

Základy biostatistiky

(MD710P09)
ak. rok 2007/2008

Karel Zvára

karel.zvara@mff.cuni.cz

<http://www.karlin.mff.cuni.cz/~zvara>

katedra pravděpodobnosti a matematické statistiky MFF UK

(naposledy upraveno 18. března 2008)



vlastnosti populačního průměru a rozptylu

srovnej s požadavky na míry polohy a míry variability

$$\mu_{\alpha+X} = \alpha + \mu_X,$$

$$\sigma_{\alpha+X}^2 = \sigma_X^2,$$

$$\sigma_{\alpha+X} = \sigma_X,$$

$$\mu_{\beta X} = \beta \cdot \mu_X,$$

$$\sigma_{\beta X}^2 = \beta^2 \cdot \sigma_X^2,$$

$$\sigma_{\beta X} = |\beta| \cdot \sigma_X,$$

pro součet náhodných veličin $X + Y$ dále platí

$$\mu_{X+Y} = \mu_X + \mu_Y$$

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY}$$

$$\sigma_{X,Y} = 0$$

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

obecně

pro nezávislé X, Y

pro nezávislé X, Y

ukázka důkazu

$$\begin{aligned}\mu_{\alpha+\beta X} &= E(\alpha + \beta X) \\ &= \sum_i (\alpha + \beta x_i^*) P(X = x_i^*) \\ &= \sum_i \alpha P(X = x_i^*) + \sum_i \beta x_i^* P(X = x_i^*) \\ &= \alpha \sum_i P(X = x_i^*) + \beta \sum_i x_i^* P(X = x_i^*) \\ &= \alpha + \beta \cdot EX = \alpha + \beta \cdot \mu_X\end{aligned}$$

normování náhodné veličiny X (populační obdoba z-skórů)

$$\begin{aligned}Z &= \frac{X - \mu_X}{\sigma_X} \quad (\text{bezrozměrné!}) \\ \Rightarrow \quad \mu_Z &= 0, \quad \sigma_Z = 1\end{aligned}$$

charakteristiky založené na normované verzi

charakteristiky X nezávislé na μ_X a σ_X ,
protějšky popisných statistik

- ▶ (populační) **korelační koeficient** [correlation coefficient]

$$\rho_{XY} = \text{cov} \left(\frac{X - \mu_X}{\sigma_X}, \frac{Y - \mu_Y}{\sigma_Y} \right) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

- ▶ (populační) **šikmost** náhodné veličiny X [skewness]

$$\gamma_1 = E \left(\frac{X - \mu_X}{\sigma_X} \right)^3 = \frac{E(X - \mu_X)^3}{\sigma_X^3}$$

- ▶ (populační) **špičatost** náhodné veličiny X (někdy bez -3)
[kurtosis]

$$\gamma_2 = E \left(\frac{X - \mu_X}{\sigma_X} \right)^4 - 3 = \frac{E(X - \mu_X)^4}{\sigma_X^4} - 3$$

charakteristiky založené na normované verzi

charakteristiky X nezávislé na μ_X a σ_X ,
protějšky popisných statistik

- ▶ (populační) **korelační koeficient** [correlation coefficient]

$$\rho_{XY} = \text{cov} \left(\frac{X - \mu_X}{\sigma_X}, \frac{Y - \mu_Y}{\sigma_Y} \right) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

- ▶ (populační) **šikmost** náhodné veličiny X [skewness]

$$\gamma_1 = E \left(\frac{X - \mu_X}{\sigma_X} \right)^3 = \frac{E(X - \mu_X)^3}{\sigma_X^3}$$

- ▶ (populační) **špičatost** náhodné veličiny X (někdy bez -3)
[kurtosis]

$$\gamma_2 = E \left(\frac{X - \mu_X}{\sigma_X} \right)^4 - 3 = \frac{E(X - \mu_X)^4}{\sigma_X^4} - 3$$

charakteristiky založené na normované verzi

charakteristiky X nezávislé na μ_X a σ_X ,
protějšky popisných statistik

- ▶ (populační) **korelační koeficient** [correlation coefficient]

$$\rho_{XY} = \text{cov} \left(\frac{X - \mu_X}{\sigma_X}, \frac{Y - \mu_Y}{\sigma_Y} \right) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

