

# Metody hodnocení a formulační dokumentace léčivých přípravků

## Statistické vyhodnocování farmaceutických studií I



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,  
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání  
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

# Základní definice a pojmy: Náhodná veličina

- **Náhodná veličina (náhodná proměnná)** jsou hodnoty, rozměry, kterými se charakterizují výsledky experimentu, (pozorování)
- Náhodná veličina nabývá určitých hodnot s určitou pravděpodobností
- Proměnlivost výsledků je důvod, proč je zapotřebí statistických metod při plánování pokusů a analyzování výsledků
- **spojitá** náhodná veličina – hodnota z intervalu
  - např. hmotnost tablet → daná citlivostí vah
- **diskrétní** náhodná veličina - jen celé číslo
  - např. počet bakteriálních kolonií

# Základní definice a pojmy:

## Náhodná veličina

- **číselná** data → plně kvantifikují popisovanou vlastnost (teplota)
- **ordinální** data → např. roztřídění bolesti na určitý stupeň (žádná bolest –0, nepatrná –1, mírná –2,..)
- **kategoriální** (nominální) data → zařazení do určité kategorie, uspořádání nemá smysl (vedlejší účinky léčby (nevolnost, bolest hlavy, ..)
- **rozdělení četnosti** → ke každé hodnotě (variantě) proměnné je uvedena odpovídající četnost
  - intervalové rozdělení četnosti - četnost přísluší intervalům hodnot
  - obecně platí, že je vhodná stejná velikost intervalů a co nejmíň intervalů má být prázdných
  - Sturgersovo pravidlo:  $k=1+3,3 \cdot \log N$ 
    - $N$  = počet různých hodnot proměnné

# Základní definice a pojmy: Vzorek a populace

- Populace = **základní soubor** → velký soubor prvků (soubor všech dat)
- Vzorek = **výběrový soubor, výběr** → relativně malý počet prvků vybraných z populace (pacienti vybraní pro klinickou studii)
- **parametry** popisné charakteristiky (populace, souboru)
  - parametry populace jsou odhadovány pomocí parametrů vzorku
  - dvě skupiny parametrů: střední hodnoty souboru a charakteristiky variability
- **výběrový průměr** -  $\bar{x}$  (střední hodnota populace  $\mu$ )
- **medián** → rozděljuje data přesně na polovinu, není ovlivnitelný extrémními hodnotami souboru
- **modus** → hodnota, která se vyskytuje v souboru nejčastěji

# Základní definice a pojmy: Vzorek a populace

- **Rozpětí**  $R$  = rozdíl mezi největší a nejmenší hodnotou souboru,

$$R = x_{\max} - x_{\min}$$

- **Rozptyl**  $s^2$ : 
$$s^2 = \frac{\sum (x - \bar{x})^2}{N - 1}$$

- **Směrodatná odchylka**  $s$  = odmocnina z rozptylu
  - v rozmezí  $\pm s$  je asi 68% dat, v rozmezí  $\pm 2s$  asi 95% dat, v rozmezí  $\pm 3s$  asi 99% dat
- **Variační koeficient**  $v$  = poměr směrodatné odchylky a průměru

# Základní definice a pojmy: Pravděpodobnost

- podmínka **náhodných jevů** → nelze dopředu určit, zda jev nastane nebo ne → **náhodný pokus**
- poměr **absolutní četnosti** jevu A a celkového počtu pokusů (pozorování) → **relativní četnost náhodného jevu A**:

$$p_A = \frac{N(A)}{N}$$

- **Pravděpodobnost výskytu jevu A**:

$$P(A) = \lim_{N \rightarrow \infty} \frac{N(A)}{N}$$

- hodnota  $100 \times P$  je odhadem, v kolika procentech všech pokusů lze očekávat výskyt jevu A

# Základní definice a pojmy: Pravděpodobnost

- Základní vlastnosti pravděpodobnosti jsou:
  - Pravděpodobnost výskytu jevu nemůže být menší než 0 a větší než 1;  $0 \leq P(A) \leq 1$
  - Pravděpodobnost jistého jevu je jedna;  $P(J) = 1$
  - Pravděpodobnost výskytu nemožného jevu je nula;  $P(\emptyset) = 0$
  - Mezi pravděpodobností výskytu jevu  $A$ , a opačného jevu  $\bar{A}$  existuje vztah  $P(\bar{A}) = 1 - P(A)$
  - Pravděpodobnost výskytu sjednocení neslučitelných jevů (tj. jevů, které nemohou nastat současně) je rovna součtu pravděpodobností těchto jevů;  $P(A_1 \ A_2 \ \dots) = P(A_1) + P(A_2) + \dots$
  - Součet pravděpodobností výskytu neslučitelných jevů zahrnujících všechny možné výsledky je jedna;  $P(A) + P(B) + P(C) + \dots = 1$ .

