Přehled

Diskrétní n.v. – střední hodnota a rozptyl

Parametry náhodných veličin

Náhodné vektory

Podmíněné rozdělení

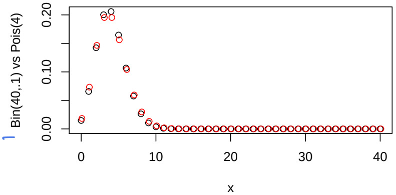
Co už víme

- Co je diskrétní n.v.
- Jak je popisovat pomocí pravděpodobnostní a/nebo distribuční funkce.
- Příklady rozdělení: Bernoulliho, binomické, hypergeometrické, Poissonovo, geometrické.
- Co je střední hodnota: dvě možné definice:
- $\mathbb{E}(X) = \sum_{x \in Im(X)} x \cdot P(X = x)$
- $\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\})$
- $\mathbb{E}(g(X)) = \sum_{x \in Im(X)} g(x) P(X = x)$ (LOTUS)
- „Kolik čekáme, že průměrně dostaneme, když budeme opakovat nezávislé pokusy s výsledkem popsaným X“ ... bude tzv. zákon velkých čísel

X ... doba běhu ; N pokusů

$N \cdot \mathbb{E}(X) = \sum_x x \cdot P(X=x) \cdot N$ ... očekávaný # pokusů, kde $X = x$

Srovnání binomického a Poissonova rozdělení: pravděpodobnostní funkce



Vygenerováno následujícím kódem v R

```
x = 0:40
bin = dbinom(x,40,0.1)
pois = dpois(x,4)
plot(x,bin, ylab="Bin(40,.1)␣vs␣Pois(4)")
points(x+.1,pois,col="red")
```

Vlastnosti $\mathbb{E}$

Věta

Nechť X, Y *jsou diskrétní n.v. a* $a, b \in \mathbb{R}$.

1. *Pokud* $P(X \geq 0) = 1$ *a* $\mathbb{E}(X) = 0$, *tak* $P(X = 0) = 1$.
2. *Pokud* $\mathbb{E}(X) \geq 0$ *tak* $P(X \geq 0) > 0$.
3. $\mathbb{E}(a \cdot X + b) = a \cdot \mathbb{E}(X) + b$.
4. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

$P(X<0)=0$

4: platí : pro součet n proměnných

4 ... pro diskr. Ω

$\mathbb{E}(X+Y) = \sum_{\omega} (X(\omega)+Y(\omega)) \cdot P(\{\omega\}) = \sum X(\omega) \cdot P(\{\omega\}) + \sum Y(\omega) \cdot P(\{\omega\})$

Dk 1 $\mathbb{E}X = \sum_{x \in \mathrm{Im} X} x \cdot P(X=x) = 0 \cdot P(X=0) + \underbrace{\sum_{\substack{x>0 \\ \& x \in \mathrm{Im} X}} x \cdot \underset{\geq 0}{P(X=x)}}_{=0} + \underbrace{\sum_{\substack{x<0 \\ \& x \in \mathrm{Im} X}} x \cdot P(X=x)}_{=0}$

$= 0 \Rightarrow \forall x>0 : P(X=x)=0$

$\Rightarrow P(X=0)=1$

2 $\mathbb{E}(X) = \sum_{x} x \cdot \underbrace{P(X=x)}_{\geq 0}$

Kdyby ne : $P(X \geq 0) = 0$, všechny členy u by byly záporné -- spor

3 $\mathbb{E}(a \cdot X + b) \overset{LOTUS}{=} \sum_{x \in \mathrm{Im} X} (a x + b) \cdot P(X=x) = a \cdot \underbrace{\sum x \cdot P(X=x)}_{\mathbb{E}X} + b \cdot \underbrace{\sum P(X=x)}_{=1}$

Alternativní formulka pro střední hodnotu

Věta

Nechť X je diskrétní n.v. nabývající jen hodnot z $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. Pak platí

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} P(X > n).$$

Dk $\mathbb{E}(X) = \sum_{k=0}^{\infty} k \cdot P(X=k) = \sum_{k=0}^{\infty} \sum_{n=0}^{k-1} P(X=k)$

$n < k$

$= \sum_{n=0}^{\infty} \sum_{k=n+1}^{\infty} P(X=k) = \sum_{n=0}^{\infty} P(X>n)$ ∎

$P(\{\omega \in \Omega : X(\omega) = k\})$

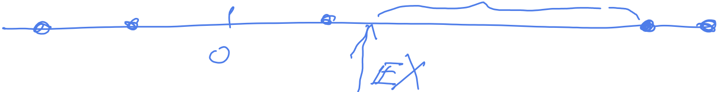
$P(\{\omega \in \Omega : X(\omega) > n\})$

