NMAI059 Pravděpodobnost a statistika 1
4. přednáška

Robert Šámal

Přehled

Diskrétní n.v. – střední hodnota a rozptyl

Parametry náhodných veličin

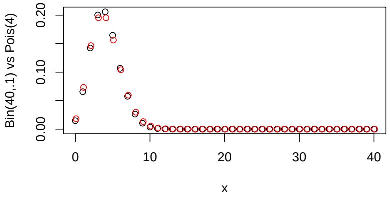
Náhodné vektory

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

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

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*

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

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).

Věta

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

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.

Rozbor všech možností

Věta

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

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

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

Rozbor všech možností

Přehled

Diskrétní n.v. – střední hodnota a rozptyl

Parametry náhodných veličin

Náhodné vektory

Podmíněné rozdělení

Parametry rozdělení – Bernoulliho

Pro $X \sim Bern(p)$ je

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

Parametry rozdělení – binomické

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 $X_i =$
- $\mathbb{E}(X_i) = P(X_i = 1) =$

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

Parametry rozdělení – hypergeometrické

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

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

- První postup: $X = \sum_{i=1}^{n} X_i$, kde $X_i =$
- $\mathbb{E}(X_i) = P(X_i = 1) =$

- Druhý postup: $X = \sum_{j=1}^{K} Y_j$, kde $Y_j =$
- $\mathbb{E}(Y_j) = P(Y_j = 1) =$

Parametry rozdělení – geometrické

Pro $X \sim Geom(p)$ je

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

Parametry rozdělení – Poissonovo

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

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

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*

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

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

Marginální rozdělení

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

Marginální rozdělení

Věta
Nechť X, Y jsou diskrétní n.v. Pak

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

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

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

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

Důkaz věty o rozptylu

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*

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

Součin nezávislých n.v.

Věta

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

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

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				