# Úvod

Aplikace se nachází na adrese

* <https://shiny-hypotheses-538722985971.europe-west6.run.app/>
* Pokud se aplikace prve načte „nepěkně“, stačí aktualizovat stránku:)

Zdrojový kód se nachází na

* <https://github.com/MichalLauer/ShinyHypotheses>

Aplikace funguje na zmíněném odkazu. Pokud by náhodou odkaz nefungoval, je aplikaci možné spustit na vlastním zařízení. Nejprve je nutné nainstalovat všechny balíčky ze soubor *global.R*. Může zde vzniknout problém s balíčkem {distr6}, který není na CRANu ale je nutné ho nainstalovat z GitHubu[[1]](#footnote-1) . Následně stačí stisknout zelené tlačítko v R Studiu v souboru *global.R.*

A blue and grey rectangular object

Description automatically generated with medium confidence

# Ovládání aplikace

Aplikace se ovládá pomocí levého panelu a jeho třech nabídek. Po zvolení parametrů stačí kliknout na tlačítko „Zapínáme pásy!“ a proběhne simulace.

## Nastavení dat

První nabídka obsahuje nastavení skutečných populací. Pokud je vyplněno rozdělení pouze pro první populaci, testy automaticky provedou pouze jednovýběrové simulace. V případě, že jsou zvolená dvě rozdělení, jsou automaticky vypočítány dvouvýběrové simulace. Rozdělení lze intuitivně definovat textem, např.:

* N(0,1) – normované normální rozdělení,
* N(10, 20) – normální rozdělení s průměrem 10 a rozptylem 20,
* LN(15, 15) – lognormální rozdělení s logmean = 15 a logvar = 15

Jako populační rozdělení lze zvolit libovolné rozdělení z balíčku {distr6} (Sonabend & Kiraly, 2024).

## Testování hypotéz

Druhá nabídka umožňuje více nastavit parametry hypotéz. Parametr značí velikost vzorku, hladinu významnosti a hodnotu . Dále lze nastavit, zda jsou data závislá a zda by se mělo uvažovat, že mají populace stejný rozptyl. Závěrem lze upravit parametry, které ovládají počet simulací. Parametr K slouží k určení počtu simulací pro chyby I. a II. typu a parametr B k počtu bootstrapových vzorků.

## Replikovatelnost

Pokud by byla potřeba replikovat výsledky, zde nastavit seed, díky kterému budou výsledky identické.

## Vzhled aplikace

Výsledky jsou zobrazené ve čtyřech různých záložkách, které se soustředí na různé typy testů. Každý panel má tři části. Hlavní graf ukazuje rozdělení, které je pro daný test vhodné. Levý dolní panel slouží k popisu výsledků a pravý dolní panel k výsledkům ze simulace. Graf je interaktivní, a tak po použití myši lze zjistit další informace o daném rozdělení či testu.

## Populace

První záložka ukazuje zvolené populace. Graf porovnává jedno či dvě zvolená rozdělení, levý spodní panel popisuje první zvolené rozdělení a pravý dolní panel druhé. V případě, že jsou zvolena dvě rozdělení, má graf dvě osy Y. To je z toho důvodu, aby šlo rozumně zobrazit různá rozdělení, která mohou nabývat mnohonásobě jiných hustot.

## Parametrický test

Druhá záložka zobrazuje výsledné parametrické testy. V případě že je zvolené jedno rozdělení se provede jednovýběrový t-test. Pokud jsou zadána dvě rozdělení, data nejsou závislá a populační rozptyl není stejný, je proveden Welchův dvouvýběrový t-test. V případě závislých dat je proveden párový t-test. Pokud jsou předpokládané populační rozptyly stejné, je proveden klasický dvouvýběrový t-test.

Hlavní obrázek ukazuje rozdělení testové statistiky, tedy studentovo rozdělení. Zelený bod ukazuje vypočítanou testovací statistiku a zvýrazněná červená plocha p-hodnotu. Po najetí myší lze zjistit o hodnotách více informací. Levý dolní panel ukazuje nulovou hypotézu, testovací statistiku a přesnou p-hodnotu. Pravý spodní roh ukazuje simulovanou chybu I. a II. typu a také sílu testu.

## Neparametrické testy

Po parametrických testech aplikace nabízí také testy neparametrické. V případě jedné zvolené populace je použit Wilcoxonův znaménkový test. U dvou populací s nezávislými daty je použit Wilcoxonův test pořadí. V případě závislých dat je znovu použit Wilcoxonův znaménkový test. Neparametrické testy nepracují vůbec s předpokladem stejného rozptylu, a proto zde není tato možnost použita.

Na grafu je díky aproximaci Wilcoxonových testů zobrazeno normované normální rozdělení. Zelený bod zobrazuje transformovanou testovací statistiku a červená plocha p-hodnotu. Levý dolní panel zobrazuje statistiky výsledného testu a pravý dolní panel pak simulované chyby a sílu testu.

