Nejpoužívanější diskrétní rozdělení pravděpodobnosti a jejich základní číselné charakteristiky

XSZD-06a-2025

Obsah

2.3 Číselné charakteristiky

Náhodnou veličinu X jednoznačně určují a plně popisují

- X diskrétní
  - pravděpodobnostní funkce
  - distribuční funkce
- X spojitá
  - hustota
  - distribuční funkce

číselné charakteristiky = shrnutí informací o X do několika čísel, které ji dostatečně charakterizují

Typy číselných charakteristik

Členění dle popisované vlastnosti

- **polohy:** „střed“, kolem kterého jsou koncentrovány hodnoty X
  - střední hodnota
  - modus
  - kvantily
- **variability:** rozptýlenost hodnot X kolem charakteristiky polohy
  - rozptyl
  - směrodatná odchylka
- **šikmost:** tvar rozdělení pravděpodobností (symetrické nebo asymetrické)
- **špičatost:** tvar rozdělení pravděpodobností (špičaté nebo zploštělé)

Střední hodnota $\mathrm{E}(X)$

- základní charakteristika polohy
- „střed“ (těžiště), kolem kterého jsou koncentrovány hodnoty X

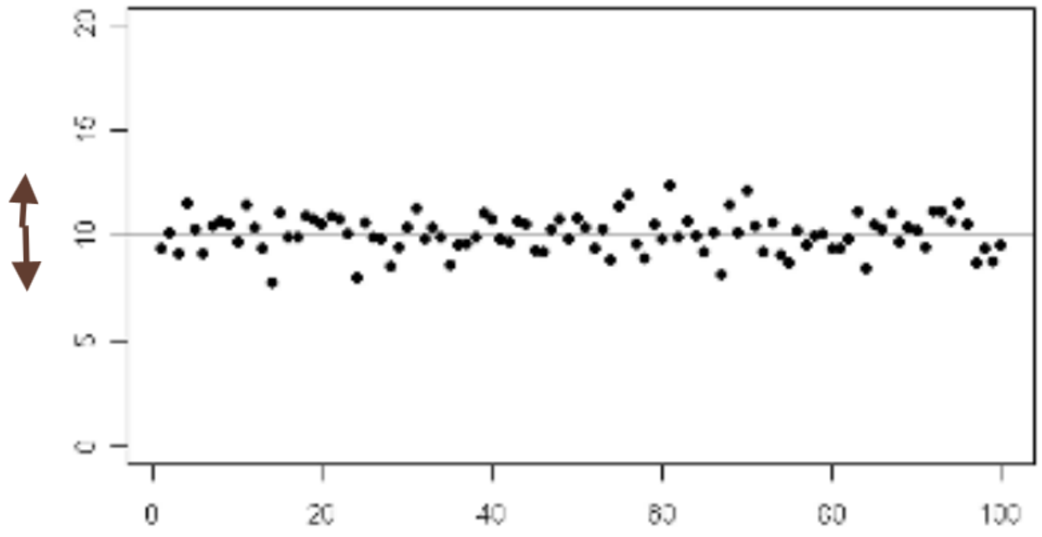
X diskrétní s oborem hodnot M

$$\mathrm{E}(X) = \sum_{x_i \in M} x_i \mathrm{P}(X = x_i)$$

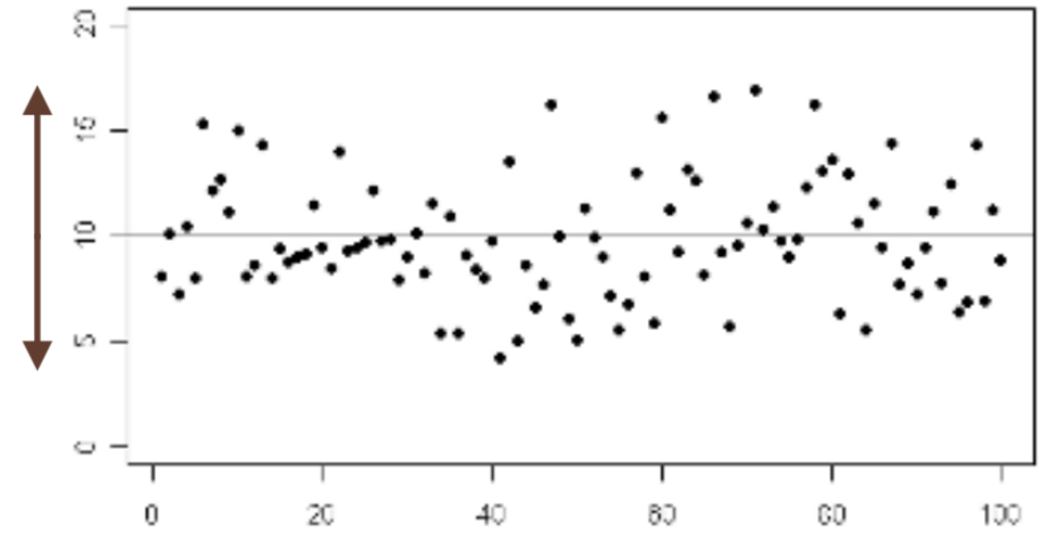
X spojitá s hustotou $f(x)$

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$$

Charakteristiky variability



- číslo udávající **koncentraci (rozptýlení)** hodnot okolo středu



Rozptyl $D(X)$

- základní charakteristika variability
- charakterizuje měřítko (šířku) rozdělení
- popisuje rozptýlenost hodnot kolem střední hodnoty

Rozptyl

$$D(X) = \mathrm{E}\left[X - \mathrm{E}(X)\right]^2, \text{ existuje-li } \mathrm{E}(X)$$