- ▶ (populační) **šikmost** náhodné veličiny X [skewness]

$$\gamma_1 = E \left(\frac{X - \mu_X}{\sigma_X} \right)^3 = \frac{E(X - \mu_X)^3}{\sigma_X^3}$$

- ▶ (populační) **špičatost** náhodné veličiny X (někdy bez -3)
[kurtosis]

$$\gamma_2 = E \left(\frac{X - \mu_X}{\sigma_X} \right)^4 - 3 = \frac{E(X - \mu_X)^4}{\sigma_X^4} - 3$$

alternativní rozdělení

nula-jedničkové, Bernoulliovo

- ▶ pouze dvě možné hodnoty: 1 (zdar), 0 (nezdar)
- ▶ $P(X = 1) = \pi, P(X = 0) = 1 - \pi$
- ▶ π je jediný parametr, $0 < \pi < 1$
- ▶ $\mu_X = EX = 1 \cdot \pi + 0 \cdot (1 - \pi) = \pi$
- ▶ $\sigma_X^2 = \text{var } X = (1 - \pi)^2 \cdot \pi + (0 - \pi)^2 \cdot (1 - \pi) = \pi(1 - \pi)$
- ▶ X – počet zdarů v jednom pokusu, kde pst zdaru je π
- ▶ $X \sim \text{alt}(\pi)$

alternativní rozdělení

nula-jedničkové, Bernoulliovo

- ▶ pouze dvě možné hodnoty: 1 (zdar), 0 (nezdar)
- ▶ $P(X = 1) = \pi, P(X = 0) = 1 - \pi$
- ▶ π je jediný parametr, $0 < \pi < 1$
- ▶ $\mu_X = E X = 1 \cdot \pi + 0 \cdot (1 - \pi) = \pi$
- ▶ $\sigma_X^2 = \text{var } X = (1 - \pi)^2 \cdot \pi + (0 - \pi)^2 \cdot (1 - \pi) = \pi(1 - \pi)$
- ▶ X – počet zdarů v jednom pokusu, kde pst zdaru je π
- ▶ $X \sim \text{alt}(\pi)$

alternativní rozdělení

nula-jedničkové, Bernoulliovo

- ▶ pouze dvě možné hodnoty: 1 (zdar), 0 (nezdar)
- ▶ $P(X = 1) = \pi, P(X = 0) = 1 - \pi$
- ▶ π je jediný parametr, $0 < \pi < 1$
- ▶ $\mu_X = E X = 1 \cdot \pi + 0 \cdot (1 - \pi) = \pi$
- ▶ $\sigma_X^2 = \text{var } X = (1 - \pi)^2 \cdot \pi + (0 - \pi)^2 \cdot (1 - \pi) = \pi(1 - \pi)$
- ▶ X – počet zdarů v jednom pokusu, kde pst zdaru je π
- ▶ $X \sim \text{alt}(\pi)$

alternativní rozdělení

nula-jedničkové, Bernoulliovo

- ▶ pouze dvě možné hodnoty: 1 (zdar), 0 (nezdar)
- ▶ $P(X = 1) = \pi, P(X = 0) = 1 - \pi$
- ▶ π je jediný parametr, $0 < \pi < 1$
- ▶ $\mu_X = EX = 1 \cdot \pi + 0 \cdot (1 - \pi) = \pi$
- ▶ $\sigma_X^2 = \text{var } X = (1 - \pi)^2 \cdot \pi + (0 - \pi)^2 \cdot (1 - \pi) = \pi(1 - \pi)$
- ▶ X – počet zdarů v jednom pokusu, kde pst zdaru je π
- ▶ $X \sim \text{alt}(\pi)$

alternativní rozdělení

nula-jedničkové, Bernoulliovo

- ▶ pouze dvě možné hodnoty: 1 (zdar), 0 (nezdar)
- ▶ $P(X = 1) = \pi, P(X = 0) = 1 - \pi$
- ▶ π je jediný parametr, $0 < \pi < 1$
- ▶ $\mu_X = EX = 1 \cdot \pi + 0 \cdot (1 - \pi) = \pi$
- ▶ $\sigma_X^2 = \text{var } X = (1 - \pi)^2 \cdot \pi + (0 - \pi)^2 \cdot (1 - \pi) = \pi(1 - \pi)$
- ▶ X – počet zdarů v jednom pokusu, kde pst zdaru je π
- ▶ $X \sim \text{alt}(\pi)$

