

Danka Haruštiaková, Institut biostatistiky a analýz

# ANALÝZA A MANAGEMENT DAT PRO ZDRAVOTNICKÉ OBORY

Učební texty pro studenty kombinovaného studia

Brno, 2017

## ZADÁVÁNÍ A PŘÍPRAVA DAT

Statistická analýza se neobejde bez zpracování dat pomocí statistických software. Předpokladem úspěchu je správné uložení dat umožňující jejich zpracování v libovolné aplikaci.

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 pacient, nebo jednotlivá vyšetření).

Pro databázovou tabulku platí následovně:

- Každý sloupec obsahuje pouze jediný typ dat, identifikovaný hlavičkou sloupce.
- Každý řádek obsahuje minimální jednotku dat (např. pacient, jedna návštěva pacienta apod.)
- Není přípustné kombinovat v jednom sloupci číselné a textové hodnoty.
- Komentáře jsou uloženy v samostatných sloupcích.
- U textových dat je nezbytné kontrolovat překlepy 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 MS Office Excel (Tab. 1).

Tabulka 1. Ukázka databázové tabulky. V řádcích jsou pacienti, v sloupcích jednotlivé proměnné. V prvním řádku jsou uvedeny názvy proměnných. ID pacienta = jednoznačný identifikační údaj pacienta. Atopický exém 1 = přítomnost lézí atopického exému před léčbou. Atopický exém 2 = přítomnost lézí atopického exému po léčbě.

Tep. frekv. 1 = tepová frekvence před léčbou. Tep. frekv. 2 = tepová frekvence po léčbě.

ID pacienta	Pohlaví	Atopický exém 1	Atopický exém 2	Tep. frekv. 1	Tep. frekv. 2
001-09-01-P	Muž	ANO	ANO	85	82
002-28-05-P	Muž	ANO	ANO	98	85
003-28-06-P	Muž	ANO	ANO	99	89
004-28-07-P	Muž	ANO	NE	79	81
005-28-08-P	Muž	ANO	ANO	86	83
007-28-10-P	Muž	NE	NE	85	86
008-28-11-P	Muž	NE	ANO	80	80
009-28-12-P	Muž	NE	NE	92	89
011-28-14-P	Muž	NE	NE	95	90
013-28-16-P	Muž	NE	NE	91	91
016-07-01-P	Muž	NE	NE	80	82
017-07-02-P	Muž	NE	NE	81	78
018-07-03-P	Žena	ANO	NE	85	81
019-07-04-P	Žena	ANO	ANO	93	89
020-07-05-P	Žena	ANO	ANO	88	89
021-02-26-P	Žena	ANO	ANO	90	91
022-02-27-P	Žena	ANO	ANO	85	82
023-02-28-P	Žena	ANO	ANO	104	95
024-02-29-P	Žena	ANO	NE	98	93
025-02-30-P	Žena	ANO	NE	97	89
026-02-31-P	Žena	ANO	ANO	85	83
027-02-32-P	Žena	ANO	ANO	95	89
028-10-26-P	Žena	ANO	ANO	97	96
029-09-14-P	Žena	NE	ANO	102	95
030-09-15-P	Žena	NE	NE	90	85
031-09-16-P	Žena	NE	NE	87	83
032-02-34-P	Žena	NE	NE	83	82
033-07-06-P	Žena	NE	NE	87	84

## TYPY DAT

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

Rozlišujeme tyto typy proměnných:

- *Kvalitativní (kategoriální) 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 a pod. Pro zobrazení zastoupení kategorií v souboru je vhodný výsečový - koláčový graf (Obr. 1a)

Kvalitativní proměnnou lze dále dělit do následujících skupin:

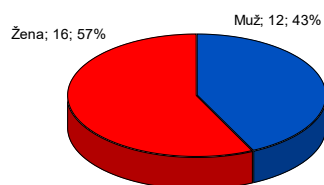
- *Binární proměnná*, která nabývá pouze dvou hodnot. Tyto většinou kódujeme pomocí číslic 1 (přítomnost sledovaného znaku) a 0 (nepřítomnost sledovaného znaku). Příkladem binární proměnné je např. diabetes (1 - ano, 0 - ne), pohlaví (1 - muž, 0 - žena).
- *Nominální proměnná* obsahuje několik kategorií, 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, a pod. Pro zobrazení zastoupení hodnot v souboru je vhodný histogram (Obr. 1b).

