině významnosti 0,05 nezamítáme nulovou hypotézu o shodě rozptylů mezi skupinami (tj. rozptyly jsou ve všech skupinách stejné).

Analysis of Variance (02_Biostatistika_Data02.sta)							
Marked effects are significant at p < .05000							
Variable	SS	df	MS	SS	df	MS	F
	Effect	Effect	Effect	Error	Error	Error	
Vek	1331,701	2	665,8507	41984,78	404	103,9227	6,407172
							0,001822

p-hodnota ANOVA

Na základě p-hodnoty ANOVA $p = 0,002$ vyhodnotíme pozorovaný rozdíl mezi průměry 64 let, 70 let, a 71 let jakožto statisticky významný výsledek a lze tedy prohlásit, že **existuje alespoň jedna dvojice terapie mozkového infarktu, která se liší v průměrném věku pacientů.**

Tukey HSD test; Variable: Vek (02_Biostatistika_Data02.sta)			
Marked differences are significant at p < .05000			
Terapie	{1}	{2}	{3}
	M=71,365	M=69,937	M=64,379
jiná farmakologická terapie {1}		0,509593	0,001251
intravenózní trombolýza rt-PA {2}	0,509593		0,032293
mechanická trombektomie {3}	0,001251	0,032293	

p-hodnoty mnohonásobného porovnání všech skupin

Mnohonásobným porovnáním jsme navíc prokázali významný rozdíl mezi trombektomií a rt-PA trombolýzou a mezi trombektomií a jinou terapií. Jinými slovy, **pacienti podstupující mechanickou trombektomii jsou významně mladší než pacienti podstupující ostatní dvě terapie.**

8. Neparametrické testy pro kvantitativní proměnné

Neparametrické testy vyžadují splnění méně předpokladů o rozložení vstupních dat. Lze je tedy použít i při asymetrickém rozložení, přítomnosti odlehlých hodnot, či nedetekovatelném rozložení. Neparametrické testy nevyužívají původní hodnoty, ale nejčastěji pouze jejich pořadí. Důsledkem redukce informační hodnoty původních dat je snížená síly těchto testů. Neparametrické testy se používají také při hodnocení souborů s nízkým počtem pozorování, kdy nejsme schopni normalitu dat spolehlivě ověřit.

8.1. Jednovýběrový Wilcoxonův a znaménkový test

Neparametrickou alternativou t-testu pro jeden výběr je Wilcoxonův test, který není testem o střední hodnotě, ale testem o mediánu. Jeho jediným předpokladem je symetrie rozdělení dat kolem mediánu. V situaci, kdy není splněn předpoklad symetrie rozdělení kolem mediánu, je možnou alternativou znaménkový test. Oba testy testují nulovou hypotézu o rovnosti mediánu jednoho výběru s referenční hodnotou.

Postup výpočtu Wilcoxonova testu je následovný.

1. Stanovíme nulovou a alternativní hypotézu. $H_0: x_{0,5} = c$, $H_A: x_{0,5} \neq c$.
2. Určíme rozdíly hodnot výběru s testovanou hodnotou mediánu.
3. Absolutní hodnoty rozdílů uspořádáme vzestupně a přiřadíme jim pořadí. Pro rozdíly stejné absolutní velikosti použijeme průměrné pořadí.
4. Spočítáme statistiky S_w^+ a S_w^- , které odpovídají součtu pořadí kladných (S_w^+) a záporných rozdílů (S_w^-). Jako finální hodnotu testové statistiky bereme minimum z S_w^+ a S_w^- . Nulovou hypotézu zamítáme, pokud je hodnota testové statistiky menší nebo rovna tabelované kritické hodnotě (při dané hladině významnosti a počtu nenulových rozdílů), nebo když příslušná p-hodnota \leq zvolená hladina významnosti.

Nebo: Pro $N > 30$ lze využít asymptotické normality statistiky S_w^+ .

V případě, že pozorované hodnoty jsou symetricky rozděleny kolem předpokládané hodnoty c , bude přibližně jedna polovina rozdílů kladná a druhá záporná. Navíc součet pořadí kladných rozdílů a součet pořadí záporných rozdílů bude přibližně stejný. Za platnosti H_0 tak lze předpokládat, že hodnoty S_w^+ a S_w^- budou zhruba vyrovnané. Na druhou stranu, kdy H_0 nebude platit, bude mezi hodnotami S_w^+ a S_w^- rozdíl.

Postup výpočtu znaménkového testu je následovný.

1. Stanovíme nulovou a alternativní hypotézu. $H_0: x_{0,5} = c$, $H_A: x_{0,5} \neq c$.
2. Spočítáme rozdíly hodnot výběru s testovanou hodnotou mediánu.
3. Spočítáme statistiku S_z^+ , která odpovídá počtu kladných rozdílů (test nevyužívá hodnot pořadí původních dat, ale pouze informaci, zda se hodnota realizuje nad nebo pod mediánem a tím dochází ke snížení síly testu).

4. Nulovou hypotézu zamítáme, pokud statistika S_z^+ realizuje v kritickém oboru hodnot $W = (0, k_1) \cup (k_2, n)$, kde n odpovídá počtu nenulových rozdílů a hodnoty k_1 a k_2 lze dohledat v matematických tabulkách; nebo když příslušná p-hodnota \leq zvolená hladina významnosti.

Nebo: Pro $N > 20$ lze využít asymptotické normality statistiky S_z^+ .

V případě, že pozorované hodnoty jsou symetricky rozděleny kolem předpokládané hodnoty c , bude přibližně jedna polovina rozdílů kladná a druhá záporná. Za platnosti H_0 tak lze předpokládat, že hodnota S_z^+ je přibližně polovinou z hodnoty počtu nenulových rozdílů.

Výpočet jednovýběrového Wilcoxonova i znaménkového testu uvádíme na příkladu pacientů čekajících u lékaře. U 15 pacientů byla vyhodnocena doba, kterou museli strávit v čekárně u lékaře. Zjistíme, zda medián čekací doby je roven půl hodině (Tabulka 9.1). Nulová hypotéza, kterou testujeme, je: $H_0: x_{0,5} = 30$.

Tabulka 8.1. Ukázka výpočtu jednovýběrového Wilcoxonova a znaménkového testu na příkladu 15 pacientů čekajících u lékaře.

ID	Doba čekání	Medián	Rozdíl	Rozdíl	Pořadí	Větší než medián?
1	1	30	-29	29	15	Ne
2	45	30	15	15	10	Ano
3	25	30	-5	5	3,5	Ne
4	15	30	-15	15	10	Ne
5	34	30	4	4	2	Ano
6	19	30	-11	11	8	Ne
7	31	30	1	1	1	Ano
8	25	30	-5	5	3,5	Ne
9	8	30	-22	22	14	Ne
10	12	30	-18	18	12	Ne
11	20	30	-10	10	6	Ne
12	15	30	-15	15	10	Ne
13	40	30	10	10	6	Ano
14	20	30	-10	10	6	Ne
15	10	30	-20	20	13	Ne

U Wilcoxonova testu spočítáme součet pořadí kladných rozdílů a součet pořadí záporných rozdílů. $S_w^+ = 19$, $S_w^- = 101$. Minimum z těchto dvou hodnot je 19, kritická hodnota pro 15 nenulových rozdílů je $w_{15(0,05)} = 25$. Porovnáním hodnoty testové statistiky a kritické hodnoty zjistíme, že testová statistika je menší než kritická hodnota, tudíž zamítáme nulovou hypotézu H_0 .

U znaménkového testu spočítáme počet kladných rozdílů: $S_z^+ = 4$. Kritický obor pro 15 nenulových rozdílů je: $W = (0,3) \cup (12,15)$. Jelikož se hodnota testové statistiky realizuje mimo kritický obor hodnot, nezamítáme nulovou hypotézu H_0 .

Vidíme, že výsledek Wilcoxonova testu a výsledek znaménkového testu se liší. Důvodem je nízký počet pozorování a menší síla znaménkového testu (tj. nižší schopnost rozpoznat neplatnou nulovou hypotézu).

Příklad: Jednovýběrový Wilcoxonův test

Zadání: V podobné zahraniční studii byla publikovaná střední hodnota indexu Barthelové na konci akutní rehabilitace po mozkovém infarktu ve výši 64,4. Zjistěte, zda výsledné dosažení stupně soběstačnosti dle BI ve vašich datech je stejné nebo jiné než v této studii.

Postup v programu Statistica

1. Ověříme normalitu rozložení hodnot indexu Barthelové na konci rehabilitace (vizuálně i Shapiro-Wilkův test). Postupujeme všemi kroky stejně jako v příkladu 6.1, tj. nastavíme provedení Shapiro-Wilkova testu u histogramu i diagnostického N-P grafu.

Postup výpočtu jednovýběrového Wilcoxonova testu (po nemožnosti použití jednovýběrový t-test)

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu $H_0: x_{0,5} = 64,4$ proti $H_A: x_{0,5} \neq 64,4$
2. Původní hodnoty Barthelové indexu převedeme na pořadí (určené podle absolutní hodnoty rozdílu oproti referenci). Do datové tabulky je potřeba přidat sloupec obsahující konstantní hodnotu reference, se kterou porovnáme naše výsledky.
3. Vypočítáme testovou statistiku S_w nebo Z a odpovídající **p-hodnotu**. V menu **Statistics** zvolíme **Nonparametrics**, vybereme **Comparing two dependent samples (groups)**. Vybereme proměnné (**Variables**), které chceme testovat (testovaný parametr a reference). Kliknutím na **Wilcoxon matched pair test** získáme výsledky. $S_w = 41\ 099$, $Z = 0,17$, $p = 0,861$.
4. Vypočítané statistiky porovnáme s kritickou hodnotou, nebo porovnáme p-hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p-hodnota** $> \alpha$, **nezamítáme H_0** . Výsledná soběstačnost pacientů v našem souboru se neliší od výsledků publikovaných v porovnávané studii.

Výsledky v programu Statistica

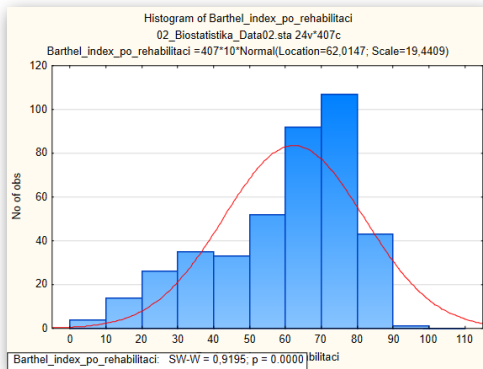
Srovnání průměru a mediánu

Variable	Descriptive Statistics (02_Biostatistik)		
	Valid N	Mean	Median
Barthel_index_po_rehabilitaci	407	62,01474	70,00000

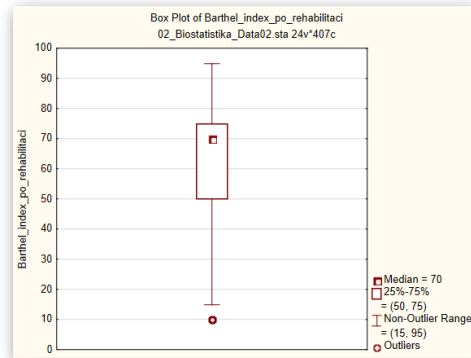
Průměr a medián se výrazně liší (průměr 62 bodů, medián 70 bodů. Data jsou nejspíše asymetrická.

Ověření normality Barthelové indexu

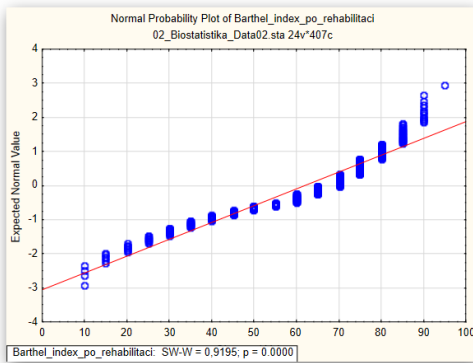
Histogram



Krabicový graf



N-P graf



Asymetrie je patrná i z krabicového grafu a histogramu. Z histogramu je navíc zřetelně vidět odlišnost od normálního rozdělení. Odchytky od normality jsou patrné i z N-P grafu. Na základě p-hodnoty $< 0,001$ zamítáme nulovou hypotézu o normalitě (tj. data nejsou normálně rozdělená).

	Rozsah výběru	Hodnota testové statistiky S_w a Z		
Pair of Variables	Valid N	T	Z	p-value
Barthel_index_po_rehabilitaci & Barthel_index_reference	407	41099,00	0,174762	0,861267

**p-hodnota
Wilcoxonova testu**

Pozorovaný výsledný medián Barthelové indexu je 70 bodů, což je oproti výsledku 64,4 bodů v porovnávané studii **lepší výsledný stav o 5,6 bodů**. P-hodnota statistické významnosti tohoto pozorovaného rozdílu je ale $p = 0,861$, což na hladině významnosti 0,05 značí **nevýznamný rozdíl**, a z dostupných dat tedy nelze prokázat, že by výsledná soběstačnost pacientů léčených s mozkovým infarktem v našem souboru byla odlišná od výsledků publikovaných v porovnávané studii.

8.2. Párový Wilcoxonův a znaménkový test

Wilcoxonův test i znaménkový test jsou neparametrickou alternativou párového t-testu. Wilcoxonův test pracuje s rozdíly párových hodnot, pak jejich absolutní hodnoty provádí na pořadí. Znamená to, že při užití testu předpokládáme, že můžeme původní hodnoty odečítat. Proto jej nemůžeme použít na ordinální data, mezi jejichž kategoriemi nejsou konstantní rozdíly. Dále Wilcoxonův test předpokládá symetrické rozdělení rozdílů kolem mediánu. Znaménkový test pracuje s rozdíly párových hodnot a uvažuje pouze počet kladných a počet záporných rozdílů. Nemá žádné předpoklady o jejich rozdělení. Dále přecházíme na design jednovýběrových testů.

Nulová a alternativní hypotéza se týká mediánu rozdílů (diferencí – D) párových hodnot:

$$H_0: D_{0,5} = 0, H_A: D_{0,5} \neq 0.$$

Dále postupujeme stejně jako u jednovýběrových testů výpočtem testové statistiky S_w^+ a S_w^- (u Wilcoxonova testu), resp. S_z^+ (u znaménkového testu) a jejich porovnáním s kritickou hodnotou, resp. s kritickým intervalem (nebo pro větší vzorky použijeme aproximaci normálním rozdělením).

Výpočet párového Wilcoxonova testu uvádíme na příkladu změny krevního parametru po podání léku u 10 pacientů (Tabulka 8.2). Testujeme nulovou hypotézu $H_0: D_{0,5} = 0$.

U Wilcoxonova testu spočítáme součet pořadí kladných rozdílů a součet pořadí záporných rozdílů: $S_w^+ = 51$, $S_w^- = 4$. Minimum z těchto dvou hodnot je 4, kritická hodnota pro 10 nenulových rozdílů je $w_{10(0,05)} = 8$. Porovnáním hodnoty testové statistiky a kritické hodnoty zjišťujeme, že testová statistika je menší než kritická hodnota, tudíž zamítáme nulovou hypotézu H_0 .

Tabulka 8.2. Ukázka výpočtu párového Wilcoxonova testu na příkladu 10 pacientů a hodnotami jejich krevního parametru před a po podání léku.

ID	Před	Po	Rozdíl	Rozdíl	Pořadí
1	142	138	4	4	4,5
2	140	136	4	4	4,5
3	144	147	-3	3	3
4	144	139	5	5	7
5	142	143	-1	1	1
6	146	141	5	5	7
7	149	143	6	6	9,5
8	150	145	5	5	7
9	142	136	6	6	9,5
10	148	146	2	2	2

Příklad: Párový Wilcoxonův test

Zadání: Pacientům hospitalizovaným s mozkovým infarktem byla na lůžku akutní péče poskytnuta terapie pro obnovu krevního oběhu v postižené části mozku. Po zvládnutí akutní fáze byl u pacientů vyhodnocen stupeň soběstačnosti v základních denních aktivitách (ADL) pomocí indexu Barthelové (BI) a byli přeloženi na rehabilitační oddělení. Po dvou týdnech byl opět vyhodnocen stupeň soběstačnosti dle BI. Zjistěte, zda poskytnutá rehabilitační péče vedla ke zlepšení soběstačnosti ADL.

Postup v programu Statistica

1. Ověříme normalitu rozložení hodnot změn indexu Barthelové.

Postup výpočtu párového Wilcoxonova testu (po nemožnosti použít párový t-test):

2. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu o diferencích párových hodnot.
 $H_0: D_{0,5} = 0$, $H_A: D_{0,5} \neq 0$.
2. Původní hodnoty vypočítaných diferencí obou měření převedeme na pořadí (určené podle jejich absolutní hodnoty).
3. Vypočítáme **testovou statistiku S_w nebo Z** a odpovídající **p-hodnotu**. V menu *Statistics* zvolíme *Nonparametrics*, vybereme *Comparing two dependent samples (groups)*. Vybereme proměnné (*Variables*), které chceme testovat. Kliknutím na *Wilcoxon matched pair test* získáme výsledky. $S_w = 198,5$, $Z = 17,29$, $p < 0,001$.
4. Vypočítané statistiky porovnáme s kritickou hodnotou, nebo porovnáme p-hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p-hodnota $\leq \alpha$** , **zamítáme H_0** . Během rehabilitace se podařilo změnit soběstačnost pacientů v denních aktivitách. Ke stejnému závěru jsme došli při použití parametrického t-testu.

Výsledky v programu Statistica

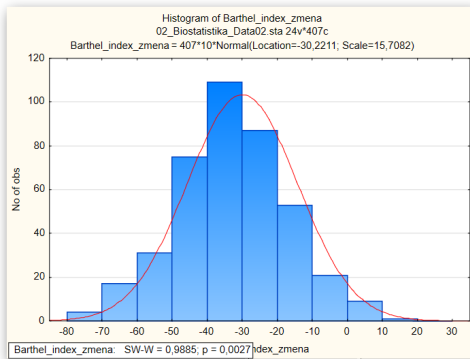
Srovnání průměru a mediánu

Variable	Descriptive Statistics (02_Biostatistik		
	Valid N	Mean	Median
Barthel_index_zmena	407	-30,2211	-30,0000

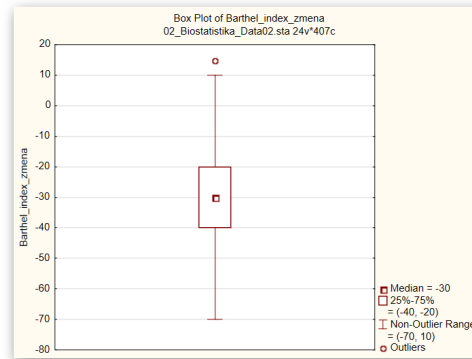
Průměr a medián jsou v podstatě shodné (cca -30) a data jsou tedy nejspíš alespoň symetrická.

Ověření normality změna Barthelové indexu

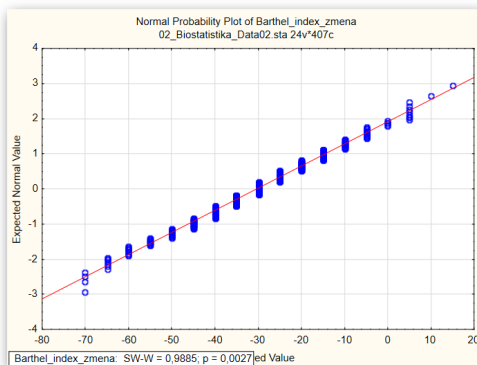
Histogram



Krabicový graf



N-P graf



Symetrie je patrná i z krabicového grafu. Navíc histogram je svým průběhem velmi podobný normálnímu rozdělení. Z N-P grafu také nejsou patrné odchylky od normality. Na základě p-hodnoty 0,003 zamítáme nulovou hypotézu o normalitě (tj. data formálně dle testu nejsou normálně rozdělená).

Pair of Variables	Valid N	T	Z	p-value
Barthel_index_pred_rehabilitaci & Barthel_index_po_rehabilitaci	402	198,5000	17,28938	0,0000

Wilcoxon Matched Pairs Test (02_Biostatistika)
 Marked tests are significant at p < .05000

p-hodnota Wilcoxonova testu

Pozorovaný medián **zlepšení Barthelové indexu** na začátku a po rehabilitaci je **30 bodů**. P-hodnota statistické významnosti této pozorované změny je $p < 0,001$, což na hladině významnosti 0,05 značí **významný rozdíl**, a lze tedy prohlásit, že **stupeň soběstačnosti v základních denních aktivitách se viditelně během péče zlepšil**.

8.3. Mannův-Whitneyův test

Mannův-Whitneyův test (označována také Mannův-Whitneyův U test, nebo zkráceně U test) je neparametrickou alternativou dvouvýběrového nepárového t-testu. Stejně jako řada jiných neparametrických testů počítá i tento test s pořadím dat ve výběrech místo s původními daty. To znamená, že původní hodnoty jsou nahrazeny jejich pořadím. Pořadí hodnot je stanoveno v obou výběrech bez ohledu na příslušnost hodnoty ke konkrétnímu výběru. V případě stejné

hodnoty dvou čísel je jako pořadí použito průměrné pořadí, které by tato čísla dostala v případě odlišné hodnoty (např. dvě stejná čísla, která by v případě nerovnosti byla na pozici 7 a 8, získají pořadí 7,5). Filozofie testu spočívá v tom, že pokud budou původní výběry podobné, potom bude součet pořadí obou výběrů podobný. Test lze použít i pro ordinální data.

Mannův-Whitneyův test testuje nulovou hypotézu, že oba výběry pocházejí ze stejného rozdělení, tedy už nikoliv hypotézu, že střední hodnoty jsou stejné. Pokud chceme interpretovat výsledek testu jako test o poloze (tedy o mediánu nebo průměru), musíme předpokládat, že tvar rozdělení je v obou skupinách stejný.

Postup výpočtu Mannova-Whitneyova testu je následovný.

1. Stanovíme nulovou a alternativní hypotézu o distribuční funkci $F(x)$.
 $H_0: F(x_1) = F(x_2)$, $H_A: F(x_1) \neq F(x_2)$.
2. Hodnoty obou výběrů (skupin) jsou sloučena a je určeno jejich pořadí bez ohledu na příslušnost ke skupině.
3. Pro oba výběry zvlášť je spočítán součet pořadí (T_1 a T_2).
4. Ze součtů pořadí ve skupinách je určena finální hodnota testové statistiky U a příslušnou p-hodnotu.

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - T_1, \quad U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - T_2, \quad U = \min(U_1, U_2).$$

5. Hodnotu testové statistiky U porovnáme s kritickou hodnotou testu, nebo p-hodnotu s hladinou významnosti α . Pokud je hodnota testové statistiky menší než kritická hodnota testu, nebo p-hodnota menší než hladina významnosti α , zamítáme nulovou hypotézu shody distribučních funkcí obou skupin.

Pro velká n_1 a n_2 (> 30) lze využít asymptotické normality statistiky U .

Výpočet Mannova-Whitneyova testu ukážeme na příkladu dvou skupin štěňat trénovaných k hygienickým návykům pomocí pozitivní (8 štěňat) nebo negativní motivace (9 štěňat) (Tabulka 9.3). Testujeme nulovou hypotézu $H_0: F(x_1) = F(x_2)$.

U Mannova-Whitneyova testu spočítáme součet pořadí v každé skupině ($T_1 = 49,5$ a $T_2 = 103,5$) a hodnotu testové statistiky ($U_1 = 58,5$ a $U_2 = 13,5$; $\min(U_1, U_2) = 13,5$). Hodnotu testové statistiky porovnáme s kritickou hodnotou U pro daný počet pozorování ve skupinách $U_{(8,9;0,05)} = 15$. Hodnota testové statistiky je menší než kritická hodnota, tudíž zamítáme H_0 ; délka výcviku se u pozitivně motivovaných a negativně motivovaných štěňat liší.

Tabulka 8.3. Ukázka výpočtu Mannova-Whitneyova U testu na příkladu 17 štěňat. Délka výcviku – počet dní, za které štěně úspěšně zvládlo výcvik k hygienickým návykům; skupina – metoda výcviku.

ID	Délka výcviku	Skupina	Pořadí
1	35	pozitivní	1
2	41	pozitivní	2
3	43	pozitivní	4
4	44	pozitivní	5
5	47	pozitivní	7,5
6	48	pozitivní	9,5
7	48	pozitivní	9,5
8	51	pozitivní	11
9	42	negativní	3
10	46	negativní	6
11	47	negativní	7,5
12	53	negativní	12
13	54	negativní	13
14	57	negativní	14
15	59	negativní	15
16	65	negativní	16
17	74	negativní	17

Příklad: Mannův-Whitneyův U test

Zadání: U pacientů hospitalizovaných pro mozkový infarkt by po úspěšné terapii a absolvování akutní rehabilitace měl následovat přesun do ambulantní péče nebo na následné lůžko k pokračování v další rehabilitaci. Při správném managementu péče by do následné lůžkové péče měli pokračovat pouze pacienti, u kterých dosud nebylo dosaženo dostatečné rekonvalescence. Zkontrolujte, zda pacienti překládaní na následné lůžko mají skutečně horší míru soběstačnosti v základních denních aktivitách (ADL) vyjádřenou indexem Barthelové určenou v době propouštění.

Postup v programu Statistica

1. Ověříme normalitu rozložení hodnot indexu Barthelové po rehabilitaci v obou skupinách.

Postup výpočtu Mannova-Whitneyova U testu (po nemožnosti použít dvouvýběrový t-test):

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu $H_0: F(x_1) = F(x_2)$ proti $H_A: F(x_1) \neq F(x_2)$.
2. Původní hodnoty Barthelové indexu převedeme na pořadí v celém souboru.

3. Vypočítáme **testovou statistiku U nebo Z** a odpovídající **p -hodnotu**. V menu **Statistics** zvolíme **Nonparametrics**, vybereme **Comparing two independent samples (groups)**. Vybereme proměnnou, kterou chceme testovat (**dependent**) a proměnnou obsahující skupiny, které srovnáváme (**grouping**). Kliknutím na **Mann-Whitney U test**, nebo na M-W U test získáme výstupy.

$U = 1\,998$, $Z = 14$, $p < 0,001$.

4. Vypočítané statistiky porovnáme s kritickou hodnotou, nebo porovnáme p -hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p -hodnota $\leq \alpha$, zamítáme H_0** . Aktuální soběstačnost pacientů je určující pro jejich další pokračování v systému zdravotní péče.

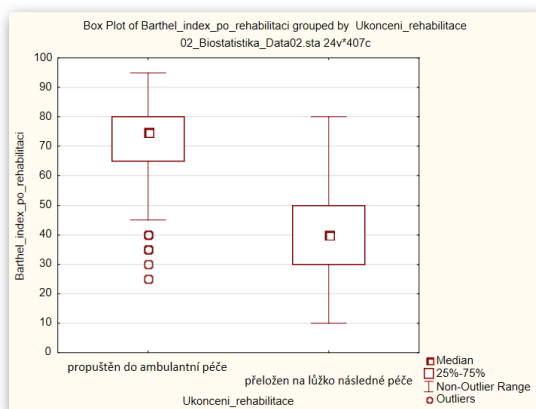
Výsledky v programu Statistica

Srovnání průměru a mediánu

Aggregate Results				
Descriptive Statistics (02_Biostatistika_Data02.sta)				
Variable	Ukonceni_rehabilitace	Valid N	Mean	Median
Barthel_index_po_rehabilitaci	propuštěn do ambulantní péče	290	71,34483	75,00000
Barthel_index_po_rehabilitaci	přeložen na lůžko následné péče	117	38,88889	40,00000

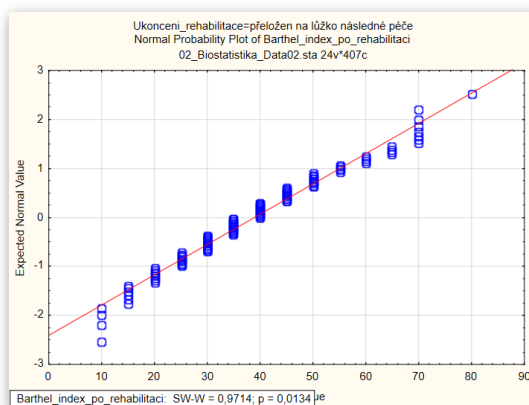
Ověření normality změna Barthelové indexu

Krabicový graf

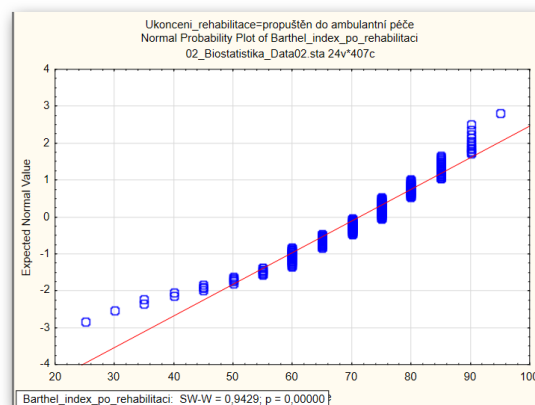


Základní popis i grafické srovnání ukazuje výrazný rozdíl mezi skupinami (soběstačnost při propuštění do ambulantní péče je v mediánu 75 bodů, ale pacienti pokračující do následné péče mají medián pouze 40 bodů).

N-P graf, pacienti přeloženi



N-P graf, pacienti propuštěni



Normalitu dat zamítáme u obou skupin ($p = 0,013$ a $p < 0,001$) a přinejmenším u pacientů propuštěných domů je výrazné porušení normality patrné graficky i z N-P grafu.

Hodnota testové statistiky U a Z **Rozsahy výběru obou skupin**

Mann-Whitney U Test (w/ continuity correction) (02_Biostatistika_Data02.sta)									
By variable Ukonceni_rehabilitace									
Marked tests are significant at p < .05000									
variable	Rank Sum propuštěn do ambulantní péče	Rank Sum přeložen na lůžko následné péče	U	Z	p-value	Z adjusted	p-value	Valid N propuštěn do ambulantní péče	Valid N přeložen na lůžko následné péče
Barthel_index_po_rehabilitaci	74127,00	8901,000	1998,000	13,93442	0,0000	14,00397	0,0000	290	117

**p-hodnota
Mannova-Whitneyova testu**

Z předchozího popisu je patrný výrazný rozdíl mezi skupinami (soběstačnost při propuštění do ambulantní péče je v mediánu 75 bodů, ale pacienti pokračující do následné péče mají medián pouze 40 bodů). P-hodnota statistické významnosti tohoto pozorovaného rozdílu je $p < 0,001$, což na hladině významnosti 0,05 značí **významný rozdíl**, a ze získaných dat tedy lze říct, že **aktuální soběstačnost pacientů souvisí s jejich dalším pokračováním v systému zdravotní péče**.

8.4. Kruskalův-Wallisův test

Kruskalův-Wallisův test je neparametrickou alternativou analýzou rozptylu (ANOVA). Je zobecněním Mannova-Whitneyova U testu pro více než dva výběry. Stejně jako předešlé neparametrické metody počítá s pořadím hodnot místo s původními daty. Podobně jako Mannův-Whitneyův U test i Kruskalův-Wallisův test testuje nulovou hypotézu o distribučních funkcích, ne o středních hodnotách. Pokud chceme interpretovat výsledek testu jako test o poloze (tedy o mediánu nebo průměru), musíme předpokládat, že tvar rozdělení je ve všech skupinách stejný. Z uvedeného plyne předpoklad stejného tvaru rozdělení ve všech skupinách. Test lze použít i pro ordinální data.

Postup výpočtu Kruskalova-Wallisova testu je následující.