alternativní rozdělení

nula-jedničkové, Bernoulliovo

- ▶ pouze dvě možné hodnoty: 1 (zdar), 0 (nezdar)
- ▶ $P(X = 1) = \pi, P(X = 0) = 1 - \pi$
- ▶ π je jediný parametr, $0 < \pi < 1$
- ▶ $\mu_X = EX = 1 \cdot \pi + 0 \cdot (1 - \pi) = \pi$
- ▶ $\sigma_X^2 = \text{var } X = (1 - \pi)^2 \cdot \pi + (0 - \pi)^2 \cdot (1 - \pi) = \pi(1 - \pi)$
- ▶ X – počet zdarů v jednom pokusu, kde pst zdaru je π
- ▶ $X \sim \text{alt}(\pi)$

alternativní rozdělení

nula-jedničkové, Bernoulliovo

- ▶ pouze dvě možné hodnoty: 1 (zdar), 0 (nezdar)
- ▶ $P(X = 1) = \pi, P(X = 0) = 1 - \pi$
- ▶ π je jediný parametr, $0 < \pi < 1$
- ▶ $\mu_X = EX = 1 \cdot \pi + 0 \cdot (1 - \pi) = \pi$
- ▶ $\sigma_X^2 = \text{var } X = (1 - \pi)^2 \cdot \pi + (0 - \pi)^2 \cdot (1 - \pi) = \pi(1 - \pi)$
- ▶ X – počet zdarů v jednom pokusu, kde pst zdaru je π
- ▶ $X \sim \text{alt}(\pi)$

binomické rozdělení

[binomial distribution]

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ n **nezávislých** pokusů takových, že
- ▶ $P(\text{zdar}) = \pi, P(\text{nezdar}) = 1 - \pi, (0 < \pi < 1)$
- ▶ Y je **počet zdarů** v těchto pokusech

$$P(Y = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}, \quad k = 0, 1, \dots, n$$

[dbinom(k,n,pst)]

- ▶ např. ze 7 vajíček se vylíhne Y slepiček, $Y \sim \text{bi}(7, 1/2)$
- ▶ např. při 60 hodech kostkou padlo Y šestek, $Y \sim \text{bi}(60, 1/6)$
- ▶ předem nevíme, kolik bude slepiček (šestek), ale v dlouhodobém průměru je relativní četnost blízká $1/2$ ($1/6$)

binomické rozdělení

[binomial distribution]

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ n **nezávislých** pokusů takových, že
- ▶ $P(\text{zdar}) = \pi, P(\text{nezdar}) = 1 - \pi, (0 < \pi < 1)$
- ▶ Y je **počet zdarů** v těchto pokusech

$$P(Y = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}, \quad k = 0, 1, \dots, n$$

[dbinom(k,n,pst)]

- ▶ např. ze 7 vajíček se vylíhne Y slepiček, $Y \sim \text{bi}(7, 1/2)$
- ▶ např. při 60 hodech kostkou padlo Y šestek, $Y \sim \text{bi}(60, 1/6)$
- ▶ předem nevíme, kolik bude slepiček (šestek), ale v dlouhodobém průměru je relativní četnost blízká $1/2$ ($1/6$)

binomické rozdělení

[binomial distribution]

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ n **nezávislých** pokusů takových, že
- ▶ $P(\text{zdar}) = \pi, P(\text{nezdar}) = 1 - \pi, (0 < \pi < 1)$
- ▶ Y je **počet zdarů** v těchto pokusech

$$P(Y = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}, \quad k = 0, 1, \dots, n$$

[dbinom(k,n,pst)]

- ▶ např. ze 7 vajíček se vylíhne Y slepiček, $Y \sim \text{bi}(7, 1/2)$
- ▶ např. při 60 hodech kostkou padlo Y šestek, $Y \sim \text{bi}(60, 1/6)$
- ▶ předem nevíme, kolik bude slepiček (šestek), ale v dlouhodobém průměru je relativní četnost blízká $1/2$ ($1/6$)

binomické rozdělení

[binomial distribution]