Kvantitativní proměnná může být:

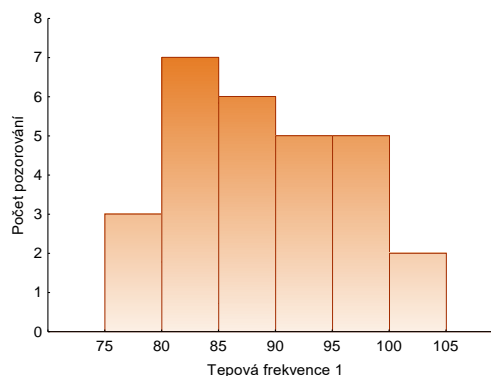
- *Spojité proměnná*, která nabývá jakýchkoliv hodnot v určitém rozmezí. Příkladem spojitých dat je výška a hmotnost osob, délka časového období od narození po výskyt nemoci, teplota a pod.
- *Diskrétní proměnná* může nabývat pouze spočetně mnoha hodnot. Příkladem diskrétních dat je počet krevních buněk v 1 ml krve, počet krvácivých epizod, počet dětí v rodině a pod.

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). *Poměrová stupnice* nám umožňuje ptát i na otázku, kolikrát je jedna hodnota větší než druhá. Příkladem poměrových dat jsou např. výška a váha osob, počet krevních buněk v 1 ml krve.

a



b



Obrázek 1. Ukázka zobrazení kvalitativní proměnné "pohlaví" ve výsečovém - koláčovém grafu (a) a kvantitativní proměnné "tepová frekvence 1" v histogramu (b).

## STATISTICKÁ SUMARIZACE

Prvním krokem jakékoliv analýzy bez ohledu na typ dat by měla být sumarizace a vizualizace pozorovaných hodnot. K sumarizaci pozorovaných hodnot slouží metody *popisné statistiky*. Tyto jsou potřebné ke shrnutí základních informací o datech. Objektivně není možné pouhým pohledem na data odečíst potřebné informace ze souboru např. 1000 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.

### Popis kvalitativních dat

U kvalitativních dat předpokládáme, že se jednotlivé hodnoty opakují. Proto je logické je sumarizovat pomocí tabulky s četnostmi možných hodnot (*tabulka četností, frekvenční tabulka*). V takové tabulce uvádíme pozorovanou četnost a relativní četnost. Tabulku četností proměnné "pohlaví" z Tabulky 1 uvádíme v Tabulce 2.

Tabulka 2. Tabulka četností proměnné "pohlaví" z Tabulky 1.

Pohlaví	Absolutní četnost (počet)	Relativní četnost (%)
Muž	12	42,9
Žena	16	57,1
Celkem	28	100

Vzhledem k tomu, že kvalitativní data nelze seřadit dle velikosti, používá se jako frekvenční charakteristika těchto dat tzv. *mód*, nebo *modus*. Je to kategorie kvalitativní proměnné s nejvyšší četností. V našem příkladu z Tab. 2 je modus 'žena' (tj. kategorie s nejvyšší četností).

### Popis kvantitativních dat

Nejznámější statistikou, která shrnuje soubor dat do jednoho čísla a představuje typickou hodnotu, kolem které mají ostatní hodnoty tendenci kolísat, je *průměr*. Průměr lze jednoduše spočítat jako součet pozorovaných hodnot dělený jejich počtem.