# Diskrétní rozdělení náhodné veličiny: Binomické rozdělení

- **Binomické rozdělení** (rozdělení pravděpodobnosti) → výsledky pokusu do jedné ze dvou kategorií:
  - LD 50 – střední letální dávka (dávka, jež usmrtí 50% zvířat; zvíře pokus přežilo nebo nepřežilo), léčba je úspěšná nebo neúspěšná
- jednoznačně určeno dvěma parametry: pravděpodobností výsledku a počtem pozorování
- Pravděpodobnost výsledků z N pokusů z binomického rozvoje:
  - $(p+q)^N$ , kde p je pravděpodobnost úspěchu a q je pravděpodobnost neúspěchu



# Diskrétní rozdělení náhodné veličiny: Binomické rozdělení

- Pravděpodobnost, že náhodná proměnná  $X$  nabude hodnoty  $x$  (v  $n$  pokusech nastal jev  $x$ -krát):

$$- \binom{n}{x} p^x (1-p)^{n-x} \quad \binom{n}{x} = \frac{n!}{x!(n-x)!}$$

- Je-li např. pravděpodobnost vyléčení pacienta určitým antibiotikem 0,75, tak pravděpodobnost, že se vyléčí 3 ze 4 pacientů je:

$$- \binom{4}{3} \cdot (0,75)^3 \cdot (0,25)^1 = \frac{4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 1} \cdot 0,42188 \cdot 0,25 = 0,42188 \quad (42,188 \%)$$

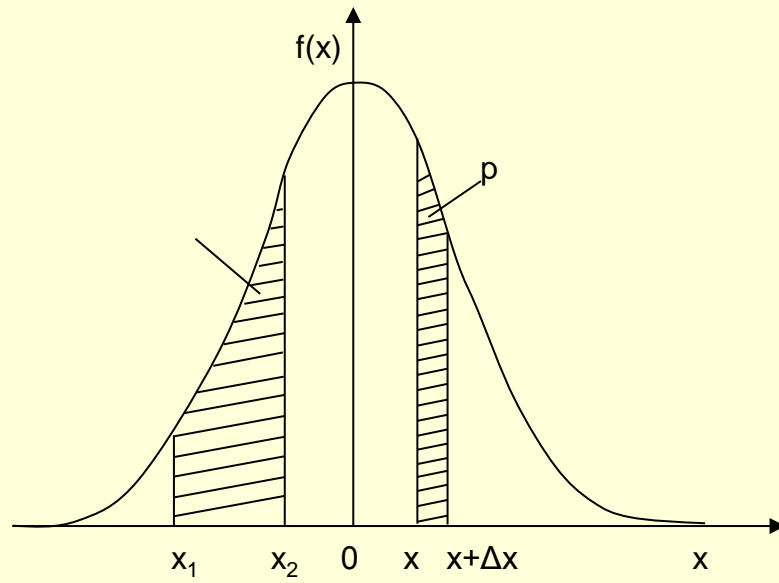
- když  $n > 30$  a  $n \cdot p \cdot q < 9$  → aproximace **Poissonovým rozdělením**
- když  $n > 30$  a  $n \cdot p \cdot q > 9$  → aproximace **Normálním rozdělením**

# Spojité rozdělení

- Rozdělení spojité náhodné veličiny  $X$  je dáno **distribuční funkcí**  $F(x)$ , případně **hustotou pravděpodobnosti** (frekvenční funkcí)  $f(x)$
- **Distribuční funkce**  $F(x)$  pro reálné číslo  $x$  udává pravděpodobnost, že náhodná veličina  $X$  nabude hodnoty menší nebo rovno než je toto číslo  $x$ :  $F(x) = P(X \leq x)$
- Znamená to, že ve  $100 \cdot F(x)$  % případů lze očekávat výskyt hodnot menších než  $x$
- Pravděpodobnost, že náhodná veličina nabude hodnoty z intervalu  $(x_1, x_2)$  je rovna rozdílu distribučních funkcí horní a dolní meze intervalu: 
$$P(x_1 \leq X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = F(x_2) - F(x_1)$$

# Spojité rozdělení

- Grafickým vyjádřením **hustoty pravděpodobnosti  $f(x)$**  je spojitá frekvenční křivka
- Plocha pod frekvenční křivkou je rovna číslu 1, protože pravděpodobnost, že veličina nabude libovolné hodnoty od mínus nekonečna do plus nekonečna je rovna jistotě ( $P=1$ )
- Hustota pravděpodobnosti umožňuje určit pro každý interval  $(x_1, x_2)$  pravděpodobnost, že náhodná veličina nabude hodnoty z tohoto intervalu jako plochu vymezenou grafem funkce  $f(x)$  nad tímto intervalem:



# Spojité rozdělení

- Při řešení statistických úloh často potřebujeme pro danou pravděpodobnost  $p$  nalézt zpětně z tabulky distribuční funkce takovou hodnotu  $x_p$ , pro kterou platí:  $P(X \leq x_p) = F(x_p) = p$
- Tato hodnota se nazývá **100\*p%-ním kvantilem**.
- Jinak řečeno: 100\*p% kvantil rozdělí reálnou osu, na které jsou rozloženy hodnoty náhodné veličiny  $x$ , na dva intervaly.
- V intervalu  $(-\infty, x_p)$  se nachází 100\*p% hodnot a v intervalu  $(x_p, \infty)$  pak 100\*(1 - p)% hodnot náhodné veličiny  $\rightarrow$  100\*p% hodnot je menších nebo rovno než je hodnota kvantilu, 100\*(1 - p)% hodnot je větších než  $x_p$ .
- pomocí kvantilů se často vyjadřují kritické hodnoty jednotlivých rozdělení