1. Stanovíme nulovou a alternativní hypotézu o distribuční funkci $F(x)$ pro k skupin.
 $H_0: F(x_1) = F(x_2) = \dots = F(x_k)$, H_A : alespoň jedna $F(x_j)$ se liší od ostatních.
2. Hodnoty všech výběrů (skupin) jsou sloučena a je určeno jejich pořadí bez ohledu na příslušnost ke skupině.
3. Pro všechny výběry zvlášť je spočítán součet pořadí (T_1, \dots, T_k).
4. Ze součtů pořadí ve skupinách je určena finální hodnota testové statistiky Q :

$$Q = \frac{12}{n(n+1)} \sum_{j=1}^k \frac{T_j^2}{n_j} - 3(n+1)$$

5. Pokud je $Q \geq \chi^2 (k-1)$, nebo když příslušná p-hodnota \leq zvolená hladina významnosti, zamítáme nulovou hypotézu. Pro malé velikosti výběrů určujeme kritický obor z tabulek pro Kruskalův-Wallisův test.
6. V případě zamítnutí nulové hypotézy pokračujeme dále hledáním lišících se dvojic pomocí metod mnohonásobného porovnávání.

Příklad: Kruskalův-Wallisův test

Zadání: Zjistěte, zda etiologie vzniku mozkového infarktu (deficit způsobený embolií, trombózou nebo neurčenou okluzí/stenózou) je potenciálním prediktivním faktorem výsledného stupně soběstačnosti v základních denních aktivitách (ADL) vyjádřeného indexem Barthelové. Tj., liší se pacienti s různým typem vzniku mozkového infarktu ve výsledné soběstačnosti?

Postup v programu Statistica

1. Ověříme normalitu rozložení hodnot indexu Barthelové po rehabilitaci ve všech skupinách.

Postup výpočtu Kruskalova-Wallisova testu (po nemožnosti použít ANOVA):

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu $H_0: F(x_1) = F(x_2) = F(x_3)$, proti H_A : alespoň jedna dvojice $F(x_j)$ se liší.
2. Původní hodnoty Barthelové indexu převedeme na pořadí v celém souboru.
3. Vypočítáme **testovou statistiku Q** a odpovídající **p-hodnotu**. V menu *Statistics* zvolíme *Nonparametrics*, vybereme *Comparing multiple indep. samples (groups)*. Vybereme proměnnou, kterou chceme testovat (*dependent*) a proměnnou obsahující skupiny, které srovnáváme (*grouping*). Kliknutím na *Multiple comparisons of mean ranks for all groups* získáme výstupy (celkové srovnání ale také mnohonásobné porovnání mezi všemi skupinami). $Q = 23,63$, $p < 0,001$.
4. Testovou statistiku porovnáme s kritickou hodnotou nebo porovnáme p-hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p-hodnota $\leq \alpha$** , **zamítáme H_0** . Existuje alespoň jedna dvojice způsobu vzniku mozkového infarktu, která se liší v následné soběstačnosti pacientů. Dvojici s rozdílnou hodnotou Barthelové indexu interpretujeme z tabulky mnohonásobného porovnání.

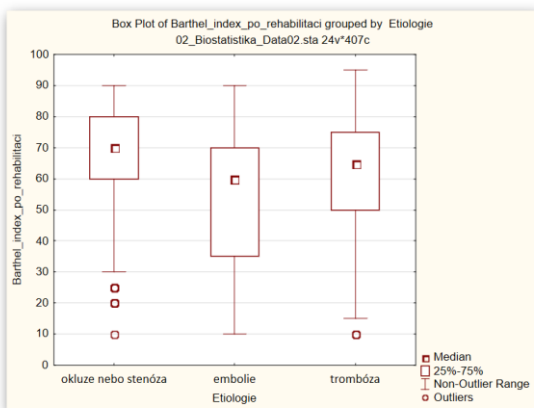
Výsledky v programu Statistica

Srovnání průměru a mediánu

Variable	Aggregate Results Descriptive Statistics (02_Biostatistika_Data02.sta)			
	Etiologie	Valid N	Mean	Median
Barthel_index_po_rehabilitaci	trombóza	128	60,89844	65,00000
Barthel_index_po_rehabilitaci	okluze nebo stenóza	201	65,99502	70,00000
Barthel_index_po_rehabilitaci	embolie	78	53,58974	60,00000

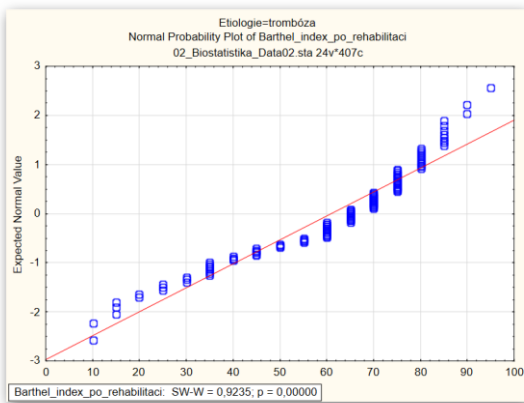
Ověření normality indexu Barthelové po rehabilitaci

Krabicový graf

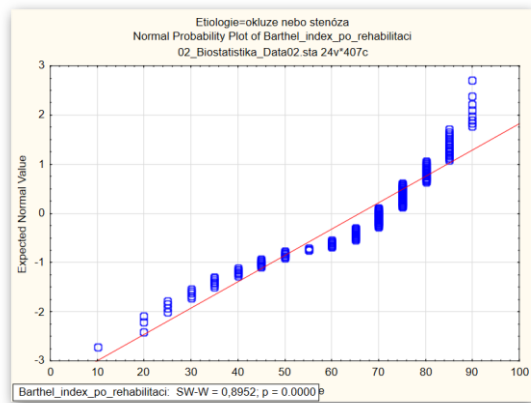


Základní popis i grafické srovnání ukazuje možný rozdíl mezi skupinami (soběstačnost po embolii je v mediánu 60 bodů, po trombóze 65 bodů a po neurčené okluzi nebo stenóze 70 bodů).

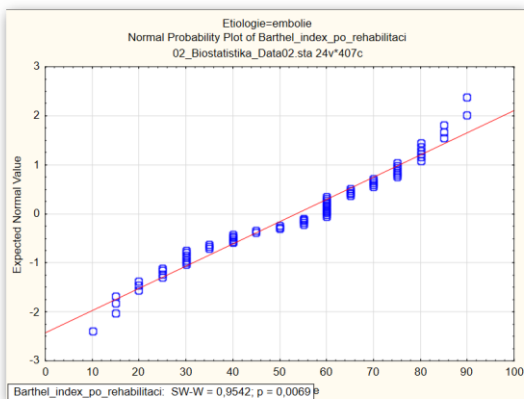
N-P graf BI pacientů s



N-P graf BI pacientů s



N-P graf BI pacientů s



Normalitu dat zamítáme u všech tří skupin ($p < 0,001$, $p < 0,001$ a $p = 0,007$) s tím, že u všech je porušení normality patrné graficky i z N-P grafu.

Depend.:	Multiple Comparisons p values (2-tailed); Barthel_index p		
Barthel_index_po_rehabilitaci	Independent (grouping) variable: Etiologie		
	Kruskal-Wallis test: H (2, N= 407) =23,6347 p =,0000		
	okluze nebo stenóza	embolie	trombóza
	R:228.81	R:154.65	R:195.11
okluze nebo stenóza		0,000007	0,033834
embolie	0,000007		0,050002
trombóza	0,033834	0,050002	

**Souhrnná p-hodnota
Kruskalova-Wallisova
testu**

**p-hodnoty mnohonásobného
porovnání všech skupin**

Z předchozího popisu je patrný možný rozdíl mezi skupinami (soběstačnost po embolii je v mediánu 60 bodů, po trombóze 65 bodů a po neurčené okluzi nebo stenóze 70 bodů).

Souhrnná p-hodnota statistické významnosti tohoto pozorovaného rozdílu je $p < 0,001$, což na hladině významnosti 0,05 značí **významný rozdíl** a ze získaných dat tedy **lze říct, že existuje alespoň jedna dvojice způsobu vzniku mozkového infarktu, která se liší v následné soběstačnosti pacientů (tj. etiologie souvisí s další soběstačností).**

Mnohonásobným porovnáním jsme navíc prokázali významný rozdíl mezi embolií a okluzí/stenózou a mezi trombózou a okluzí/stenózou (rozdíl mezi embolií a trombózou významný není). Jinými slovy, **výsledný stupeň soběstačnosti je významně lepší u pacientů s okluzí/stenózou oproti embolii i trombóze.**

9. Testy pro kvalitativní proměnné

Předchozí část byla věnována hodnocení kvantitativních proměnných, u nichž předpokládáme, že mohou nabývat mnoha rozdílných hodnot. V biologii i medicíně se často setkáváme s kvalitativními (kategoriálními) proměnnými, které mohou nabývat pouze omezený počet hodnot (např. pohlaví: muž/žena, výskyt atopického ekzému: ano/ne, krevní skupina: A/B/AB/0).

Stejně jako u kvantitativních, tak i u kvalitativních proměnných můžeme hodnotit, zda je hodnota vybrané charakteristiky rovna zvolené hodnotě, nebo zda spolu souvisí výskyt dvou proměnných.

Pro kategoriální data používáme testy dobré shody a používáme kritéria (statistiky) χ^2 :

$\chi^2 = \sum_{i=1}^k \frac{(O-E)^2}{E}$, kde O je pozorovaná četnost v i -té kategorii (z angličtiny observed) a E je očekávaná četnost (expected). Výpočet testu dobré shody si můžeme přiblížit na příkladu hození mincí. Představme si, že 10 000 lidí hází mincí a v 4 000 případech padne rub, v 6 000 případech padne líc. Lze výsledek považovat za statisticky významně odlišný od očekávaného poměru 1 : 1? Nulová hypotéza předpokládá, že výskyt jevů rub a líc nastává v poměru 1 : 1. Očekávanými četnostmi obou jevů je tedy hodnota 5 000. Dosazením do vzorce získáme hodnotu testové statistiky χ^2 .

$$\chi^2 = \frac{(4000-5000)^2}{5000} + \frac{(6000-5000)^2}{5000} = 200 + 200 = 400$$

Vypočítanou hodnotu testové statistiky porovnáme s tabulkovou hodnotou pro zvolenou hladinu významnosti $\alpha = 0,05$ a jeden stupeň volnosti, která je: $\chi^2_{(1)}(0,95) = 3,84$. Vypočítaná hodnota testové statistiky je větší, než tabulková hodnota pro zvolenou hladinu významnosti a jeden stupeň volnosti a proto zamítáme nulovou hypotézu. Námi pozorovaný výskyt rubové a lícové strany mince se liší od poměru 1:1.

Test dobré shody se používá také pro srovnání pozorovaných četností proti očekávaným četnostem daným určitým pravidlem (např. Hardy-Weinbergova rovnováha v genetice).

9.1. Kontingenční tabulka

Vztah dvou nebo více kategoriálních proměnných lze sumarizovat v kontingenční tabulce. Nejčastěji sledujeme vztah dvou proměnných, který hodnotíme pomocí dvourozměrné tabulky. Řádky (r) jsou tvořeny hodnotami (kategoriemi) prvního znaku (x_1, \dots, x_r), sloupce (c) hodnotami druhého znaku (y_1, \dots, y_c). Nezáleží na tom, která proměnná tvoří řádky a která proměnná tvoří sloupce tabulky. V příslušné buňce tabulky je uveden počet případů s hodnotou prvního znaku odpovídající příslušnému řádku a druhého znaku s hodnotou odpovídající příslušnému sloupci (n_{11}, \dots, n_{rc}). V posledním sloupci jsou uvedeny řádkové součty ($n_{1.}, \dots, n_{r.}$), v posledním řádku jsou uvedeny sloupcové součty ($n_{.1}, \dots, n_{.c}$) – oboje označujeme jako tzv. marginální četnosti. (Tabulka 9.1)

Tabulka 9.1. Kontingenční tabulka dvou kategoriálních proměnných X a Y – obecný zápis.

	y_1	...	y_c	
x_1	n_{11}	...	n_{1c}	$n_{1.}$
...
x_r	n_{r1}	...	n_{rc}	$n_{r.}$
	$n_{.1}$...	$n_{.c}$	N

Speciálním případem kontingenční tabulky je čtyřpolní tabulka s dvěma řádky a dvěma sloupci pro hodnocení vztahu kategoriálních proměnných popisujících dva binární znaky, např. pohlaví a nemocnost (Tabulka 9.2, 9.3).

Tabulka 9.2. Kontingenční tabulka dvou kategoriálních proměnných: pohlaví a nemocnost – obecný zápis.

	Nemocný	Zdravý	Celkem
Muž	a	b	a + b
Žena	c	d	c + d
Celkem	a + c	b + d	a + b + c + d

Tabulka 9.3. Kontingenční tabulka dvou kategoriálních proměnných: pohlaví a nemocnost 87 pacientů.

	Nemocný	Zdravý	Celkem
Muž	45	11	56
Žena	25	6	31
Celkem	70	17	87

U kontingenčních tabulek máme možnost analyzovat vazbu mezi dvěma kategoriálními proměnnými. Základním způsobem testování je tzv. chí-kvadrát test, který srovnává pozorované četnosti kombinací kategorií oproti očekávaným četnostem, které vychází z teoretické situace, kdy je vztah mezi proměnnými náhodný. Očekávané četnosti odhadujeme na základě marginálních četností. Očekávaná četnost dané kombinace kategorií je součinem marginálních četností příslušného sloupce a řádku poděleným celkovým počtem pozorování.

Kontingenční tabulka umožňuje testování následujících hypotéz:

1. *Hypotéza o nezávislosti.* Pomocí testu nezávislosti můžeme rozhodnout, zda spolu souvisí výskyt dvou kategoriálních proměnných (např. pohlaví a výskyt atopického exému). V tomto případě sledujeme dvě kategoriální proměnné u jednoho výběru. Testem nezávislosti pro kontingenční tabulku je Pearsonův chí-kvadrát test nebo Fisherův exaktní test.
2. *Hypotéza o shodnosti struktury (test homogenity).* O testování homogenity mluvíme v situaci, kdy nás zajímá výskyt kategoriální proměnné u r nezávislých výběrů z r různých populací (např. typy nežádoucích účinků v několika r nemocnicích). V tomto případě sledujeme jednu kategoriální proměnnou u více

výběrů. K hodnocení shodnosti struktury používáme znovu Pearsonův chí-kvadrát test nebo Fisherův exaktní test.

3. *Hypotéza o symetrii.* V případě, že uvažujeme opakované měření jedné proměnné na jednom výběrovém souboru a zajímá nás hodnocení změny v jejích hodnotách, mluvíme o testování symetrie. Jde o obdobu párového testování kvantitativních proměnných. Příkladem testování symetrie je např. hodnocení výskytu lézí atopického exému před a po léčbě, výskyt bolesti před a po léčbě apod. K testování symetrie používáme McNemarův test.

9.2. Pearsonův chí-kvadrát test

Pearsonův chí-kvadrát test je základním a nejpoužívanějším testem nezávislosti i homogenity. Nulovou hypotézou je zde tvrzení, že proměnné jsou nezávislé, resp. homogenní. Test je založen na srovnání pozorovaných četností a tzv. očekávaných četností jednotlivých kombinací kategorií dvou proměnných. Když označíme n_{ij} počet pozorování, u nichž byla zjištěna i -tá kategorie první proměnné a j -tá kategorie druhé proměnné, můžeme označit marginální četnosti i -té varianty první proměnné jako $n_{i.} = \sum_{j=1}^c n_{ij}$ a j -té varianty druhé proměnné jako $n_{.j} = \sum_{i=1}^r n_{ij}$. Očekávané četnosti jednotlivých kombinací kategorií pak vypočítáme podle vzorce $e_{ij} = np_{ij} = np_i p_j = n \frac{n_{i.}}{n} \frac{n_{.j}}{n} = \frac{n_{i.} n_{.j}}{n}$. Testová statistika

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

má za platnosti nulové hypotézy o nezávislosti resp. o homogenitě rozdělení pravděpodobnosti s parametrem $(r-1)(c-1)$. Nulovou hypotézu zamítáme, když

$$\chi^2 \geq \chi_{(r-1)(c-1)}^2(1 - \alpha).$$

Pearsonův chí-kvadrát test není možné použít ve všech případech, ale pouze za splnění předpokladů testu:

1. Jednotlivá pozorování shrnutá v kontingenční tabulce jsou nezávislá, tedy každý prvek je zahrnut pouze v jedné buňce kontingenční tabulky.
2. Podmínka dobré aproximace. Alespoň 80 % buněk kontingenční tabulky má očekávanou četnost větší než 5 a všechny buňky tabulky (tedy 100 % buněk) mají očekávanou četnost větší než 2. Na tomto místě upozorňujeme, že očekávaná četnost není četnost pozorovaná a je součástí výsledků uváděných většinou statistických softwarů. V případě, že není tento předpoklad splněn, je vhodné sloučit kategorie s nízkými četnostmi. V případě čtyřpolní tabulky, tj. tabulky se dvěma kategoriemi u obou proměnných, kde již není možné další slučování kategorií a při neplnění podmínky dobré aproximace není možné použít Pearsonův chí-kvadrát test. Neparametrickou alternativou je Fisherův exaktní test.

Postup výpočtu chí-kvadrát testu je následovný (po ověření jeho předpokladů).

1. Stanovíme nulovou a alternativní hypotézu o nezávislosti kategoriálních proměnných:

H_0 : Proměnné X a Y jsou nezávislé náhodné veličiny.

H_A : Proměnné X a Y jsou závislé náhodné veličiny.

2. Určíme pozorované i očekávané četnosti všech kombinací kategorií.
3. Z očekávaných a pozorovaných četností vypočítáme finální hodnotu testové statistiky χ^2 .
4. Hodnotu testové statistiky χ^2 porovnáme s kritickou hodnotou testu. Pokud je tato hodnota větší než kritická hodnota testu, nebo p-hodnota menší než hladina významnosti α , zamítáme nulovou hypotézu o nezávislosti.

Výpočet chí-kvadrát testu ukážeme na příkladu výskytu nemoci u mužů a žen z Tabulky 9.3. Testujeme nulovou hypotézu H_0 : Pohlaví a výskyt nemoci jsou nezávislé veličiny.

U chí-kvadrát testu určíme pozorované a očekávané četnosti (Tabulka 9.3, 9.4). Vypočítáme hodnotu testové statistiky $\chi^2 = 0,001$, určíme počet stupňů volnosti $df = 1$ a příslušnou p-hodnotu $p = 0,974$. Hodnotu testové statistiky porovnáme s kritickou hodnotou $\chi^2_{(1)}(0,95) = 3,84$ pro danou hladinu významnosti a daný počet stupňů volnosti. Hodnota testové statistiky je menší než kritická hodnota, tudíž nezamítáme H_0 .

Tabulka 9.4. Očekávané četnosti pacientů v kategoriích definovaných pohlavím a nemocností.

	Nemocný	Zdravý	Celkem
Muž	45,1 (70·56)/87	10,9 (17·56)/87	56
Žena	24,9 (70·31)/87	6,1 (17·31)/87	31
Celkem	70	17	87

Pearsonův chí-kvadrát test používáme i k testování hypotézy o shodě struktury. Výpočet probíhá stejně jako u hypotézy o nezávislosti pomocí testového kritéria χ^2 založeného na pozorovaných a očekávaných četnostech. Rozdíl je ve formulaci nulové hypotézy H_0 , která tvrdí, že pravděpodobnostní rozdělení kategoriální proměnné je stejné v různých populacích. Zajímá nás totiž výskyt kategoriální proměnné u r nezávislých výběrů. Jako příklad hypotézy o shodě struktury můžeme uvést hodnocení typologie zaznamenaných nežádoucích účinků u pacientů s infarktem myokardu v několika (r) nemocnicích.

Příklad: Pearsonův chí-kvadrát test

Zadání: Stupeň soběstačnosti pacientů po mozkovém infarktu lze pomocí indexu Barthelové vyjádřit také kategoriálně. Např. pro definici vysoce závislých pacientů bylo stanoveno rozmezí 0 až 40 bodů. Zjistěte, zda je u žen a mužů stejné procento alespoň částečně soběstačných pacientů (45 až 100 bodů) a zda je tento rozdíl statisticky významný.

Postup výpočtu Pearsonova chí-kvadrát testu v programu Statistica

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu
 H_0 : Stupeň soběstačnosti nezávisí na pohlaví
 H_A : Stupeň soběstačnosti a pohlaví jsou závislé veličiny.

2. Vypočítáme očekávané a pozorované četnosti v kategoriích.
3. Vypočítáme **testovou statistiku** a odpovídající **p-hodnotu**. V menu *Statistics* zvolíme *Basic statistics*, vybereme *Tables and banners* (v češtině *Kontingenční tabulky*). Na záložce *Stub-and-banner* vybereme **proměnné**, které chceme testovat, a potvrdíme *OK*. Na záložce *Options* zaškrtneme *Expected frequencies* (*Očekávané četnosti* potřebné k ověření podmínek dobré aproximace) a *Pearsonův chí-kvadrát*. Poté se vrátíme na záložku *Advanced* a přes volbu *Detailed two-way tables* získáme výsledky.

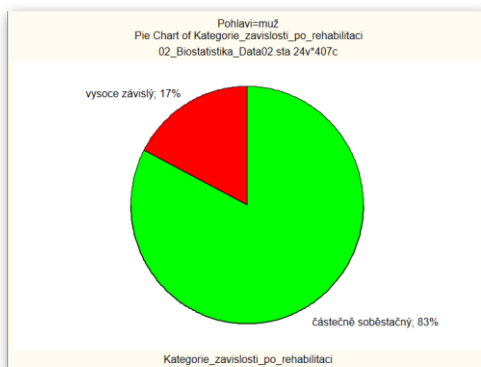
$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - e_{ij})^2}{e_{ij}} = \frac{(205-200)^2}{200} + \frac{(123-128)^2}{128} + \frac{(43-48)^2}{48} + \frac{(36-31)^2}{31}$$

$$\chi^2 = 1,74, p = 0,187.$$

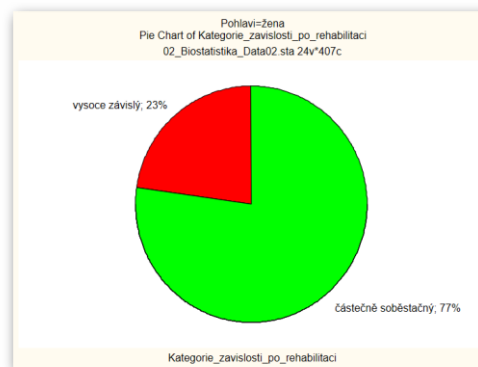
4. Testovou statistiku porovnáme s kritickou hodnotou nebo porovnáme p-hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p-hodnota** $> \alpha$, **nezamítáme H_0** . Stupeň soběstačnosti nezávisí na pohlaví (tj. výsledná míra soběstačnosti se u žen a u mužů neliší).

Výsledky v programu Statistica

Zastoupení částečně soběstačných a vysoce závislých mužů



Zastoupení částečně soběstačných a vysoce závislých žen



Ze základního popisu je patrný mírný rozdíl v procentu částečně soběstačných pacientů na konci hospitalizace. U žen je podíl těchto pacientů 77 % oproti 83 % u mužů.

Pozorované četnosti

2-Way Summary Table: Observed Frequencies (02_Biostat)			
Marked cells have counts > 10			
Pohlaví	Kategorie_zavislosti_po_rehabilitaci částečně soběstačný	Kategorie_zavislosti_po_rehabilitaci vysoce závislý	Row Totals
muž	205	43	248
žena	123	36	159
Totals	328	79	407

Z předchozího popisu je patrný mírný rozdíl mezi muži a ženami (u žen je podíl částečně soběstačných pacientů 77 % oproti 83 % u mužů).

Očekávané četnosti

	Kategorie_zavislosti_ po_rehabilitaci částečně soběstačný	Kategorie_zavislosti_ po_rehabilitaci vysoce závislý	Row Totals
Pohlavi			
muž	199,8624	48,13759	248,0000
žena	128,1376	30,86241	159,0000
Totals	328,0000	79,00000	407,0000

Očekávané četnosti jsou 200, 48, 128 a 31, což jsou dostatečně vysoké počty a podmínka dobré aproximace pro použití chí-kvadrát testu je tedy splněna.

Statistic	Chi-square	df	p
Pearson Chi-square	1,741617	df=1	p=,18693
M-L Chi-square	1,720067	df=1	p=,18808

**p-hodnota
Pearsonova
chí-kvadrát testu**

P-hodnota statistické významnosti pozorované závislosti je $p = 0,187$, což na hladině významnosti 0,05 značí **nevýznamný výsledek** a ze získaných dat tedy **nelze říct, že by míra soběstačnosti souvisela s pohlavím**.

9.3. Fisherův exaktní test

V případech, kdy nemůžeme použít Pearsonův chí-kvadrát test z důvodu nesplnění jeho předpokladu, testujeme hypotézu o nezávislosti i hypotézu o shodě struktury pomocí Fisherova exaktního testu. Jeho použití je tedy vhodné v případě, kdy máme kontingenční tabulku s malými očekávanými četnostmi. Tento test je ovšem v běžných statistických programech k dispozici pouze pro čtyřpolní tabulky. Fisherův exaktní test patří mezi neparametrické testy pracující s daty na nominální škále, v nejjednodušší podobě s binárními proměnnými. Hlavní myšlenkou Fisherova exaktního testu je výpočet pravděpodobnosti, se kterou bychom získali čtyřpolní tabulky stejně nebo více vzdálené od nulové hypotézy při zachování pozorovaných marginálních četností. Slovo exaktní (přímý) znamená, že test přímo vypočítává přesnou p-hodnotu jako pravděpodobnost, s jakou dostaneme za předpokladu platnosti nulové hypotézy tabulku stejně nebo více odlišnou od nulové hypotézy.

Ukázka podrobného výpočtu exaktní p-hodnoty u Fisherova testu není cílem této učebnice. Zvědavého čitatele můžeme nasměrovat např. na učebnici Biostatistika autorů Pavlík, Dušek (2012), kde je postup výpočtu Fisherova exaktního testu uveden i s řešením vzorového příkladu na str. 109-110.

Stručný postup výpočtu Fisherova exaktního testu je následovný.

1. Stanovíme nulovou a alternativní hypotézu o nezávislosti dvou kategoriálních proměnných:
 H_0 : Proměnné X a Y jsou nezávislé náhodné veličiny.
 H_A : Proměnné X a Y jsou závislé náhodné veličiny.
2. Určíme pozorované četnosti všech kombinací kategorií.
3. Vypočítáme exaktní p-hodnotu.

4. Exaktní p-hodnotu porovnáme s hladinou významnosti α , na které hodnotíme nulovou hypotézu. Pokud je vypočítaná p-hodnota menší než hladina významnosti, zamítáme nulovou hypotézu o nezávislosti.

Příklad: Fisherův exaktní test

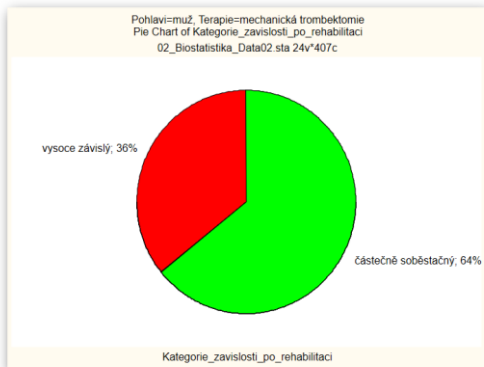
Zadání: Stupeň soběstačnosti pacientů po mozkovém infarktu lze pomocí indexu Barthelové vyjádřit také kategoriálně. Např. pro definici vysoce závislých pacientů bylo stanoveno rozmezí 0 až 40 bodů. Zjistěte, zda je u žen a mužů léčených mechanickou trombektomií stejné procento alespoň částečně soběstačných pacientů (45 až 100 bodů) a zda je tento rozdíl statisticky významný.

Postup výpočtu Fisherova exaktního testu v programu Statistica (po nemožnosti použít Pearsonův chí-kvadrát test)

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu
 H_0 : Stupeň soběstačnosti nezávisí na pohlaví
 H_A : Stupeň soběstačnosti a pohlaví jsou závislé veličiny.
2. Spočítá se parciální pravděpodobnost (p_a) všech možných tabulek při zachování marginálních četností. Výsledná p-hodnota je součtem p_a menších nebo stejných jako pravděpodobnost, která přísluší naší pozorované tabulce. V menu *Statistics* zvolíme *Basic statistics*, vybereme *Tables and banners* (v češtině *Kontingenční tabulky*). Na záložce *Stub-and-banner* vybereme **proměnné**, které chceme testovat, a potvrdíme *OK*. Na záložce *Options* zaškrtneme *Expected frequencies* (*Očekávané četnosti* potřebné k ověření podmínek dobré aproximace) a *Fisher exact*. V nastavení *By Group* vybereme jako třídící proměnnou terapii (analýza se tak provede pro všechny druhy terapie samostatně). Poté se vrátíme na záložku *Advanced* a přes volbu *Detailed two-way tables* získáme výsledky (zde nás budou zajímat pouze výsledky pro skupinu pacientů s mechanickou trombektomií).
 $p = 0,700$
3. Vypočítané p porovnáme s hladinou významnosti $\alpha = 0,05$.
4. Je-li **p-hodnota** $> \alpha$, **nezamítáme H_0** . Stupeň soběstačnosti nezávisí na pohlaví (tj. výsledná míra soběstačnosti se u žen a u mužů podstupujících mechanickou trombektomii neliší).

Výsledky v programu Statistica

Zastoupení částečně soběstačných a vysoce závislých mužů



Zastoupení částečně soběstačných a vysoce závislých žen



Ze základního popisu je patrný mírný rozdíl v procentu částečně soběstačných pacientů na konci hospitalizace. U žen je podíl těchto pacientů 73 % oproti 64 % u mužů.

Pozorované četnosti

Terapie=mechanická trombektomie 2-Way Summary Table: Observed Frequencies (02_Biostatistika_Data02 sta 24v*407c)			
Pohlavi	Kategorie_zavislosti_po_rehabilitaci		Row Totals
	vysoce závislý	částečně soběstačný	
muž	5	9	14
žena	4	11	15
Totals	9	20	29

Z předchozího popisu je patrný mírný rozdíl mezi muži a ženami (u žen je podíl částečně soběstačných pacientů 73 % oproti 64 % u mužů).

Očekávané četnosti

Terapie=mechanická trombektomie 2-Way Summary Table: Expected Frequencies (02_Biostatistika_Data02 sta 24v*407c)			
Pohlavi	Kategorie_zavislosti_po_rehabilitaci		Row Totals
	vysoce závislý	částečně soběstačný	
muž	4,344828	9,65517	14,00000
žena	4,655172	10,34483	15,00000
Totals	9,000000	20,00000	29,00000

Očekávané četnosti jsou 4, 10, 5 a 10, což nejsou dostatečně vysoké počty a místo chí-kvadrát testu je tedy vhodné použít Fisherův exaktní test.

Statistic	Terapie=mechanická trombektomie Statistics: Pohlavi(2) x Kategorie_zavislosti_po_rehabilitaci		
	Chi-square	df	p
Pearson Chi-square	,2769577	df=1	p=,59870
M-L Chi-square	,2771859	df=1	p=,59855
Yates Chi-square	,0155357	df=1	p=,90081
Fisher exact, one-tailed			p=,44998
two-tailed			p=,69985
McNemar Chi-square (A/D)	1,562500	df=1	p=,21130
(B/C)	1,230769	df=1	p=,26726

**p-hodnota
Fisherova
exaktního testu**

P-hodnota statistické významnosti pozorované závislosti je $p = 0,700$, což na hladině významnosti 0,05 značí **nevýznamný výsledek** a ze získaných dat tedy **nelze říct, že by míra soběstačnosti souvisela s pohlavím**.

9.4. McNemarův test

McNemarův test je test pro kontingenční tabulku v případě párového uspořádání experimentu, kdy sledujeme výskyt kategoriální proměnné na stejném výběrovém souboru dvakrát po sobě. Jde o obdobu párového t-testu. McNemarovým testem hodnotíme, zda se mezi oběma opakováními experimentu (opakovaným sledováním) liší pravděpodobnosti výskytu jednotlivých variant proměnné.