$X \sim \mathrm{Geom}(p) \Rightarrow \mathbb{E}(X) = \sum (1-p)^n$

$P(X>n) = (1-p)^n$

$= \frac{1}{1-(1-p)} = \frac{1}{p}$

Rozptyl



Definice

Rozptyl (variance) n.v. X *nazveme číslo* $\mathbb{E}((X-\mathbb{E}X)^2)$*.*
Značíme jej $var(X)$*.*

- *Směrodatná odchylka (standard deviation)* $\sigma_X = \sqrt{var(X)}$ – „stejné jednotky jako X".
- Měří, jak je daleko „typicky" je X od $\mathbb{E}(X)$. Mohli bychom to měřit i jinak (např. $\mathbb{E}(|X-\mathbb{E}(X)|)$, ale rozptyl je výhodnější).

$var(X) \geq 0$ ← obvykle
$= 0$ -- spec. n.v.
< 0 ✗

$\mathbb{E}\,g(X) = g(\mathbb{E}X)$
platí pro
g lineární

Věta

$$var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Dk $\mu := \mathbb{E}(X)$

$$var(X) = \mathbb{E}((X-\mu)^2) = \mathbb{E}(X^2 - 2\mu X + \mu^2) \overset{\text{linearita}}{=} \mathbb{E}(X^2) - 2\mu\,\mathbb{E}(X) + \mu^2 = \mathbb{E}(X^2) - \mu^2$$

Podmíněná střední hodnota

Definice

Pokud X je diskrétní n.v. a $P(B) > 0$, tak podmíněná střední hodnota X za předpokladu B (conditional expectation of X given B) je

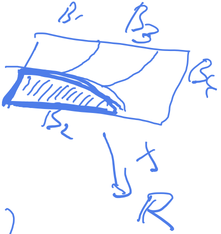
$$\mathbb{E}(X \mid B) = \sum_{x \in Im(X)} x \cdot P(X = x \mid B),$$

pokud součet má smysl.

$$x \cdot \frac{P(X=x \,\&\, B)}{P(B)}$$

Rozbor všech možností

Věta o úplné střed. hodnotě X je d. n.v.



Pokud $B_1, B_2, \ldots$ je rozklad Ω a ~~$A \in \mathcal{F}$~~, tak

$$\mathbb{E}(X) = \sum_i \underbrace{\mathbb{E}(X \mid B_i)}_{\text{střed. hodnota}} P(B_i),$$

X zúžené na B_i

kdykoliv má součet smysl. (Sčítance s $P(B_i) = 0$ považujeme za 0.)

$$\mathbb{E}(X) = \sum_i P(B_i)\, \mathbb{E}(X \mid B_i)$$

$$= \sum_i P(B_i) \sum_x x\, P(X = x \mid B_i)$$

$$= \sum_x x \underbrace{\sum_i P(B_i) \cdot P(X = x \mid B_i)}_{= P(X = x)}$$

Rozbor všech možností

$X \sim \text{Geom}(p)$

$B_1 = S$ ... první pokus úspěšný –

$B_2 = B_1^c = F$ ... — " — neúspěšný

$$E(X) = P(S) \cdot E(X \mid S) + P(F) \cdot \underline{E(X \mid F)}$$

$$= \underline{p} \cdot 1 + \underline{(1-p)} \cdot \underbrace{\left(E(X+1)\right)}_{(1+\underline{EX})}$$

předp., že EX ex. a je $< \infty$

$$p \cdot E(X) = p + (1-p) = 1$$

$$EX = \frac{1}{p}$$

Přehled

Parametry rozdělení – Bernoulliho

Pro $X \sim Bern(p)$ je

- $\mathbb{E}(X) = p$
- $var(X) = p(1-p)$

$Im(X) = \{0, 1\}$

$E(X) = 0 \cdot P(X=0) + 1 \cdot P(X=1) = P(X=1) = p$

$var(X) = E(X-p)^2 = (0-p)^2 \cdot P(X=0) + (1-p)^2 P(X=1)$

$\boxed{Yes}$ $= p^2(1-p) + (1-p)^2 \cdot p = p(1-p)\underbrace{(p + (1-p))}_{1}$

$var\, X = EX^2 - (EX)^2 = p - p^2 = p(1-p)$

$\boxed{No}$

$X^2 = X$

Parametry rozdělení – binomické

bude později: (nezplatí vždy!)