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ n **nezávislých** pokusů takových, že
- ▶ $P(\text{zdar}) = \pi, P(\text{nezdar}) = 1 - \pi, (0 < \pi < 1)$
- ▶ Y je **počet zdarů** v těchto pokusech

$$P(Y = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}, \quad k = 0, 1, \dots, n$$

[dbinom(k,n,pst)]

- ▶ např. ze 7 vajíček se vylíhne Y slepiček, $Y \sim \text{bi}(7, 1/2)$
- ▶ např. při 60 hodech kostkou padlo Y šestek, $Y \sim \text{bi}(60, 1/6)$
- ▶ předem nevíme, kolik bude slepiček (šestek), ale v dlouhodobém průměru je relativní četnost blízká $1/2$ ($1/6$)

binomické rozdělení

[binomial distribution]

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ n **nezávislých** pokusů takových, že
- ▶ $P(\text{zdar}) = \pi, P(\text{nezdar}) = 1 - \pi, (0 < \pi < 1)$
- ▶ Y je **počet zdarů** v těchto pokusech

- ▶
$$P(Y = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}, \quad k = 0, 1, \dots, n$$

[dbinom(k,n,pst)]

- ▶ např. ze 7 vajíček se vylíhne Y slepiček, $Y \sim \text{bi}(7, 1/2)$
- ▶ např. při 60 hodech kostkou padlo Y šestek, $Y \sim \text{bi}(60, 1/6)$
- ▶ předem nevíme, kolik bude slepiček (šestek), ale v dlouhodobém průměru je relativní četnost blízká $1/2$ ($1/6$)

binomické rozdělení

[binomial distribution]

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ n **nezávislých** pokusů takových, že
- ▶ $P(\text{zdar}) = \pi, P(\text{nezdar}) = 1 - \pi, (0 < \pi < 1)$
- ▶ Y je **počet zdarů** v těchto pokusech

- ▶
$$P(Y = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}, \quad k = 0, 1, \dots, n$$

[dbinom(k,n,pst)]

- ▶ např. ze 7 vajíček se vylíhne Y slepiček, $Y \sim \text{bi}(7, 1/2)$
- ▶ např. při 60 hodech kostkou padlo Y šestek, $Y \sim \text{bi}(60, 1/6)$
- ▶ předem nevíme, kolik bude slepiček (šestek), ale v dlouhodobém průměru je relativní četnost blízká $1/2$ ($1/6$)

binomické rozdělení

[binomial distribution]

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ n **nezávislých** pokusů takových, že
- ▶ $P(\text{zdar}) = \pi, P(\text{nezdar}) = 1 - \pi, (0 < \pi < 1)$
- ▶ Y je **počet zdarů** v těchto pokusech

- ▶
$$P(Y = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}, \quad k = 0, 1, \dots, n$$

[dbinom(k,n,pst)]

- ▶ např. ze 7 vajíček se vylíhne Y slepiček, $Y \sim \text{bi}(7, 1/2)$
- ▶ např. při 60 hodech kostkou padlo Y šestek, $Y \sim \text{bi}(60, 1/6)$
- ▶ předem nevíme, kolik bude slepiček (šestek), ale v dlouhodobém průměru je relativní četnost blízká $1/2$ ($1/6$)

binomické rozdělení

[binomial distribution]

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ n **nezávislých** pokusů takových, že
- ▶ $P(\text{zdar}) = \pi, P(\text{nezdar}) = 1 - \pi, (0 < \pi < 1)$
- ▶ Y je **počet zdarů** v těchto pokusech

- ▶
$$P(Y = k) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}, \quad k = 0, 1, \dots, n$$

[dbinom(k,n,pst)]