V biologii a medicíně je relativně častá situace, kdy je opakovaně sledovaná kategoriální proměnná nabývá dvou hodnot (binární proměnná). V tomto případě je kontingenční tabulka čtyřpolní, její obecná forma je uvedena v Tabulce 9.5. Příkladem hodnocení symetrie binární proměnné může být zhodnocení výskytu atopického ekzému před léčbou a po léčbě, nebo výskyt bolesti před a po užití léku. Nulová hypotéza je definována následovně. H_0 : Pravděpodobnost nastání první varianty při prvním měření a druhé varianty při druhém měření je stejná jak nastání druhé varianty při prvním měření a první varianty při druhém měření. Při definování pozorovaných četností v čtyřpolní tabulce jako a, b, c, d (Tabulka 9.5) lze nulovou hypotézu stručně zapsat jako $p_b = p_c$.

Tabulka 9.5. Kontingenční tabulka dvou kategoriálních proměnných: pohlaví a nemocnost – obecný zápis.

		Po léčbě		Celkem
		Bolest ANO	Bolest NE	
Před léčbou	Bolest ANO	a	b	a + b
	Bolest NE	c	d	c + d
Celkem		a + c	b + d	a + b + c + d

Testová statistika McNemarova testu χ^2 má tvar: $\chi^2 = \frac{(|b-c|-1)^2}{b+c}$. Pokud je vypočítaná hodnota testové statistiky $\chi^2 \geq \chi^2_{(1)}(1 - \alpha)$, pak zamítáme nulovou hypotézu.

Postup výpočtu McNemarova testu je následovný.

1. Stanovíme nulovou a alternativní hypotézu o symetrii kategoriální proměnné:
 H_0 : Počet případů b je stejný jako počet případů c .
 H_A : Počet případů b není stejný jako počet případů c .
2. Určíme pozorované četnosti všech kombinací kategorií.
3. Vypočítáme hodnotu testové statistiky $\chi^2 = \frac{(|b-c|-1)^2}{b+c}$ a příslušnou p-hodnotu.
4. Hodnotu testové statistiky χ^2 porovnáme s kritickou hodnotou testu, nebo p-hodnotu s hladinou významnosti. Pokud je hodnota testové statistiky větší než kritická hodnota testu, nebo p-hodnota menší než hladina významnosti α , zamítáme nulovou hypotézu o symetrii.

Příklad: McNemarův test

Zadání: Pacientům hospitalizovaným s mozkovým infarktem byla na lůžku akutní péče poskytnuta terapie pro obnovu krevního oběhu v postižené části mozku. Po zvládnutí akutní fáze byl u pacientů vyhodnocen stupeň soběstačnosti pomocí indexu Barthelové (BI) jako vysoce závislý (0 až 40 bodů) nebo částečně soběstačný (45 až 100 bodů) a byli přeloženi na rehabilitační oddělení. Po dvou týdnech byl stejně vyhodnocen stupeň soběstačnosti dle BI. Zjistěte, zda poskytnutá rehabilitační péče vedla ke zvýšení podílu alespoň částečně soběstačných pacientů.

Postup výpočtu McNemarova testu v programu Statistica

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu
 H_0 : „Počet zhoršených případů je stejný jako počet zlepšení“ proti
 H_A : „Počet zhoršených případů není stejný jako počet zlepšení.“
2. Vypočítáme pozorované četnosti měnících se stavů.
3. Vypočítáme **testovou statistiku χ^2** a odpovídající **p-hodnotu**. V menu *Statistics* zvolíme *Basic statistics*, vybereme *Tables and banners* (v češtině *Kontingenční tabulky*). Na záložce *Stub-and-banner* vybereme **proměnné**, které chceme testovat, a potvrdíme **OK**. Na záložce *Options* zaškrtneme *McNemar (2x2)*. Poté se vrátíme na záložku *Advanced* a přes volbu *Detailed two-way tables* získáme výsledky.
$$\chi^2 = \frac{(|b-c|-1)^2}{b+c} = \frac{(|280-0|-1)^2}{280+0} = 278, p < 0,001$$
4. Testovou statistiku porovnáme s kritickou hodnotou nebo porovnáme p-hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p-hodnota $\leq \alpha$** , **zamítáme H_0** . Během rehabilitace se podařilo změnit míru soběstačnosti pacientů.

Výsledky v programu Statistica

Pozorované četnosti

Kategorie_zavislosti_pred_rehabilitaci	Kategorie_zavislosti_po_rehabilitaci_částečně soběstačný	Kategorie_zavislosti_po_rehabilitaci_vysoce závislý	Row Totals
vysoce závislý	A 280	B 79	359
částečně soběstačný	C 48	D 0	48
Totals	328	79	407

Počet pacientů, u kterých došlo ke změně z vysoce závislého stavu do částečně soběstačného je 280. Naopak ke zhoršení nedošlo u žádného pacienta. Počty změn jsou v kontingenční tabulce na pozicích A a D.

Statistics: Kategorie_zavislosti_pred			
Statistic	Chi-square	df	p
Pearson Chi-square	13,10673	df=1	p=,00029
M-L Chi-square	22,21371	df=1	p=,00000
Yates Chi-square	11,73772	df=1	p=,00061
Fisher exact, one-tailed			p=,00002
two-tailed			p=,00002
McNemar Chi-square (A/D)	278,0036	df=1	p=0,0000
(B/C)	7,086614	df=1	p=,00777

**p-hodnota
McNemarova testu**

Dvě hodnoty testových statistik a p-hodnoty podle toho, kde jsou ve výstupní kontingenční tabulce uloženy četnosti, u kterých jsme při opakovaném měření zaznamenali rozdílné výsledky (A/D nebo B/C).

P-hodnota statistické významnosti pozorované změny je $p < 0,001$, což na hladině významnosti 0,05 značí **významný výsledek** a ze získaných dat jsme **prokázali, že během rehabilitace se podařilo změnit míru soběstačnosti pacientů v denních aktivitách.**

10. Závislost dvou kvantitativních proměnných: korelace

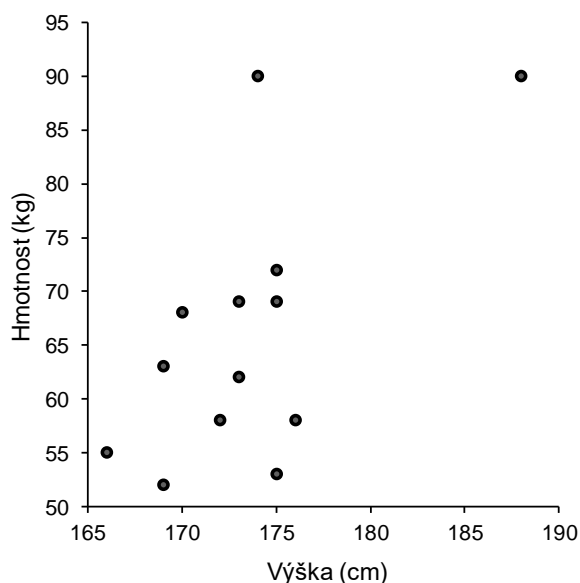
V této kapitole představíme hodnocení vztahu mezi dvěma spojitými proměnnými, které je základem tzv. korelační a regresní analýzy.

Úkoly, které můžeme těmito metodami řešit, jsou následující:

1. Chceme zjistit, jestli mezi proměnnými existuje potenciální vztah, např. jestli vyšší hodnoty jedné proměnné znamenají nižší hodnoty druhé proměnné (např. zda vyšší hladina krevní glukózy souvisí s vyšší hladinou jiné látky v krevní plazmě).
2. Chceme predikovat hodnoty jedné proměnné na základě znalosti hodnot druhé proměnné (např. na základě znalostí koncentrace určité látky v prostředí predikovat koncentrace jiné látky, která je těžko měřitelná nebo její měření je finančně velice náročné).
3. Chceme kvantifikovat vztah mezi dvěma spojitými proměnnými, např. pro použití jedné proměnné na místo druhé proměnné jako diagnostického testu.

První úkol je typickým příkladem pro korelační analýzu, jež je využívána pro hodnocení míry vztahu spojitých proměnných. Na rozdíl od ní, regresní analýza vytváří model vztahu spojitých proměnných a zabývá se závislostí tzv. vysvětlované proměnné na vysvětlující proměnné (prediktor).

V této části se budeme věnovat korelační analýze. U ní nepředpokládáme, že nutně existuje funkční závislost jedné proměnné na druhé proměnné. Dvě proměnné jsou pouze korelovány a obě proměnné jsou zatíženy náhodnou variabilitou. Nejjednodušším způsobem, jak vizualizovat vztah dvou spojitých proměnných je jejich zobrazení v bodovém grafu, který nám naznačí, zda hodnoty jedné proměnné stoupají nebo klesají se zvyšujícími se hodnotami druhé proměnné. Zobrazení vztahu dvou proměnných uvádíme na příkladu výšky a hmotnosti skupiny studentů v jarním semestru 2010 (Obrázek 10.1).



Obrázek 10.1. Bodový graf výšky (cm) a hmotnosti (kg) skupiny studentů.

10.1. Pearsonův korelační koeficient

Obdobně jako jiné statistické metody, i korelační koeficient rozlišujeme parametrický a neparametrický. Předpokladem pro použití parametrického korelačního koeficientu je dvourozměrné normální rozdělení proměnných, tzn., že pro každou hodnotu X má proměnná Y normální rozdělení a pro každou hodnotu Y má proměnná X normální rozdělení. Z tohoto předpokladu plyne linearita vztahu. Pro jeho kvantifikaci slouží Pearsonův korelační koeficient r (když uvedeme zkráceně pojem korelační koeficient, myslíme tím obvykle Pearsonův korelační koeficient). Nabývá hodnot od -1 do 1 , jeho hodnota je kladná když vyšší hodnoty proměnné X souvisí s vyššími hodnotami proměnné Y , naopak korelační koeficient je záporný, když nižší hodnoty proměnné X souvisí s vyššími hodnotami proměnné Y (Obrázek 10.2a, 10.2b). V případě, že proměnné nejsou nezávislé, ale jejich vztah není lineární, je jejich korelační koeficient nulový (Obrázek 10.2c).

Výpočet korelačního koeficientu r výběru se řídí vzorcem:

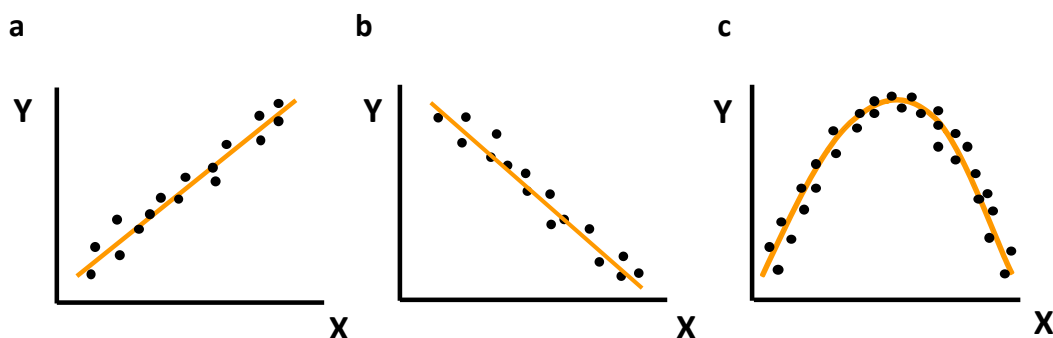
$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}},$$

kde \bar{x} , \bar{y} jsou výběrové průměry, x_i , y_i jsou hodnoty proměnných X a Y i -tého pozorování.

Hodnotu r počítáme pro výběr pozorování. Často nás zajímá, zda je korelační koeficient statisticky významný, konkrétně testujeme nulovou hypotézu, že korelační koeficient je nulový. Testování probíhá prostřednictvím testové statistiky t :

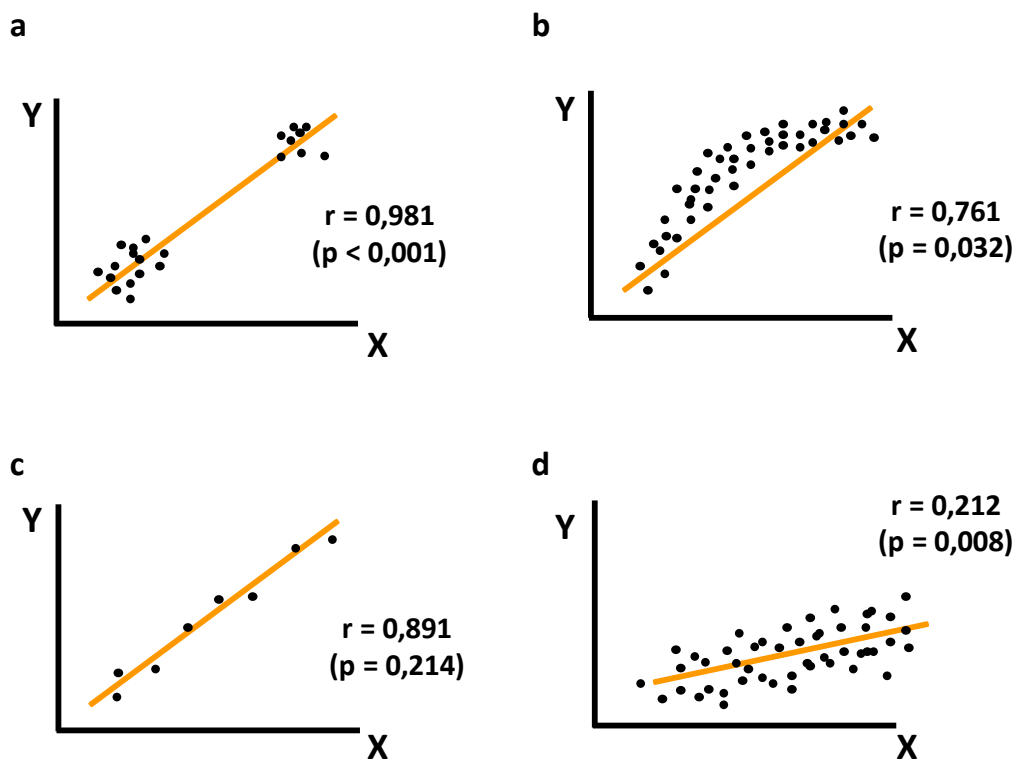
$$t = r \sqrt{\frac{n-2}{1-r^2}}.$$

Nulovou hypotézu zamítáme na hladině α , když hodnota testové statistiky přesáhne v absolutní hodnotě kvantil $t_{1-\alpha/2}^{(n-2)}$, nebo příslušná p -hodnota je nižší než hladina významnosti α .



Obrázek 10.2. Ukázka kladně korelovaných proměnných (a), záporně korelovaných proměnných (b) a proměnných, které jsou závislé nelineárně a jejichž korelace je nulová (c).

Obrázek 10.3 znázorňuje situace, kde je výpočet Pearsonova korelačního koeficientu problematický, resp. nesmyslný. První případ zobrazuje situaci, kdy soubor obsahuje dvě skupiny pozorování s odlišnými hodnotami proměnných X a Y. Jejich korelační koeficient ukazuje na silnou kladnou korelaci, vysoce statisticky významnou. Problémem je ovšem výpočet korelačního koeficientu pro celý soubor. Správným řešením by bylo rozdělení souboru na dva podsoubory a výpočet korelačního koeficientu pro každou skupinu zvlášť (Obrázek 10.3a). Druhý případ ukazuje situaci, kdy je mezi proměnnými X a Y nelineární vztah. Korelační koeficient znovu poukazuje na kladnou korelaci, jež je statisticky významná. Neodpovídá ovšem skutečnosti (Obrázek 10.3b). Poslední dva případy poukazují na problém s velikostí souboru, kdy v případě nízkého počtu pozorování je vysoká hodnota korelačního koeficientu statisticky nevýznamná (Obrázek 10.3c) a v případě vysokého počtu pozorování je relativně nízká hodnota korelačního koeficientu je statisticky významná (Obrázek 10.3d).



Obrázek 10.3. Ukázky možných problematických situací pro výpočet Pearsonova korelačního koeficientu: problém se dvěma skupinami (a), nelineárním vztahem mezi proměnnými (b), malým výběrovým souborem (c) a velkým výběrovým souborem (d).

Příklad: Pearsonův korelační koeficient

Zadání: U pacientů hospitalizovaných s mozkovým infarktem bylo při propuštění vyhodnoceno zlepšení míry soběstačnosti vyjádřené diferencí hodnot indexu Barthelové. Zjistěte, zda má věk vliv na úspěšnost terapeutické a rehabilitační péče. Jinými slovy, určete, zda věk koreluje s diferencí indexu Barthelové.

Postup v programu Statistica

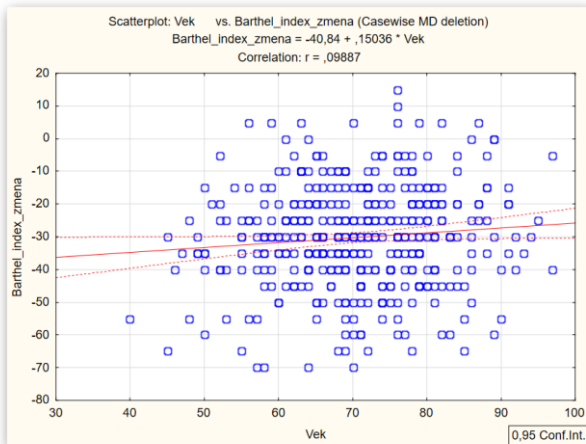
1. Ověříme předpoklady použití Pearsonova korelačního koeficientu: normalita rozložení věku a diferencí BI (vizuálně i Shapiro-Wilkovým testem). Postupujeme všemi kroky stejně jako v příkladu 6.1, tj. nastavíme provedení Shapiro-Wilkova testu u histogramu i diagnostického N-P grafu.

Postup výpočtu Pearsonova korelačního koeficientu (po ověření předpokladu normality)

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu
 $H_0: r = 0$ proti $H_A: r \neq 0$.
2. Graficky znázorníme závislost obou proměnných pomocí bodového XY grafu.
3. Vypočítáme hodnotu **korelačního koeficientu r** a odpovídající **p-hodnotu**. V menu *Statistics* zvolíme *Basic statistics*, vybereme *Correlation matrices*. Vybereme obě proměnné, které chceme testovat (*Two lists*). V záložce *Advanced* kliknutím na *2D scatterplots* získáme grafické znázornění závislosti vybraných proměnných. Poté v záložce *Options* zvolíme možnost *Display r, p-values, and N's* a přes *Summary* zobrazíme výsledky.
 $r = 0,099$, $p = 0,046$
4. Testovou statistiku porovnáme s kritickou hodnotou nebo porovnáme p-hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p-hodnota $\leq \alpha$** , **zamítáme H_0** . Věk pacienta má vliv na zlepšení míry soběstačnosti po léčbě mozkového infarktu. Pozitivní korelace značí, že u starších pacientů je zlepšení menší (diference jsou vypočítány tak, že nižší hodnoty odpovídají většímu zlepšení).

Výsledky v programu Statistica

Bodový graf



Correlations (02_Biostatist	
Marked correlations are si	
N=407 (Casewise deletion	
Variable	Barthel_index_zmena
Vek	,0989 p = ,046

**Korelační koeficient
a p-hodnota**

Z grafu sice není nikterak výrazná závislost přímo patrná, nicméně je možné, že je přítomen mírně pozitivní trend.

P-hodnota statistické významnosti korelace je $p = 0,046$, což na hladině významnosti 0,05 značí **významný výsledek** a ze získaných dat jsme tedy **prokázali, že věk pacienta má vliv na zlepšení míry soběstačnosti po léčbě mozkového infarktu**. Přesto je potřeba výsledky interpretovat s opatrností, neboť samotná korelace je velmi slabá (0,099).

10.2. Spearmanův korelační koeficient

V situacích, kdy není možné použít Pearsonův korelační koeficient zejména z důvodu porušení jeho předpokladu, máme k dispozici neparametrický Spearmanův korelační koeficient. Spearmanův korelační koeficient je robustní vůči odlehlým hodnotám a odchylkám od normality, neboť pracuje pouze s pořadími pozorovaných hodnot dvou spojitých proměnných. Označujeme jej r_s a lze jej použít i v případě ordinálních proměnných. Výpočet Spearmanova korelačního koeficientu je vlastně výpočtem Pearsonova korelačního koeficientu na pořadích pozorovaných hodnot, tj. Spearmanův korelační koeficient spočítáme podle vzorce Pearsonova korelačního koeficientu zadáním pořadí hodnot místo pozorovaných hodnot a průměrných pořadí místo výběrových průměrů.

Hodnoty Spearmanova korelačního koeficientu r_s se pohybují stejně jako hodnoty Pearsonova korelačního koeficientu r od -1 do 1. Hodnot -1 a 1 nabývá v situacích, že jedna z proměnných je monotónní funkcí druhé proměnné (monotónní znamená rostoucí, klesající, konstantní, nerostoucí nebo neklesající; nemusí být ovšem lineární). Hodnoty Spearmanova korelačního koeficientu blízké nule znamenají, že pořadí hodnot x_i a y_i jsou náhodně zpřeházená a mezi proměnnými není žádný vztah.

Statistickou významnost Spearmanova korelačního koeficientu lze stejně jako u Pearsonova korelačního koeficientu testovat pomocí testové statistiky t ; předpokladem je dostatečná velikost souboru alespoň 30 pozorování.

Příklad: Spearmanův korelační koeficient

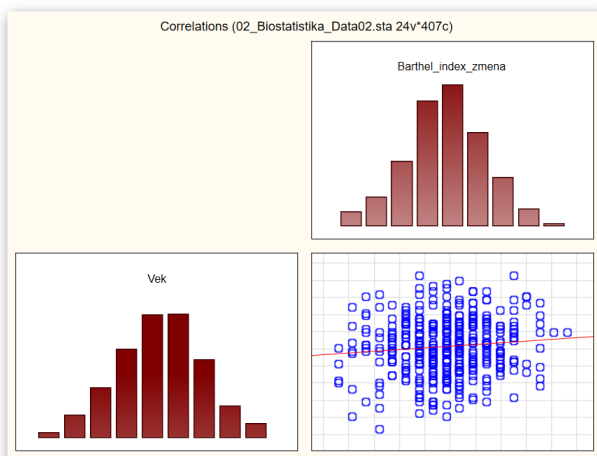
Zadání: U pacientů hospitalizovaných s mozkovým infarktem bylo při propuštění vyhodnoceno zlepšení míry soběstačnosti vyjádřené diferencí hodnot indexu Barthelové. Zjistěte, zda má věk vliv na úspěšnost terapeutické a rehabilitační péče. Jinými slovy, určete, zda věk koreluje s diferencí indexu Barthelové.

Postup výpočtu Spearmanova korelačního koeficientu v programu Statistica (po nemožnosti použít Pearsonův korelační koeficient)

1. Na hladině významnosti $\alpha = 0,05$ testujeme hypotézu
 $H_0: r = 0$ proti $H_A: r \neq 0$.
2. Graficky znázorníme závislost obou proměnných pomocí bodového XY grafu.
3. Vypočítáme hodnotu **korelačního koeficientu r_s** a odpovídající **p-hodnotu**. V menu *Statistics* zvolíme *Nonparametrics*, vybereme *Correlation (Spearman, ...)*. V možnostech *Compute*: vybereme *Detailed report*. Vybereme jednotlivé proměnné, které chceme testovat (*Variables*). V záložce *Advanced* kliknutím na *Scatterplot matrix* získáme grafické znázornění závislosti vybraných proměnných. Poté přes *Spearman rank R* zobrazíme výsledky. $r_s = 0,074$, $p = 0,136$
4. Testovou statistiku porovnáme s kritickou hodnotou nebo porovnáme p-hodnotu s hladinou významnosti $\alpha = 0,05$.
5. Je-li **p-hodnota $> \alpha$** , **nezamítáme H_0** . Neprokázali jsme, že by věk pacienta měl vliv na zlepšení míry soběstačnosti po léčbě mozkového infarktu.

Výsledky v Statistica

Bodový graf



Z grafu není nikterak výrazná závislost patrná, nicméně je možné, že je přítomen mírně pozitivní trend.

Spearman Rank Order Correlations (02_Biostatis)				
MD pairwise deleted				
Marked correlations are significant at p < .05000				
Pair of Variables	Valid N	Spearman R	t(N-2)	p-value
Vek & Barthel_index_zmena	40	0,073994	1,493195	0,136165

Korelační koeficient a p-hodnota

P-hodnota statistické významnosti korelace je $p = 0,136$, což na hladině významnosti 0,05 značí **nevýznamný výsledek** a ze získaných dat jsme tedy **neprokázali, že by věk pacienta měl vliv na zlepšení míry soběstačnosti po léčbě mozkového infarktu.**

Seznam použité literatury

- Anděl, Jiří. Základy matematické statistiky. MatfyzPress, Praha, 2011. 358 pp. ISBN: 978-80-7378-162-0.
- Bencko, Vladimír., Zvárová, Jana, Malý, Marek. Statistické metody v epidemiologii. Nakladatelství Karolinum, Praha, 2003. ISBN: 80-246-0763-8.
- Ferguson, George, Takane, Yoshio. Statistical Analysis In Psychology and Education, 6. vydání. McGraw-Hill, 1989. 587 pp.
- Heiberger, Richard, Holland, Burt, Statistical Analysis and Data Display, 2. vydání. Springer, New York, 2015. 898 pp. ISBN: 978-14-939-2121-8.
- Jarkovský, Jiří. Analýza a management dat pro zdravotnické obory, Analýza klinických dat. In Holčík, Jiří, Komenda, Martin (eds.) a kol. Matematická biologie: e-learningová učebnice. 1. vydání. Brno: Masarykova univerzita, 2015. ISBN 978-80-210-8095-9.
- Lepš, Jan, Šmilauer, Petr. Biostatistika. Jihočeská univerzita v Českých Budějovicích, 2016. 438 pp. ISBN 978-80-7394-587-9.
- Litschmannová, Martina. Úvod do statistiky. Elektronická skripta a doplňkové interaktivní materiály, Ostrava, 2011. 379 pp.
- Nisbet, Robert, Miner, Gary, Elder John. Handbook of Statistical Analysis and Data Mining Applications, 1. vydání. Academic Press, 2009. pp 864. ISBN: 978-00-809-1203-5.
- Pavlík, Tomáš. Biostatistika pro matematickou biologii. In Holčík, Jiří, Komenda, Martin (eds.) a kol. Matematická biologie: e-learningová učebnice. 1. vydání. Brno: Masarykova univerzita, 2015. ISBN 978-80-210-8095-9.
- Pavlík, Tomáš, Dušek Ladislav. Biostatistika. Akademické nakladatelství CERM, s.r.o. Brno, 2012. 131 pp. ISBN 978-80-7204-782-6.
- Ott, Lyman, Longnecker, Michael. An introduction to statistical methods and data analysis, 7. vydání. Cengage Learning, 2015. ISBN: 978-13-052-6947-7.
- Procházka, Bohumír. Biostatistika pro lékaře. Nakladatelství Karolinum, Praha, 2015. 344 pp. ISBN: 978-80-246-2782-3.
- StatSoft 2011. STATISTICA™ Quick Reference. TIBCO Software Inc. 304 pp.
- Zar, Jerrold. Biostatistical analysis, 5. vydání. Person Prentice-Hall, New Jersey, 2010. 960 pp. ISBN: 978-03-216-5686-5.
- Zvára, Karel. Biostatistika. Nakladatelství Karolinum, Praha, 2008. 213 pp. ISBN: 978-08-246-0739-9.
- Zvárová, Jana. Biomedicínská statistika I. – Základy statistiky pro biomedicínské obory. Nakladatelství Karolinum, Praha, 2016. 220 pp. ISBN: 978-80-246-3416-6.

ČÁST DRUHÁ

Cvičení 1. Práce s datovým souborem

V celé sbírce příkladů budeme pracovat s datovým souborem *Zdravotni_setreni*. V tomto datovém souboru se nachází záznamy **100 osob**, které se zúčastnili výběrového šetření o zdraví. V rámci tohoto šetření byly u každé osoby zaznamenány základní charakteristiky (*pohlaví, věk, vzdělání*) a změřeny fyziologické parametry (výška, váha, obvod pasu, obvod boků, systolický tlak, diastolický tlak, cholesterol, glykemie). Každá osoba následně vyplnila krátký dotazník týkající se životního stylu. Všechny 100 osob poté navštívilo kurz zaměřený na zdravý životní styl. Po šesti měsících od absolvování kurzu proběhlo u všech účastníků kontrolní měření, v rámci kterého byly změřeny stejné fyziologické parametry a vyplněn stejný dotazník odrážející životní styl.

V poslední části této kapitoly se věnujeme úvodní přípravě datového souboru, která je nezbytným prvním krokem při každé statistické analýze dat. Zaměříme se na kontrolu a čištění datového souboru a následně na výpočet pomocných proměnných odvozených z původních dat.

V celé kapitole budeme pracovat se souborem *Zdravotni_setreni_pripava_dat.xlsx*. Pro přípravu dat využijeme tabulkový procesor *Microsoft Excel*.

1.1. Data z dotazníkového šetření

Dotazník týkající se životního stylu, který účastníci vyplnili před kurzem a následně 6 měsíců po absolvování kurzu, zahrnoval následující otázky:

1. Kouříte v současnosti nějaké tabákové výrobky?
 - a. Ano
 - b. Ne
2. V kolika dnech v typickém týdnu obvykle pijete alkohol?
 - a. Každý den nebo téměř každý den
 - b. 5 až 6 dní v týdnu
 - c. 3 až 4 dny v týdnu
 - d. 1 až 2 dny v týdnu
 - e. Nikdy
3. V kolika dnech v typickém týdnu se věnujete sportu nejméně 30 minut v kuse?
 - a. Každý den nebo téměř každý den
 - b. 5 až 6 dní v týdnu
 - c. 3 až 4 dny v týdnu
 - d. 1 až 2 dny v týdnu
 - e. Nikdy
4. Jak často konzumujete ovoce?
 - a. Jednou či vícekrát denně
 - b. 4 až 6 dní v týdnu
 - c. 1 až 3 dny v týdnu
 - d. Méně než jedenkrát týdně

- e. Nikdy
- 5. Jak často konzumujete zeleninu?
 - a. Jednou či vícekrát denně
 - b. 4 až 6 dní v týdnu
 - c. 1 až 3 dny v týdnu
 - d. Méně než jedenkrát týdně
 - e. Nikdy
- 6. Jak často pijete slazené nealkoholické nápoje?
 - a. Jednou či vícekrát denně
 - b. 4 až 6 dní v týdnu
 - c. 1 až 3 dny v týdnu
 - d. Méně než jedenkrát týdně
 - e. Nikdy

1.2. Popis proměnných

V datovém souboru *Zdravotni_setreni* se nacházejí následující proměnné:

- **ID** – unikátní identifikátor osoby
- **Pohlavi** – pohlaví, 0 = muž, 1 = žena
- **Vek** – věk osoby v době konání kurzu
- **Vzdelani** – nejvyšší dosažené vzdělání, 1 = základní, 2 = středoškolské, 3 = vysokoškolské
- **Vyska** – výška [cm]
- **Vaha** – váha [kg]
- **Obvod_pasu** – obvod pasu [cm]
- **Obvod_boku** – obvod boků [cm]
- **WHR** – poměr obvodu pasu ku obvodu boku („*Waiste to Hip Ratio*“)
- **WHR_riziko** – riziko odvozené z hodnoty WHR, 1 = nízké (muži do 0,95, ženy do 0,8), 2 = střední (muži 0,95–1,00, ženy 0,8–0,85), 3 = vysoké (muži nad 1,00, ženy nad 0,85)
- **Syst_tlak** – hodnota systolického tlaku [mmHg]
- **Diast_tlak** – hodnota diastolického tlaku [mmHg]
- **Cholesterol** – hodnota cholesterolu v krvi [mmol/l]
- **Glykemie** – hodnota glukózy v krvi na lačno [mmol/l]
- **Koureni** – odpověď na otázku č. 1 z dotazníku, 0 = Ne, 1 = Ano
- **Alkohol** – odpověď na otázku č. 2 z dotazníku, 1 = a, 2 = b, 3 = c, 4 = d, 5 = e
- **Sport** – odpověď na otázku č. 3 z dotazníku, 1 = a, 2 = b, 3 = c, 4 = d, 5 = e
- **Ovoce** – odpověď na otázku č. 4 z dotazníku, 1 = a, 2 = b, 3 = c, 4 = d, 5 = e
- **Zelenina** – odpověď na otázku č. 5 z dotazníku, 1 = a, 2 = b, 3 = c, 4 = d, 5 = e
- **Slazene_napoje** – odpověď na otázku č. 6 z dotazníku, 1 = a, 2 = b, 3 = c, 4 = d, 5 = e
- **Vaha_po** – váha [kg] v kontrolním měření po 6 měsících

- *Obvod_pasu_po* – obvod pasu [cm] v kontrolním měření po 6 měsících
- *Obvod_boku_po* – obvod boků [cm] v kontrolním měření po 6 měsících
- *WHR_po* – poměr obvodu pasu ku obvodu boku v kontrolním měření po 6 měsících
- *WHR_riziko_po* – riziko odvozené z hodnoty WHR v kontrolním měření po 6 měsících
- *Syst_tlak_po* – hodnota systolického tlaku [mmHg] v kontrolním měření po 6 měsících
- *Diast_tlak_po* – hodnota diastolického tlaku [mmHg] v kontrolním měření po 6 měsících
- *Cholesterol_po* – hodnota cholesterolu v krvi [mmol/l] v kontrolním měření po 6 měsících
- *Glykemie_po* – hodnota glukózy v krvi na lačno [mmol/l] v kontrolním měření po 6 měsících
- *Koureni_po* – odpověď na otázku č. 1 z dotazníku po 6 měsících
- *Alkohol_po* – odpověď na otázku č. 2 z dotazníku po 6 měsících
- *Sport_po* – odpověď na otázku č. 3 z dotazníku po 6 měsících
- *Ovoce_po* – odpověď na otázku č. 4 z dotazníku po 6 měsících
- *Zelenina_po* – odpověď na otázku č. 5 z dotazníku po 6 měsících
- *Slazene_napoje_po* – odpověď na otázku č. 6 z dotazníku po 6 měsících

1.3. Práce s daty

Řešené příklady

1. příklad: Seřadíte datový soubor na základě proměnné ID.