Směrodatná odchylka

$$\sqrt{D(X)}$$

Modus $\hat{x}$

- charakteristika polohy
- X diskrétní: nejpravděpodobnější hodnota
- X spojitá: bod x, ve kterém hustota pravděpodobností $f(x)$ nabývá lokálního maxima

Rozdělení pravděpodobností náhodné veličiny X

k	1	2	4	5
$P(X=k)$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$

Střední hodnota $E(X) = \sum_n x_n p(x_n)$

- $E(X) = 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{4} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{4} = \frac{11}{4} = 2{,}75$

Rozptyl a směrodatná odchylka

$D(X) = E\left[X - E(X)\right]^2 = E(X^2) - [E(X)]^2$

$$E(X^2) = \sum_n x_n^2 p(x_n) = 1^2 \cdot \frac{1}{3} + 2^2 \cdot \frac{1}{4} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{4} = \frac{41}{4} = 10{,}25$$

- $D(X) = E(X^2) - [E(X)]^2 = \frac{41}{4} - \left(\frac{11}{4}\right)^2 = \frac{43}{16} \doteq 2{,}69$
- $\sqrt{D(X)} = \sqrt{\frac{43}{16}} \doteq 1{,}64$

Modus: nejpravděpodobnější hodnota: $\hat{x} = 1$

$$f(x) = \begin{cases} 6x(1-x) & 0 < x \leq 1 \\ 0 & \text{jinak} \end{cases}$$

- **Střední hodnota:**

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

$$\begin{aligned} \mathrm{E}(X) &= \int_{-\infty}^{\infty} xf(x)dx = \int_{0}^{1} x \cdot 6x(1-x)dx = \\ &= \int_{0}^{1} (6x^2 - 6x^3)dx = \left[\frac{6x^3}{3} - \frac{6x^4}{4}\right]_0^1 = \frac{6}{3} - \frac{6}{4} = \frac{24-18}{12} \\ &= \frac{6}{12} = \frac{1}{2} \end{aligned}$$

$$f(x) = \begin{cases} 6x(1-x) & 0 < x \leq 1 \\ 0 & \text{jinak} \end{cases}, \quad \mathrm{E}(X) = \frac{1}{2}$$

