

ANALÝZA DAT V R

Náhodná veličina

- R má pro vybraná pravděpodobnostní rozdělení implementovanou hustotu, resp. pravděpodobnostní funkci, distribuční funkci, kvantilovou funkci a generátor náhodných čísel
- nechť X je náhodná veličina

ZKRATKY FUNKCÍ		
zkratka	význam	argumenty
d	hustota $f_X(x)$, resp. pravěpodobnostní funkce $P(X = x)$ (z angl. density)	x , param.rozdeleni
p	distribuční funkce $F_X(q) = P(X \leq q)$ (z angl. probability)	q , param.rozdeleni
q	kvantilová funkce $F_X^{-1}(p) = \min\{x : F_X(q) \geq p\}$ (z angl. quantile)	p , param.rozdeleni
r	náhodný výběr (z angl. random)	n , param.rozdeleni

Poznámka: Argumenty **x**, **q** a **p** mohou být i vektory.

PRAVDĚPODOBNOSTNÍ ROZDĚLENÍ			
zkratka	rozdělení	parametry	$P(X = x)$, resp. $f(x)$
binom	binomické	size , prob (n, p)	$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, x = 0, 1, \dots, n$
geom	geometrické	prob (p)	$P(X = x) = p(1 - p)^x, x = 0, 1, \dots$
hyper	hypergeometrické	m , n , k (m, n, k)	$P(X = x) = \frac{\binom{m}{x} \binom{n}{k-x}}{\binom{m+n}{k}}, \max(0, k - n) \leq x \leq \min(k, m)$
pois	Poissonovo	lambda (λ)	$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, \dots$
unif	rovnorné	min , max (a, b)	$f(x) = \frac{1}{b-a}, x \in (a, b)$
exp	exponenciální	rate (λ)	$f(x) = \lambda e^{-\lambda x}, x \in (0, \infty)$
norm	normální	mean , sd (μ, σ)	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}$
t	Studentovo t	df (n)	$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}, x \in \mathbb{R}$
chisq	χ^2	df (n)	$f(x) = \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} x^{n/2-1} e^{-x/2}, x \in (0, \infty)$

- kombinace zkratky funkce a zkratky rozdělení plus zadání konkrétních hodnot argumentů, např.
`> pbinom(x=5, size=10, prob=0.2) ... P(X ≤ 5)`, kde X má binomické rozdělení s par. $n = 10$ a $p = 0.2$
`> qgeom(q=0.7, p=0.6) ... F_X^{-1}(0.7)`, kde X má geometrické rozdělení s par. $p = 0.6$

```

> rpois(n=20, lambda=6) ... náhodný výběr o rozsahu 20 z Poissonova rozdělení s par.  $\lambda = 6$ 
> dexp(x=2, rate=1/4) ...  $f_X(2)$ , kde  $X$  má exponenciální rozdělení s par.  $\lambda = \frac{1}{4}$ 
> pnorm(x=1.96, mean=1, sd=0.5) ...  $P(X \leq 1.96)$ , kde  $X$  má normální rozdělení se střední hodnotou  $\mu = 1$  a směrodatnou odchylkou  $\sigma = 0.5$ 
> runif(n=100, min=0, max=1) ... náhodný výběr o rozsahu 100 z rovnoměrného rozdělení s par.  $a = 0, b = 1$ 
> qt(q=0.9, df=15) ...  $F_X^{-1}(0.9)$ , kde  $X$  má Studentovo  $t$  rozdělení s počtem stupňů volnosti  $n = 15$ 

```

Příklad 1

V osudí jsou dvě bílé a tři černé koule. Náhodně vytáhneme s vracením třikrát vždy jednu kouli.

- Jaká je pravděpodobnost, že právě dvě z vytažených koulí jsou bílé?
- Vygenerujte 5 čísel z rozdělení, které má náhodná veličina X udávající počet vytažených bílých koulí.
- Nakreslete graf pravděpodobnostní funkce příslušné náhodné veličiny.

Důležitá poznámka: Nastavením počáteční hodnoty generátoru náhodných čísel pomocí příkazu `set.seed()` docílíme stejných a především opakovatelných hodnot.

Příklad 2

V osudí jsou dvě bílé a tři černé koule. Náhodně vytáhneme bez vracení tři koule.

- Jaká je pravděpodobnost, že právě dvě z vytažených koulí jsou bílé?

- b) Jaká je pravděpodobnost, že nejvýše jedna z vytažených koulí je bílá?

Příklad 3

V osudí jsou dvě bílé a tři černé koule. Náhodně vytahujeme s vracením jednu kouli tak dlouho, dokud nevytáhneme bílou kouli. Jaká je pravděpodobnost, že poprvé bílou kouli vytáhneme později než ve druhém tahu?

Příklad 4

Bylo zjištěno, že pojišťovací agent získá v průměru jednoho klienta za čtyři pracovní dny.

- a) Jaká je pravděpodobnost, že se mu následující pracovní den podaří získat alespoň jednoho?
- b) Nakreslete graf distribuční funkce příslušné náhodné veličiny.

Příklad 5

Doba čekání operátorky na první telefonní hovor má přibližně exponenciální rozdělení se střední hodnotou 20 minut.

- a) Jaká je pravděpodobnost, že operátorka bude na první hovor čekat déle než hodinu?
- b) Nakreslete graf hustoty příslušné náhodné veličiny.

- c) S pravděpodobností 0,05 překročí příchod prvního hovoru jakou časovou mez? (neboli: určete 95% kvantil tohoto rozdělení).

Příklad 6

Určete pravděpodobnosti $P(|U| < 1)$, $P(|U| > 3)$, jestliže U je normovaná normální náhodná veličina.

Příklad 7

Výška desetiletých chlapců je náhodná veličina s normálním rozdělením $N(\mu = 136.1, \sigma^2 = 6.4^2)$.

- (a) Nakreslete graf hustoty a distribuční funkce příslušné náhodné veličiny.
- (b) S jakou pravděpodobností má náhodně vybraný desetiletý chlapec méně než 135 cm?
- (c) S jakou pravděpodobností má náhodně vybraný desetiletý chlapec více než 150 cm?
- (d) S jakou pravděpodobností má náhodně vybraný desetiletý chlapec výšku v rozmezí 130 cm až 140 cm?
- (e) V obchodě prodávají konfekční velikosti oblečení na chlapce ve výškovém rozmezí 120 cm až 150 cm. S jakou pravděpodobností si náhodně vybraný desetiletý chlapec v obchodě nebude moci koupit oblečení, protože mu bude moc malé nebo moc velké?

(f) Jakou výšku by musel mít Váš desetiletý bratr, aby patřil k 5 % nejvyšších v populaci?

(g) Jakou výšku by musel mít Váš desetiletý bratr, aby patřil k 10 % nejnižších v populaci?