

# Testování hypotéz

# Statistická hypotéza

- důležitý postup statistického usuzování
- na základě informací získaných z náhodných výběrů, zda přijmeme nebo zamítneme určitou hypotézu týkající se základního souboru
- **Statistickou hypotézou** rozumíme jakékoliv tvrzení, které se může týkat neznámých parametrů, tvaru rozdělení a dalších vlastností základního souboru

# Příklad statistické hypotézy

Daný náhodný výběr pochází z normálního rozdělení

2 náhodné výběry pocházejí ze **stejného rozdělení**

2 náhodné výběry jsou z rozdělení, která mají stejnou střední hodnotu, rozptyl, apod.

# Statistický test

- rozhodovací pravidlo, kterým přiřadíme rozhodnutí o platnosti či neplatnosti hypotézy
- **parametrický test**
  - statistické hypotézy se týkají neznámých parametrů a při provádění testů hypotéz vycházíme ze známého rozdělení sledované náhodné veličiny v základním souboru (nejčastěji Gaussovo normální rozdělení)
- **neparametrický test**
  - statistické hypotézy se týkají obecných vlastností základního souboru a příslušný test nevyžaduje znalost typu rozdělení v základním souboru

# Formulace statistické hypotézy

1. **nulová hypotéza** (označená  $H_0$ ) – tvrzení, které obvykle vyjadřuje „žádný neboli nulový rozdíl“ mezi testovanými soubory dat
2. **alternativní hypotéza** (označená  $H_1$ ) - popírá platnost nulové hypotézy  $H_0$ . Obvykle se vyjadřuje jako „existence difference“ mezi soubory nebo „existence závislosti“ mezi proměnnými

# Příklad

- Máme v experimentu 2 skupiny osob, jednu pokusnou ( $P$ ) a druhou kontrolní ( $K$ ). U pokusné skupiny sledujeme působení léku na onemocnění, jímž jsou postiženy stejně obě skupiny.
- Kdybychom lék nepoužili, měly by výsledky měření v obou skupinách být zhruba stejné (v průměru). V tomto případě bychom tvrdili, že obě skupiny osob patří do téhož základního souboru, a že rozdíl mezi nimi je nulový (platí nulová hypotéza o shodě středních hodnot obou souborů).
- Dostaneme-li v experimentu u ošetřené skupiny výsledky výrazně odlišné (v průměru) oproti skupině kontrolní (neošetřené), pak nulovou hypotézu zamítáme a přijmeme alternativní hypotézu, že skupina  $P$  patří do jiného základního souboru než skupina  $K$ , tzn., že účinek zkoumaného léku je prokazatelný.
- V případě, že výsledky se prakticky nebudou lišit (budou se vyskytovat jen náhodné rozdíly, způsobené variabilitou biologického materiálu), přijmeme  $H_0$ , tzn. prohlásíme lék za neúčinný.