Rozptyl:

$$D(X) = \mathrm{E}\left[X - \mathrm{E}(X)\right]^2 = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$$

$$\begin{aligned} \mathrm{E}(X^2) &= \int_{-\infty}^{\infty} x^2 \cdot f(x) dx = \int_0^1 x^2 \cdot 6x(1-x) dx = \\ &= \int_0^1 6x^3 - 6x^4 dx = \left[\frac{6x^4}{4} - \frac{6x^5}{5}\right]_0^1 = \frac{6}{4} - \frac{6}{5} = \frac{30-24}{20} = \frac{6}{20} \end{aligned}$$

$$D(X) = \frac{6}{20} - \left(\frac{1}{2}\right)^2 = \frac{6}{20} - \frac{1}{4} = \frac{6-5}{20} = \frac{1}{20}$$

$$f(x) = \begin{cases} 6x(1-x) & 0 < x \leq 1 \\ 0 & \text{jinak} \end{cases}$$

Modus: bod, v němž je lokální maximum hustoty

$$f'(x) = 6 - 12x$$

$$6 - 12x = 0 \quad \Leftrightarrow \quad x = 0{,}5$$

$$f''(x) = -12 < 0 \quad \Rightarrow \quad \text{maximum} \quad \Rightarrow \quad \hat{x} = 0{,}5$$

Kvantily spojité náhodné veličiny

Definice

Nechť $\alpha \in (0, 1)$. α-kvantilem spojité náhodné veličiny X rozumíme kterékoli reálné číslo x_α, které splňuje

$$\mathrm{P}(X \leq x_\alpha) = \alpha.$$

- X spojitá: α-kvantil určen jednoznačně vztahem

$$F(x_\alpha) = \alpha$$

- X diskrétní: kvantily nejsou určeny jednoznačně, nebudeme uvažovat

Speciální názvy kvantilů

- $x_{0,5}$ – medián
- $x_{0,25}$ – dolní kvartil
- $x_{0,75}$ – horní kvartil
- $x_{k/10}$, $k = 1, \ldots, 9$ – k-tý decil
- $x_{k/100}$, $k = 1, \ldots, 99$ – k-tý percentil

Základní diskrétní rozdělení pravděpodobností

- Alternativní rozdělení
- Binomické rozdělení
- Hypergeometrické rozdělení
- Diskrétní rovnoměrné rozdělení
- Poissonovo rozdělení

Alternativní rozdělení $X \sim \mathrm{Alt}(p)$

Model: **pokus**, ve kterém nastávají pouze **2 různé (dichotomické) výsledky** a sledujeme, co nastane

- indikátor náhodného jevu (jev nastal/nenastal)
- pravdivostní hodnota výroku (pravda/lež)
- úspěšnost pokusu (úspěch/neúspěch)
- kvalita výrobku (vadný/ bez chyby)
- ...

Parametr rozdělení

- p ... pravděpodobnost úspěchu

Alternativní rozdělení $X \sim \mathrm{Alt}(p)$

$$X = \begin{cases} 1 & \text{úspěch} \\ 0 & \text{neúspěch} \end{cases}$$

Pravděpodobnostní funkce

$$\mathrm{P}(X = 1) = p, \quad \mathrm{P}(X = 0) = 1 - p, \quad p \in (0, 1)$$

Číselné charakteristiky

$$E(X) = p, \quad D(X) = p(1 - p), \; \sigma_X = \sqrt{p(1 - p)}$$

Binomické rozdělení $X \sim \mathrm{Bi}(n, p)$

Model: **počet úspěchů**, které nastanou při provedení n **nezávislých dichotomických pokusů**

Příklady

- počet děvčat v rodině s n dětmi
- počet zmetků mezi n výrobky
- počet úspěšných zkoušek přístroje z celkem n zkoušek

Parametry rozdělení

- n ... počet nezávislých dichotomických pokusů
- p ... pravděpodobnost úspěchu v jednom pokusu

Binomické rozdělení $X \sim \mathrm{Bi}(n, p)$

Pravděpodobnostní funkce (Bernoulliho schéma)

$$\mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \ k = 0, \dots, n, \quad n \geq 1, \ p \in (0, 1)$$

Číselné charakteristiky

$$E(X) = np, \quad D(X) = np(1-p), \quad \sigma_X = \sqrt{np(1-p)}$$

Excel: BINOM.DIST(k;n;p;logická proměnná)

- logická proměnná = 0 pro výpočet $\mathrm{P}(X = k)$
- logická proměnná = 1 pro výpočet $F(k)$

Příklad (Barva květů hrachu)

Je křížen bělokvětý hrách s fialovým, přičemž předpokládáme, že rostliny, na nichž je pokus prováděn, nebyly dosud kříženy. Podle pravidel dědičnosti lze očekávat, že $\frac{3}{4}$ nově vzniklých rostlin (potomků) pokvetou fialově a $\frac{1}{4}$ bíle. Zatím vzklíčilo 10 nových rostlin. ***Jaká je pravděpodobnost, že 9 pokvete fialově?***

Nějaký nápad?

