

Testování hypotéz

Základní pojmy

Zdroje

- Discovering statistics with SPSS (Andy Field)
 - Nejlepší učebnice ever
 - Srozumitelné, s příklady, včetně vysvětlení, jak počítat v SPSS
 - Existuje i verze pro R
 - Cca 800 stran, nečtěte najednou
 - <http://www.soc.univ.kiev.ua/sites/default/files/library/elopen/andy-field-discovering-statistics-using-spss-third-edition-20091.pdf>
 - Dnes nejnovější 5. edice
- Khan academy:
 - Videá
 - Jasně vysvětlené, ale hodně orientované na matematickou teorii
 - <https://www.khanacademy.org/math/statistics-probability>

Pojmy potřebné pro testování hypotéz

- Standardní chyba
- Nulová a alternativní hypotéza
- Chyba 1. a 2. druhu
- P hodnota
- Kritický a normální obor, hladina/hranice kritického oboru

Standardní chyba

- Odhad odchyly parametru vzorku od parametru populace
- Např. odhad odchyly průměrného platu vzorku od průměrného platu celé populace

- $SE = \frac{\text{Směrodatná odchylnka}}{\sqrt{\text{Počet pozorování}}} = \sqrt{\frac{SD^2}{N}}$

- Vyjadřuje „statistický šum“ ve vzorku

Nulová a alternativní hypotéza

- Hypotézy nelze potvrzovat, pouze zamítat (protože Popper)
- Nulová hypotéza (H_0) = tvrzení, které se testováním snažíme zamítnout
- Alternativní hypotéza (H_1) = opak/negace nulové hypotézy
- Př.:
 - H_0 = Neexistuje rozdíl v průměrném příjmu mužů a žen
 - H_1 = Existuje rozdíl v průměrném příjmu mužů a žen

Chyba prvního a druhého druhu

- Chyba 1. druhu
 - Zamítnutí platné nulové hypotézy -> falešně pozitivní výsledek
 - Pravděpodobnost výskytu nastavitelné
 - Vyjádřená hladinou spolehlivosti (α)
- Chyba 2. druhu
 - Nezamítnutí neplatné nulové hypotézy -> falešně negativní výsledek
 - Pravděpodobnost výskytu pouze odhadnutelná, závislá na použitém test
 - Vyjádřená silou testu (β)

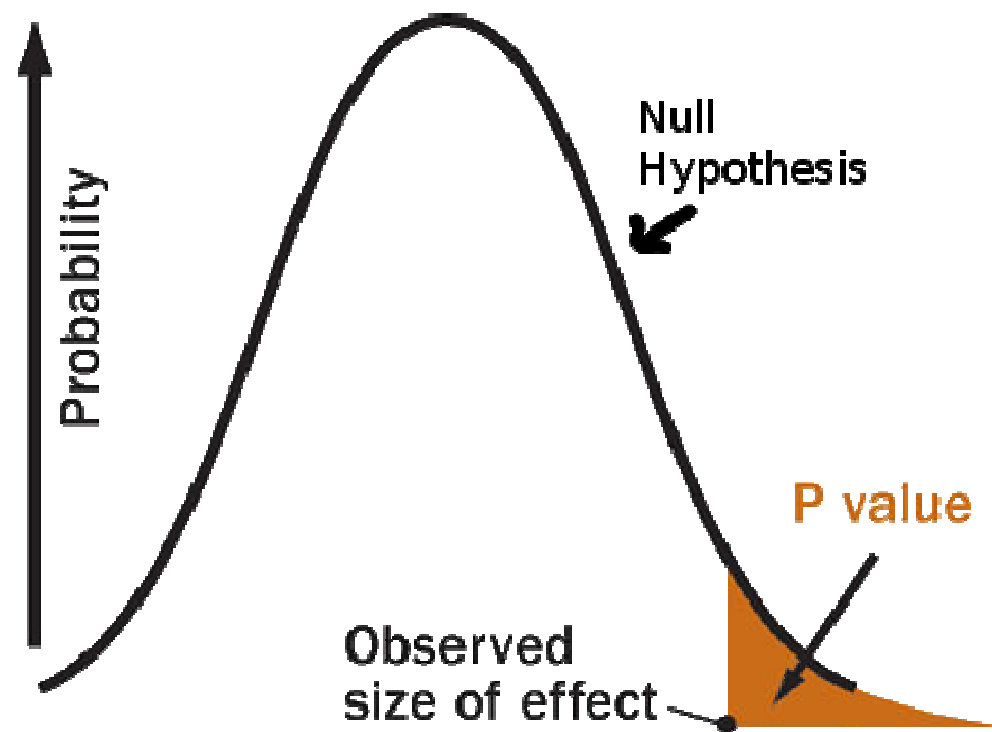
Chyba prvního a druhého druhu - příklady