Řešení:

1. Klikneme kurzorem myši kamkoli do oblasti tabulky.
2. Karta Domů – Seřadit a filtrovat – Vlastní řazení.
3. Seřadit podle – vybereme proměnou ID; Řazení – vybereme Hodnoty buněk; Pořadí – vybereme Od nejmenšího k největšímu.

2. příklad: Ukotvíte ID respondentů a názvy proměnných ve sloupcích.

Řešení:

1. Klikneme kurzorem myši do buňky B2.
2. Karta Zobrazení – Ukotvit příčky – vybereme možnost Ukotvit příčky.

3. příklad: Zkontrolujte, zda se v datech nenachází duplicitní záznamy na základě proměnné ID (jednoznačný identifikátor osoby). V případě, že naleznete duplicitní záznam, smažte ho.

Řešení:

1. Označíme sloupec ID.

2. Karta Domů – Podmíněné formátování – Pravidla zvýraznění buněk – Duplicitní hodnoty
3. Vybereme možnost duplicitní a vybereme libovolný formát zvýraznění.
4. Zjistíme, že respondent s ID 67 je v datovém souboru dvakrát.
5. Jeden z řádků s ID 67 odstraníme (označíme řádek – pravé tlačítko myši – odstranit).

4. příklad: Zkontrolujte, zda se v datech nenachází chybné hodnoty nebo překlepy. Pokud je to možné, opravte je.

Řešení:

1. Klikneme kurzorem myši kamkoli do oblasti tabulky.
2. Zapneme automatický filtr – Karta Domů – Seřadit a filtrovat – Filtr
3. Pomocí filtrů (případně jiných nástrojů) prověříme všechny sloupce, zda se v nich nenacházejí chybné hodnoty nebo překlepy.
4. Zjistíme, že respondent s ID 74 má v proměnné *pohlavi* hodnotu Z – zřejmě se jedná o ženu – přepíšeme hodnotu Z na 1.
5. Zjistíme, že respondent s ID 22 má v proměnné *vyska* hodnotu 1,71 – zřejmě se jedná o výšku uvedenou v metrech – přepíšeme hodnotu 1,71 na 171.
6. Zjistíme, že respondent s ID 46 má v proměnné *koureni* hodnotu 11 – zřejmě se jedná o překlep – přepíšeme hodnotu 11 na 1.

5. příklad: V datovém souboru vytvořte a vypočtete následující proměnné:

- a) *Zvyseny_cholesterol* – hodnota cholesterolu v krvi (proměnná *cholesterol*) nad 5.
- b) *Diabetes* – hodnota glukózy v krvi na lačno (proměnná *glykemie*) nad 7.
- c) *BMI* – index tělesné hmotnosti; výpočetní vzorec: $BMI = \text{hmotnost [kg]} / (\text{výška [m]})^2$, vypočtené hodnoty zaokrouhlete na 1 desetinné místo.
- d) *Nadvaha* – hodnota BMI nad 25.
- e) *Hypertenze* – hodnota systolického tlaku (proměnná *syst_tlak*) nad 140 nebo hodnota diastolického tlaku (proměnná *diast_tlak*) nad 90.

Řešení:

- a) *Zvyseny_cholesterol*
 1. Vytvoříme nový sloupec za proměnnou *cholesterol* (označíme sloupec *glykemie* – pravé tlačítko myši – Vložit buňky) nazveme ho *zvyseny_cholesterol*.
 2. Do první buňky nového sloupce napíšeme vzorec:
=KDYŽ(M2>5;1;0)
 3. Zkopírujeme vzorec do všech buněk sloupce (dvojklik do levého dolního rohu).
- b) *Diabetes*
 1. Vytvoříme nový sloupec za proměnnou *glykemie* (označíme sloupec *koureni* – pravé tlačítko myši – Vložit buňky) nazveme ho *diabetes*.
 2. Do první buňky nového sloupce napíšeme vzorec:

=KDYŽ(O2>7;1;0)

3. Zkopírujeme vzorec do všech buněk sloupce (dvojklik do levého dolního rohu).

c) *BMI*

1. Vytvoříme nový sloupec za proměnnou *vaha* (označíme sloupec *obvod_pasu* – pravé tlačítko myši – Vložit buňky) nazveme ho *BMI*.

2. Do první buňky nového sloupce napíšeme vzorec:

=ZAOKROUHLIT(F2/(E2/100*E2/100);1)

3. Zkopírujeme vzorec do všech buněk sloupce (dvojklik do levého dolního rohu).

d) *nadvaha*

1. Vytvoříme nový sloupec za proměnnou *BMI* (označíme sloupec *obvod_pasu* – pravé tlačítko myši – Vložit buňky) nazveme ho *nadvaha*.

2. Do první buňky nového sloupce napíšeme vzorec:

=KDYŽ(G2>25;1;0)

3. Zkopírujeme vzorec do všech buněk sloupce (dvojklik do levého dolního rohu).

e) *hypertenze*

1. Vytvoříme nový sloupec za proměnnou *diast_tlak* (označíme sloupec *cholesterol* – pravé tlačítko myši – Vložit buňky) nazveme ho *hypertenze*.

2. Do první buňky nového sloupce napíšeme vzorec:

=KDYŽ(NEBO(M2>140; N2>90);1;0)

3. Zkopírujeme vzorec do všech buněk sloupce (dvojklik do levého dolního rohu).

Příklady k procvičení

1. příklad. V datovém souboru vytvořte a vypočtete stejné proměnné, jako v 5. příkladu pro kontrolní měření:

a. *Zvyseny_cholesterol_po* – hodnota proměnné *cholesterol_po* nad 5.

b. *Diabetes_po* – hodnota proměnné *glykemie_po* nad 7.

c. *BMI_po* – výpočet z proměnné *vaha_po* a *vyska*.

d. *Nadvaha_po* – hodnota proměnné *BMI_po* nad 25.

e. *Hypertenze_po* – hodnota proměnná *syst_tlak_po* nad 140 nebo hodnota proměnné *diast_tlak_po* nad 90.

V případě, že byly všechny kroky čištění a přípravy dat provedeny správně, výsledný datový soubor by měl být identický souboru *Zdravotni_setreni.xlsx*. S tímto datovým souborem budeme pracovat ve všech následujících kapitolách.

Cvičení 2. Popisná statistika

Následující kapitola bude věnována popisné statistice, která nám slouží k prvotnímu náhledu na jednotlivé proměnné datového souboru. V rámci této kapitoly se zaměříme na popisné statistiky a grafickou vizualizaci pro kategoriální (kvalitativní) proměnné a také pro spojitě (kvantitativní) proměnné.

2.1. Typy proměnných

U každé z následujících otázek vyberte správnou odpověď (vždy 0 až všechny správné možnosti):

1. Proměnná udávající hodnotu glukózy v krvi na lačno (proměnná *glykemie*) představuje:
 - a) Spojitou proměnnou
 - b) Kategoriální proměnnou
 - c) Binární proměnnou
 - d) Nominální proměnnou
 - e) Ordinální proměnnou
2. Proměnná určující, zda respondent v současnosti kouří či nekouří tabákové výrobky (proměnná *kourení*) představuje:
 - a) Spojitou proměnnou
 - b) Kategoriální proměnnou
 - c) Binární proměnnou
 - d) Nominální proměnnou
 - e) Ordinální proměnnou
3. Proměnná určující, jaké je maximální dosažené vzdělání respondenta (proměnná *vzdelani*) představuje:
 - a) Spojitou proměnnou
 - b) Kategoriální proměnnou
 - c) Binární proměnnou
 - d) Nominální proměnnou
 - e) Ordinální proměnnou
4. Která z následujících proměnných představuje spojitou proměnnou:
 - a) Frekvence konzumace ovoce (proměnná *ovoce*)
 - b) Diastolický tlak (proměnná *diast_tlak*)
 - c) Přítomnost/nepřítomnost nadváhy (proměnná *nadvaha*)
 - d) Hodnota cholesterolu v kontrolním měření (proměnná *cholesterol_po*)
 - e) BMI – index tělesné hmotnosti (proměnná *BMI*)
5. Která z následujících proměnných představuje nominální proměnnou:
 - a) Frekvence sportování (proměnná *sport*)
 - b) Nejvyšší dosažené vzdělání respondenta (proměnná *vzdelani*)
 - c) Přítomnost/nepřítomnost vysokého krevního tlaku (proměnná *hypertenze*)
 - d) Hodnota glukózy v krvi na lačno v kontrolním měření (proměnná *glykemie_po*)
 - e) Obvod boků (proměnná *obvod_boku*)
6. Která z následujících proměnných představuje ordinální proměnnou:
 - a) Hodnota systolického tlaku (proměnná *syst_tlak*)

- b) Frekvence konzumace alkoholu v kontrolním měření (proměnná *alkohol_po*)
- c) Obvod pasu v kontrolním měření (proměnná *obvod_pasu*)
- d) Proměnná udávající riziko odvozené od poměru obvodu pasu ku obvodu boků (proměnná *WHR_riziko*)
- e) Přítomnost/nepřítomnost zvýšeného cholesterolu (proměnná *zvyseny_cholesterol*)

Řešení: 1. a; 2. b, c; 3. b, e; 4. b, d, e; 5 žádná odpověď není správná, 6. b, d.:

2.2. Popisné statistiky kategoriálních proměnných

Řešené příklady

1. příklad: Pro proměnnou nejvyšší dosažené vzdělání (proměnná *vzdelani*) určete: modus absolutní četnost, absolutní kumulativní četnost, relativní četnost a relativní kumulativní četnost.

Řešení:

Vzdělání	Absolutní četnost	Absolutní kumulativní četnost	Relativní četnost	Relativní kumulativní četnost
Základní	16	16	16 %	16 %
Středoškolské	57	73	57 %	73 %
Vysokoškolské	27	100	27 %	100 %

Modus (nejčastější hodnota datového souboru): středoškolské vzdělání.

Návod pro software Statistica: *Statistics – Basic Statistics – Frequency tables.*

2. příklad: Pro proměnnou udávající frekvenci konzumace alkoholu (proměnná *alkohol*) vykreslete koláčový graf samostatně pro muže a ženy (proměnná *pohlavi*).

Řešení:



- Každý den nebo téměř každý den
- 5 až 6 dní v týdnu
- 3 až 4 dny v týdnu
- 1 až 2 dny v týdnu
- Nikdy

Návod pro software Statistica: *Graphs – 2D – Pie Charts; samostatně pro muže a ženy pomocí By Group dle pohlaví.*

3. příklad: Zjistěte, kolik procent ze všech respondentů mělo váhu 80 kg nebo méně.

Řešení:

V datovém souboru se nachází **76 %** respondentů s váhou 80 kg nebo méně.

Návod pro software Statistica:

- I. *Vytvoření nové proměnné vaha_kat udávající, zda má respondent váhu pod 80 kg či nikoli:*
 - i. *Vytvoření nové proměnné vaha_kat za proměnnou vaha (klikneme pravým tlačítkem ve sloupci BMI – Add variables)*
 - ii. *Klikneme na sloupec vaha_kat – Data – Recode – specifikace 2 kategorií.*
- II. *Určení relativní četnosti pro proměnnou vaha_kat: Statistics – Basic Statistics – Frequency tables.*

Příklady k procvičení

- 1. příklad:** Pro proměnnou udávající riziko odvozené od poměru obvodu pasu ku obvodu boků (proměnná *WHR_riziko*) určete: modus, absolutní četnost, absolutní kumulativní četnost, relativní četnost a relativní kumulativní četnost.
- 2. příklad:** Pro proměnnou udávající frekvenci sportování (proměnná *sport*) vykreslete koláčový graf samostatně pro respondenty s nadváhou a bez nadváhy (proměnná *nadvaha*).
- 3. příklad:** Zjistěte, kolik procent ze všech respondentů mělo hodnotu systolického tlaku (proměnná *syst_tlak*) vyšší než 130?

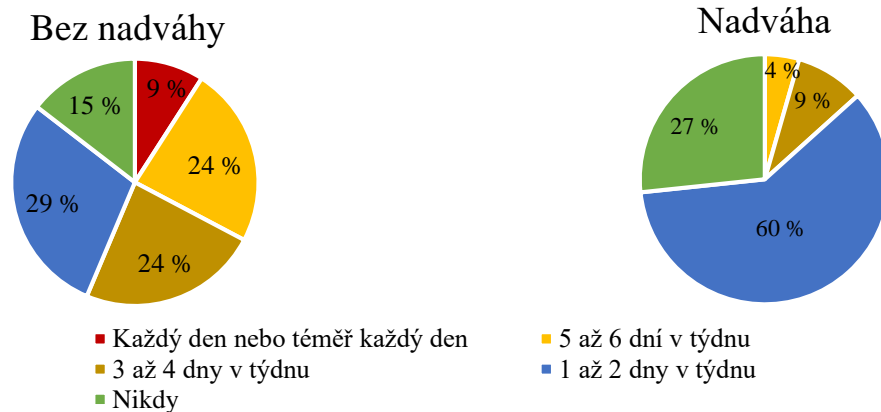
Řešení:

1. příklad:

WHR riziko	Absolutní četnost	Absolutní kumulativní četnost	Relativní četnost	Relativní kumulativní četnost
Nízké	34	34	34 %	34 %
Střední	30	64	30 %	64 %
Vysoké	36	100	36 %	100 %

Modus: vysoké riziko.

2. příklad:



3. příklad: 43 %.

2.3. Popisné statistiky spojitých proměnných

Řešené příklady

1. příklad: Pro proměnnou udávající hodnotu glukózy v krvi na lačno (proměnná *glykemie*) určete: průměr, medián, minimum, maximum, směrodatnou odchylku, rozptyl, dolní a horní kvartil.

Řešení:

Charakteristika	Glykemie
N	100
Průměr	5,8
Medián	4,9
Minimum–maximum	2,8–25,0
Směrodatná odchylka	3,2
Rozptyl	10,0
Dolní – horní kvartil	4,6–5,3

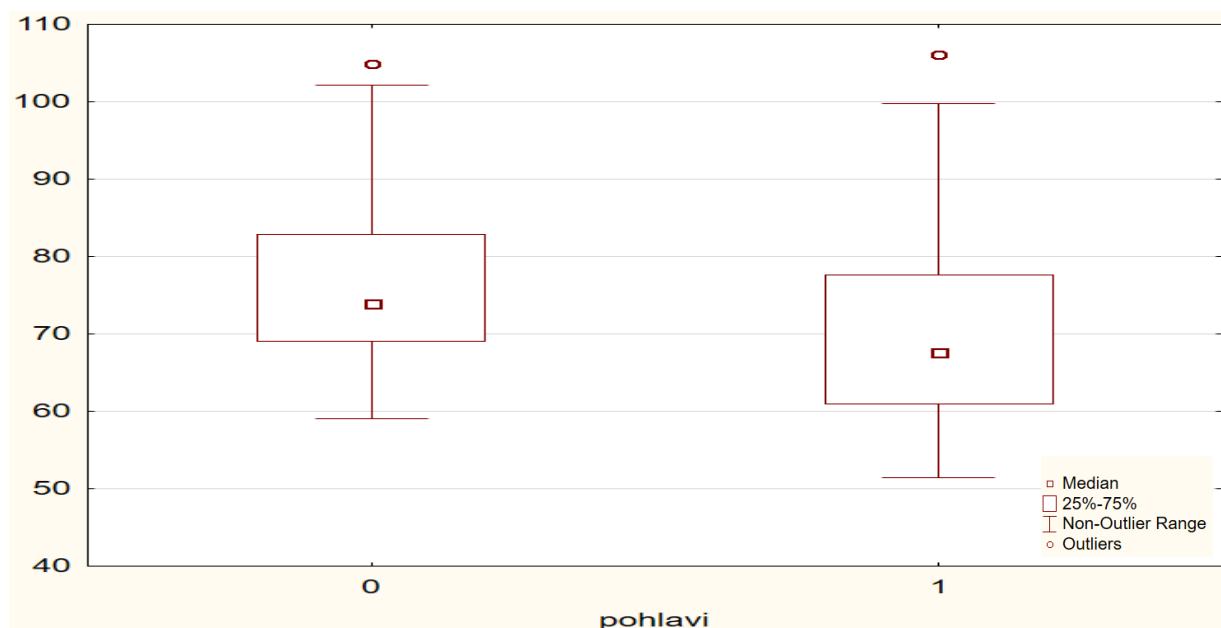
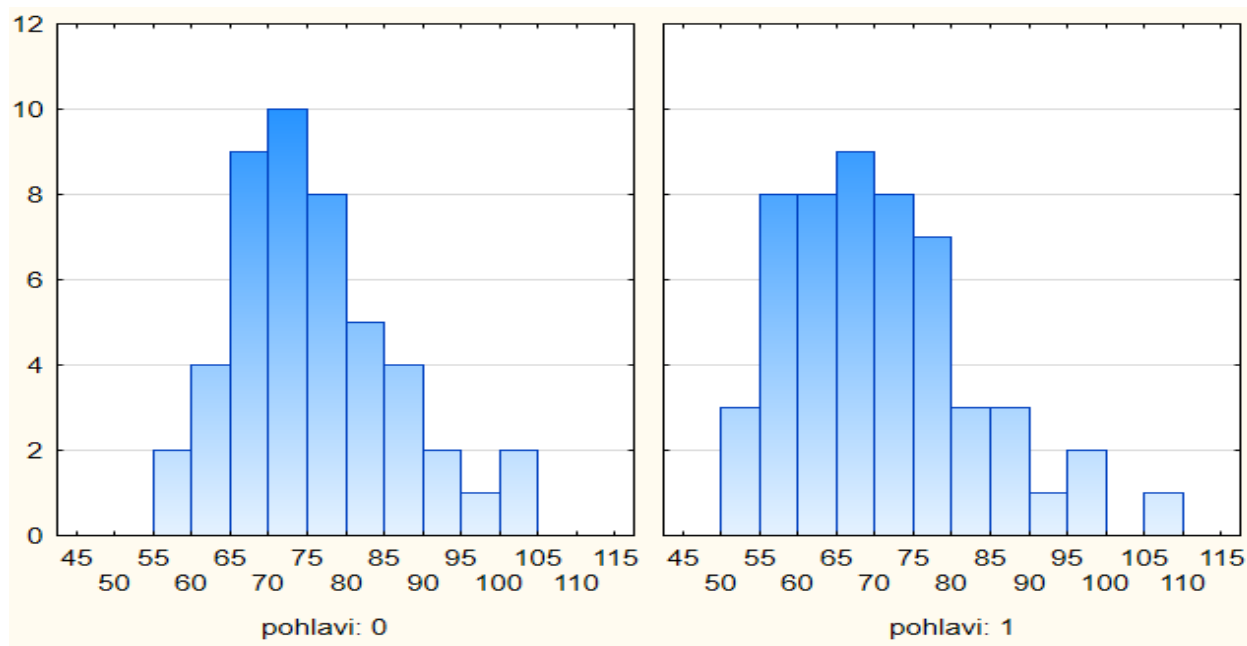


Pozn.: Odlišné hodnoty průměru a mediánu značí asymetrické rozložení – vhodnější charakteristika středu je v tomto případě **MEDIÁN**.

Návod pro software Statistica: *Statistics – Basic Statistics – Descriptive statistics*.

2. příklad: Pro proměnnou udávající váhu respondentů (proměnná *vaha*) vykreslete histogram a krabicový graf samostatně pro muže a ženy.

Řešení:



U mužů i u žen vidíme **asymetrické rozložení dat** (výrazněji u žen; pohlavi = 1). Vhodnější charakteristikou středu by v tomto případě byl opět medián než průměr stejně jako v předchozím příkladě.

Návod pro software Statistica:

- I. *histogram: Graphs – Categorized – Histogram; rozdělení dle pohlaví pomocí X-Category.*
- II. *krabicový graf: Graphs – Box Plots; rozdělení dle pohlaví pomocí Grouping variable v základním výběru proměnných.*

3. příklad: Určete, medián, minimální a maximální hodnotu BMI – indexu tělesné hmotnosti (proměnná *BMI*) pro ženy s věkem nad 50 let a pro ženy s věkem 50 let a méně.

Řešení:

BMI	Ženy 50 let a méně (N = 36)	Ženy nad 50 let (N = 17)
Medián	24,1	25,5
Minimum–maximum	18,4–33,8	20,6–36,4

Návod pro software Statistica:

- I. Vytvoření nové proměnné *vek_kat* udávající, zda má respondent věk nad 50 let či nikoli:
 - i. Vytvoření nové proměnné *vek_kat* za proměnnou *vek* (klikneme pravým tlačítkem ve sloupci *vzdelani* – *Add variables*)
 - ii. Klikneme na sloupec *vek_kat* – *Data* – *Recode* – *specifikace 2 kategorií*.
- II. Selekcce pouze žen – pomocí *Select cases* (*pohlavi = 1*).
- III. Určení mediánu, minima a maxima dle proměnné *vek_kat*: *Statistics* – *Basic Statistics* – *Descriptive statistics*; rozdělení dle *vek_kat* např. pomocí *By Group*.

Příklady k procvičení

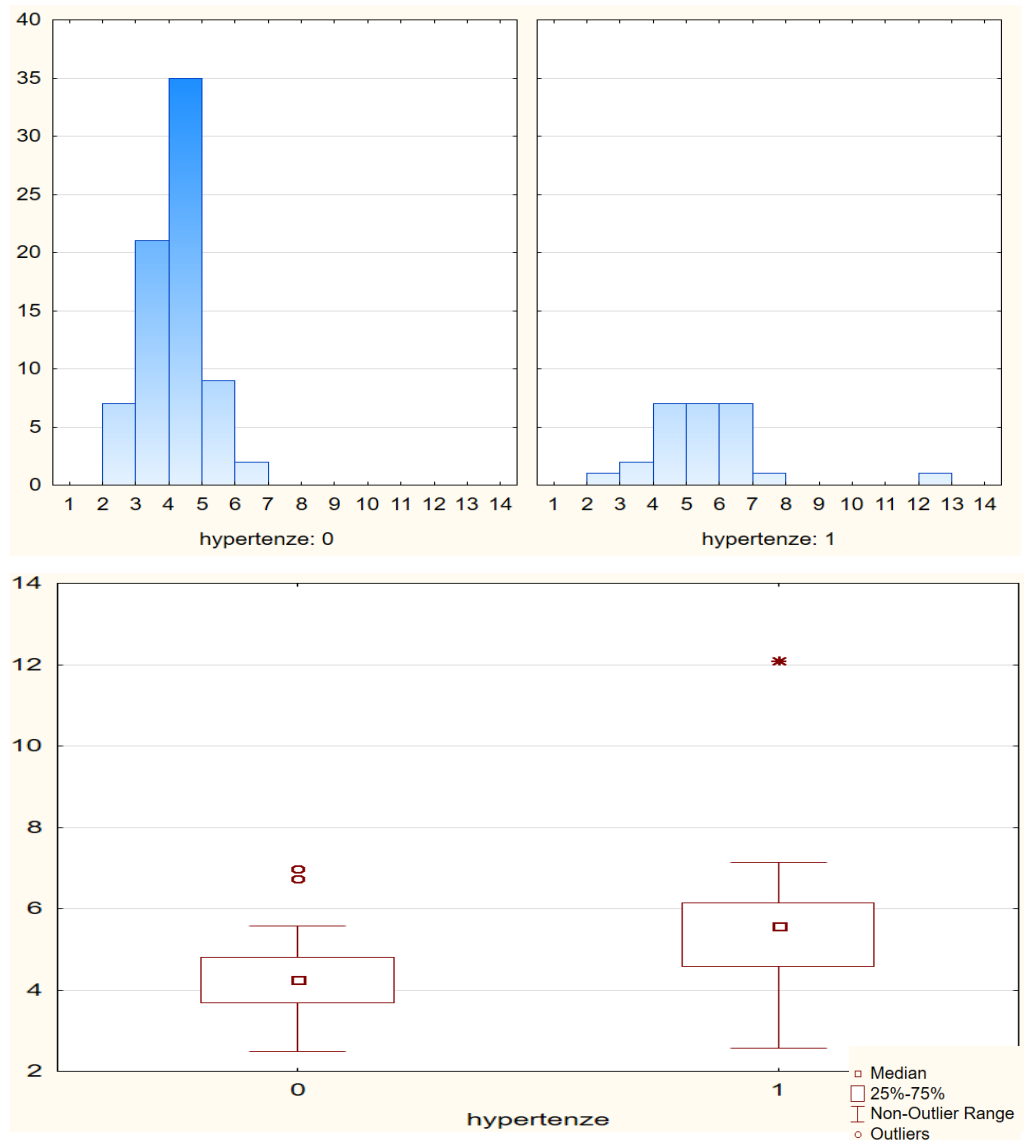
- 1. příklad:** Pro proměnnou udávající hodnotu systolického tlaku (proměnná *syst_tlak*) určete: průměr, medián, minimum, maximum, směrodatnou odchylku, rozptyl, dolní a horní kvartil.
- 2. příklad:** Pro proměnnou udávající hodnotu cholesterolu v krvi (proměnná *cholesterol*) vykreslete histogram a krabicový graf pro respondenty se zvýšeným krevním tlakem a bez zvýšeného krevního tlaku (proměnná *hypertenze*).
- 3. příklad:** Určete, průměr a směrodatnou odchylku výšky (proměnná *vyska*) pro muže a ženy.

Řešení:

1. příklad:

Charakteristika	Systolický tlak
N	100
Průměr	127,2
Medián	125,0
Minimum–maximum	90,0–170,0
Směrodatná odchylka	17,8
Rozptyl	317,8
Dolní– horní kvartil	114,0–140,5

2. příklad:



3. příklad:

Výška	Muži (N = 47)	Ženy (N = 53)
Průměr	180,4	167,7
Směrodatná odchylka	7,6	8,4

Cvičení 3. Rozložení dat

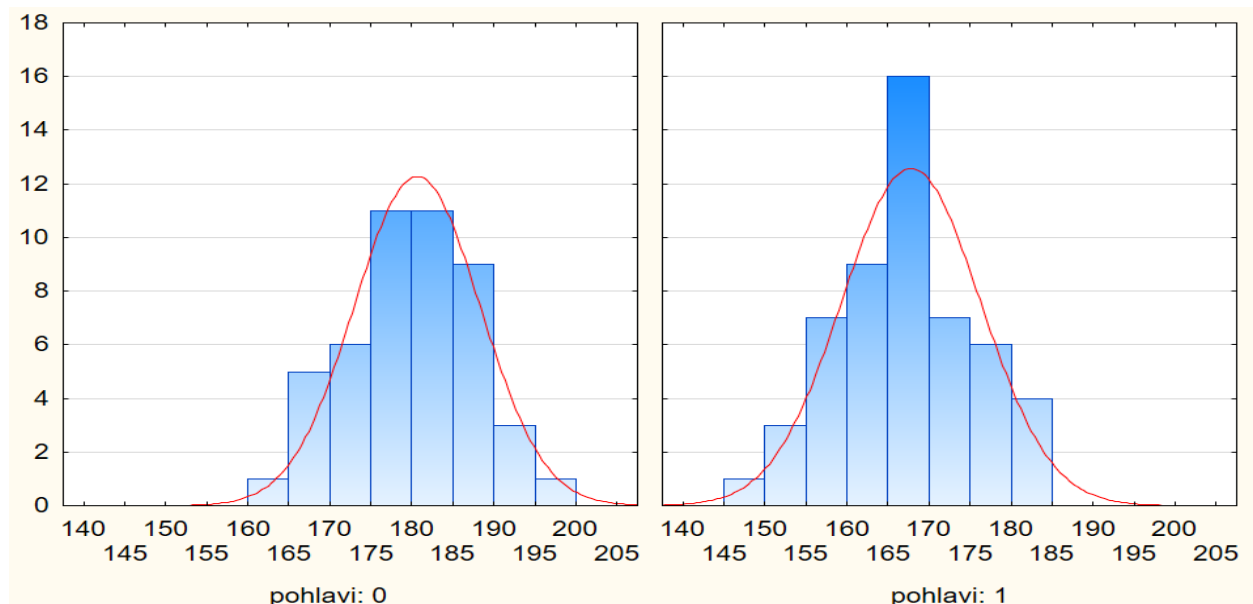
Tato kapitola je věnována modelovým rozložením dat, přičemž největší důraz je kladen na normální (Gaussovo) rozdělení, které je klíčovým předpokladem pro celou řadu statistických testů. V rámci této kapitoly se budeme věnovat jak vizuálnímu ověření normality, tak také ověření normality pomocí statistických testů.