## Boostrap test

Poslední dostupný test je bootstrapové testování hypotéz. Pro takový test je nutné nejprve vytvořit jeden vzorek, se kterým se bude pracovat. V případě jedné populace je použit vzorek , v případě dvou populací je vytvořen vzorek . Následně je pořízeno bootstrapových vzorků o velikost , ze kterých je pořízen průměr . Rozdělení těchto průměrů je zobrazeno na hlavním panelu pomocí histogramu.

Levý dolní panel zobrazuje nulovou hypotézu. Testovací statistika je v tomto případě pozorovaný průměr a p hodnota je vypočtena jako procento případů, ve kterých je absolutní odchylka generovaných průměrů od jejich celkového průměru větší než absolutní odchylka mezi pozorovaným průměrem a generovanými průměry . Pravý dolní panel zobrazuje simulované chyby I. a II. typu a také sílu testu.

# Výpočty

Před všemi výpočty jsou vytvořeny vzorky a (pokud je zadaná druhá populace), se kterými pracuje každý test. Výsledky jsou proto porovnatelné, jelikož vycházejí ze stejných dat.

## Parametrické testy

Pro parametrické a testy je použita funkce t.test ze základního R balíčku {stats} (R Core Team, 2024).

## Neparametrické testy

Pro neparametrické a testy je použita funkce wilcox.test ze základního R balíčku {stats} (R Core Team, 2024). Při výpočtu je použit parametr exact = FALSE, který místo exaktních testů používá normální aproximaci. Druhý parametr correct = FALSE vypíná opravou na spojitost.

Výsledky je ale dále nutné transformovat tak, aby měly normované normální rozdělení pro účely vizualizace. Pro normalizaci jednovýběrového testu nebo dvouvýběrového závislého testu je použita parametrizace (Bellera et al., 2010)

V transformaci značí počet vzorků různých od . V párovém testu značí počet rozdílu (), které jsou různé od .

U závislého dvouvýběrového testu je pak použito (Bellera et al., 2010)

Testové statistiky, které vrací R funkce, jsou následně transformovaný jako

a díky centrální limitní větě má transformace s rostoucím n normované normální rozdělení.

## Bootstrap testy

Pro otestování pomocí neparametrického bootstrapu je data nutné v případě dvou výběrů přetransformovat tak, aby byl vytvořen jeden vzorek jako . V případě jedné populace je použit vytvořený vzorek . Ze nového vzorku je vytvořeno bootstrapovaných vzorků, ze kterých je vypočítán průměr . Z průměrů a bootstrapovaného vzorku je vytvořena testovací statistika i p-hodnota.

## Chyba I. typu

Chyba I. typu určuje *pravděpodobnost, že je zamítnuta H0, která ve skutečnosti platí.* V případě parametrického i neparametrického testu je vypočítána tak, že se do funkcí t.test a wilcox.test zadá skutečný populační rozdíl a je sledováno, v kolika případech z  možných nových vzorků je zamítnuta.

V případě bootstrapu je vytvořen nový bootstrapový vzorek a jeho vycentrovaná verze jako

ve kterém platí . Z tohoto vzorku je vytvořeno nových vzorků, na kterých je vypočítán průměr. Následně je sledována pravděpodobnost, že absolutní odchylka vypočtených průměrů od jejich celkového průměru je větší než odchylka mezi pozorovaným průměrem a bootstrapovanými průměry. Pokud je taková pravděpodobnost menší než , nulová hypotéze je nekorektně zamítnuta. Tento proces se opakuje -krát.

## Chyba II. typu

Chyba II. typu hovoří o *pravděpodobnosti nezamítnutí nulové hypotézy, která ve skutečnosti neplatí.* U těchto výpočtů se předpokládá, že nulová hypotéza neplatí a je tedy na uživateli, aby ji správně nastavil. Ve funkcích t.test a wilcox.test je nastavena nulová hypotéza na hypotézu zvolenou uživatelem a je zkoumáno, v jakém procentu případů není nulová hypotéza, která neplatí, zamítnuta.

V případě bootstrapového testování se postupuje podobně. Je vytvořen nový bootstrapový vzorek jako (v případě jedné populace) nebo (v případě dvou populací). Následně je vytvořeno bootstrapovaných vzorků, ze kterých je vypočítán průměr . P-hodnota je zjištěna jako procento případů, ve kterých je absolutní odchylka vygenerovaných průměrů a jejich celkového průměru větší než absolutní odchylka mezi průměry a hypotetizovanou Toto je opakováno -krát.

## Síla testu

Síla testu je vypočítána jako .

# Reference

Sonabend, R., & Kiraly, F. (2024). distr6: The Complete R6 Probability Distributions Interface (R package version 1.8.4) [Computer software]. <https://xoopr.github.io/distr6/>

R Core Team. (2024). R: A Language and Environment for Statistical Computing [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>

Bellera, C. A., Julien, M., & Hanley, J. A. (2010). Normal Approximations to the Distributions of the Wilcoxon Statistics: Accurate to What *N*? Graphical Insights. *Journal of Statistics Education*, *18*(2), 1. <https://doi.org/10.1080/10691898.2010.11889486>

1. <https://github.com/xoopR/distr6> [↑](#footnote-ref-1)