$var X = \sum var(X_i)$

indik. n.v. I_{A_i}

Pro $X \sim Bin(n, p)$ je

- $\mathbb{E}(X) = np$
- $var(X) = np(1-p)$

- První postup: $X = \sum_{i=1}^{n} X_i$, kde $X_i = \underbrace{[\text{i-tý hod úspěl}]}_{A_i}$
- $\mathbb{E}(X_i) = P(X_i = 1) = p$

lin. stř. hodn. $\mathbb{E}X = \sum_{i=1}^{n} \mathbb{E}X_i = np$

stačí $k = 1, 2, \ldots, n$

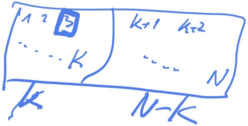
- Druhý postup:
  $\mathbb{E}(X) = \sum_{k=0}^{n} k \cdot P(X = k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$

$\mathbb{E}X^2 = \sum k^2 \cdot P(X=k)$

$$\sum_{k=1}^{n} p\, n \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

pr. $(p + (1-p))^{n-1} = np$

Parametry rozdělení – hypergeometrické



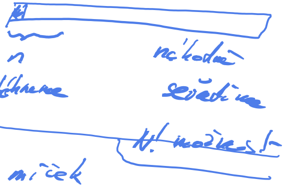
Pro $X \sim Hyper(N, K, n)$

- $\mathbb{E}(X) = n\frac{K}{N}$ $= n \cdot \frac{K}{N} = K \cdot \frac{n}{N}$
- $var(X) = n\frac{K}{N}(1 - \frac{K}{N})\frac{N-n}{N-1}$

- První postup: $X = \sum_{i=1}^{n} X_i$, kde $X_i =$ [i-tý míček červený]
- $\mathbb{E}(X_i) = P(X_i = 1) =$ $\frac{K}{N}$

je stejné pro $i=1$
platí pro všechna i

N koulí



- Druhý postup: $X = \sum_{j=1}^{K} Y_j$, kde $Y_j =$ [byl vytažen míček s číslem j] v n tazích
- $\mathbb{E}(Y_j) = P(Y_j = 1) =$ $\boxed{\frac{n}{N}}$

máme vytaženo n míčků

$$= \frac{\binom{N-1}{n-1}}{\binom{N}{n}} \quad \begin{array}{l}\#\text{ dobrých}\\ \#\text{ všech}\end{array} = \frac{n}{N}$$

Parametry rozdělení – geometrické

Pro $X \sim Geom(p)$ je

- $\mathbb{E}(X) = 1/p$
- $var(X) = \frac{1-p}{p^2}$

$\sum P(X > n)$

rozklad vlastnost

$$\sum k \cdot P(X=k) = \sum_{k=1}^{\infty} k (1-p)^{k-1} \cdot p$$

(vynecháme)

$\mathbb{E}(X^2) - (\mathbb{E}X)^2$

Parametry rozdělení – Poissonovo

Pro $X \sim Pois(\lambda)$ je

- $\mathbb{E}(X) = \lambda$
- $var(X) = \lambda$

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\mathbb{E}X = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \underbrace{\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda}}_{=1} = \lambda$$

stačí $k = 1, 2, \ldots$

$g(X)$

$$\mathbb{E}(X(X-1)) = \sum_{k=0}^{\infty} k(k-1) \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} e^{-\lambda} = \lambda^2$$

$$\mathbb{E}(X^2 - X) = \mathbb{E}X^2 - \mathbb{E}X$$

$$var X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - (\mathbb{E}X)^2 = \lambda^2 + \lambda - \lambda^2 = \underline{\lambda}$$

Přehled

Základní popis náhodných vektorů

- X, Y – náhodné veličiny na stejném pravděpodobnostním prostoru $(\Omega, \mathcal{F}, P)$.
- Budeme chtít uvažovat (X, Y) jako jeden objekt – náhodný vektor.
- Jak to udělat?
- Příklad: házíme dvakrát čtyřstěnnou kostkou, $X =$ první hod, $Y =$ druhý hod.

Sdružené rozdělení

Definice

Pro diskrétní n.v. X, Y na pravděpodobnostním prostoru $(\Omega, \mathcal{F}, P)$ definujeme jejich sdruženou pravděpodobnostní funkci (joint pmf) $p_{X,Y} : \mathbb{R}^2 \to [0, 1]$ předpisem

$$p_{\underline{X,Y}}(\underline{x, y}) = P(\{\omega \in \Omega : X(\omega) = x \& Y(\omega) = y\}).$$

$= P(X = x$
$\& Y = y)$

- Mohli bychom definovat i pro více než dvě n.v. $p_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$.