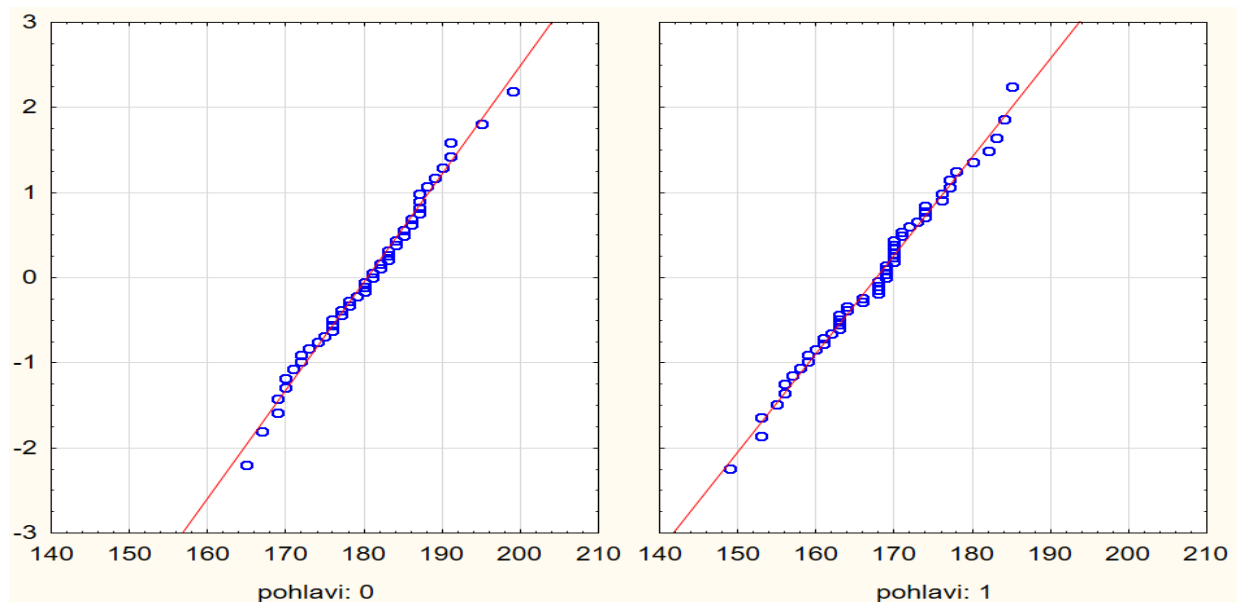
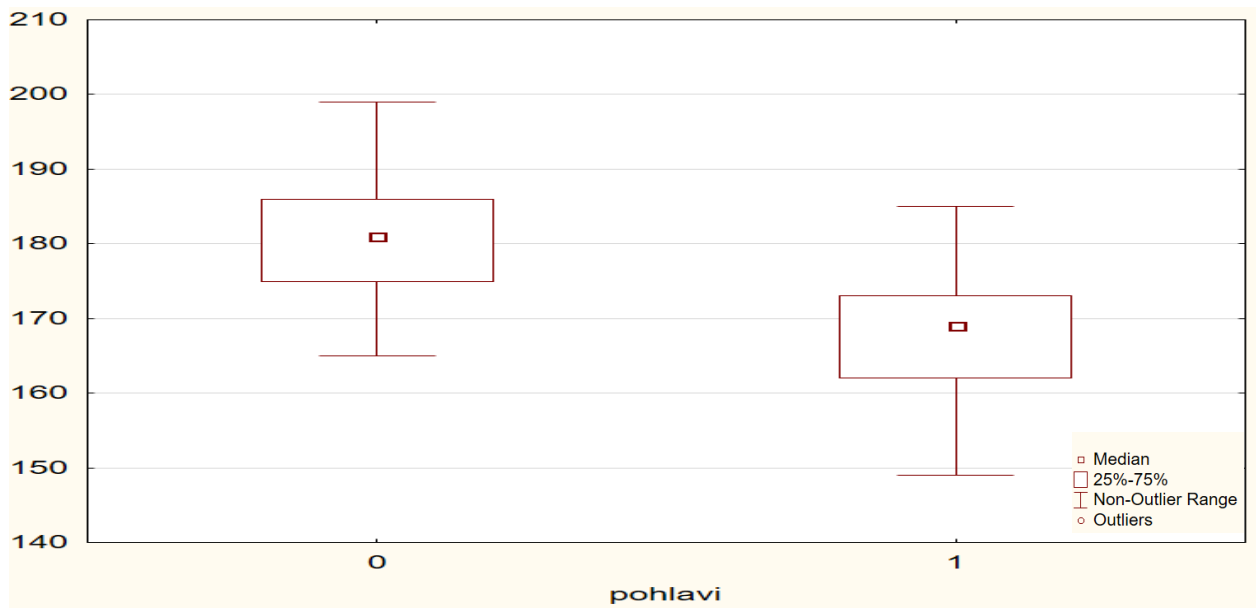
Řešené příklady

1. příklad: Ověřte normalitu proměnné výška samostatně pro muže a ženy pomocí grafických nástrojů (histogram, krabicový graf, N-P graf) a pomocí Shapirova-Wilkova testu.

Řešení:

1. Vizuální ověření normality pomocí grafických nástrojů (histogram, krabicový graf, N-P graf).





Na základě grafických nástrojů **není patrné porušení normality** pro výšku u mužů ani u žen.

Návod pro software Statistica:

- I. *histogram: Graphs – Categorized – Histogram; rozdělení dle pohlaví. pomocí X-Category.*
- II. *krabicový graf: Graphs – Box Plots; rozdělení dle pohlaví. pomocí Grouping variable v základním výběru proměnných.*
- III. *N-P graf: Graphs – Categorized– Normal Probability Plots; rozdělení dle pohlaví. pomocí X-Category.*

2. Ověření normality pomocí Shapiro-Wilkova testu.

H_0 : Výška u mužů/žen se řídí normálním rozdělením.

H_A : Výška u mužů/žen se neřídí normálním rozdělením.

Výška	Muži (N = 47)	Ženy (N = 53)
P-hodnota Shapiro-Wilkova testu	0,918	0,777

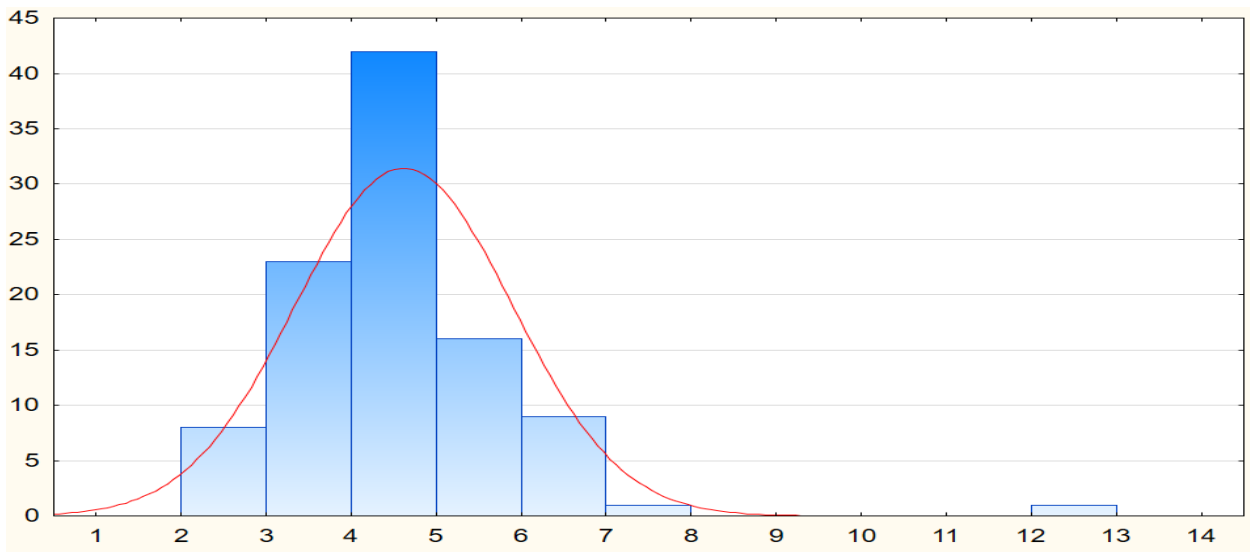
Závěr: Na základě Shapiro-Wilkova testu **nezamítáme nulovou hypotézu** ($p > 0,05$), že rozdělení výšky u mužů/žen se řídí normálním rozdělením.

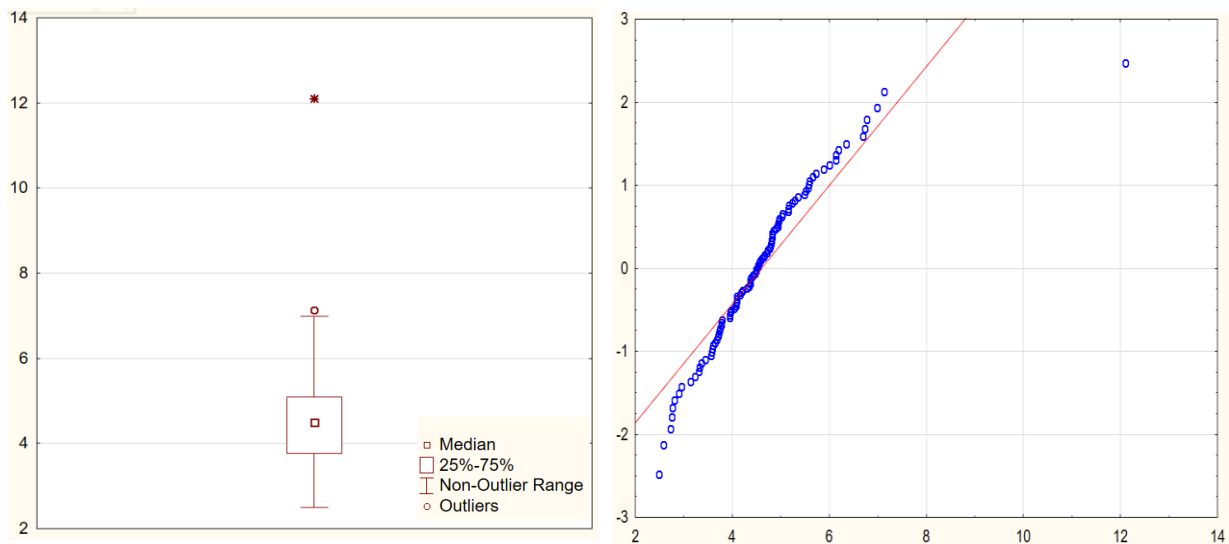
Návod pro software Statistica: Výpočet Shapiro-Wilkova testu např. při tvorbě histogramu: Graphs – Histogram – záložka Advanced – zatrhnutí Shapiro-Wilkova testu.

2. příklad: Ověřte normalitu proměnné *cholesterol* pomocí grafických nástrojů (histogram, krabicový graf, N-P graf) a pomocí Shapiro-Wilkova testu.

Řešení:

1. *Vizuální ověření normality pomocí grafických nástrojů (histogram, krabicový graf, N-P graf).*





Na základě grafických nástrojů je patrná přítomnost **odlehle hodnoty** v datovém souboru.

2. Ověření normality pomocí Shapiro-Wilkova testu.

H_0 : Proměnná cholesterol se řídí normálním rozdělením.

H_A : Proměnná cholesterol se neřídí normálním rozdělením.

P-hodnota Shapiro-Wilkova testu

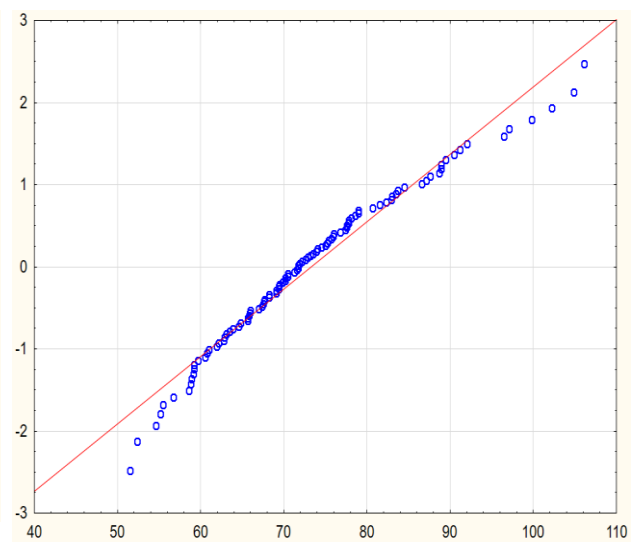
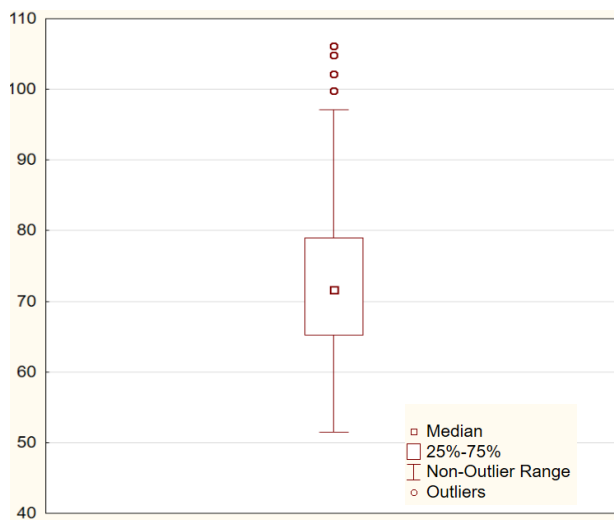
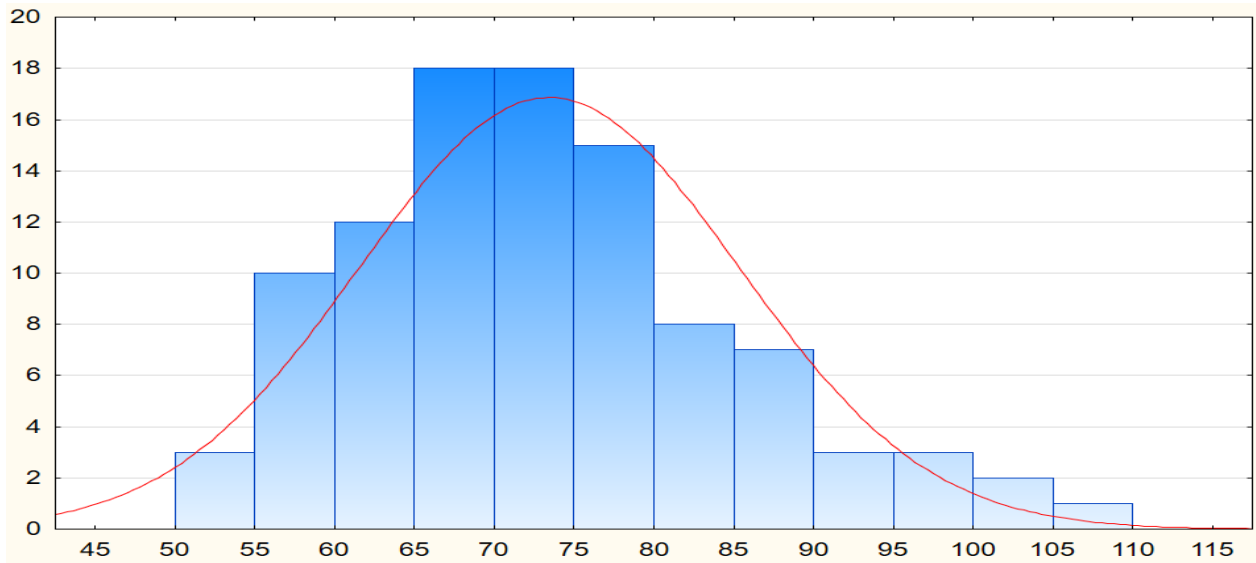
< 0,001

Závěr: Na základě Shapiro-Wilkova testu **zamítáme nulovou hypotézu** ($p < 0,05$), rozdělení proměnné cholesterol se neřídí normálním rozdělením. V důsledku přítomnosti odlehle hodnoty je zamítnut předpoklad normality. Odlehlá hodnota představuje subjekt s ID 31, u kterého byla naměřena hodnota cholesterolu 12,10 mmol/l. Vzhledem k tomu, že u tohoto subjektu je hodnota cholesterolu v kontrolním měření rovněž velmi vysoká (10,84 mmol/l), nebude se v tomto případě s velkou pravděpodobností jednat o chybný údaj v datech, ale o skutečnou, reálně naměřenou hodnotu. Není tedy vhodné tuto hodnotu z dat odstraňovat. Pro následnou analýzu však musíme mít na paměti, že v důsledku přítomnosti odlehle hodnoty není splněn předpoklad normality, a nebude tedy možné pro testování hypotéz využít parametrických testů.

3. příklad: Ověřte normalitu proměnné *vaha* pomocí grafických nástrojů (histogram, krabicový graf, N-P graf) a pomocí Shapiro-Wilkova testu.

Řešení:

1. Vizuální ověření normality pomocí grafických nástrojů (histogram, krabicový graf, N-P graf).



Na základě grafických nástrojů je patrné **asymetrické rozložení dat** (kladné zešikmení).

2. Ověření normality pomocí Shapiro-Wilkova testu.

H_0 : Váha se řídí normálním rozdělením.

H_A : Váha se neřídí normálním rozdělením.

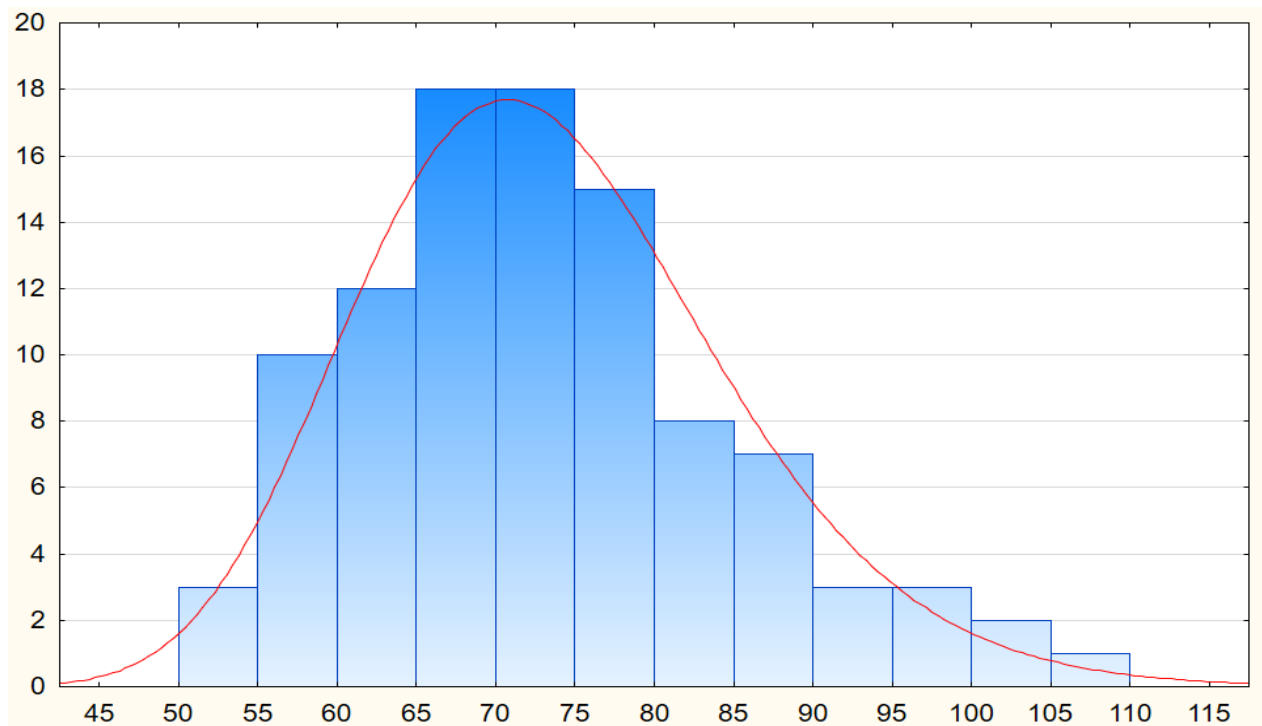
P-hodnota Shapiro-Wilkova testu

0,015

Závěr: Na základě Shapiro-Wilkova testu **zamítáme nulovou hypotézu** ($p < 0,05$), rozdělení váhy se neřídí normálním rozdělením.

3. Proložení log-normálního rozdělení.

Vzhledem ke kladnému zešikmení dat zkusíme proložit histogramem **log-normální rozdělení**.



Návod pro software Statistica: Graphs – Histogram – záložka Advanced – zatrhneme Lognormal v oblasti Fit type.

Z obrázku je patrné, že log-normální rozdělání odpovídá pozorovaným datům lépe než normální rozdělání (viz předchozí obrázek).

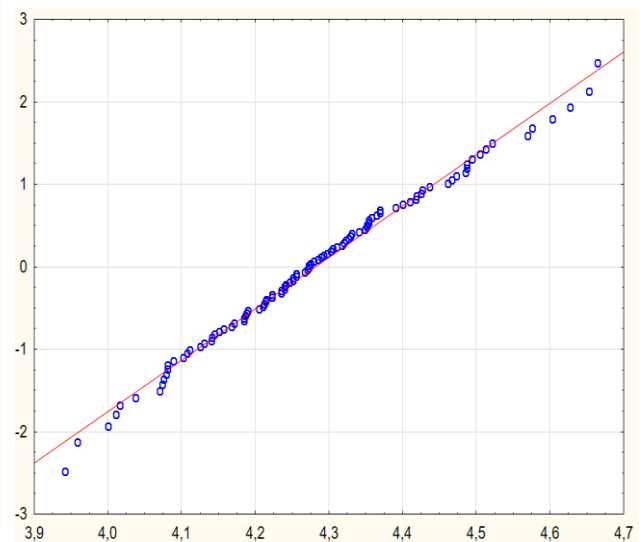
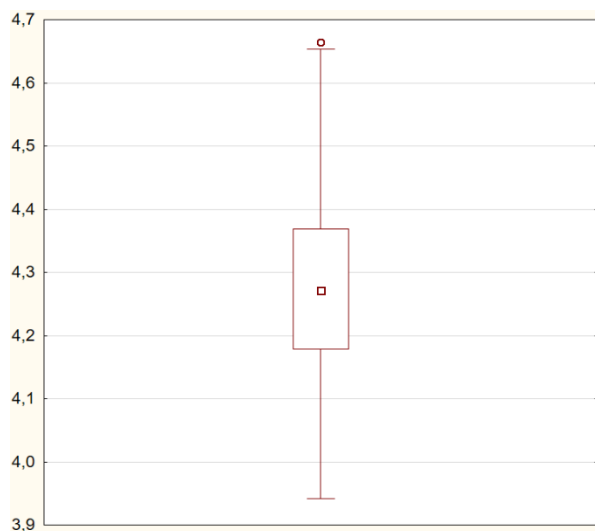
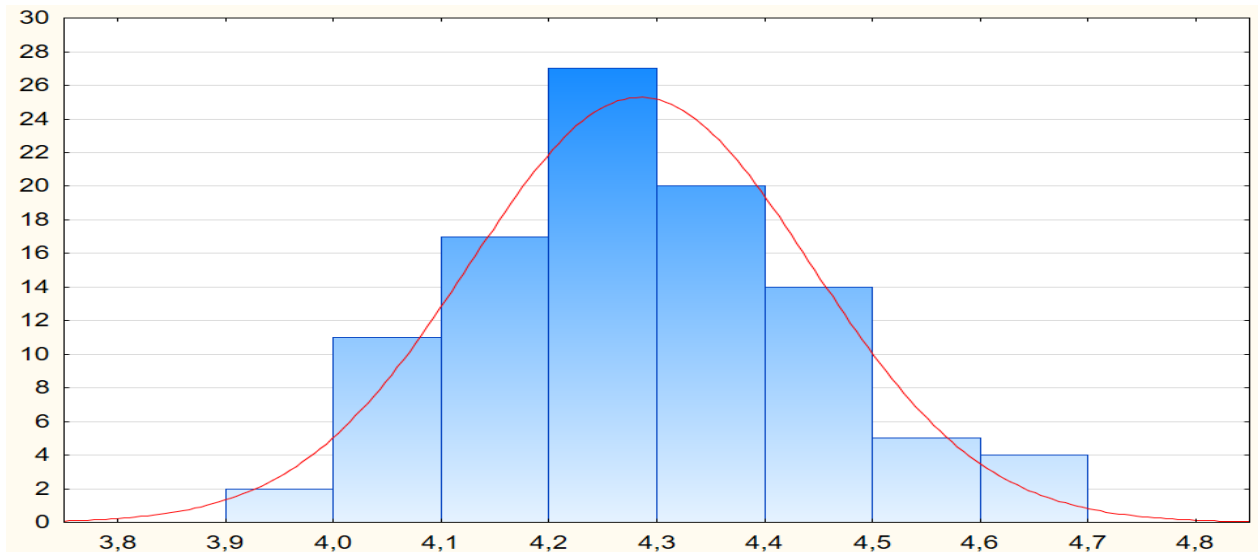
4. Transformace proměnné váha pomocí logaritmu.

Vzhledem k tomu, že log-normální rozdělání odpovídá pozorovaným datům lépe než normální rozdělání, zkusíme provést transformaci váhy pomocí **přirozeného logaritmu** a pro transformovanou proměnnou opětovně otestujeme normalitu.

Návod pro software Statistica:

- I. Vytvoření nové proměnné vaha_log za proměnnou vaha (klikneme pravým tlačítkem ve sloupci BMI – Add variables)*
- II. Do kolonky Long name napíšeme vzorec: =log(vaha).*

5. Vizuální ověření normality transformované proměnné (*vaha_log*) pomocí grafických nástrojů.



Na základě grafických nástrojů **není patrné porušení normality** pro transformovanou proměnnou.

6. Ověření normality transformované proměnné (*vaha_log*) pomocí Shapiro-Wilkova testu.

H_0 : Transformovaná proměnná (*vaha_log*) se řídí normálním rozdělením.

H_A : Transformovaná proměnná (*vaha_log*) se neřídí normálním rozdělením.

P-hodnota Shapiro-Wilkova testu	0,635
---------------------------------	-------

Závěr: Na základě Shapiro-Wilkova testu **nezamítáme nulovou hypotézu** ($p > 0,05$), že rozdělení transformované proměnné (*vaha_log*) se řídí normálním rozdělením.

Pozn.: Vidíme, že transformace pomocí přirozeného logaritmu z hlediska normality pomohla, na druhou stranu po transformaci je **obtížnější interpretace** jednotek pro tuto proměnnou.

Příklady k procvičení

- 1. příklad:** Ověřte normalitu proměnné *obvod_pasu* zvlášť pro muže a ženy (proměnná *pohlavi*) pomocí grafických nástrojů (histogram – zde je pro lepší vizualizaci vhodné nastavení hranic 50–120 s krokem 10, krabicový graf, N-P graf) a pomocí Shapirova-Wilkova testu.
- 2. příklad:** Ověřte normalitu diastolického tlaku v kontrolním měření (proměnná *diast_tlak_po*) u žen pomocí grafických nástrojů (histogram, krabicový graf, N-P graf) a pomocí Shapirova-Wilkova testu. V případě nesplnění normality vyzkoušejte logaritmickou transformaci.

Řešení:

1. příklad: Na základě diagnostických grafů není patrné porušení normality pro obvod pasu u mužů ani u žen. Na základě Shapirova-Wilkova testu **nezamítáme nulovou hypotézu**, že rozdělení obvodu pasu u mužů/žen se řídí normálním rozdělením (p-hodnota muži: 0,752; p-hodnota ženy: 0,970).

2. příklad: Na základě diagnostických grafů je patrné lehké kladné zešikmení pro proměnnou diastolický tlak v kontrolním měření u žen. Na základě Shapirova-Wilkova testu **zamítáme nulovou hypotézu**, že rozdělení diastolického tlaku v kontrolním měření u žen se řídí normálním rozdělením (p-hodnota: 0,036). Po logaritmické transformaci: Na základě Shapirova-Wilkova testu **nezamítáme nulovou hypotézu**, že rozdělení logaritmicky transformovaného diastolického tlaku v kontrolním měření u žen se řídí normálním rozdělením (p-hodnota: 0,299).

Cvičení 4. Parametrické testy pro kvantitativní proměnné

Následující kapitola bude věnována parametrickým testům. V rámci této kapitoly komplexně pokryjeme všechny typy běžně používaných parametrických testů, konkrétně se budeme věnovat jednovýběrovému t-testu, dvouvýběrovému t-testu, párovému t-testu a analýze rozptylu. Kromě samotného testování hypotéz se u každého příkladu zaměříme rovněž na nezbytnou a často opomínanou část, a to ověření předpokladů vstupních dat.

4.1. Jednovýběrový t-test

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda je systolický tlak respondentů (proměnná *syst_tlak*) statisticky významně odlišný od fyziologické hodnoty 120 mmHg.

Řešení:

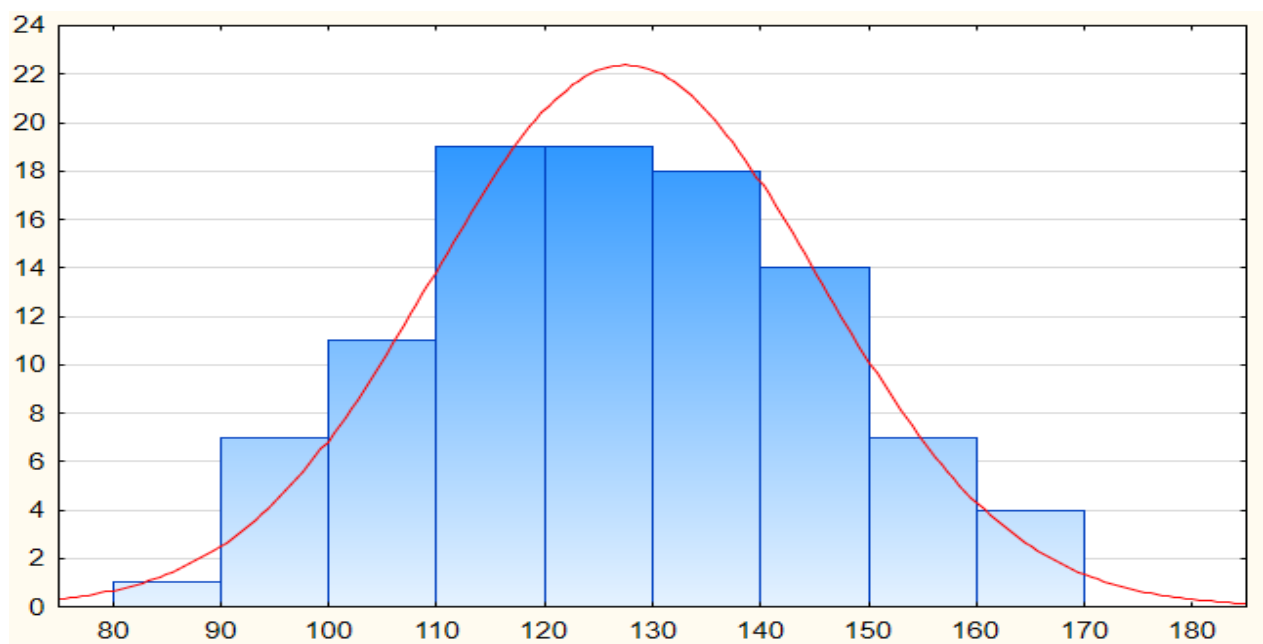
1. Stanovení nulové a alternativní hypotézy.

H_0 : Systolický tlak respondentů se **rovná** hodnotě 120 mmHg.

H_A : Systolický tlak respondentů se **nerovná** hodnotě 120 mmHg.

2. Ověření předpokladů parametrického testu.

Pro jednovýběrový t-test je nutné ověřit **předpoklad normality** náhodného výběru (v našem případě normální rozdělení systolického tlaku). Tento předpoklad ověříme např. pomocí histogramu a Shapiro-Wilkova testu.



Na základě histogramu a Shapirova-Wilkova testu nezamítáme hypotézu, že rozdělení systolického tlaku se řídí normálním rozdělením (p-hodnota: 0,409) – **předpoklad normality je splněn.**

3. Testování nulové hypotézy (H_0 : Systolický tlak respondentů se rovná hodnotě 120 mmHg) pomocí jednovýběrového t-testu.

Vzhledem k tomu, že předpoklad parametrického jednovýběrového t-testu je splněn, můžeme tento test využít k testování nulové hypotézy.

P-hodnota jednovýběrového t-testu	< 0,001
-----------------------------------	---------

Závěr: Na základě jednovýběrového t-testu **zamítáme nulovou hypotézu** ($p < 0,05$), systolický tlak respondentů se nerovná hodnotě 120 mmHg. Průměrný systolický tlak u respondentů je 127,2 mmHg.