Příklad (Barva květů hrachu)

Je křížen bělokvětý hrách s fialovým, přičemž předpokládáme, že rostliny, na nichž je pokus prováděn, nebyly dosud kříženy. Podle pravidel dědičnosti lze očekávat, že $\frac{3}{4}$ nově vzniklých rostlin (potomků) pokvetou fialově a $\frac{1}{4}$ bíle. Zatím vzklíčilo 10 nových rostlin. ***Jaká je pravděpodobnost, že 9 pokvete fialově?***

Řešení

X . . . počet fialově kvetoucích rostlin z 10 rostlin

$$X \sim \mathrm{Bi}(10; 0{,}75)$$

$$P(X = 9) = \binom{10}{9} 0{,}75^9 0{,}25^1 = 0{,}188$$

Excel: BINOM.DIST(9;10;3/4;0)

Hypergeometrické rozdělení $X \sim \mathrm{Hg}(N, M, n)$

Model: počet úspěchů, které nastanou při provedení n **závislých dichotomických pokusů** (výběr bez vracení)

Příklady

- počet vybraných zmetků, vybíráme-li n výrobků z N výrobků, přičemž M z nich jsou zmetky
- počet vybraných nenaučených otázek, losujeme-li n otázek z N, přičemž na M z nich nejsme naučeni

Parametry rozdělení

- n ... počet vybraných jednotek (závislých dichotomických pokusů)
- N ... počet všech jednotek
- M ... počet jednotek, které mají sledovanou vlastnost

Hypergeometrické rozdělení $X \sim \mathrm{Hg}(N, M, n)$

Pravděpodobnostní funkce

$$\mathrm{P}(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \; k = 0, \ldots, n, \quad N \geq n \geq 1, \; 0 \leq M \leq N$$

Číselné charakteristiky

$$E(X) = \frac{n \cdot M}{N}, \quad D(X) = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Excel: HYPGEOM.DIST(k;n;M;N;logická proměnná)

- logická proměnná = 0 pro výpočet $\mathrm{P}(X = k)$
- logická proměnná = 1 pro výpočet $F(k)$

Příklad (Vadné výrobky)

Mezi 100 výrobky je 20 zmetků. Vybereme 10 výrobků a sledujeme počet zmetků mezi vybranými výrobky. ***Jaká je pravděpodobnost, že jsme vybrali právě 3 zmetky?***

Nějaký nápad?

Příklad (Vadné výrobky)

Mezi 100 výrobky je 20 zmetků. Vybereme 10 výrobků a sledujeme počet zmetků mezi vybranými výrobky. ***Jaká je pravděpodobnost, že jsme vybrali právě 3 zmetky?***

Řešení

X . . . počet vybraných zmetků

$N = 100,\ M = 20,\ n = 10$

$$X \sim \mathrm{Hg}(100; 20; 10)$$

$$P(X = 3) = \frac{\binom{20}{3}\binom{80}{7}}{\binom{100}{10}} = 0{,}209$$

Excel: HYPGEOM.DIST(3;10;20;100;0)

Diskrétní rovnoměrné rozdělení $X \sim \mathrm{Ro}(n)$

Model: **pokus**, který má n **různých výsledků**, které jsou **stejně pravděpodobné** a sledujeme, co nastane

Příklady

- hod pravidelnou kostkou $M = \{1, 2, 3, 4, 5, 6\}$, $p(i) = \frac{1}{6}$
- hod férovou mincí $M = \{0, 1\}$, $p(i) = \frac{1}{2}$
- padnutí čísla v ruletě v Monte Carlo $M = \{0, 1, 2, \ldots, 36\}$, $p(i) = \frac{1}{37}$
- padnutí čísla v ruletě v Las Vegas $M = \{0, 00, 1, 2, \ldots, 36\}$, $p(i) = \frac{1}{38}$