Marginální rozdělení

$P_X(1) = P_{X,Y}(1,1) + \dots$

- Máme-li dáno $p_{X,Y}$, jak zjistit rozdělení jednotlivých složek, tj. p_X a p_Y?

$$\boxed{P(X=1) = P(X=1 \,\&\, Y=1) + P(X=1, Y=2) + P(X=1 \,\&\, Y=3) + P(X=1, Y=4)}$$

X\Y	1	2	3	4	
1	1/16	1/16	1/16	1/16	4/16 = 1/4
2	1/16	...			1/4
3					1/4
4					1/4
	1/4	1/4	1/4	1/4	

dvojice nez. kostek

$P(Y=2) = P(Y=2 \,\&\, X=1) + \dots$

X\Y	1	2	3	4	
1	1/4	0	0	0	1/4
2	0	1/4	0	0	1/4
3	0	0	1/4	0	1/4
4	0	0	0	1/4	1/4
	1/4	1/4	1/4	1/4	

dvoj-kost = prům-kost

Marginální rozdělení

Věta

Nechť X, Y jsou diskrétní n.v. Pak

def. P_X P disj. sjedn. def. $P_{X,Y}$

$$p_X(x) = P(X = x) = \sum_{y \in Im(Y)} P(X = x \& Y = y) = \sum_{y \in Im(Y)} p_{X,Y}(x, y)$$

$$p_Y(y) = P(Y = y) = \sum_{x \in Im(X)} P(X = x \& Y = y) = \sum_{x \in Im(X)} p_{X,Y}(x, y)$$

Funkce náhodného vektoru

Věta

Nechť X, Y jsou n.v. na $(\Omega, \mathcal{F}, P)$, nechť $g : \mathbb{R}^2 \to \mathbb{R}$ je funkce.

- *Pak $Z = g(X, Y)$ je n.v. na $(\Omega, \mathcal{F}, P)$*
- *a platí pro ni*

$$\mathbb{E}(g(X, Y)) = \sum_{x \in ImX} \sum_{y \in ImY} g(x, y) P(X = x, Y = y).$$

Věta

Pro X, Y n.v. a $a, b \in \mathbb{R}$ platí

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$

Nezávislost náhodných veličin

Definice

Diskrétní n.v. X, Y jsou nezávislé (independent) pokud pro každé $x, y \in \mathbb{R}$ jsou jevy $\{X = x\}$ a $\{Y = y\}$ nezávislé. To nastane, právě když

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

pravd. průniku

$$P_{X,Y}(x,y) = P_X(x) \cdot P_Y(y)$$

pokud X, Y nez., tak $P_{X,Y}$ jde určit z P_X a P_Y.

Součin nezávislých n.v.

Věta

Pro nezávislé diskrétní n.v. X, Y *platí*

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Dk

$$\mathbb{E}(XY) \overset{?}{=} \sum_{\substack{x \in \operatorname{Im} X \\ y \in \operatorname{Im} Y}} (x y) \cdot \underline{P(X = x, Y = y)}$$

$$= \sum_{x, y} x y \, P(X = x) \cdot P(Y = y)$$

$$= \sum_{x} x \, P(X = x) \cdot \sum_{y} y \, P(Y = y) = \mathbb{E}X \cdot \mathbb{E}Y.$$

Součet nezávislých n.v.

- Máme-li dáno $p_{X,Y}$, jak zjistit rozdělení součtu, $Z = X + Y$?

Součet nezávislých n.v. – konvoluce

Věta

Pokud X, Y *jsou diskrétní nezávislé náhodné veličiny (zkráceně n.n.v.), tak jejich součet* $Z = X + Y$ *má pravděpodobnostní funkci*

$$P(Z = z) = \sum_{x \in ImX} P(X = x) P(Y = z - x).$$

Přehled

Podmíněné rozdělení

X, Y – diskrétní náhodné veličiny na $(\Omega, \mathcal{F}, P)$, $A \in \mathcal{F}$

- $p_{X|A}(x) := P(X = x \mid A)$
  příklad: X je výsledek hodu kostkou, $A =$ padlo sudé číslo
- $p_{X|Y}(x|y) = P(X = x \mid Y = y)$ příklad: X, Z jsou výsledky dvou nezávislých hodů kostkou, $Y = X + Z$.

$$p_{X|Y}(6|10) =$$

- $p_{X|Y}$ z $p_{X,Y}$:

Sdružené vs. podmíněné rozdělení

$p_{X,Y}$	...	10	11	12
1				
2				
3				
4				
5				
6				

$p_{X\|Y}$	...	10	11	12
1				
2				
3				
4				
5				
6				