*Pozn.: V případě, že by předpoklad normality nebyl splněn, nemůžeme k testování využít jednovýběrový t-test, ale musíme využít jeho **neparametrickou alternativu** – **jednovýběrový Wilcoxonův** nebo **znaménkový test** (viz kapitola 6.1).*

Návod pro software Statistica:

- I. *Statistics – Basic statistics – t-test, single sample.*
- II. *Do políčka test all means against napíšeme hodnotu 120 (testovaná fyziologická hodnota systolického tlaku).*

Příklady k procvičení

1. **příklad:** Na hladině významnosti 0,05 zjistěte, zda je diastolický tlak (proměnná *diast_tlak*) u žen statisticky významně odlišný od fyziologické hodnoty 80 mmHg.
2. **příklad:** Na hladině významnosti 0,05 zjistěte, zda je glukóza v krvi na lačno (proměnná *glykemie*) u respondentů, kteří netrpí cukrovkou (proměnná *diabetes*), statisticky významně odlišná od fyziologické hodnoty 4,8 mmol/l.

Řešení:

1. příklad: Předpoklad normality je splněn (p-hodnota S-W testu: 0,316). Na základě jednovýběrového t-testu **nezamítáme nulovou hypotézu**, že diastolický tlak u žen se rovná hodnotě 80 mmHg (p-hodnota: 0,052). Průměrný diastolický tlak u žen je 82,3 mmHg.

2. příklad: Předpoklad normality není splněn (p-hodnota S-W testu: < 0,001). K testování hypotézy není možné využít jednovýběrový t-test, ale je nutné

využít neparametrické alternativy – jednovýběrový Wilcoxonův nebo znaménkový test (viz kapitola 6.1).

4.2. Dvouvýběrový t-test

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda se systolický tlak (proměnná *syst_tlak*) statisticky významně liší u respondentů se zvýšeným cholesterolem a bez zvýšeného cholesterolu (proměnná *zvyseny_cholesterol*).

Řešení:

1. Stanovení nulové a alternativní hypotézy.

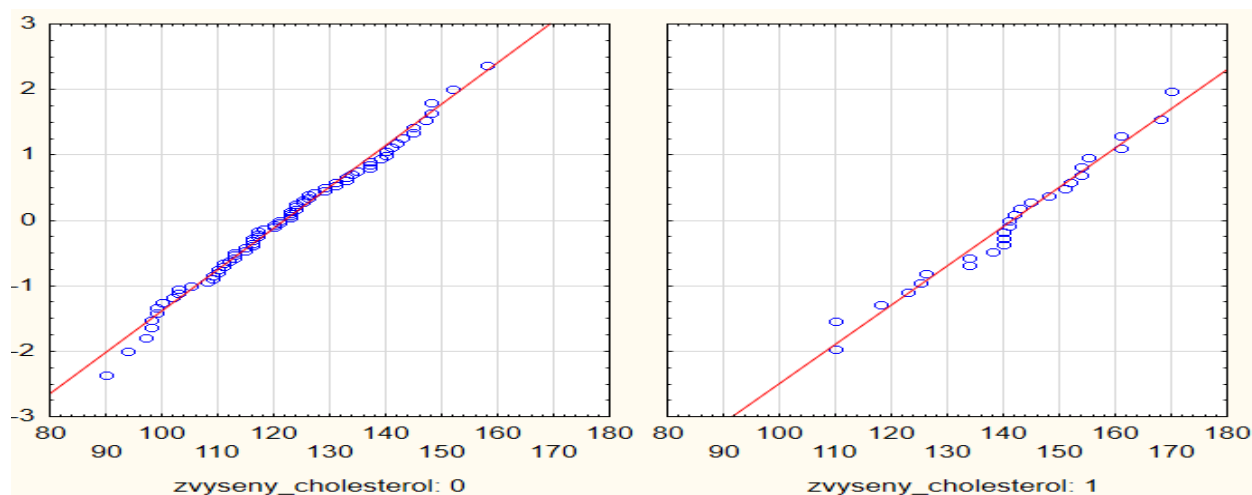
H_0 : Systolický tlak se u respondentů se zvýšeným a bez zvýšeného cholesterolu **neliší**.

H_A : Systolický tlak se u respondentů se zvýšeným a bez zvýšeného cholesterolu **liší**.

2. Ověření předpokladů parametrického testu.

Dvouvýběrový t-test má 2 předpoklady, které je nutné před jeho využitím ověřit. Prvním předpokladem je **normalita** náhodného výběru v **rámci testovaných skupin** (v našem případě normální rozdělení systolického tlaku ve skupině respondentů se zvýšeným a bez zvýšeného cholesterolu). Druhým předpokladem je **homogenita rozptylů v rámci skupin** (v našem případě to znamená, že rozptyl hodnot systolického tlaku ve skupině respondentů se zvýšeným cholesterolem by měl být přibližně shodný jako u respondentů bez zvýšeného cholesterolu).

První předpoklad můžeme ověřit např. pomocí N-P grafu a Shapirova-Wilkova testu.



Na základě N-P grafu a Shapirova-Wilkova testu nezamítáme hypotézu, že rozdělení systolického tlaku u respondentů se zvýšeným cholesterolem a bez zvýšeného cholesterolu se řídí normálním rozdělením (p-hodnota zvýšený cholesterol: 0,5931; p-hodnota bez zvýšeného cholesterolu: 0,611) – **předpoklad normality v rámci skupin je splněn**.

Pozn.: Všimněme si, že normalita celého náhodného výběru předpokladem není, a tedy není nutné, aby náš datový soubor splňoval předpoklad normality v rámci celku ale pouze v rámci uvažovaných skupin. V mnohých případech normalita v celém datovém souboru ani není možná, např. když je rozdíl mezi skupinami velmi velký, rozdělení v celém výběru bude dvouvrcholové.

Druhý předpoklad ověříme pomocí F-testu.

P-hodnota F-testu	0,835
-------------------	-------

Na základě provedeného F-testu nezamítáme hypotézu, že rozptyl systolického tlaku u respondentů se zvýšeným cholesterolem a bez zvýšeného cholesterolu se neliší – **předpoklad homogenity rozptylů v rámci skupin je splněn.**

Návod pro software Statistica: Hodnotu testové statistiky a p-hodnotu pro F-test nalezneme v rámci výstupu dvouvýběrového t-testu (poslední 2 sloupce).

3. Testování nulové hypotézy (H_0 : Systolický tlak u respondentů se zvýšeným cholesterolem a bez zvýšeného cholesterolu se neliší) pomocí dvouvýběrového t-testu.

Vzhledem k tomu, že oba předpoklady parametrického dvouvýběrového t-testu jsou splněny, můžeme tento test využít k testování nulové hypotézy.

P-hodnota dvouvýběrového t-testu	< 0,001
----------------------------------	---------

Závěr: Na základě dvouvýběrového t-testu **zamítáme nulovou hypotézu** ($p < 0,05$), systolický tlak respondentů se zvýšeným cholesterolem a bez zvýšeného cholesterolu se liší. Průměrný systolický tlak u respondentů se zvýšeným cholesterolem je 141,6 mmHg, u respondentů bez zvýšeného cholesterolu 121,9 mmHg.

Pozn.: V případě, že by kterýkoli z předpokladů (normalita v rámci skupin, homogenita rozptylů) nebyl splněn, nemůžeme k testování využít dvouvýběrový t-test, ale musíme využít jeho neparametrickou alternativu – Mannův-Whiteyův test (viz kapitola 6.2).

Návod pro software Statistica: Statistics – Basic statistics – t-test, independent, by groups.

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 zjistěte, zda se systolický tlak (proměnná *syst_tlak*) statisticky významně liší u respondentů s nadváhou a bez nadváhy (proměnná *nadvaha*).
- 2. příklad:** Na hladině významnosti 0,05 zjistěte, zda se WHR – poměr pasu ku bokům (proměnná *WHR*) statisticky významně liší u respondentů, kteří konzumují zeleninu alespoň 4krát týdně a respondentů, kteří konzumují zeleninu méně než 4krát týdně (kategorizace proměnné *zelenina* – 1 + 2, vs. 3 + 4 + 5).

Návod pro kategorizaci proměnné zelenina v softwaru Statistica:

- i. *Vytvoření nové proměnné zelenina_kat za proměnnou zelenina (klikneme pravým tlačítkem ve sloupci slazene_napoj – Add variables)*
- ii. *Klikneme na sloupec zelenina_kat – Data – Recode – specifikace 2 kategorií.*

3. příklad: Na hladině významnosti 0,05 zjistíte, zda se obvod pasu (proměnná *obvod_pasu*) statisticky významně liší u respondentů, kteří pijí slazené nápoje alespoň 4krát týdně a respondentů, kteří pijí slazené nápoje méně než 4krát týdně (proměnná *slazene_napoj*).

4. příklad: Na hladině významnosti 0,05 zjistíte, zda se cholesterol (proměnná *cholesterol*) statisticky významně liší u respondentů s vysokým krevním tlakem a bez vysokého krevního tlaku (proměnná *hypertenze*).

Řešení:

1. příklad: Předpoklad normality v rámci skupin je splněn (p-hodnota S-W testu: nadváha: 0,471; bez nadváhy: 0,151). Předpoklad homogenity rozptylů v rámci skupin je splněn (p-hodnota F-testu: 0,182). Na základě dvouvýběrového t-testu **zamítáme nulovou hypotézu**, systolický tlak u respondentů s nadváhou a bez nadváhy se liší (p-hodnota < 0,001). Průměrný systolický tlak u respondentů s nadváhou je 135,0 mmHg, u respondentů bez nadváhy 120,9 mmHg.

2. příklad: Předpoklad normality v rámci skupin je splněn (p-hodnota S-W testu: zelenina alespoň 4krát týdně: 0,085; p-hodnota zelenina méně než 4krát týdně: 0,305). Předpoklad homogenity rozptylů v rámci skupin je splněn (p-hodnota F-testu: 0,129). Na základě dvouvýběrového t-testu **nezamítáme nulovou hypotézu**, že WHR se neliší u respondentů, kteří konzumují zeleninu alespoň 4krát týdně a respondentů, kteří konzumují zeleninu méně než 4krát týdně (p-hodnota: 0,142). Průměrná hodnota WHR u respondentů, kteří konzumují zeleninu alespoň 4krát týdně je 0,87, u respondentů, kteří konzumují zeleninu méně než 4krát týdně 0,91.

3. příklad: Předpoklad normality v rámci skupin je splněn (p-hodnota S-W testu: slazené nápoje alespoň 4krát týdně: 0,927; slazené nápoje méně než 4krát týdně: 0,899). Předpoklad homogenity rozptylů v rámci skupin je splněn (p-hodnota F-testu: 0,921). Na základě dvouvýběrového t-testu **zamítáme nulovou hypotézu**, obvod pasu u respondentů, kteří pijí slazené nápoje alespoň 4krát týdně a respondentů, kteří pijí slazené nápoje méně než 4krát týdně se liší (p-hodnota < 0,001). Průměrný obvod pasu u respondentů, kteří pijí slazené nápoje alespoň 4krát týdně je 88,7 cm, u respondentů, kteří pijí slazené nápoje méně než 4krát týdně je 79,6 cm.

4. příklad: Předpoklad normality v rámci skupin není splněn (p-hodnota S-W testu: hypertenze: $< 0,001$; bez hypertenze: $0,081$). K testování hypotézy není možné využít dvouvýběrový t-test, ale je nutné využít neparametrickou alternativu – Mann-Whitneyův test (viz kapitola 6.2).

4.3. Párový t-test

Řešený příklad

Na hladině významnosti $0,05$ zjistěte, zda je váha respondentů statisticky významně odlišná před absolvováním kurzu zaměřeného na zdravý životní styl (proměnná *vaha*) a po absolvování tohoto kurzu (proměnná *vaha_po*).

Řešení:

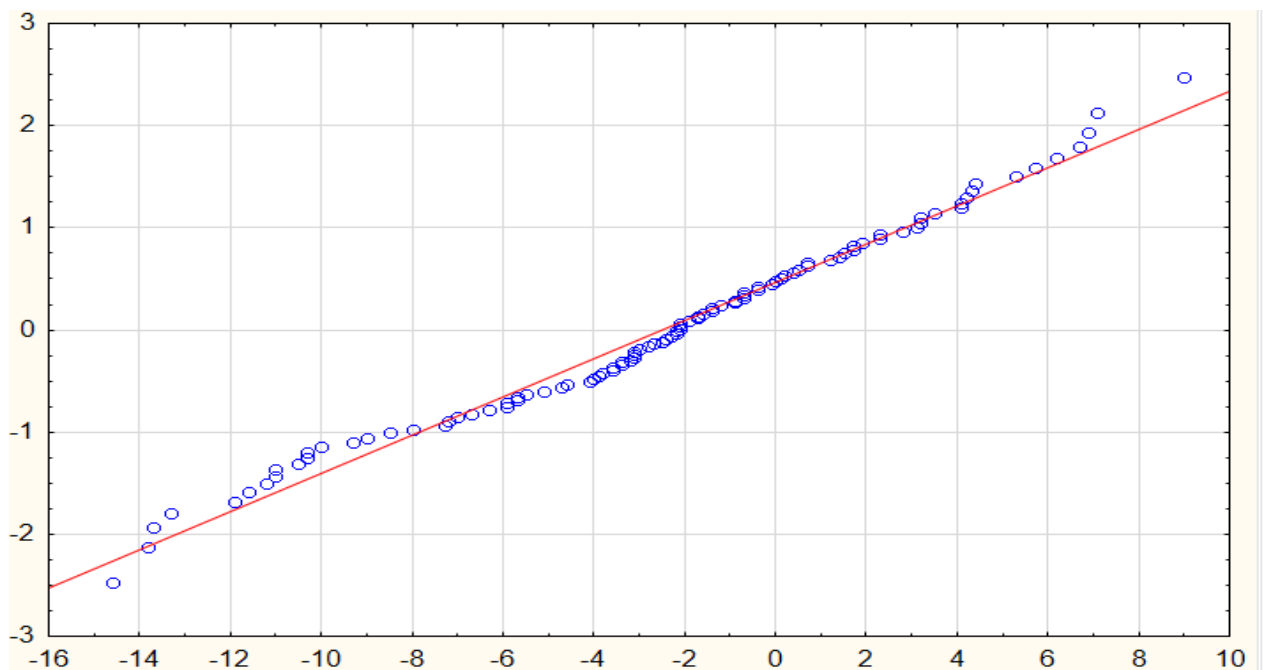
1. Stanovení nulové a alternativní hypotézy.

H_0 : Váha před absolvováním kurzu a po absolvování kurzu se **neliší**.

H_A : Váha před absolvováním kurzu a po absolvování kurzu se **liší**.

2. Ověření předpokladů parametrického testu.

Pro párový t-test je nutné ověřit **předpoklad normality difference** (v našem případě rozložení hodnot rozdílu váhy před kurzem a váhy po kurzu). Tento předpoklad ověříme např. pomocí N-P grafu a Shapirova-Wilkova testu.



Na základě N-P grafu a Shapirova-Wilkova testu nezamítáme hypotézu, že rozdělení difference váhy před a po kurzu se řídí normálním rozdělením (p-hodnota: $0,200$) – **předpoklad normality difference je splněn**.

Návod pro software Statistica: Abychom mohly ověřit předpoklad normality difference, musíme si nejprve vytvořit novou proměnnou udávající rozdíl mezi váhou před kurzem a váhou po kurzu.

- I. Vytvoření nové proměnné *vaha_rozdil* za proměnnou *vaha_po* (klikneme pravým tlačítkem ve sloupci *BMI_po* – Add variables)
- II. Do kolonky Long name napíšeme vzorec: = *vaha_po* – *vaha*.

3. Testování nulové hypotézy (H_0 : Váha před absolvováním kurzu a po absolvování kurzu se neliší) pomocí párového t-testu.

Vzhledem k tomu, že předpoklad parametrického párového t-testu je splněn, můžeme tento test využít k testování nulové hypotézy.

P-hodnota párového t-testu

< 0,001

Závěr: Na základě párového t-testu **zamítáme nulovou hypotézu** ($p < 0,05$), váha před absolvováním kurzu a po absolvování kurzu se liší. Průměrně došlo ke snížení váhy o 2,5 kg.

*Pozn.: V případě, že by předpoklad normality difference nebyl splněn, nemůžeme k testování využít párový t-test, ale musíme využít jeho **neparametrickou alternativu** (viz kapitola 6.3).*

Návod pro software Statistica: Statistics – Basic statistics – t-test, dependent samples.

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 zjistěte, zda je obvod pasu respondentů statisticky významně odlišný před absolvováním kurzu zaměřeného na zdravý životní styl (proměnná *obvod_pasu*) a po absolvování tohoto kurzu (proměnná *obvod_pasu_po*).
- 2. příklad:** Na hladině významnosti 0,05 zjistěte, zda došlo ke statisticky významné změně systolického tlaku respondentů po absolvování kurzu zaměřeného na zdravý životní styl (proměnná *syst_tlak* a *syst_tlak_po*).
- 3. příklad:** Na hladině významnosti 0,05 zjistěte, zda je WHR – poměr pasu ku bokům statisticky významně odlišný před absolvováním kurzu zaměřeného na zdravý životní styl (proměnná *WHR*) a po absolvování tohoto kurzu (proměnná *WHR_po*).

Řešení:

1. příklad: Předpoklad normality difference je splněn (p-hodnota S-W testu: 0,401). Na základě párového t-testu **zamítáme nulovou hypotézu**, obvod pasu před kurzem a po kurzu se liší (p-hodnota $< 0,001$). Průměrně došlo ke snížení obvodu pasu o 2,9 cm.

2. příklad: Předpoklad normality difference je splněn (p-hodnota S-W testu: 0,114). Na základě párového t-testu **nezamítáme nulovou hypotézu**, že systolický tlak před kurzem a po kurzu se neliší (p-hodnota: 0,359). Průměrně došlo ke zvýšení systolického tlaku o 0,9 mmHg.

3. příklad: Předpoklad normality difference není splněn (p-hodnota S-W testu: 0,002). K testování hypotézy není možné využít párový t-test, ale je nutné využít neparametrické alternativy – párový Wilcoxonův nebo znaménkový test (viz kapitola 6.3).

4.4. Analýza rozptylu (ANOVA)

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda se WHR – poměr pasu ku bokům (proměnná *WHR*) statisticky významně liší u respondentů s rozdílným vzděláním (proměnná *vzdelani*).

Řešení:

1. Stanovení nulové a alternativní hypotézy.

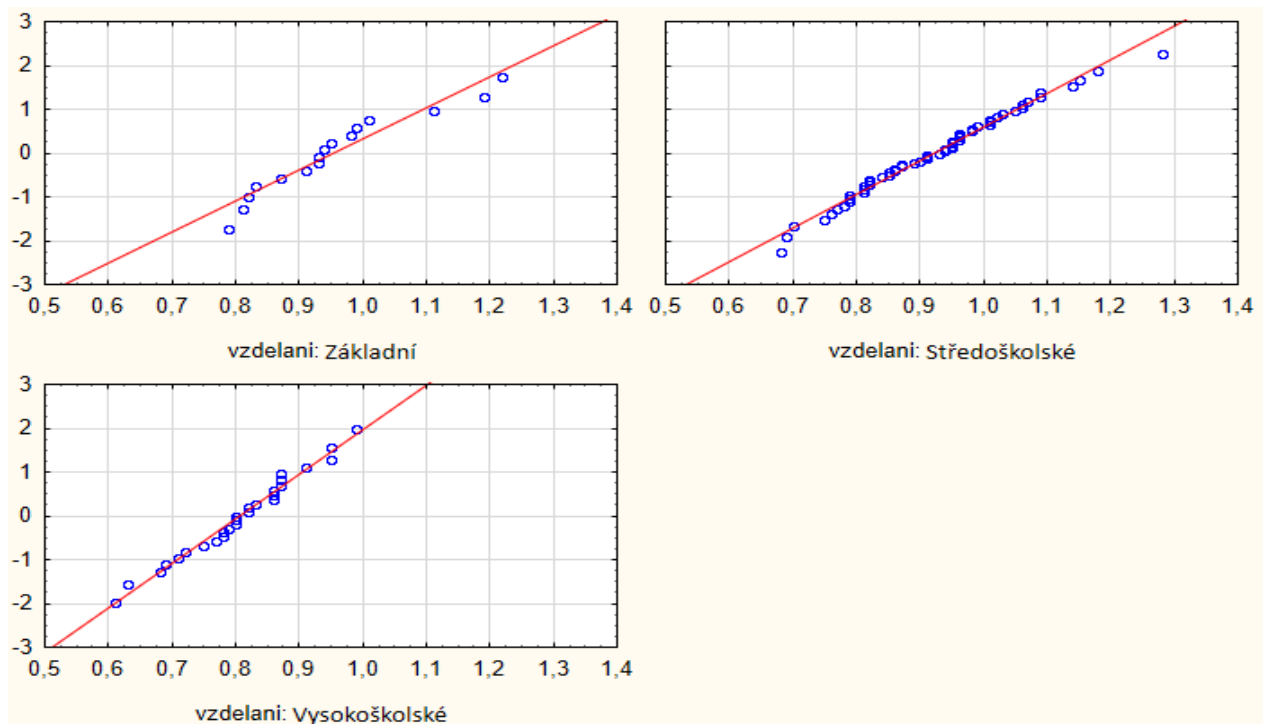
H_0 : WHR se **neliší** u respondentů s rozdílným vzděláním.

H_A : WHR se **liši** u respondentů s rozdílným vzděláním.

2. Ověření předpokladů parametrického testu.

Analýza rozptylu (ANOVA) má 2 předpoklady, které je nutné před jejím využitím ověřit. Prvním předpokladem je **normalita** náhodného výběru v **rámci testovaných skupin** (v našem případě rozložení WHR ve skupině respondentů se základním, středoškolským a vysokoškolským vzděláním). Druhým předpokladem je **homogenita rozptylů v rámci skupin** (v našem případě to znamená, že rozptyly WHR ve skupině respondentů se základním, středoškolským a vysokoškolským vzděláním by měly být přibližně shodné).

První předpoklad můžeme ověřit např. pomocí N-P grafu a Shapirova-Wilkova testu.



Na základě N-P grafu a Shapirova-Wilkova testu nezamítáme hypotézu, že rozdělení WHR u respondentů se základním, středoškolským a vysokoškolským vzděláním se řídí normálním rozdělením (p-hodnota základní vzdělání: 0,153; středoškolské vzdělání: 0,667; vysokoškolské vzdělání: 0,856) – **předpoklad normality v rámci skupin je splněn.**

Pozn.: Stejně jako v případě dvouvýběrového t-testu (viz kapitola 5.2) není nutné, aby datový soubor splňoval předpoklad normality v rámci celku, ale pouze v rámci uvažovaných skupin.

Druhý předpoklad ověříme pomocí Levenova testu.

P-hodnota Levenova testu	0,212
--------------------------	-------

Na základě provedeného Levenova testu nezamítáme hypotézu, že rozptyl WHR u respondentů se základním, středoškolským a vysokoškolským vzděláním se neliší – **předpoklad homogenity rozptylů v rámci skupin je splněn.**

Návod pro software Statistica: Statistics – Basic statistics – Breakdown & one-way ANOVA – OK – Levene tests.

3. Testování nulové hypotézy (H_0 : Hodnota WHR se neliší u respondentů s rozdílným vzděláním) pomocí analýzy rozptylu.

Vzhledem k tomu, že oba předpoklady analýzy rozptylu jsou splněny, můžeme tento test využít k testování nulové hypotézy.

P-hodnota analýzy rozptylu	< 0,001
----------------------------	---------

Závěr: Na základě provedené analýzy rozptylu **zamítáme nulovou hypotézu** ($p < 0,05$), WHR se u respondentů s rozdílným vzděláním liší. Průměrná hodnota WHR u respondentů se základním vzděláním je 0,96, se středoškolským vzděláním 0,92 a s vysokoškolským 0,81.

*Pozn.: V případě, že by kterýkoli z předpokladů (normalita v rámci skupin, homogenita rozptylů) nebyl splněn, nemůžeme k testování využít analýzu rozptylu, ale musíme využít její **neparametrickou alternativu** (viz kapitola 6.4).*

Návod pro software Statistica: Statistics – Basic statistics – Breakdown & one-way ANOVA.

4. Mnohonásobné porovnání.

V případě, že pomocí analýzy rozptylu dostáváme statisticky významný výsledek, je nutné následně provést testy mnohonásobného porovnávání, abychom zjistili, které skupiny se od sebe statisticky významně odlišují.

*Pozn.: V případě, že pomocí analýzy rozptylu nedostáváme statisticky významný výsledek ($p > 0,05$), testy mnohonásobného porovnávání dále **neprovádíme**.*

Pro mnohonásobné porovnávání využijeme Tukeyův test (v tabulce jsou uvedeny p-hodnoty pro všechny dvojice skupin):

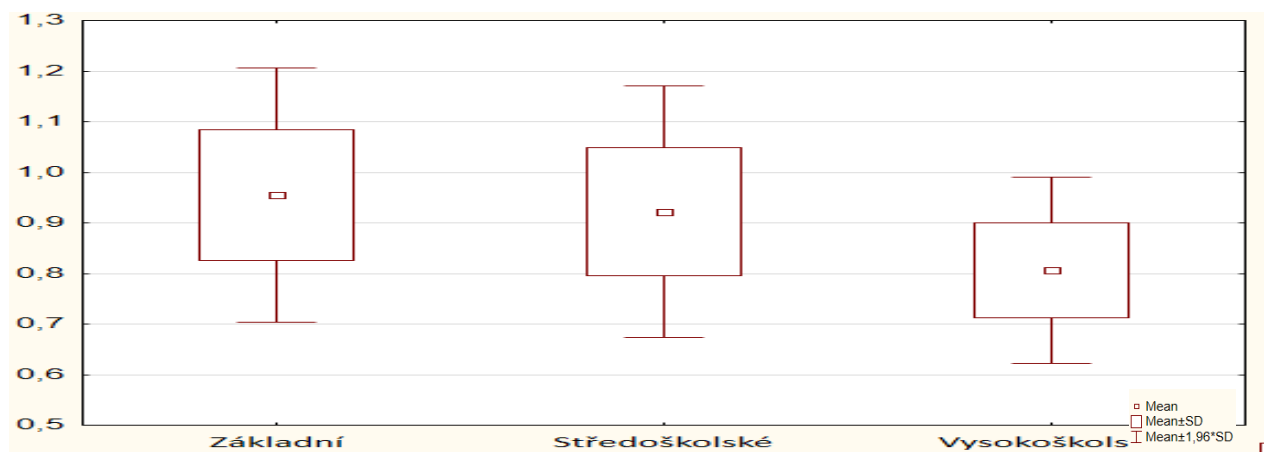
	Základní	Středoškolské	Vysokoškolské
Základní	–	0,717	0,002
Středoškolské	0,717	–	0,002
Vysokoškolské	0,002	0,002	–

Na základě testů pro mnohonásobné porovnávání je patrný statisticky významný rozdíl ($p < 0,05$) mezi skupinou respondentů se základním vzděláním a vysokoškolským vzděláním a mezi skupinou se středoškolským a vysokoškolským vzděláním. Mezi skupinou respondentů se základním a středoškolským vzděláním statisticky významný rozdíl zaznamenán nebyl ($p > 0,05$).

Návod pro software Statistica: Statistics – Basic statistics – Breakdown & one-way ANOVA – OK – záložka Post-hoc – Tukey HSD for unequal N.

*Pozn.: Může nastat situace, kdy pomocí ANOVA zamítneme nulovou hypotézu, ale metodami mnohonásobného porovnávání **nenajdeme významný rozdíl u žádné dvojice skupin**. K tomu dochází především v případě, když p-hodnota pro ANOVA je na hranici významnosti, tedy jen o málo nižší než zvolená hladina významnosti. Důvodem je, že testy mnohonásobného porovnávání mají obecně **nižší sílu** než ANOVA, a proto nemusí odhalit žádný rozdíl.*

Pro lepší představu o pozorovaném rozdílu mezi skupinami si můžeme vykreslit krabicový graf.



Z grafického srovnání vidíme výrazný rozdíl mezi respondenty se základním a vysokoškolským vzděláním a mezi respondenty se středoškolským a vysokoškolským vzděláním. Mezi respondenty se základním a středoškolským vzděláním není z grafů patrný výrazný rozdíl, což koresponduje se závěry testů mnohonásobného porovnávání.

Návod pro software Statistica: Statistics – Basic statistics – Breakdown & one-way ANOVA – OK – záložka Descriptives – Categorized box and whisker.

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 zjistěte, zda se systolický tlak (proměnná *syst_tlak*) statisticky významně liší u respondentů s rozdílným rizikem odvozeným od poměru obvodu pasu ku obvodu boků (proměnná *WHR_riziko*).
- 2. příklad:** Na hladině významnosti 0,05 zjistěte, zda se výška (proměnná *vyska*) statisticky významně liší u respondentů s rozdílným vzděláním (proměnná *vzdelani*).
- 3. příklad:** Na hladině významnosti 0,05 zjistěte, zda se obvod pasu (proměnná *obvod_pasu*) statisticky významně liší u respondentů s rozdílnou frekvencí sportování – kategorizace proměnné *sport* do 3 kategorií (1 + 2; 3; 4 + 5).
- 4. příklad:** Na hladině významnosti 0,05 zjistěte, zda se WHR – poměr pasu ku bokům (proměnná *WHR*) statisticky významně liší u respondentů s rozdílnou frekvencí konzumace alkoholu – kategorizace proměnné *alkohol* do 3 kategorií (1 + 2; 3; 4 + 5).

Řešení:

1. příklad: Předpoklad normality v rámci skupin je splněn (p-hodnota S-W testu: nízké riziko: 0,315; střední riziko: 0,370; vysoké riziko: 0,488). Předpoklad homogenity rozptylů v rámci skupin je splněn (p-hodnota Levenova testu:

0,298). Na základě ANOVA **zamítáme nulovou hypotézu**, systolický tlak u respondentů s nízkým, středním a vysokým rizikem se liší (p-hodnota $< 0,001$). Na základě testů pro mnohonásobné porovnávání je patrný statisticky významný rozdíl ($p < 0,05$) mezi všemi srovnávanými skupinami (p-hodnota skupin nízké vs. střední riziko: 0,012; nízké vs. vysoké: $< 0,001$; střední vs. vysoké: 0,002).

2. příklad: Předpoklad normality v rámci skupin je splněn (p-hodnota S-W testu: základní vzdělání: 0,734; středoškolské vzdělání: 0,830; vysokoškolské vzdělání: 0,831). Předpoklad homogenity rozptylů v rámci skupin je splněn (p-hodnota Levenova testu: 0,346). Na základě ANOVA **nezamítáme nulovou hypotézu**, že výška u respondentů se základním, středoškolským a vysokoškolským vzděláním se neliší (p-hodnota 0,406). Jelikož pomocí analýzu rozptylu nedostáváme statisticky významný výsledek, testy mnohonásobného porovnávání neprovádíme.

3. příklad: Předpoklad normality v rámci skupin je splněn (p-hodnota S-W testu: sport kategorie 1: 0,819; kategorie 2: 0,527; kategorie 3: 0,832). Předpoklad homogenity rozptylů v rámci skupin je splněn (p-hodnota Levenova testu: 0,779). Na základě ANOVA **zamítáme nulovou hypotézu**, obvod pasů u respondentů s rozdílnou frekvencí sportování se liší (p-hodnota 0,040). Na základě testů pro mnohonásobné porovnávání není patrný statisticky významný rozdíl ($p < 0,05$) mezi žádnou dvojicí srovnávaných skupin (p-hodnota: skupina sport kat. 1 vs. kat. 2: 0,792; kat. 1 vs. kat. 3: 0,129; kat. 2 vs. kat. 3: 0,485).

4. příklad: Předpoklad normality v rámci skupin není splněn (p-hodnota S-W testu: alkohol kategorie 1: 0,997; kategorie 2: 0,330; kategorie 3: 0,001). K testování hypotézy není možné využít analýzu rozptylu, ale je nutné využít neparametrickou alternativu – Kruskalův-Wallisův test (viz kapitola 6.4).