- ▶ např. ze 7 vajíček se vylíhne Y slepiček, $Y \sim \text{bi}(7, 1/2)$
- ▶ např. při 60 hodech kostkou padlo Y šestek, $Y \sim \text{bi}(60, 1/6)$
- ▶ předem nevíme, kolik bude slepiček (šestek), ale v dlouhodobém průměru je relativní četnost blízká $1/2$ ($1/6$)

binomické rozdělení pomocí alternativního

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ Y je celkový počet zdarů v n pokusech, tedy
- ▶ $Y = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$, kde X_i je počet zdarů v i -tém pokusu
- ▶ z vlastností střední hodnoty (očekávaný počet zdarů)

$$\mu_Y = E Y = E \sum_{i=1}^n X_i = \sum_{i=1}^n E X_i = \sum_{i=1}^n \pi = n\pi$$

- ▶ protože jsou pokusy nezávislé

$$\sigma_Y^2 = \text{var} \sum_{i=1}^n X_i = \sum_{i=1}^n \text{var} X_i = \sum_{i=1}^n \pi(1 - \pi) = n\pi(1 - \pi)$$

binomické rozdělení pomocí alternativního

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ Y je celkový počet zdarů v n pokusech, tedy
- ▶ $Y = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$, kde X_i je počet zdarů v i -tém pokusu
- ▶ z vlastností střední hodnoty (očekávaný počet zdarů)

$$\mu_Y = E Y = E \sum_{i=1}^n X_i = \sum_{i=1}^n E X_i = \sum_{i=1}^n \pi = n\pi$$

- ▶ protože jsou pokusy nezávislé

$$\sigma_Y^2 = \text{var} \sum_{i=1}^n X_i = \sum_{i=1}^n \text{var} X_i = \sum_{i=1}^n \pi(1 - \pi) = n\pi(1 - \pi)$$

binomické rozdělení pomocí alternativního

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ Y je celkový počet zdarů v n pokusech, tedy
- ▶ $Y = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$, kde X_i je počet zdarů v i -tém pokusu
- ▶ z vlastností střední hodnoty (očekávaný počet zdarů)

$$\mu_Y = E Y = E \sum_{i=1}^n X_i = \sum_{i=1}^n E X_i = \sum_{i=1}^n \pi = n\pi$$

- ▶ protože jsou pokusy nezávislé

$$\sigma_Y^2 = \text{var} \sum_{i=1}^n X_i = \sum_{i=1}^n \text{var} X_i = \sum_{i=1}^n \pi(1 - \pi) = n\pi(1 - \pi)$$

binomické rozdělení pomocí alternativního

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ Y je celkový počet zdarů v n pokusech, tedy
- ▶ $Y = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$, kde X_i je počet zdarů v i -tém pokusu
- ▶ z vlastností střední hodnoty (očekávaný počet zdarů)

$$\mu_Y = E Y = E \sum_{i=1}^n X_i = \sum_{i=1}^n E X_i = \sum_{i=1}^n \pi = n\pi$$

- ▶ protože jsou pokusy nezávislé

$$\sigma_Y^2 = \text{var} \sum_{i=1}^n X_i = \sum_{i=1}^n \text{var} X_i = \sum_{i=1}^n \pi(1 - \pi) = n\pi(1 - \pi)$$

binomické rozdělení pomocí alternativního

- ▶ $Y \sim \text{bi}(n, \pi)$
- ▶ Y je celkový počet zdarů v n pokusech, tedy
- ▶ $Y = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$, kde X_i je počet zdarů v i -tém pokusu
- ▶ z vlastností střední hodnoty (očekávaný počet zdarů)

$$\mu_Y = E Y = E \sum_{i=1}^n X_i = \sum_{i=1}^n E X_i = \sum_{i=1}^n \pi = n\pi$$

- ▶ protože jsou pokusy nezávislé

$$\sigma_Y^2 = \text{var} \sum_{i=1}^n X_i = \sum_{i=1}^n \text{var} X_i = \sum_{i=1}^n \pi(1 - \pi) = n\pi(1 - \pi)$$

příklad: kuřáci

- ▶ mezi dvacetiletými muži je (řekněme) 35 % kuřáků ($\pi = 0,35$)
- ▶ je-li dvacetiletých 70 tisíc ($m = 70000$), pak je kuřáků asi $m\pi = 70\,000 \cdot 0,35 = 24\,500$, ale nevíme, kteří to jsou
- ▶ vyberme náhodně $n = 60$ dvacetiletých mužů, označme jako Y počet kuřáků mezi nimi, je tedy $Y \sim \text{bi}(60, 0,35)$
- ▶ střední hodnota (očekávaný počet), rozptyl

$$\mu_Y = 60 \cdot 0,35 = 21 \quad \sigma_Y^2 = 60 \cdot 0,35 \cdot 0,65 = 13,65 \doteq (3,7)^2$$