- Chyba 1. druhu:
 - Prohlásíme, že existuje vztah mezi vakcinací dětí a onemocněním autismem, i když ve skutečnosti žádný není.
- Chyba 2. druhu:
 - Prohlásíme, že neexistuje vztah mezi vzděláním rodičů a příjmem, i když ve skutečnosti existuje.

P hodnota

- Pravděpodobnost, že testová statistika vyjde tak, jak vyšla nebo extrémněji, za předpokladu, že nulová hypotéza je pravdivá.
- Dá se interpretovat jako:
 - Míra toho, jak moc data sedí na nulovou hypotézu
 - (Čím menší hodnota, tím méně data sedí na nulovou hypotézu)
 - Míra toho, jak překvapující by byl konkrétní výsledek, kdybychom si mysleli, že nulová hypotéza je pravdivá
 - (Čím menší hodnota, tím více je výsledek překvapující)

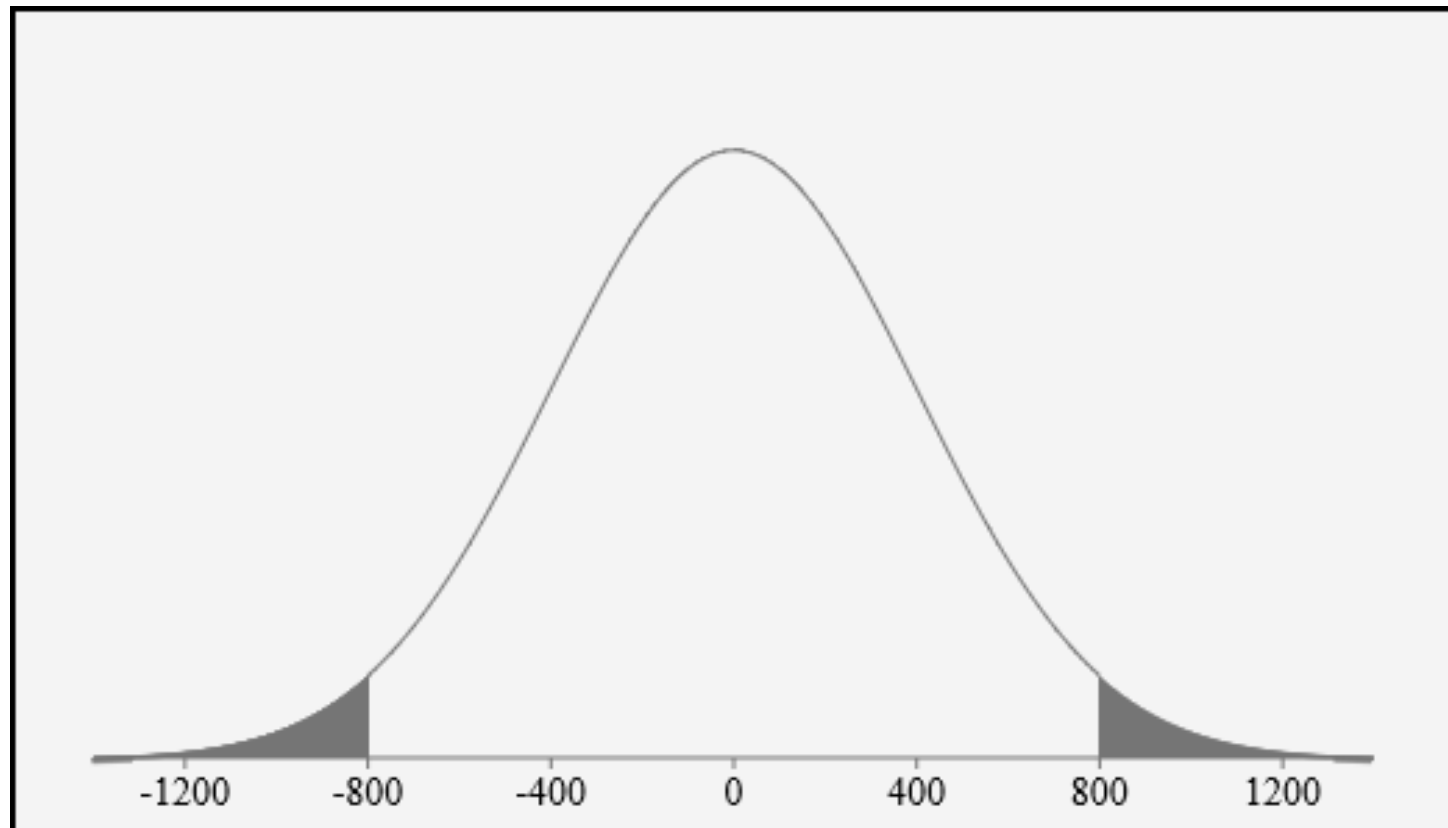
P hodnota



P hodnota – (zjednodušený) příklad

- Chceme zjistit zda se liší průměrný plat mužů a žen
- H_0 :
 - Rozdíl mezi průměrným platem mužů a žen (v populaci) je nulový
 - Prům. plat mužů – prům. plat žen = 0
- V našem vzorku:
 - Průměrný plat mužů = 33 000 Kč
 - Průměrný plat žen = 32 200 Kč