Cvičení 5. Neparametrické testy pro kvantitativní proměnné

Následující kapitola bude věnována neparametrickým testům, které slouží k testování statistických hypotéz v případě, kdy nejsou splněny předpoklady testů parametrických (viz kapitola 5). Konkrétně se budeme zabývat jednovýběrovým Wilcoxonovým a znaménkovým testem, jakožto neparametrickým alternativám jednovýběrového t-testu, Mannovým-Whitneyovým testem, jakožto neparametrické alternativě dvouvýběrového t-testu, párovým Wilcoxonovým a znaménkovým testem, jakožto neparametrickým alternativám párového t-testu a poslední část bude věnována Kruskalovu-Wallisovu testu, jakožto neparametrické alternativě analýzy rozptylu.

5.1. Jednovýběrový Wilcoxonův a znaménkový test

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda je glukóza v krvi na lačno (proměnná *glykemie*) u respondentů, kteří netrpí cukrovkou (proměnná *diabetes*), statisticky významně odlišná od fyziologické hodnoty 4,8 mmol/l.

Řešení:

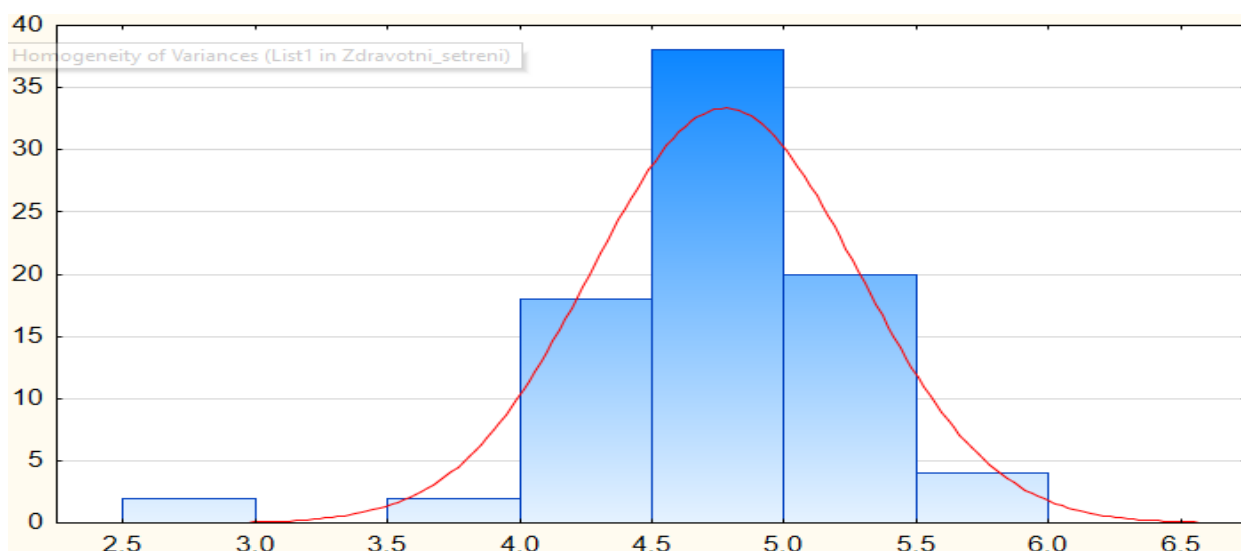
1. Stanovení nulové a alternativní hypotézy.

H_0 : Glukóza v krvi na lačno u respondentů, kteří netrpí cukrovkou, se **rovná** hodnotě 4,8 mmol/l.

H_A : Glukóza v krvi na lačno u respondentů, kteří netrpí cukrovkou, se **nerovná** hodnotě 4,8 mmol/l.

2. Důvod využití neparametrického testu.

Tento typ hypotézy bychom testovali pomocí parametrického jednovýběrového t-testu, jehož předpokladem je **normalita** náhodného výběru (v našem případě normální rozložení glukózy v krvi). Tento předpoklad ověříme např. pomocí histogramu a Shapirova-Wilkova testu.



Na základě histogramu a Shapiro-Wilkova testu zamítáme hypotézu, že rozdělení glukózy v krvi se řídí normálním rozdělením (p-hodnota: $< 0,001$) – **předpoklad normality není splněn.**

3. Testování nulové hypotézy (H_0 : Glukóza v krvi na lačno u respondentů, kteří netrpí cukrovkou se rovná hodnotě 4,8 mmol/l) pomocí jednovýběrového Wilcoxonova a znaménkového testu.

Vzhledem k tomu, že předpoklad parametrického jednovýběrového t-testu není splněn, musíme k testování využít jeho neparametrické alternativy – jednovýběrový Wilcoxonův nebo znaménkový test.

*Pozn.: V případě, že by předpoklad normality byl splněn, je vhodnější k testování využít jednovýběrový t-test (viz kapitola 5.1), neboť parametrický test má oproti neparametrickému **větší sílu** (schopnost zamítnout neplatnou nulovou hypotézu).*

P-hodnota Wilcoxonova testu	0,057
P-hodnota znaménkového testu	0,363

Závěr: Na základě jednovýběrového Wilcoxonova i znaménkového testu **nezamítáme nulovou hypotézu** ($p < 0,05$), že glukóza v krvi u respondentů, kteří netrpí cukrovkou, se rovná hodnotě 4,8 mmol/l. Medián glukózy v krvi u respondentů, kteří netrpí cukrovkou, je 4,795 mmol/l.

Pozn.: Rozdíl mezi Wilcoxonovým a znaménkovým testem je, že Wilcoxonův test má předpoklad symetrického rozložení dat kolem mediánu. V případě výrazného porušení tohoto předpokladu je vhodnější využití znaménkového testu.

Návod pro software Statistica:

- I. Vytvoření nové proměnné *glykemie_median* obsahující hodnotu 4,8 ve všech řádcích:
 - i. Klikneme pravým tlačítkem ve sloupci *diabetes* – *Add variables*.
 - ii. Do kolonky *Long name* napíšeme vzorec: $=4,8$.
- II. Provedení párového Wilcoxonova a znaménkového testu: *Statistics – Nonparametrics – Comparing two dependent samples (variables) – Wilcoxon matched pair test / Sign test*

Příklady k procvičení

1. příklad: Na hladině významnosti 0,05 zjistěte, zda je cholesterol u respondentů (proměnná *cholesterol*), statisticky významně odlišný od mediánu v běžné české populaci – 5,2 mmol/l.

Řešení:

1. příklad: Předpoklad normality není splněn (p-hodnota S-W testu: $< 0,001$), k testování hypotézy je tedy nutné využít neparametrický test. Na základě jednovýběrového Wilcoxonova i znaménkového testu **zamítáme nulovou hypotézu**, cholesterol u respondentů se nerovná hodnotě 5,2 mmol/l (p-hodnota Wilcoxonova testu $< 0,001$; znaménkového testu $< 0,001$). Medián cholesterolu u respondentů je 4,51 mmol/l.

5.2. Mannův-Whitneyův test

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda se cholesterol (proměnná *cholesterol*) statisticky významně liší u respondentů s vysokým krevním tlakem a bez vysokého krevního tlaku (proměnná *hypertenze*).

Řešení:

1. Stanovení nulové a alternativní hypotézy.

H_0 : Cholesterol se u respondentů s vysokým krevním tlakem a bez vysokého krevního tlaku **neliší**.

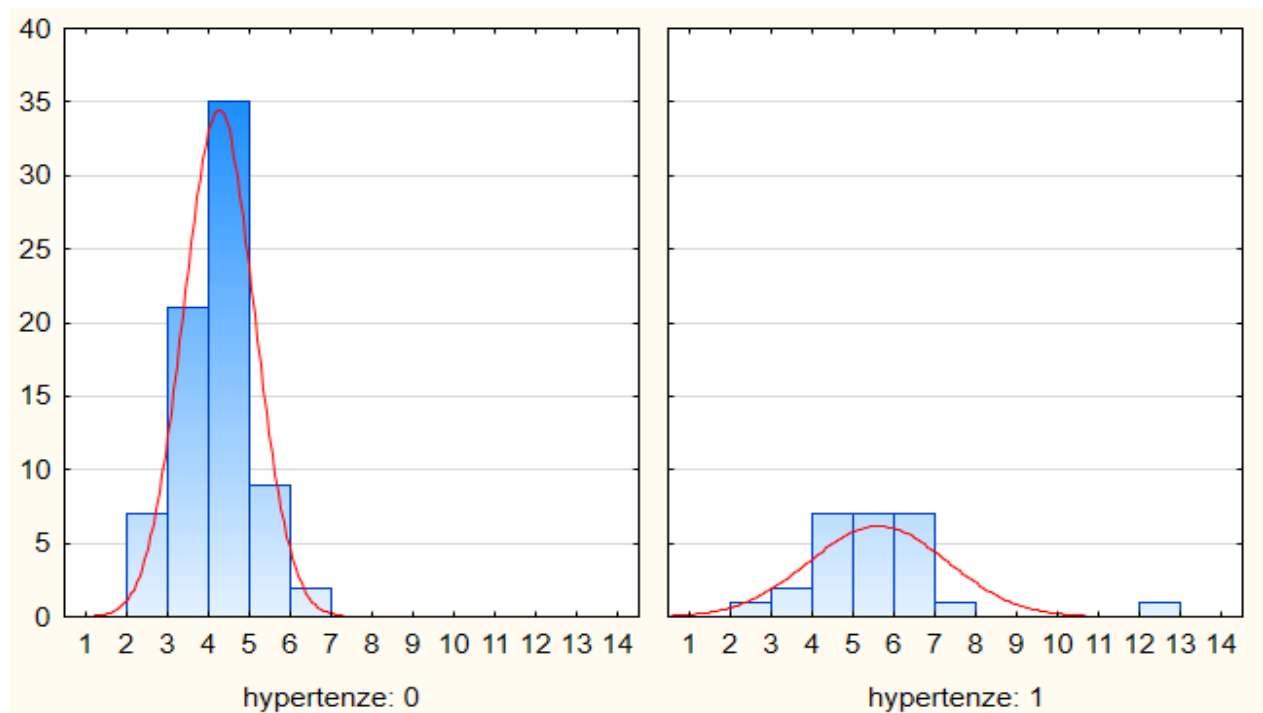
H_A : Cholesterol se u respondentů s vysokým krevním tlakem a bez vysokého krevního tlaku **liši**.

2. Důvod využití neparametrického testu.

Tento typ hypotézy bychom testovali pomocí parametrického dvouvýběrového t-testu. Jedním z jeho předpokladů je **normalita** náhodného výběru v **rámci testovaných skupin** (v našem

případě normální rozdělení cholesterolu ve skupině respondentů s vysokým krevním tlakem i ve skupině bez vysokého krevního tlaku).

Tento předpoklad ověříme např. pomocí histogramu a Shapiro-Wilkova testu.



Na základě histogramu a Shapiro-Wilkova testu zamítáme hypotézu, že rozdělení cholesterolu ve skupině respondentů s vysokým krevním tlakem se řídí normálním rozdělením (p -hodnota: $< 0,001$) – **předpoklad normality v rámci skupin není splněn.**

3. Testování nulové hypotézy (H_0 : Cholesterol se u respondentů s vysokým krevním tlakem a bez vysokého krevního tlaku neliší) pomocí Mannova-Whitneyova testu.

Vzhledem k tomu, že předpoklady parametrického dvouvýběrového t -testu nejsou splněny, musíme k testování využít jeho neparametrickou alternativu – Mannův-Whitneyův test.

*Pozn.: V případě, že by oba předpoklady dvouvýběrového t -testu byly splněny, je vhodnější k testování využít dvouvýběrový t -test (viz kapitola 5.2), neboť parametrický test má oproti neparametrickému **větší sílu** (schopnost zamítnout neplatnou nulovou hypotézu).*

P-hodnota Mannova-Whitneyova testu

$< 0,001$

Závěr: Na základě Mannova-Whitneyova testu **zamítáme nulovou hypotézu** ($p < 0,05$), cholesterol se u respondentů s vysokým krevním tlakem a bez vysokého krevního tlaku liší. Medián cholesterolu u respondentů s vysokým krevním tlakem je 5,58 mmol/l, u respondentů bez vysokého krevního tlaku zvýšeného cholesterolu 4,27 mmol/l.

Návod pro software Statistica: Statistics – Nonparametrics – Comparing two independent samples (groups).

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 zjistěte, zda se BMI – index tělesné hmotnosti (proměnná *BMI*) statisticky významně liší u respondentů, kteří sportují alespoň 3krát týdně a respondentů, kteří sportují méně než 3krát týdně (kategorizace proměnné *sport* – 1 + 2 + 3, vs. 4 + 5).
- 2. příklad:** Na hladině významnosti 0,05 zjistěte, zda se cholesterol (proměnná *cholesterol*) statisticky významně liší u respondentů s nadváhou a bez nadváhy (proměnná *nadváha*).
- 3. příklad:** Na hladině významnosti 0,05 zjistěte, zda se glukóza v krvi na lačno (proměnná *glykemie*) statisticky významně liší u respondentů, kteří konzumují ovoce alespoň 4krát týdně a respondentů, kteří konzumují ovoce méně než 4krát týdně (kategorizace proměnné *ovoce* – 1 + 2, vs. 3 + 4 + 5).

Řešení:

1. příklad: Předpoklad normality v rámci skupin není splněn (p-hodnota S-W testu pro sportování méně než 3krát týdně: 0,011), k testování hypotézy je tedy nutné využít neparametrický test. Na základě Mannova-Whitneyova testu **zamítáme nulovou hypotézu**, tj. BMI respondentů, kteří sportují alespoň 3krát týdně a respondentů, kteří sportují méně než 3krát týdně se liší (p-hodnota < 0,001). Medián BMI u respondentů sportujících alespoň 3krát týdně je 21,8, u respondentů sportujících méně než 3krát týdně je 25,4.
2. příklad: Předpoklad normality v rámci skupin není splněn (p-hodnota S-W testu pro nadváhu < 0,001), k testování hypotézy je tedy nutné využít neparametrický test. Na základě Mannova-Whitneyova testu **zamítáme nulovou hypotézu**, tj. hladina cholesterolu se u respondentů s nadváhou a respondentů bez nadváhy liší (p-hodnota < 0,001). Medián cholesterolu respondentů s nadváhou je 5,16 mmol/l, u respondentů bez nadváhy je 4,08 mmol/l.
3. příklad: Předpoklad normality v rámci skupin není splněn (p-hodnota S-W testu pro konzumaci ovoce alespoň 4krát týdně i pro konzumaci ovoce méně než 4krát týdně < 0,001), k testování hypotézy je tedy nutné využít neparametrický test. Na základě Mannova-Whitneyova testu **nezamítáme nulovou hypotézu**, že glykemie u respondentů konzumujících ovoce alespoň 4krát týdně a respondentů konzumujících ovoce méně než 4krát týdně se neliší (p-hodnota: 0,469). Medián glykemie u respondentů konzumujících ovoce alespoň 4krát týdně je 4,97 mmol/l, u respondentů konzumujících ovoce méně než 4krát týdně je 4,83 mmol/l.

5.3. Párový Wilcoxonův a znaménkový test

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda je WHR – poměr pasu ku bokům statisticky významně odlišný před absolvováním kurzu zaměřeného na zdravý životní styl (proměnná *WHR*) a po absolvování tohoto kurzu (proměnná *WHR_po*).

Řešení:

1. Stanovení nulové a alternativní hypotézy.

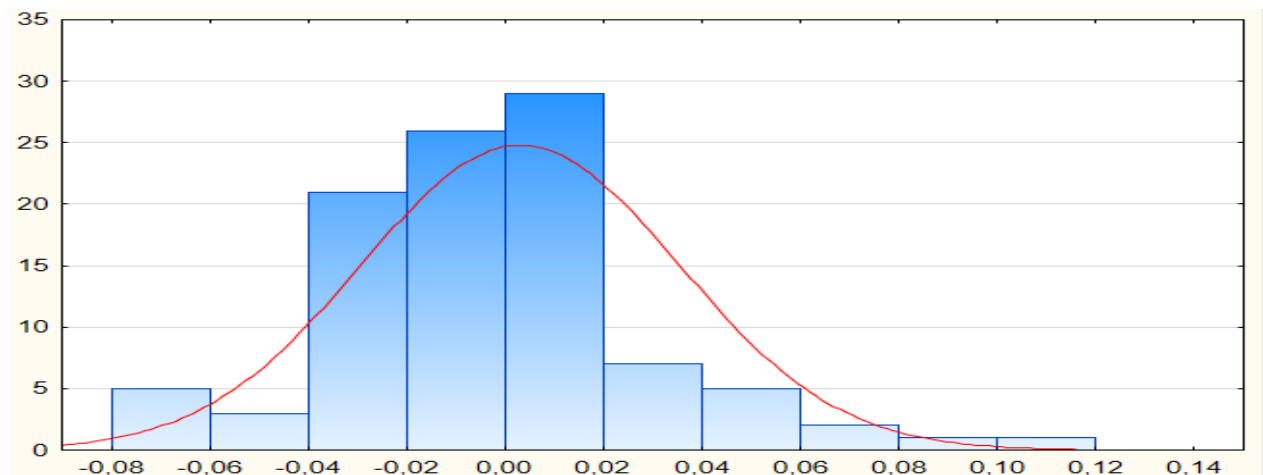
H_0 : WHR před absolvováním kurzu a po absolvování kurzu se **neliší**.

H_A : WHR před absolvováním kurzu a po absolvování kurzu se **liší**.

2. Důvod využití neparametrického testu.

Tento typ hypotézy bychom testovali pomocí parametrického párového t-testu, jehož předpokladem je **normalita difference** (v našem případě rozložení rozdílu hodnot WHR před kurzem a po kurzu).

Tento předpoklad ověříme např. pomocí histogramu a Shapiro-Wilkova testu.



Na základě N-P grafu a Shapiro-Wilkova testu zamítáme hypotézu, že rozdělení difference WHR před a po kurzu se řídí normálním rozdělením (p-hodnota: 0,002) – **předpoklad normality difference není splněn**.

3. Testování nulové hypotézy (H_0 : WHR před absolvováním kurzu a po absolvování kurzu se neliší) pomocí párového Wilcoxonova a znaménkového testu.

Vzhledem k tomu, že předpoklad parametrického párového t-testu není splněn, musíme k testování využít jeho neparametrické alternativy – párový Wilcoxonův nebo znaménkový test.

Pozn.: V případě, že by předpoklad normality byl splněn, je vhodnější k testování využít párový *t*-test (viz kapitola 5.3), neboť parametrický test má oproti neparametrickému **větší sílu** (schopnost zamítnout neplatnou nulovou hypotézu).

P-hodnota Wilcoxonova testu	0,647
P-hodnota znaménkového testu	0,664

Závěr: Na základě párového Wilcoxonova i znaménkového testu **nezamítáme nulovou hypotézu** ($p < 0,05$), že WHR před absolvováním kurzu a po absolvování kurzu se neliší. Medián rozdílu WHR před kurzem a po kurzu 0,0.

Návod pro software Statistica: Statistics – Nonparametrics – Comparing two dependent samples (variables) – Wilcoxon matched pair test / Sign test.

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 zjistěte, zda je cholesterol u respondentů před absolvováním kurzu (proměnná *syst_tlak*) a po absolvování tohoto kurzu (proměnná *syst_tlak_po*) statisticky významně odlišný.
- 2. příklad:** Na hladině významnosti 0,05 zjistěte, zda došlo ke statisticky významné změně glukózy v krvi respondentů po absolvování kurzu zaměřeného na zdravý životní styl (proměnná *glykemie* a *glykemie_po*)
- 3. příklad:** Na hladině významnosti 0,05 zjistěte, zda je systolický tlak u respondentů, kteří měli před absolvováním kurzu zvýšený cholesterol (proměnná *zvyseny_cholesterol*) statisticky významně odlišný před absolvováním kurzu (proměnná *syst_tlak*) a po absolvování tohoto kurzu (proměnná *syst_tlak_po*).

Řešení:

1. příklad: Předpoklad normality difference není splněn (p-hodnota S-W testu: $< 0,001$), k testování hypotézy je tedy nutné využít neparametrický test. Na základě párového Wilcoxonova a znaménkového testu **nezamítáme nulovou hypotézu**, že cholesterol před kurzem a po kurzu se neliší (p-hodnota Wilcoxonova testu: 0,312; znaménkového testu: 0,920). Medián snížení cholesterolu po absolvování kurzu je $-0,01$ mmol/l.

2. příklad: Předpoklad normality difference není splněn (p-hodnota S-W testu: $< 0,001$), k testování hypotézy je tedy nutné využít neparametrický test. Na základě párového Wilcoxonova i znaménkového testu **zamítáme nulovou hypotézu**, tj. glykemie před kurzem a po kurzu se liší (p-hodnota Wilcoxonova testu: $< 0,001$; znaménkového testu: $< 0,001$). Medián snížení glykemie po absolvování kurzu je $-0,23$ mmol/l.

3. příklad: Předpoklad normality difference není splněn (p-hodnota S-W testu: 0,021), k testování hypotézy je tedy nutné využít neparametrický test. Na základě párového Wilcoxonova i znaménkového testu **zamítáme nulovou hypotézu**, tj. systolický tlak u respondentů, kteří měli před absolvováním kurzu zvýšený cholesterol, se před kurzem a po kurzu liší (p-hodnota Wilcoxonova testu: 0,002; znaménkového testu: 0,021). Medián zvýšení systolického tlaku po absolvování kurzu je 7 mmHg.

5.4. Kruskalův-Wallisův test

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda se WHR – poměr pasu ku bokům (proměnná *WHR*) statisticky významně liší u respondentů s rozdílnou frekvencí konzumace alkoholu – kategorizace proměnné *alkohol* do 3 kategorií (1 + 2; 3; 4 + 5).

Řešení:

1. Stanovení nulové a alternativní hypotézy.

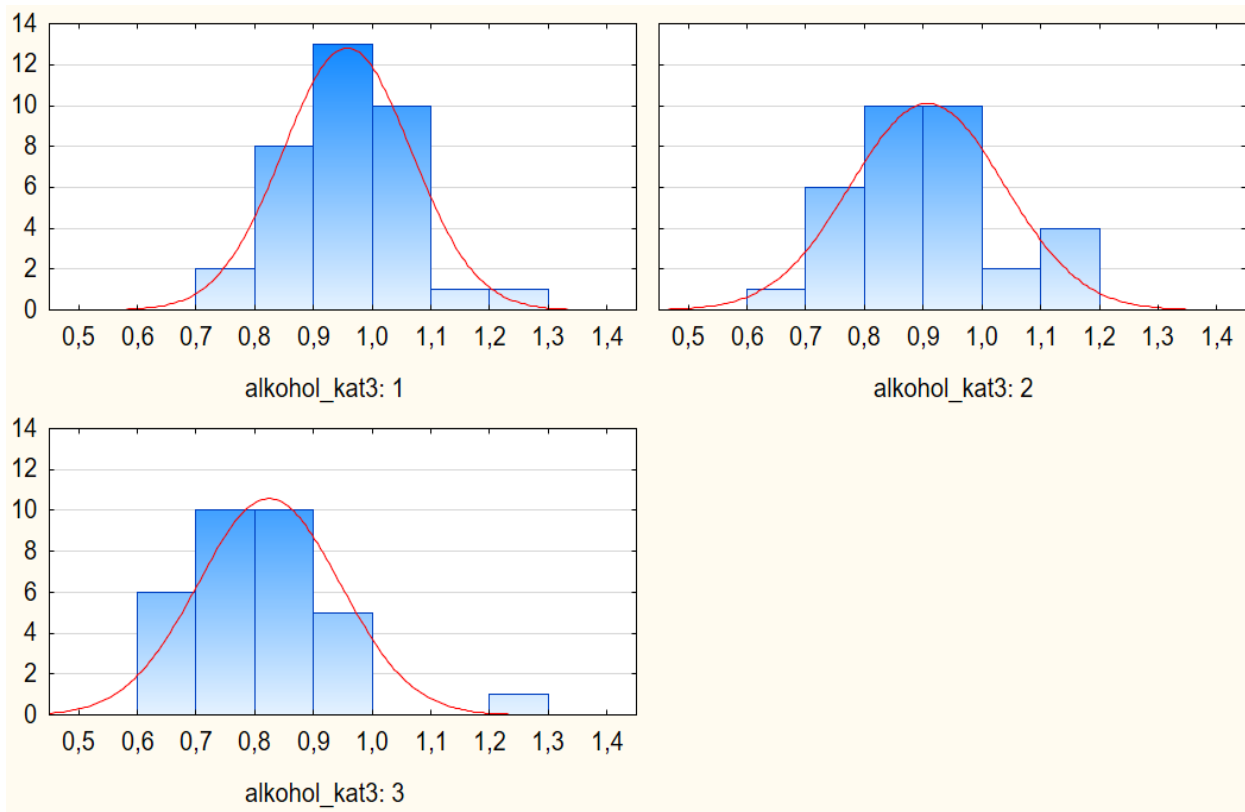
H_0 : WHR se **neliší** u respondentů s rozdílnou frekvencí konzumace alkoholu.

H_A : WHR se **liší** u respondentů s rozdílnou frekvencí konzumace alkoholu.

2. Důvod využití neparametrického testu.

Tento typ hypotézy bychom testovali pomocí analýzy rozptylu (ANOVA). Jedním z předpokladů ANOVA je **normalita** náhodného výběru v **rámci testovaných skupin** (v našem případě normální rozdělení WHR ve skupinách respondentů s rozdílnou frekvencí konzumace alkoholu).

Tento předpoklad můžeme ověřit např. pomocí histogramu a Shapirova-Wilkova testu.



Na základě histogramu a Shapiro-Wilkova testu zamítáme hypotézu, že rozdělení WHR u respondentů, kteří konzumují alkohol maximálně 2krát týdně se řídí normálním rozdělením (p -hodnota: 0,001) – **předpoklad normality v rámci skupin není splněn.**

3. Testování nulové hypotézy (H_0 : WHR se neliší u respondentů s rozdílnou frekvencí konzumace alkoholu) pomocí Kruskalova-Wallisova testu.

Vzhledem k tomu, že předpoklady ANOVA nejsou splněny, musíme k testování využít její neparametrické alternativy – Kruskalova-Wallisova testu.

*Pozn.: V případě splnění předpokladu normality rozdělení hodnot ve všech skupinách je vhodnější k testování využít ANOVA (viz kapitola 5.4), neboť parametrický test má oproti neparametrickému **větší sílu** (schopnost zamítnout neplatnou nulovou hypotézu).*

P-hodnota Kruskalova-Wallisova testu

< 0,001

Závěr: Na základě provedeného Kruskalova-Wallisova testu **zamítáme nulovou hypotézu** ($p < 0,05$), WHR se u respondentů s rozdílnou frekvencí konzumace alkoholu liší. Medián WHR u respondentů konzumujících alkohol alespoň 4krát týdně je 0,95, respondentů konzumujících alkohol 3krát týdně je 0,90 a respondentů konzumujících alkohol maximálně 2krát týdně je 0,81.

Návod pro software Statistica: Statistics – Nonparametrics – Comparing multiple indep. samples (groups).

4. Mnohonásobné porovnání.

Podobně, jako u analýzy rozptylu, také zde platí, že v případě, kdy pomocí Kruskalova-Wallisova testu dostáváme statisticky významný výsledek, je nutné následně provést testy mnohonásobné porovnávání, abychom zjistili, které skupiny se od sebe statisticky významně odlišují.

*Pozn.: V případě, že pomocí Kruskalova-Wallisova testu nedostáváme statisticky významný výsledek (p -hodnota $> 0,05$), testy mnohonásobného porovnávání dále **neprovádíme**.*

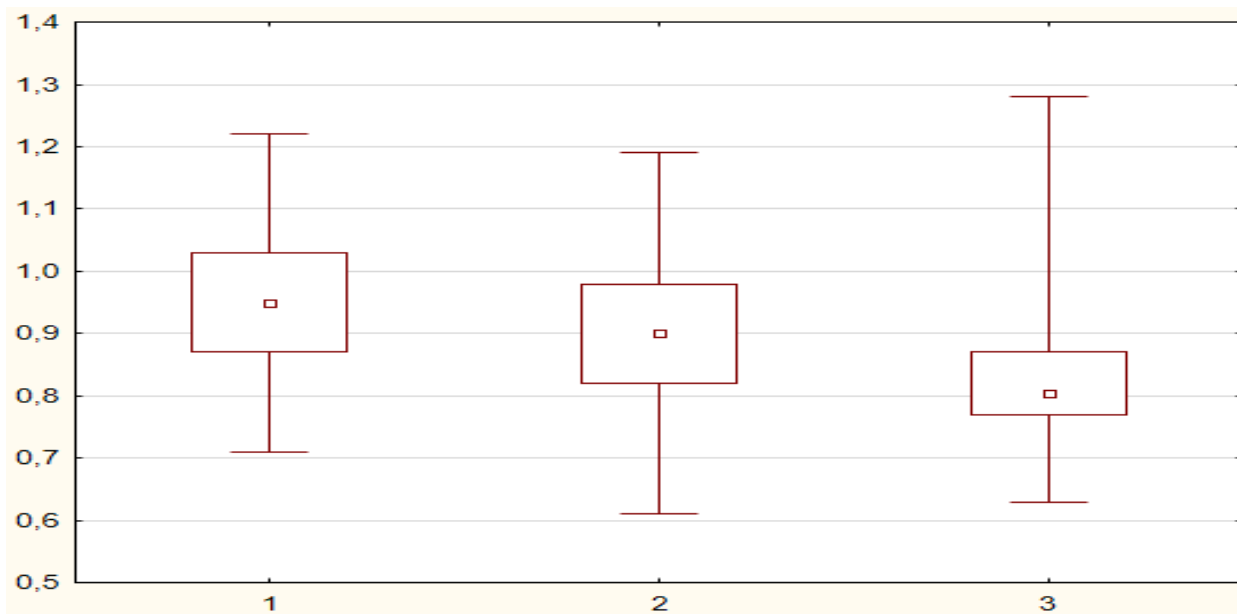
	Alkohol alespoň 4krát týdně	Alkohol 3krát týdně	Alkohol max. 2krát týdně
Alkohol alespoň 4krát týdně	–	0,185	< 0,001
Alkohol 3krát týdně	0,185	–	0,017
Alkohol max. 2krát týdně	< 0,001	0,017	–

Na základě testů pro mnohonásobné porovnávání je patrný statisticky významný rozdíl ($p < 0,05$) mezi skupinou respondentů konzumujících alkohol maximálně 2krát týdně vs. alespoň 4krát týdně a mezi skupinou respondentů konzumujících alkohol maximálně 2krát týdně vs. 3krát týdně. Mezi skupinou konzumující alkohol 3krát týdně vs. alespoň 4krát týdně statisticky významný výsledek prokázán nebyl ($p > 0,05$).

Návod pro software Statistica: Statistics – Nonparametrics – Comparing multiple indep. samples (groups) – Multiple comparisons of mean ranks for all groups.

*Pozn.: Může nastat situace, kdy pomocí Kruskalova-Wallisova testu zamítneme nulovou hypotézu, ale metodami mnohonásobného porovnávání **nenajdeme významný rozdíl u žádné dvojice skupin**. K tomu dochází především v případě, když p -hodnota pro Kruskalův-Wallisův test je na hranici významnosti, tedy jen o málo nižší než zvolená hladina významnosti. Důvodem je, že testy mnohonásobného porovnávání mají obecně **nižší sílu** než Kruskalův-Wallisův test, a proto nemusí odhalit žádný rozdíl.*

Pro lepší představu o pozorovaném rozdílu mezi skupinami si můžeme vykreslit krabicový graf.



Vidíme, že grafické srovnání koresponduje s výsledky, které jsme obdržely pomocí testů mnohonásobného porovnávání.

Návod pro software Statistica: Statistics – Nonparametrics – Comparing multiple indep. samples (groups) – Box and whisker.

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 zjistěte, zda se BMI – index tělesné hmotnosti (proměnná *BMI*) statisticky významně liší u respondentů s rozdílným vzděláním (proměnná *vzdelani*).
- 2. příklad:** Na hladině významnosti 0,05 zjistěte, zda se cholesterol (proměnná *cholesterol*) statisticky významně liší u respondentů s rozdílným rizikem odvozeným od poměru obvodu pasu ku bokům (proměnná *WHR_riziko*).
- 3. příklad:** Na hladině významnosti 0,05 zjistěte, zda se glykemie (proměnná *glykemie*) statisticky významně liší u respondentů s rozdílnou frekvencí pití slazených nápojů – kategorizace proměnné *slazene_napoje* do 3 kategorií (1 + 2; 3; 4 + 5).

