

## 2. Pravděpodobnost a matematická statistika

---

### Náhodná veličina: diskrétní a spojitě rozdělení, distribuční funkce, pravděpodobnostní funkce, hustota

---

Náhodná veličina je proměnná, jejíž hodnota závisí na výsledku náhodného jevu.

Rozlišujeme dva typy:

- Diskrétní náhodná veličina  
Nabývá jen určitých (napočítatelných) hodnot, např. počet vržených šestek při 10 hodech kostkou.
- Spojitá náhodná veličina  
Může nabývat libovolných hodnot v intervalu, např. tělesná výška člověka.

### Pravděpodobnostní funkce (PMF) a hustota pravděpodobnosti (PDF)

- Pravděpodobnostní funkce  
(  $p(x) = P(X = x)$  )  
Je definovaná pro diskrétní veličiny. U každé hodnoty (  $x$  ) říká, jaká je pravděpodobnost, že ji veličina nabude.
- Hustota pravděpodobnosti  
(  $f(x)$  )  
Používá se pro spojitě veličiny. Pravděpodobnost, že veličina spadne do intervalu ( $[a, b]$ ), je  
(  $\int_a^b f(x) dx$  ).  
Pravděpodobnost pro přesnou hodnotu je vždy nulová.

### Distribuční funkce

- Distribuční funkce  
(  $F(x) = P(X \leq x)$  )  
Udává kumulativní pravděpodobnost, že veličina je menší nebo rovna (  $x$  ).
  - Pro diskrétní veličiny má tvar schodovité funkce (skoky odpovídají jednotlivým hodnotám).
  - Pro spojitě veličiny je spojitá a plynule roste.

# Kvantily, střední hodnota, rozptyl

---

## Kvantily

- **Kvantil** je hodnota, která rozděluje uspořádaný soubor dat na zvolené části.
- **Medián** je 50% kvantil (polovina hodnot leží níže, polovina výše).
- **Kvartily** (Q1, Q2, Q3) dělí soubor na čtvrtiny, **percentily** na 100 částí.

## Střední hodnota (průměr, očekávaná hodnota)

- Pro diskrétní veličiny: ( $E[X] = \sum x \cdot p(x)$ )
- Pro spojité veličiny: ( $E[X] = \int x \cdot f(x) dx$ )
- Udává průměrnou (očekávanou) hodnotu při nekonečném opakování náhodného pokusu.

## Rozptyl

- Měří, jak moc jsou hodnoty rozptýleny kolem střední hodnoty.
- ( $\text{Var}(X) = E[(X - E[X])^2]$ )
- **Standardní odchylka** je odmocnina z rozptylu a má stejnou jednotku jako původní veličina.

## Bodové a intervalové odhady

---

### Bodový odhad

- Konkrétní číselná hodnota, kterou stanovíme na základě výběru jako odhad parametru populace (např. výběrový průměr jako odhad průměru celé populace).

### Intervalový odhad

- Rozsah hodnot, který s určitou pravděpodobností obsahuje skutečný parametr populace (například 95% interval spolehlivosti pro průměr).
- Zohledňuje nejistotu v odhadu, je praktičtější než bodový odhad.
- Interval se zužuje s větším počtem měření a s menší rozptýleností dat.

## Obecné principy testování statistických hypotéz

---

- Testování hypotéz je proces, kterým na základě vzorku dat rozhodujeme, zda nějaký předpoklad o populaci platí.

### Formulace hypotéz

- **Nulová hypotéza** ( $H_0$ ): většinou "není efekt", "data odpovídají předpokladu".
- **Alternativní hypotéza** ( $H_1$ ): existuje efekt, rozdíl apod.

## Postup testování

1. Výběr testu podle typu dat a hypotézy (např. t-test, z-test, chi-kvadrát test).
  2. Stanovení hladiny významnosti ( $\alpha$ ), často 0,05 (5 %).
  3. Výpočet testové statistiky a p-hodnoty.
  4. Rozhodnutí: Pokud ( $p < \alpha$ ), zamítáme ( $H_0$ ).
- **Chyba I. druhu:** Pravděpodobnost, že zamítneme správnou ( $H_0$ ) (falešně pozitivní výsledek, velikost je ( $\alpha$ )).
  - **Chyba II. druhu:** Pravděpodobnost, že nezamítneme nepravdivou ( $H_0$ ) (falešně negativní výsledek, velikost je ( $\beta$ )).
  - **Síla testu:** ( $1 - \beta$ ), pravděpodobnost správného zamítnutí nepravdivé nulové hypotézy.

## Testy parametrů normálního rozdělení

---

- Testují vlastnosti normálně rozdělených populací, zejména střední hodnotu:
- **Z-test:** Pro velké výběry nebo pokud známe rozptyl populace.
- **T-test:** Pro malé výběry nebo neznámý rozptyl. Existují:
  - Jednovýběrový t-test (porovnání výběrového průměru s hypotetickou hodnotou).
  - Dvouvýběrový t-test (porovnání průměrů dvou skupin, nezávislých).
  - Párový t-test (porovnání hodnot na stejných subjektech před a po zásahu).

## Chi-kvadrát test dobré shody

---

- Testuje, zda pozorované četnosti odpovídají očekávaným.
- Používá se pro kategorická data, např. pro ověření, zda rozdělení krevních skupin v populaci odpovídá teoretickému rozdělení.

## Formulace hypotéz

- ( $H_0$ ): pozorované frekvence odpovídají očekávaným.
- ( $H_1$ ): neodpovídají.

## Variabilita a normálnost dat

---

- **Variabilita:** měří rozptyl hodnot v datech – základní ukazatele jsou rozptyl, směrodatná odchylka, interkvartilové rozpětí (IQR).
- **Normálnost dat:** ověřuje se pomocí statistických testů (Shapiro-Wilk, Kolmogorov-Smirnov), případně graficky (histogram, Q-Q plot).
- Pro normální rozdělení je typická zvonovitá křivka; většina statistických testů předpokládá normalitu dat.

# Typy experimentálních a observačních studií

---

## Experimentální studie

- **Randomizovaná kontrolovaná studie (RCT):**
  - Účastníci jsou náhodně rozděleni do skupin, jedna dostává intervenci, druhá slouží jako kontrola. Minimalizuje vliv náhody a zkreslení.
- **Crossover studie:**
  - Každý účastník dostává postupně různé léčby; slouží k eliminaci individuálních rozdílů.

## Observační studie

- **Kohortové studie:**
  - Sledování skupiny lidí v čase, zjišťování výskytu určitého jevu (nemoci) podle různých faktorů.
  - **Prospektivní:** sledují se dopředu.
  - **Retrospektivní:** analyzují se již existující záznamy.
- **Případově-kontrolní studie:**
  - Porovnání skupiny s onemocněním (případy) a bez onemocnění (kontroly), hledání rozdílů v expozici rizikovým faktorům.

## Medicína založená na důkazech (Evidence Based Medicine, EBM)

---

EBM je systematický přístup, který spojuje:

- Nejlepší dostupné vědecké důkazy,
- klinické zkušenosti lékaře,
- preference a hodnoty pacienta.

## Kroky v EBM

1. **Formulace klinické otázky** (často podle PICO: pacient, intervence, komparátor, outcome/výsledek).
2. **Vyhledání důkazů** (PubMed, Cochrane atd.).
3. **Kritické zhodnocení důkazů** (například pomocí nástroje GRADE).
4. **Aplikace důkazů** do konkrétní klinické praxe s ohledem na individuální situaci pacienta.
5. **Hodnocení výsledků a zpětná vazba.**