 - Rozdíl mezi platy = 33 000 – 32 200 = 800 Kč

P hodnota – (zjednodušený) příklad



P hodnota – (zjednodušený) příklad

- P hodnota je pravděpodobnost, že testová statistika vyjde tak, jak vyšla nebo extrémněji, za předpokladu, že nulová hypotéza je pravdivá.
- Existuje 4,6% pravděpodobnost, že rozdíl v platech vyjde 800* nebo více, za předpokladu, že rozdíl v průměrném platu mužů a žen v populaci je nulový.

*Ve skutečnosti se zohledňuje i standardní chyba, ale princip zůstává stejný

P hodnota

- P hodnota **není**:
 - Pravděpodobnost, že nulová hypotéza je pravdivá.
 - Pravděpodobnost, že zamítnutím nulové hypotézy uděláme chybu.
 - Pravděpodobnost, že vztah vznikl náhodou
 - Nezvratný důkaz o (ne)existenci vztahu
 - Ukazatel velikosti vztahu/rozdílu
- Nevýhody p hodnoty:
 - Velikost závislá na počtu pozorování:
 - U velkých vzorků vychází menší, než by měla
 - U malých vzorků vychází větší, než by měla
 - Zohledňuje pouze výběrovou chybu, tzn. ne bias
 - Nevypovídá o síle efektu, tzn. nerozlišuje mezi statistickou signifikancí a reálnou významností