Pravděpodobnostní funkce

$$\mathrm{P}(X = x) = \frac{1}{n}, \ x = 1, \ldots, n, \quad n \in \mathbb{N}$$

Poissonovo rozdělení $X \sim \mathrm{Po}(\lambda)$

Model: **počet událostí**, které nastanou při provedení **velkého počtu nezávislých dichotomických pokusů**, přičemž **pravděpodobnost** vzniku události **je velmi malá**

- výskyt jevu v daném časovém/prostorovém intervalu nezávisí co se stalo jindy/jinde
- v daném okamžiku/bodě nemohou nastat 2 jevy současně
- pro každý časový okamžik/bod je pravděpodobnost výskytu jevu v malém časovém/prostorovém intervalu stejná
- λ ... intenzitu výskytu události za jednotku času/délky/objemu

Příklady

- # částic emit. za jedn. času radioaktivní látkou dané hmotnosti
- # signálů, které dojdou do telefonní ústředny za jednotku času
- # jednob. organizmů na ploše velikosti t v zorném poli mikroskopu
- # dopravních nehod za jednotku času
- # typografických chyb na stránce
- # volných dnů ve firmě během měsíce

Poissonovo rozdělení $X \sim \mathrm{Po}(\lambda)$

Pravděpodobnostní funkce

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \; k = 0, 1, 2, \dots, \quad \lambda > 0$$

Číselné charakteristiky

$$E(X) = \lambda, \quad D(X) = \lambda, \quad \sigma_X = \sqrt{\lambda}$$

Excel: POISSON.DIST(k;λ;logická proměnná)

- logická proměnná = 0 pro výpočet $P(X = k)$
- logická proměnná = 1 pro výpočet $F(k)$

Příklad (Poruchy obráběcího stroje)

Při práci obráběcího stroje dochází náhodně k výpadkům. Průměrně jsou 2 výpadky za 24 hodin.

a) *Jaká je pst, že za 24 hodin dojde alespoň k jednomu výpadku?*

b) *Jaká je pst, že za týden nebudou více než 3 výpadky?*

Nějaký nápad?

Příklad (Poruchy obráběcího stroje)

Při práci obráběcího stroje dochází náhodně k výpadkům. Průměrně jsou 2 výpadky za 24 hodin.

a) ***Jaká je pst, že za 24 hodin dojde alespoň k jednomu výpadku?***

b) *Jaká je pst, že za týden nebudou více než 3 výpadky?*

Řešení

$X \ldots$ počet výpadků za 1 den

$$\text{průměrně 2 výpadky za 24 hodin} \quad \Rightarrow \quad \lambda = 2$$

$$X \sim \mathrm{Po}(2)$$

$$\mathrm{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots$$

$$\mathrm{P}(X \geq 1) = 1 - \mathrm{P}(X < 1) = 1 - \mathrm{P}(X = 0) = 1 - e^{-2} = 0{,}865$$

Excel: 1-POISSON.DIST(0;2;0)

Příklad (Poruchy obráběcího stroje)

Při práci obráběcího stroje dochází náhodně k výpadkům. Průměrně jsou 2 výpadky za 24 hodin.

a) *Jaká je pst, že za 24 hodin dojde alespoň k jednomu výpadku?*

b) ***Jaká je pst, že za týden nebudou více než 3 výpadky?***

Řešení

Y ... počet výpadků za 1 týden

$$\begin{aligned} & \text{průměrně 2 výpadky/24 hodin} \\ \Leftrightarrow \quad & \text{průměrně } 2 \cdot 7 = 14 \text{ výpadků/týden} \end{aligned}$$

$$Y \sim \mathrm{Po}(14)$$

$$\mathrm{P}(Y \leq 3) = e^{-14}\left(1 + 14 + \frac{14^2}{2!} + \frac{14^3}{3!}\right) = 0{,}000474$$

Excel: POISSON.DIST(3;14;1)