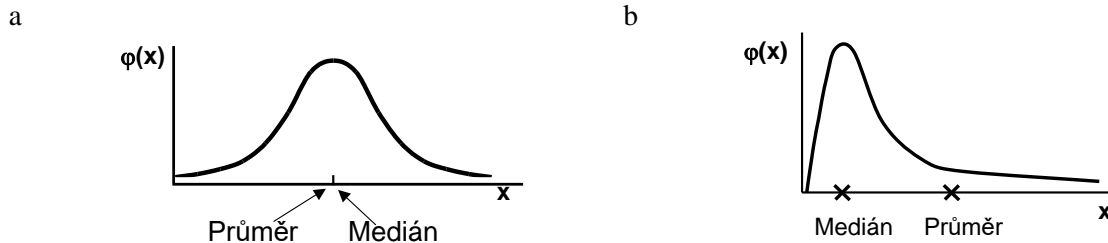
Dalším ukazatelem střední hodnoty souboru je *medián*, který definujeme jako prostřední pozorovanou hodnotu souboru. Medián dělí celý soubor na dvě poloviny, tedy polovina souboru je menší než medián a polovina souboru je větší než medián.

Medián je dobrým odhadem frekvenčního středu dat. Průměr ním 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 nimi ovlivněn není. V případě symetrického rozdělení jsou průměr a medián v podstatě shodné (Obr. 2a), v případě asymetrického rozdělení však nikoliv (Obr. 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.

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

Popisné statistiky proměnných "tepová frekvence 1" a "tepová frekvence 2" z příkladu v Tabulce 1 uvádíme v Tabulce 3.



Obrázek 2. Ukázka pozice průměru a mediánu u proměnné se symetrickým rozdělením (a) a proměnné s vysokým zastoupením nízkých hodnot v histogramu (b).

Tabulka 3. Popisné statistiky proměnných "tepová frekvence 1" a "tepová frekvence 2" z Tabulky 1.

Proměnná	Počet hodnot	Průměr	Medián	Minimum	Maximum	Sm. odch.
Tepová frekvence 1	28	89,9	89,0	79	104	7,0
Tepová frekvence 2	28	86,5	85,5	78	96	5,0

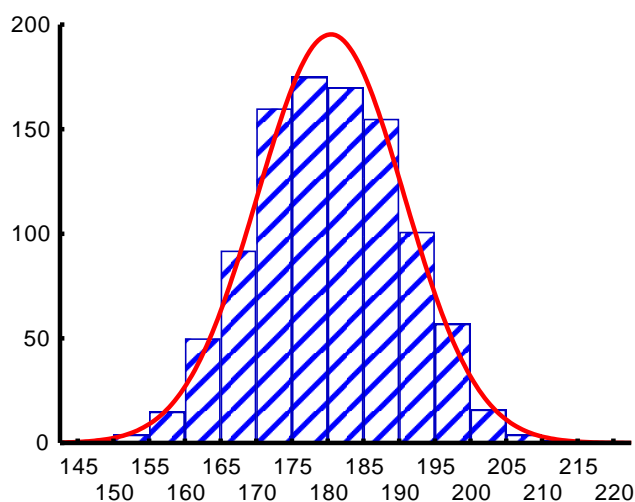
## NORMÁLNÍ ROZDĚLENÍ

Mnoho znaků v přírodě i medicíně se chová podle jistých pravidel. Toto chování proměnných lze popsat rozdělením pravděpodobnosti. Nejklasičtější modelovým rozdělením, od něhož je odvozena celá řada statistických analýz je tzv. *normální rozdělení*, známé též jako Gaussova křivka. Normální rozdělení se vyznačuje symetrií a nepřítomností odlehlých hodnot (Obr. 3). Je kompletně popsáno dvěma parametry - průměrem a rozptylem.

Normalita je klíčovým předpokladem mnoha statistických metod. Proto je v biostatistice vždy potřebné zjistit, zda sledovaná proměnná má nebo nemá normální rozdělení.

Pro ověření normality existuje řada testů a grafických metod. Nejčastěji se k testování normality rozdělení používají *Kolmogorov-Smirnov test* a *Shapiro-Wilkov test*. Výhodou Shapiro-Wilkova testu je, že jde o neparametrický test použitelný i při velmi malých počtech hodnot  $n$  s dobrou silou testu. Je zaměřený na testování symetrie.