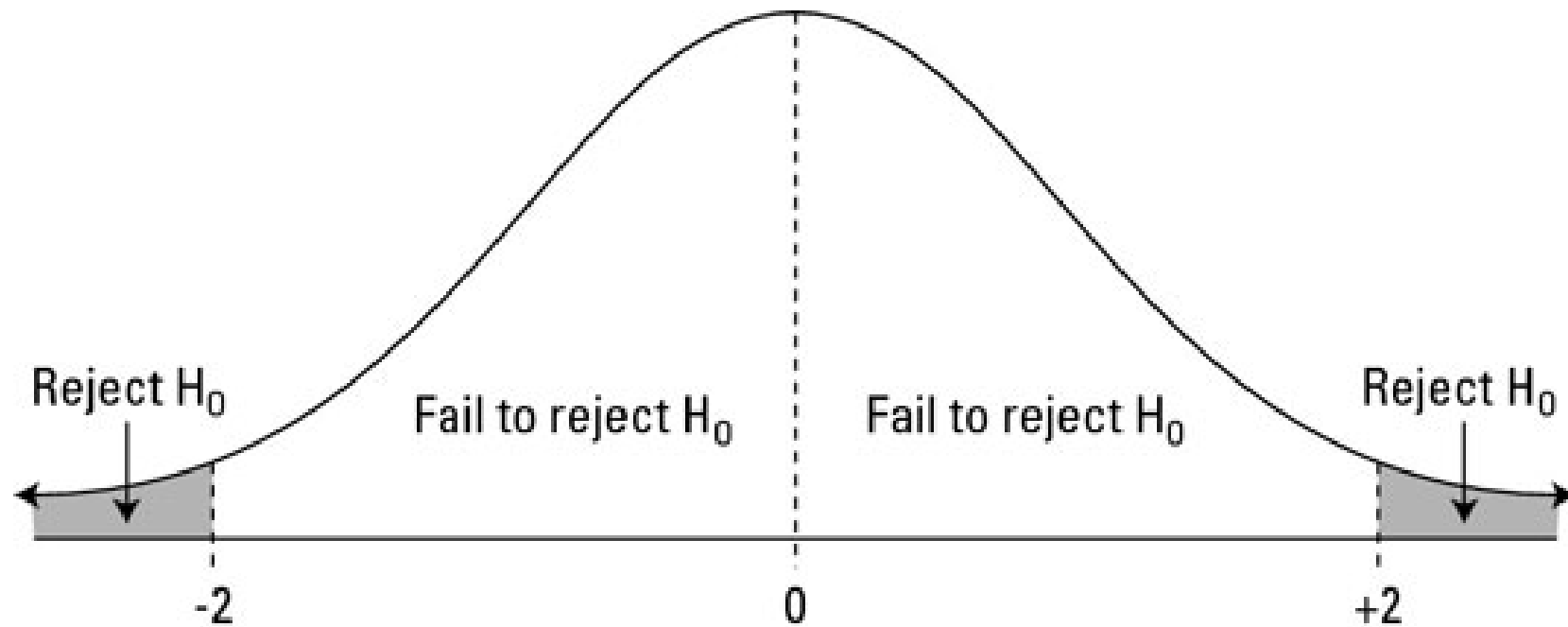
P hodnota

- Často špatně interpretovaná
- Nestačí spoléhat pouze na ní
- V analýze zohlednit:
 - P hodnoty
 - Intervaly spolehlivosti
 - Ukazatele síly efektu, např. korelační koeficient
 - Počet pozorování a způsob výběru
 - Kvalitu modelu

Kritický a normální obor

- Kritický obor = interval hodnot vedoucí k zamítnutí nulové hypotézy
- Normální obor = interval hodnot **ne**vedoucí k zamítnutí nulové hypotézy
- Hladina/hranice kritického oboru = hodnota rozdělující kritický a normální obor
 - Často označována jako hladina významnosti

Kritický a normální obor



Kritický a normální obor

- Hladina kritického oboru tradičně 0,05 (5%), případně 0,01 (1%)
- Hladina kritického oboru je určena arbitrárně, tzn. neexistuje objektivní důvod, proč právě 5%.
- Opatrně s interpretací p hodnot okolo hladiny kritického oboru
 - V praxi je respektovaná hladina 5%, ale technicky p hodnoty 5,5% a 4,5% jsou obě víceméně stejně špatné.