Řešení:

1. příklad: Předpoklad normality v rámci skupin není splněn (p-hodnota S-W testu: základní vzdělání: 0,026; středoškolské vzdělání: 0,007, vysokoškolské vzdělání: 0,016), k testování hypotézy je tedy nutné využít neparametrický test. Na základě Kruskalova-Wallisova testu **nezamítáme nulovou hypotézu**, že BMI se u respondentů s rozdílným vzděláním neliší (p-hodnota: 0,966). Jelikož pomocí Kruskalova-Wallisova testu

nedostáváme statisticky významný výsledek, testy mnohonásobného porovnávání neprovádíme.

2. příklad: Předpoklad normality pro skupinu respondentů s vysokým rizikem není splněn (p-hodnota S-W testu: $< 0,001$), k testování hypotézy je tedy nutné využít neparametrický test. Na základě Kruskalova-Wallisova testu **zamítáme nulovou hypotézu**, tj. BMI se u respondentů s rozdílným vzděláním liší (p-hodnota: 0,010). Na základě testů pro mnohonásobné porovnávání je patrný statisticky významný rozdíl mezi respondenty s nízkým rizikem vs. vysokým rizikem (p-hodnota: 0,008).

3. příklad: Předpoklad normality v rámci skupin není splněn (p-hodnota S-W testu pro všechny hodnocené skupiny $< 0,001$), k testování hypotézy je tedy nutné využít neparametrický test. Na základě Kruskalova-Wallisova testu **nezamítáme nulovou hypotézu**, že glykemie se u respondentů s rozdílnou frekvencí pití slazených nápojů neliší (p-hodnota: 0,763). Jelikož pomocí Kruskalova-Wallisova testu nedostáváme statisticky významný výsledek, testy mnohonásobného porovnávání neprovádíme.

Cvičení 6. Testy pro kategoriální proměnné

V rámci této kapitoly se zaměříme na analýzu kontingenčních tabulek, jejíž cílem je hodnocení vztahu mezi dvěma kategoriálními proměnnými. V rámci kapitoly se budeme věnovat testům, které jsou při analýze kontingenčních tabulek používány nejčastěji. Konkrétně se bude jednat o Personův chí-kvadrát test a jeho neparametrickou alternativu – Fisherův exaktní test. V závěru se budeme zabývat McNemarovým testem, který se využívá v případě párového designu experimentu.

6.1. Pearsonův chí-kvadrát test

Řešený příklad

Na hladině významnosti 0,05 testujte hypotézu, že zvýšený cholesterol (proměnná *zvyseny_cholesterol*) a nadváha (proměnná *nadvaha*) jsou nezávislé náhodné veličiny.

Řešení:

1. Stanovení nulové a alternativní hypotézy.

H_0 : Zvýšený cholesterol a nadváha **jsou nezávislé** náhodné veličiny (výskyt zvýšeného cholesterolu **nesouvisí** s výskytem nadváhy).

H_A : Zvýšený cholesterol a nadváha **jsou závislé** náhodné veličiny (výskyt zvýšeného cholesterolu **souvisí** s výskytem nadváhy).

2. Ověření předpokladů parametrického testu.

Pro Personův chí-kvadrát test je nutné ověřit **podmínky dobré aproximace** (tzn. očekávané četnosti musí být aspoň v 80 % případů větší nebo rovné 5 a ve 100 % případů větší než 2).

Pozorované četnosti:

	Bez zvýšeného cholesterolu	Zvýšený cholesterol	Celkem
Normální váha	53	2	55
Nadváha	20	25	45
Celkem	73	27	100

Očekávané četnosti:

	Bez zvýšeného cholesterolu	Zvýšený cholesterol	Celkem
Normální váha	40,15	14,85	55
Nadváha	32,85	12,15	45
Celkem	73	27	100

Všechny hodnoty očekávaných četností jsou větší nebo rovny 5 – **podmínky dobré aproximace jsou splněny**.

Návod pro software Statistica: Statistics – Basic statistics – Tables and banners – záložka Options – zatrhnutí políčka Expected frequencies.

Pozn.: V případě, že by podmínky dobré aproximace nebyly splněny, je nutné **sloučit kategorie** s nízkými četnostmi (lze provést pouze tehdy, jsou-li tyto kategorie slučitelné). Jestliže není možné sloučení kategorií, je nutné k testování využít **neparametrickou alternativu** (viz kapitola 7.2).

3. Testování nulové hypotézy (H_0 : Zvýšený cholesterol a nadváha jsou nezávislé náhodné veličiny) pomocí Personova chí-kvadrát testu.

Vzhledem k tomu, že předpoklad Personova chí-kvadrát testu je splněn, můžeme tento test využít k testování nulové hypotézy.

P-hodnota Personova chí-kvadrát testu

< 0,001

Závěr: Na základě provedeného Personova chí-kvadrát testu **zamítáme nulovou hypotézu** ($p < 0,05$), výskyt zvýšeného cholesterolu souvisí s výskytem nadváhy. Pacienti s nadváhou mají častěji zvýšený cholesterol než pacienti s normální váhou.

Návod pro software Statistica: Statistics – Basic statistics – Tables and banners – záložka Options – zatrhnutí políčka Pearson and M-L Chi-square.

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 testujte hypotézu, že pohlaví (proměnná *pohlavi*) a vzdělání (proměnná *vzdelani*) jsou nezávislé náhodné veličiny.
- 2. příklad:** Na hladině významnosti 0,05 testujte hypotézu, že kouření (proměnná *koureni*) a konzumace alkoholu (proměnná *alkohol*) spolu nesouvisí.
- 3. příklad:** Na hladině významnosti 0,05 testujte hypotézu, že konzumace ovoce (proměnná *ovoce*) a konzumace zeleniny (proměnná *zelenina*) spolu nesouvisí.
- 4. příklad:** Na hladině významnosti 0,05 testujte hypotézu, že přítomnost cukrovky (proměnná *diabetes*) a zvýšený cholesterol (proměnná *zvyseny_cholesterol*) jsou nezávislé náhodné veličiny.

Řešení:

1. příklad: Všechny očekávané četnosti jsou větší než 5 – podmínky dobré aproximace jsou splněny. Na základě Personova chí-kvadrát testu **nezamítáme nulovou hypotézu** (p-hodnota: 0,101), že pohlaví a vzdělání jsou nezávislé náhodné veličiny.

2. příklad: 8 z 10 (tj. 80 %) očekávaných četností je větších než 5, žádná není menší než 2 – podmínky dobré aproximace jsou splněny. Na základě Personova

chí-kvadrát testu **zamítáme nulovou hypotézu** (p -hodnota $< 0,001$), kouření a konzumace alkoholu spolu statisticky významně souvisí.

3. příklad: 19 z 25 (tj. 76 %) očekávaných četnosti je menších než 5, 3 z 25 (tj. 12 %) jsou menší než 2 – podmínky dobré aproximace nejsou splněny. Je nutné sloučení některých kategorií – např. kategorizace proměnné *ovoce* do 2 kategorií (1 + 2 vs. 3 + 4 + 5) a kategorizace proměnné *zelenina* do 2 kategorií (1 + 2 vs. 3 + 4 + 5). Na základě Personova chí-kvadrát testu **zamítáme nulovou hypotézu** (p -hodnota $< 0,001$), konzumace ovoce a konzumace zeleniny spolu statisticky významně souvisí.

4. příklad: 1 ze 4 (tj. 75 %) očekávaných četnosti je menších než 5 – podmínky dobré aproximace nejsou splněny. K testování hypotézy není možné využít Personův chí-kvadrát test, ale je nutné využít neparametrickou alternativu – Fisherův exaktní test (viz kapitola 7.2).

6.2. Fisherův exaktní test

Řešený příklad

Na hladině významnosti 0,05 testujte hypotézu, že přítomnost cukrovky (proměnná *diabetes*) a zvýšený cholesterol (proměnná *zvyseny_cholesterol*) jsou nezávislé náhodné veličiny.

Řešení:

1. Stanovení nulové a alternativní hypotézy.

H_0 : Přítomnost cukrovky a zvýšený cholesterol **jsou nezávislé** náhodné veličiny (výskyt cukrovky **nesouvisí** s výskytem zvýšeného cholesterolu).

H_A : Přítomnost cukrovky a zvýšený cholesterol **jsou závislé** náhodné veličiny (výskyt cukrovky **souvisí** s výskytem zvýšeného cholesterolu).

2. Důvod využití neparametrického testu.

Tento typ hypotézy bychom testovali pomocí Personova chí-kvadrát testu, jehož předpokladem je splnění podmínek dobré aproximace

Pozorované četnosti:

	Bez zvýšeného cholesterolu	Zvýšený cholesterol	Celkem
Bez cukrovky	64	20	84
Cukrovka	9	7	16
Celkem	73	27	100

Očekávané četnosti:

	Bez zvýšeného cholesterolu	Zvýšený cholesterol	Celkem
Bez cukrovky	61,32	22,68	84
Cukrovka	11,68	4,32	16
Celkem	73	27	100

1 ze 4 (tj. 75 %) očekávaných četnosti je menších než 5 – **podmínky dobré aproximace nejsou splněny.**

3. Testování nulové hypotézy (H_0 : Zvýšený cholesterol a nadváha jsou nezávislé náhodné veličiny) pomocí Fisherova exaktního testu.

Vzhledem k tomu, že předpoklad Personova chí-kvadrát testu není splněn a kategorie nelze nijak sloučit, musíme k testování využít neparametrickou alternativu – Fisherův exaktní test.

P-hodnota Fisherova exaktního testu	0,126
-------------------------------------	-------

Závěr: Na základě provedeného Fisherova exaktního testu **nezamítáme nulovou hypotézu** ($p > 0,05$), že přítomnost cukrovky a zvýšeného cholesterolu jsou nezávislé náhodné veličiny.

Návod pro software Statistica: Statistics – Basic statistics – Tables and banners – záložka Options – zatrhnutí políčka Fisher exact, Yates, McNemar (2×2).

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 testujte hypotézu, že vysoký krevní tlak (proměnná *hypertenze*) a přítomnost cukrovky (proměnná *diabetes*) spolu nesouvisí.
- 2. příklad:** Na hladině významnosti 0,05 testujte hypotézu, že u respondentů, kteří kouří tabákové výrobky (proměnná *kourení*), jsou nadváha (proměnná *nadvaha*) a sportování alespoň 3krát týdně – kategorizace proměnné *nadvaha* do 2 kategorií (1 + 2 + 3 vs. 4 + 5) nezávislé náhodné veličiny.
- 3. příklad:** Na hladině významnosti 0,05 testujte hypotézu, že u respondentů, kteří mají nadváhu (proměnná *nadvaha*), jsou zvýšený cholesterol (proměnná *zvyseny_cholesterol*) a konzumace zeleniny alespoň 4krát týdně – kategorizace proměnné *zelenina* do 2 kategorií (1 + 2 vs. 3 + 4 + 5) nezávislé náhodné veličiny.
- 4. příklad:** Na hladině významnosti 0,05 testujte hypotézu, že vysoký krevní tlak (proměnná *hypertenze*) a zvýšený cholesterol (proměnná *zvyseny_cholesterol*) u mužů jsou nezávislé náhodné veličiny.

Řešení:

1. příklad: 1 ze 4 (tj. 75 %) očekávaných četností je menší než 5 – podmínky dobré aproximace nejsou splněny a kategorie nelze nijak sloučit, k testování

hypotézy je tedy nutné využít neparametrický test. Na základě Fisherova exaktního testu **zamítáme nulovou hypotézu** (p-hodnota: 0,028), vysoký krevní tlak a přítomnost cukrovky spolu statisticky významně souvisí.

2. příklad: 2 ze 4 (tj. 50 %) očekávaných četností je menších než 5 – podmínky dobré aproximace nejsou splněny a kategorie nelze nijak sloučit, k testování hypotézy je tedy nutné využít neparametrický test. Na základě Fisherova exaktního testu **nezamítáme nulovou hypotézu** (p-hodnota: 0,260), že nadváha a sportování alespoň 3krát týdně jsou u kuřáků nezávislé náhodné veličiny.

3. příklad: 2 ze 4 (tj. 50 %) očekávaných četností je menších než 5 – podmínky dobré aproximace nejsou splněny a kategorie nelze nijak sloučit, k testování hypotézy je tedy nutné využít neparametrický test. Na základě Fisherova exaktního testu **zamítáme nulovou hypotézu** (p-hodnota: 0,015), zvýšený cholesterol a konzumace zeleniny alespoň 4krát týdně spolu statisticky významně souvisí.

4. příklad: 1 ze 4 (tj. 75 %) očekávaných četností je menší než 5 – podmínky dobré aproximace nejsou splněny a kategorie nelze nijak sloučit, k testování hypotézy je tedy nutné využít neparametrický test. Na základě Fisherova exaktního testu **zamítáme nulovou hypotézu** (p-hodnota: 0,039), vysoký krevní tlak a zvýšený cholesterol u mužů spolu statisticky významně souvisí.

6.3. McNemarův test

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda absolvování kurzu o zdravém životním stylu vede ke změně postoje respondentů ke sportování, tzn. provozují sport alespoň 3krát týdně (kategorizace proměnné *sport* a *sport_po* – 1 + 2 + 3, vs. 4 + 5).

Řešení:

1. Stanovení nulové a alternativní hypotézy.

H_0 : Absolvování kurzu o zdravém životním stylu **nemá vliv** na postoj respondentů ke sportování. (Počet respondentů provozujících sport alespoň 3krát týdně se před absolvováním kurzu a po absolvování kurzu se **neliší**.)

H_A : Absolvování kurzu o zdravém životním stylu **má vliv** na postoj respondentů ke sportování. (Počet respondentů provozujících sport alespoň 3krát týdně se před absolvováním kurzu a po absolvování kurzu se **liší**.)

2. Testování nulové hypotézy (H_0 : Absolvování kurzu o zdravém životním stylu nemá vliv na postoj respondentů ke sportování) pomocí McNemarova testu.

Souhrnná sumarizace, zda došlo ke změně ve sportování po absolvování kurzu může být znázorněna např touto tabulkou:

Sport před kurzem	Sport po kurzu		Celkem
	Alespoň 3krát týdně	Méně než 3krát týdně	
Alespoň 3krát týdně	34	3	37
Méně než 3krát týdně	22	41	63
Celkem	56	44	100

Vidíme, že důsledku kurzu začalo sportovat alespoň 3krát týdně 22 respondentů, kteří předtím sportovali méně než 3krát týdně. Naopak 3 respondenti, kteří před kurzem sportovali alespoň 3krát týdně, po kurzu sportují méně než 3krát týdně.

Testování nulové hypotézy pomocí McNemarova testu:

P-hodnota McNemarova testu	< 0,001
----------------------------	-------------------

Závěr: Na základě McNemarova testu **zamítáme nulovou hypotézu** ($p < 0,05$), absolvování kurzu o zdravém životním stylu mělo vliv na postoj respondentů ke sportování.

Návod pro software Statistica: Statistics – Basic statistics – Tables and banners – záložka Options – zatrhnutí políčka Fisher exact, Yates, McNemar (2×2).

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 zjistěte, zda absolvování kurzu o zdravém životním stylu vede ke změně postoje respondentů ke kouření (proměnná *koureni* a *koureni_po*).
- 2. příklad:** Na hladině významnosti 0,05 zjistěte, zda absolvování kurzu o zdravém životním stylu vede ke změně počtu respondentů, kteří mají nadváhu (proměnná *nadvaha* a *nadvaha_po*).
- 3. příklad:** Na hladině významnosti 0,05 zjistěte, zda absolvování kurzu o zdravém životním stylu vede ke změně počtu respondentů, kteří mají vysoký krevní tlak (proměnná *hypertenze* a *hypertenze_po*).
- 4. příklad:** Na hladině významnosti 0,05 zjistěte, zda absolvování kurzu o zdravém životním stylu vede ke změně postoje respondentů v pití slazených nápojů, tzn. pijí slazené nápoje alespoň 4krát týdně (proměnné *slazene_napoje* a *slazene_napoje_po* kategorizované do dvou kategorií).

Řešení:

1. příklad: Na základě párového McNemarova testu **nezamítáme nulovou hypotézu** (p-hodnota: 0,724), že absolvování kurzu nemělo vliv na kouření respondentů.

2. příklad: Na základě párového McNemarova testu **zamítáme nulovou hypotézu** (p-hodnota $< 0,001$), absolvování kurzu mělo vliv na nadváhu respondentů. Celkově bylo 18 respondentů, kteří měli před absolvováním kurzu nadváhu a po kurzu nadváhu neměli.
3. příklad: Na základě párového McNemarova testu **nezamítáme nulovou hypotézu** (p-hodnota: 803), že absolvování kurzu nemělo vliv na hypertenzi respondentů.
4. příklad: Na základě párového McNemarova testu **zamítáme nulovou hypotézu** (p-hodnota: 0,003), že absolvování kurzu nemělo vliv na pití slazených nápojů u respondentů. V důsledku kurzu začalo pít slazené nápoje maximálně 3krát týdně 21 respondentů, kteří před kurzem pili slazené nápoje alespoň 4krát týdně.

Cvičení 7. Základy korelační analýzy

V poslední kapitole se zaměříme na korelační analýzu, která slouží k hodnocení vztahu mezi dvěma spojitými veličinami. Konkrétně se budeme věnovat Pearsonovu korelačnímu koeficientu a následně jeho neparametrické alternativě – Spearmanovu korelačnímu koeficientu. Oba tyto koeficienty hodnotí míru lineární závislosti mezi dvěma spojitými veličinami.

7.1. Pearsonův korelační koeficient

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda existuje vztah mezi obvodem pasu respondentů (proměnná *obvod_pasu*) a obvodem boků respondentů (proměnná *obvod_boku*).

Řešení:

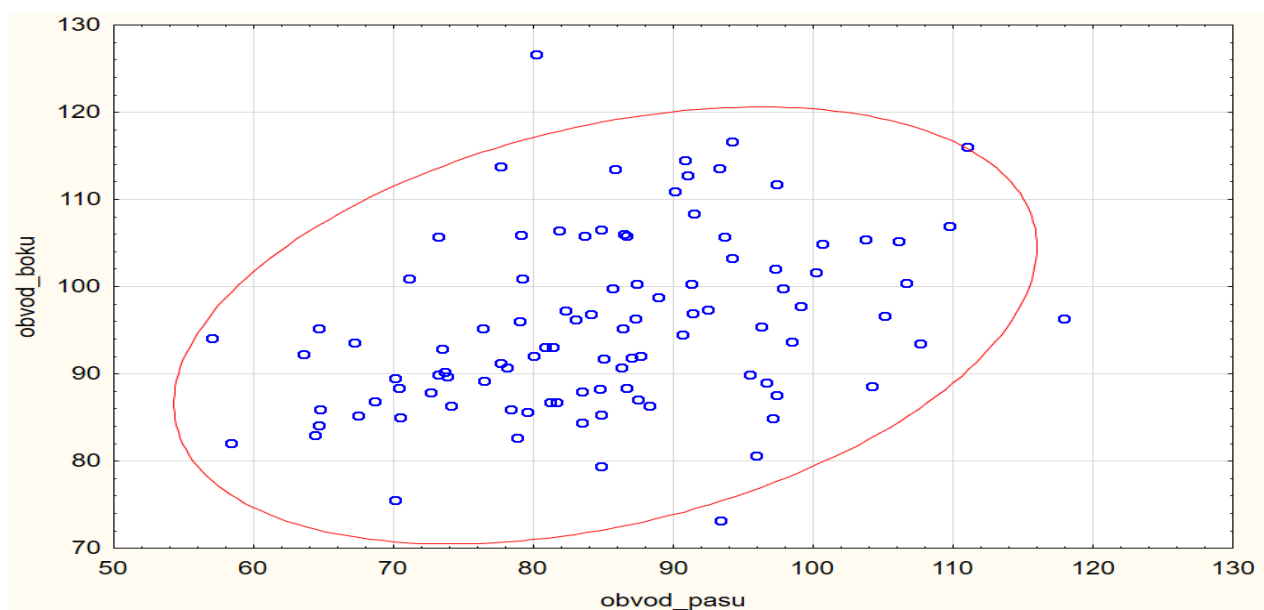
1. Stanovení nulové a alternativní hypotézy.

H_0 : Obvod pasu a obvod boků u respondentů jsou **nezávislé** náhodné veličiny. (Korelační koeficient se rovná 0.)

H_A : Obvod pasu a obvod boků u respondentů jsou **závislé** náhodné veličiny. (Korelační koeficient se nerovná 0.)

2. Ověření předpokladů Pearsonova korelačního koeficientu.

Předpokladem Pearsonova korelačního koeficientu je **dvourozměrné normální rozdělení**. Tento předpoklad můžeme orientačně ověřit např. pomocí bodového grafu a 95% konfidenční elipsy.



Na základě bodového grafu a 95% konfidenční elipsy není patrné výrazné porušení dvourozměrné normality – Pearsonův korelační koeficient využít můžeme.

*Pozn.: Pearsonův korelační koeficient je citlivý hlavně na **zešikmení rozložení dat** a na přítomnost **odlehle hodnoty**, při ověřování podmínek, by tedy měl být kladen důraz především na tyto předpoklady. Při výrazném porušení těchto podmínek, je vhodnější využít **neparametrickou alternativu** – Spearmanův korelační koeficient (viz kapitola 8.2).*

Návod pro software Statistica: Graphs – Scatterplot – záložka Advanced – zatrhneme políčko Normal u Elipse.

3. Testování nulové hypotézy (H_0 : Obvod pasu a obvod boků u respondentů jsou nezávislé náhodné veličiny) pomocí Pearsonova korelačního koeficientu.

Pearsonův korelační koeficient	P-hodnota Pearsonova korelačního koeficientu
0,359	< 0,001

Závěr: Na základě Pearsonova korelačního koeficientu **zamítáme nulovou hypotézu** ($p < 0,05$), obvod pasu a obvod boků jsou závislé náhodné veličiny. Hodnota Pearsonova korelačního koeficientu značí kladnou korelaci (čím je větší obvod pasu, tím je větší obvod boků).

Návod pro software Statistica: Statistics – Basic statistics – Correlation matrices.

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 zjistěte, zda existuje vztah mezi výškou (proměnná *vyska*) a vahou respondentů (proměnná *vaha*).
- 2. příklad:** Na hladině významnosti 0,05 zjistěte, zda existuje vztah mezi hodnotami WHR – poměr obvodu pasu ku bokům (proměnná *WHR*) a BMI – index tělesné hmotnosti (proměnná *BMI*).
- 3. příklad:** Na hladině významnosti 0,05 zjistěte, zda existuje vztah mezi hodnotami systolického tlaku (proměnná *syst_tlak*) a diastolického tlaku (proměnná *diast_tlak*).
- 4. příklad:** Na hladině významnosti 0,05 zjistěte, zda existuje vztah mezi věkem respondentů (proměnná *vek*) a hodnotou cholesterolu (proměnná *cholesterol*).

Řešení:

1. příklad: Z bodového grafu a 95% konfidenční elipsy není patrné výrazné porušení dvourozměrné normality. Na základě Pearsonova korelačního koeficientu **zamítáme nulovou hypotézu** (p -hodnota $< 0,001$), výška a váha jsou

závislé náhodné veličiny. Hodnota Pearsonova korelačního koeficientu (0,379) značí kladnou korelaci (čím je vyšší výška respondenta, tím je vyšší váha respondenta).

2. příklad: Z bodového grafu a 95% konfidenční elipsy není patrné výrazné porušení dvourozměrné normality. Na základě Pearsonova korelačního koeficientu **nezamítáme nulovou hypotézu** (p-hodnota: 0,094), že WHR a BMI jsou nezávislé náhodné veličiny. Hodnota Pearsonova korelačního koeficientu je 0,169.

3. příklad: Z bodového grafu a 95% konfidenční elipsy není patrné výrazné porušení dvourozměrné normality. Na základě Pearsonova korelačního koeficientu **zamítáme nulovou hypotézu** (p-hodnota < 0,001), systolický tlak a diastolický tlak jsou závislé náhodné veličiny. Hodnota Pearsonova korelačního koeficientu (0,656) značí kladnou korelaci (čím je vyšší hodnota systolického tlaku, tím je vyšší hodnota diastolického tlaku).

4. příklad: Z bodového grafu a 95% je patrná přítomnost odlehle hodnoty. Využití Pearsonova korelačního koeficientu tedy není vhodné, neboť interpretace by mohla být zkreslená. V tomto případě je vhodnější využít neparametrickou alternativu – Spearmanův korelační koeficient (viz kapitola 8.2).

7.2. Spearmanův korelační koeficient

Řešený příklad

Na hladině významnosti 0,05 zjistěte, zda existuje vztah mezi věkem respondentů (proměnná *vek*) a hodnotou cholesterolu (proměnná *cholesterol*).

Řešení:

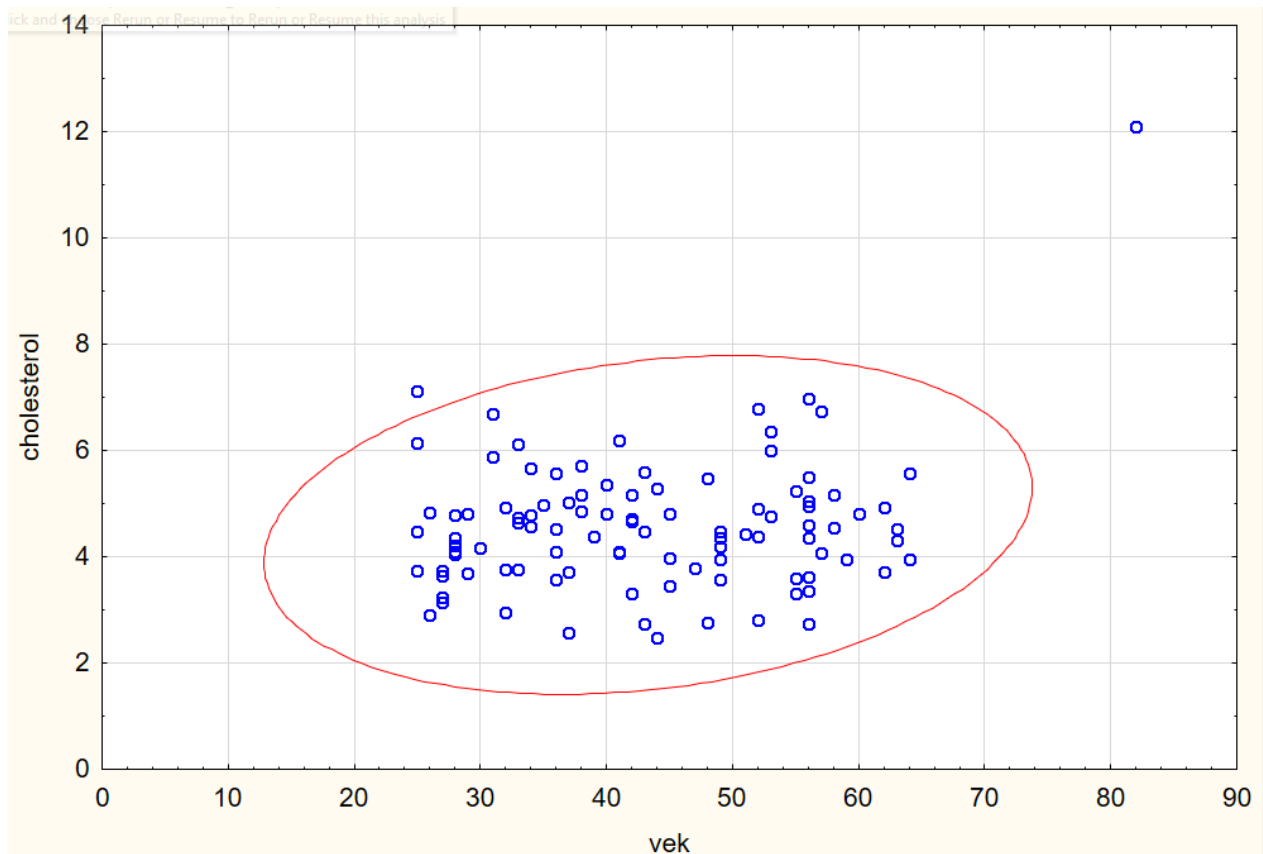
1. Stanovení nulové a alternativní hypotézy.

H_0 : Věk respondentů a hodnota cholesterolu jsou **nezávislé** náhodné veličiny. (Korelační koeficient **se rovná 0**.)

H_A : Věk respondentů a hodnota cholesterolu jsou **závislé** náhodné veličiny. (Korelační koeficient **se nerovná 0**.)

2. Důvod využití Spearmanova korelačního koeficientu.

Personův korelační koeficient je velmi citlivý na zešikmení rozložení dat a přítomnost odlehle hodnoty.



Z bodového grafu je patrná přítomnost odlehlé hodnoty. Pro testování závislosti mezi věkem respondentů a hodnotou cholesterolu je v tomto případě vhodnější Spearmanův korelační koeficient.

3. Testování nulové hypotézy (H_0 : Věk respondentů a hodnota cholesterolu jsou nezávislé náhodné veličiny) pomocí Spearmanova korelačního koeficientu.

Spearmanův korelační koeficient	P-hodnota Spearmanova korelačního koeficientu
0,081	0,424

Závěr: Na základě Spearmanova korelačního koeficientu **nezamítáme nulovou hypotézu** ($p > 0,05$), že věk respondentů a hodnota cholesterolu jsou nezávislé náhodné veličiny.

Pozn.: Kdybychom v tomto případě využili Pearsonův korelační koeficient, nulovou hypotézu bychom zamítli ($r = 0,224$; p -hodnota: $0,025$).

Návod pro software Statistica: Statistics – Nonparametrics– Correlations.

Příklady k procvičení

- 1. příklad:** Na hladině významnosti 0,05 zjistěte, zda existuje vztah mezi hodnotou glukózy v krvi (proměnná *glykemie*) a systolickým tlakem respondentů (proměnná *syst_tlak*).

- 2. příklad:** Na hladině významnosti 0,05 zjistěte, zda existuje vztah mezi hodnotou WHR – poměr obvodu pasu ku bokům (proměnná *WHR*) a hodnotou cholesterolu (proměnná *cholesterol*).
- 3. příklad:** Na hladině významnosti 0,05 zjistěte, zda existuje vztah mezi BMI – index tělesné hmotnosti (proměnná *BMI*) a hodnotou cholesterolu (proměnná *cholesterol*).

Řešení:

1. příklad: V tomto případě není vhodné využít Personův korelační koeficient, neboť rozložení glukózy v krvi je značně zešikmené. Na základě Spearmanova korelačního koeficientu **nezamítáme nulovou hypotézu** (p-hodnota: 0,918), že glukóza v krvi a systolický tlak jsou nezávislé náhodné veličiny. Hodnota Spearmanova korelačního koeficientu je 0,011.
2. příklad: V tomto případě není vhodné využít Personův korelační koeficient, neboť je přítomná odlehlá hodnota. Na základě Spearmanova korelačního koeficientu **zamítáme nulovou hypotézu** (p-hodnota: 0,028), WHR a cholesterol jsou závislé náhodné veličiny. Hodnota Spearmanova korelačního koeficientu (0,220) značí kladnou korelaci (čím je vyšší hodnota WHR, tím je vyšší hodnota cholesterolu).
3. příklad: V tomto případě není vhodné využít Personův korelační koeficient, neboť je přítomná odlehlá hodnota. Na základě Spearmanova korelačního koeficientu **zamítáme nulovou hypotézu** (p-hodnota < 0,001), BMI a cholesterol jsou závislé náhodné veličiny. Hodnota Spearmanova korelačního koeficientu (0,532) značí kladnou korelaci (čím je vyšší hodnota BMI, tím je vyšší hodnota cholesterolu).