Všechny testy normality pracují s nulovou hypotézou, že není rozdíl mezi pozorovaným rozdělením a normálním rozdělením.



Obrázek 3. Ukázka normálního rozdělení.

## 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:

- 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 a pod.),
- porovnat výběrovou charakteristiku dvou náhodných výběrů mezi sebou,
- zhodnotit změnu v hodnotách sledované proměnné,
- rozhodnout o nezávislosti dvou kvalitativních proměnných,
- rozhodnout o charakteru rozdělení pravděpodobnosti proměnné.

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 rozdílu mezi skupinami. Jinak řečeno, nulová hypotéza odráží fakt, že se něco nestalo nebo neprojevalo, 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  $\alpha$ ) a chyba II. druhu (její pravděpodobnost značíme jako  $\beta$ ). 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.

Rozhodnutí o zamítnutí nebo nezamítnutí nulové hypotézy činíme dle tzv. *p-hodnoty*. Jednoduše řečeno, vyjde-li nám *p-hodnota* nižší než stanovená hladina významnosti  $\alpha$  (standardně je přijímána hranice 5%), znamená to, že nulovou hypotézu můžeme zamítnout. Je-li *p-hodnota* větší než stanovená hladina významnosti  $\alpha$ , pak nulovou hypotézu nezamítáme. *P-hodnotu* získáme při testování hypotéz ve statistickém softwaru.

Statistické testování probíhá následovně:

1. Formulujeme nulovou hypotézu  $H_0$ , která tvrdí, že sledovaný efekt je nulový. Např. v našem případě  $H_0$ : tepová frekvence u mužů a žen se neliší. Zároveň zvolíme hladinu významnosti testu  $\alpha$  (nejčastěji  $\alpha = 0,05$ ), která představuje pravděpodobnost získání falešně pozitivního výsledku.
2. Formulujeme alternativní hypotézu  $H_A$ . V našem případě  $H_A$ : tepová frekvence 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  $\alpha$ , zamítáme nulovou hypotézu  $H_0$ . Pokud není *p-hodnota* menší než  $\alpha$ , pak  $H_0$  nezamítáme. V našem příkladu o tepové frekvenci mužů a žen byla výsledná *p-hodnota* větší než 0,05, tudíž nezamítáme nulovou hypotézu, tj. tepová frekvence mužů a žen se neliší.

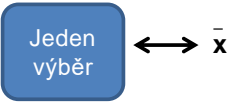
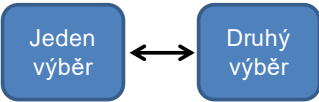
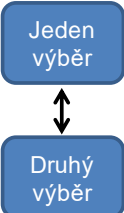
## PŘEHLED STATISTICKÝCH TESTŮ

Nejdříve se budeme 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 spojitě proměnné. Podobně můžeme za spojitou proměnnou považovat i tepovou frekvenci z našeho příkladu.

Statistické testy můžeme rozdělit na dvě velké skupiny: *parametrické testy* a *neparametrické testy* (Obr. 4). 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 odlehkým hodnotám. Nevýhodou neparametrických testů je ovšem jejich snížená síla testu, tj. snížená schopnost zamítnout neplatnou nulovou hypotézu, což je následně nutné kompenzovat větší velikostí vzorku. To může být problém hlavně v případě menších souborů.

Dále rozlišujeme testy jednovýběrové a dvouvýběrové. Testy jednovýběrové srovnávají jeden výběr s referenční hodnotou. Dvouvýběrové testy srovnávají dva výběry 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. 4).

	Parametrický test (data mají normální rozdělení)	Neparametrický test (data nemají normální rozdělení)
<b>Jeden výběr</b> (srovnání s referenční hodnotou) 	Jednovýběrový t-test	
<b>Dva závislé výběry</b> (párově uspořádaná data) 	Dvouvýběrový t-test pro závislé výběry (párový t-test)	Wilcoxonův test Znaménkový test
<b>Dva nezávislé výběry</b> (nepárově uspořádaná data) 	Dvouvýběrový t-test pro nezávislé výběry (nepárový t-test)	Mannův-Whitneyův U test

Obrázek 4. Schéma statistických testů pro hodnocení kvantitativních dat.



## Jednovýběrový t-test

V případě jednovýběrových testů jde o srovnání výběru dat (tedy jedné proměnné, jednoho sloupce datové tabulky) s cílovou populací (tj. s referenční hodnotou). Jednovýběrový test se nazývá t-test a je parametrický, tj. sledovaná proměnná musí mít normální rozdělení.

Podívejme se na data z příkladu v Tabulce 1, kde nás může zajímat tepová frekvence před léčbou a její srovnání s referenční hodnotou 85 tepů za minutu. Nulová a alternativní hypotéza jsou v našem případě definovány následovně:

$H_0$ : Průměrná hodnota tepové frekvence před léčbou se neliší od referenční hodnoty 85 ( $\text{min}^{-1}$ ).

$H_A$ : Průměrná hodnota tepové frekvence před léčbou se liší od referenční hodnoty 85 ( $\text{min}^{-1}$ ).

## Dvouvýběrový nepárový t-test

Pro srovnání dvou nezávislých souborů dat se využívá dvouvýběrový t-test. Je to parametrický test a k jeho předpokladům patří:

- nezávislost obou srovnávaných výběrů
- normální rozdělení obou srovnávaných výběrů (normalita je testována testy normality)
- přibližně shodný rozptyl v obou výběrech (shodnost rozptylu měříme tzv. F-testem)

Jako příklad použití nepárového t-testu můžeme použít srovnání hodnot tepové frekvence před léčbou u mužů a žen (Tab. 1). Nulová a alternativní hypotéza jsou definovány následovně:

$H_0$ : Průměrná hodnota tepové frekvence před léčbou se neliší mezi muži a ženami.

$H_A$ : Průměrná hodnota tepové frekvence před léčbou se liší mezi muži a ženami.

## Dvouvýběrový párový t-test

Pro srovnání dvou závislých souborů dat, kdy jsou skupiny dat spojeny přes objekt měření, se využívá dvouvýběrový párový t-test. Oba soubory musí mít shodný počet hodnot, 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 v obou souborech. Párový t-test je parametrický test. Nemá žádné předpoklady o rozdělení vstupních dat, protože je počítán až na základě jejich diferencí. Tyto difference by měly mít normální rozdělení a otázkou v párovém t-testu je, zda se průměrná hodnota diferencí rovná nějakému číslu, typicky jde o srovnání s nulou jako důkaz neexistence změny mezi oběma spárovanými skupinami.

Příkladem použití párového t-testu je srovnání změny tepové frekvence v závislosti od léčby, kdy srovnáme proměnnou tepová frekvence 1 a tepová frekvence 2 (Tab. 1). Nulová a alternativní hypotéza jsou definovány následovně:

$H_0$ : V průběhu léčby nedošlo ke změně tepové frekvence.

$H_A$ : V průběhu léčby došlo ke změně tepové frekvence.

## Mannův-Whitneyův U test

Mannův-Whitneyův U test je neparametrickou alternativou nepárového t-testu. Stejně jako řada jiných neparametrických testů počítá i tento test s pořadím dat ve výběrech namísto 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 průměru 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 by měla pořadí 7,5). Filozofie testu spočívá v tom, že pokud budou původní výběry podobné, potom bude součet pořadí podobný pro oba výběry.

Nulová a alternativní hypotéza jsou definovány stejně jako u parametrického nepárového t-testu.



## Wilcoxonův test a Znaménkový test

Wilcoxonův test i znaménkový test jsou neparametrickou alternativou párového t-testu. Wilcoxonův test má větší sílu testu, je ovšem použitelný pouze pro dostatečně velký výběr (alespoň 30 objektů, tj. např. pacientů). Znaménkový test můžeme použít pro data s  $n > 10$ . Z párových testů má nejmenší sílu testu.

Nulová a alternativní hypotéza jsou definovány stejně jako u parametrického párového t-testu.

## TESTOVÁNÍ HYPOTÉZ O KVALITATIVNÍCH PROMĚNNÝCH

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 proměnnými, které mohou nabývat pouze omezený počet hodnot (např. pohlaví: muž/žena, výskyt atopického exému: ano/ne, krevní skupina: A/B/AB/O).

Stejně jako v případě kvantitativních proměnných, tak i u kvalitativních proměnných můžeme hodnotit vztahy mezi nimi, např. zda spolu souvisí výskyt dvou proměnných.

Vztah dvou kategoriálních proměnných lze sumarizovat v kontingenční tabulce, která zobrazuje četnosti kombinací dvou kategoriálních proměnných (Tab. 4). Speciálním případem kontingenční tabulky je čtyřpolní tabulka pro hodnocení vztahu proměnných popisujících dva binární znaky.

Tabulka 4. Kontingenční tabulka proměnných "pohlaví" a "atopický exém 1" z Tabulky 1.

Pohlaví	Atopický exém 1 ANO	Atopický exém 1 NE	Celkem
Muž	5	7	12
Žena	11	5	16
Celkem	16	12	28

Kontingenční tabulka umožňuje testování následujících hypotéz:

1. *Hypotéza o nezávislosti.* Pomocí testu nezávislosti můžeme rozhodnout, zda spolu souvisí výskyt dvou kvalitativních proměnných (např. pohlaví a výskyt atopického exému). Hlavním testem nezávislosti pro kontingenční tabulku je *Pearsonův chí-kvadrát test*.
2. *Hypotéza o shodnosti struktury (test homogeneity).* O testování homogenity mluvíme v situaci, kdy nás zajímá výskyt kvalitativní proměnné u  $r$  nezávislých výběrů z  $r$  různých populací (např. typy nežádoucích účinků v několika  $r$  nemocnicích). K hodnocení shodnosti struktury používáme znovu *Pearsonův chí-kvadrát 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 stavu stromů ve dvou po sobě jdoucích sezónách, nebo i hodnocení výskytu lézí atopického exému před a po léčbě. K testování symetrie používáme *McNemarův test*.

## Pearsonův chí-kvadrát test, Fisherův exaktní test

Pearsonův chí-kvadrát test je základním a nejpoužívanějším testem nezávislosti i homogenity. Nulovou hypotézou je zde tvrzení, že proměnné jsou nezávislé, resp. homogenní. Pearsonův chí-kvadrát test není možné použít ve všech případech, ale pouze za splnění předpokladů testu:

- Jednotlivá pozorování shrnutá v kontingenční tabulce jsou nezávislá, tedy každý prvek je zahrnut pouze v jedné buňce kontingenční tabulky.
- 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 podmínka dobré aproximace není splněna, máme alternativu v podobě *Fisherova exaktního testu*, který lze použít i při velmi malých pozorovaných i očekávaných četnostech. Tento test je ovšem v běžných statistických softwarech k dispozici pouze pro čtyřpolní tabulky.

Jako příklad hodnocení nezávislosti dvou kvalitativních proměnných můžeme uvažovat hodnocení nezávislosti pohlaví a výskytu atopického ekému (Tab. 4). Nulová a alternativní hypotéza jsou definovány následovně:

$H_0$ : Pohlaví a výskyt atopického ekému jsou nezávislé proměnné.

$H_A$ : Pohlaví a výskyt atopického ekému jsou závislé proměnné.

### McNemarův test

McNemarův test je test pro kontingenční tabulku v případě párového uspořádání experimentu, kdy sledujeme výskyt kvalitativní 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.

Příkladem hodnocení symetrie kvalitativní proměnné bude v našem případě zhodnocení výskytu atopického ekému před léčbou a po léčbě. Nulová a alternativní hypotéza jsou definovány následovně:

$H_0$ : Výskyt atopického ekému je před a po léčbě stejný.

$H_A$ : Výskyt atopického ekému není před a po léčbě stejný.