- ▶ ukázky pravděpodobností možných hodnot

k	15	17	19	21	23	25
$P(Y = k)$	0,029	0,062	0,095	0,107	0,091	0,059

- ▶ psti počítány pomocí `[dbinom(0:60,60,0.35)]`

příklad: kuřáci

- ▶ mezi dvacetiletými muži je (řekněme) 35 % kuřáků ($\pi = 0,35$)
- ▶ je-li dvacetiletých 70 tisíc ($m = 70000$), pak je kuřáků asi $m\pi = 70\,000 \cdot 0,35 = 24\,500$, ale nevíme, kteří to jsou
- ▶ vyberme náhodně $n = 60$ dvacetiletých mužů, označme jako Y počet kuřáků mezi nimi, je tedy $Y \sim \text{bi}(60, 0,35)$
- ▶ střední hodnota (očekávaný počet), rozptyl

$$\mu_Y = 60 \cdot 0,35 = 21 \quad \sigma_Y^2 = 60 \cdot 0,35 \cdot 0,65 = 13,65 \doteq (3,7)^2$$

- ▶ ukázky pravděpodobností možných hodnot

k	15	17	19	21	23	25
$P(Y = k)$	0,029	0,062	0,095	0,107	0,091	0,059

- ▶ psti počítány pomocí `[dbinom(0:60,60,0.35)]`

příklad: kuřáci

- ▶ mezi dvacetiletými muži je (řekněme) 35 % kuřáků ($\pi = 0,35$)
- ▶ je-li dvacetiletých 70 tisíc ($m = 70000$), pak je kuřáků asi $m\pi = 70\,000 \cdot 0,35 = 24\,500$, ale nevíme, kteří to jsou
- ▶ vyberme náhodně $n = 60$ dvacetiletých mužů, označme jako Y počet kuřáků mezi nimi, je tedy $Y \sim \text{bi}(60, 0,35)$
- ▶ střední hodnota (očekávaný počet), rozptyl

$$\mu_Y = 60 \cdot 0,35 = 21 \quad \sigma_Y^2 = 60 \cdot 0,35 \cdot 0,65 = 13,65 \doteq (3,7)^2$$

- ▶ ukázky pravděpodobností možných hodnot

k	15	17	19	21	23	25
$P(Y = k)$	0,029	0,062	0,095	0,107	0,091	0,059

- ▶ psti počítány pomocí `[dbinom(0:60,60,0.35)]`

příklad: kuřáci

- ▶ mezi dvacetiletými muži je (řekněme) 35 % kuřáků ($\pi = 0,35$)
- ▶ je-li dvacetiletých 70 tisíc ($m = 70000$), pak je kuřáků asi $m\pi = 70\,000 \cdot 0,35 = 24\,500$, ale nevíme, kteří to jsou
- ▶ vyberme náhodně $n = 60$ dvacetiletých mužů, označme jako Y počet kuřáků mezi nimi, je tedy $Y \sim \text{bi}(60, 0,35)$
- ▶ střední hodnota (očekávaný počet), rozptyl

$$\mu_Y = 60 \cdot 0,35 = 21 \quad \sigma_Y^2 = 60 \cdot 0,35 \cdot 0,65 = 13,65 \doteq (3,7)^2$$

- ▶ ukázky pravděpodobností možných hodnot

k	15	17	19	21	23	25
$P(Y = k)$	0,029	0,062	0,095	0,107	0,091	0,059

- ▶ psti počítány pomocí `[dbinom(0:60,60,0.35)]`

příklad: kuřáci

- ▶ mezi dvacetiletými muži je (řekněme) 35 % kuřáků ($\pi = 0,35$)
- ▶ je-li dvacetiletých 70 tisíc ($m = 70000$), pak je kuřáků asi $m\pi = 70\,000 \cdot 0,35 = 24\,500$, ale nevíme, kteří to jsou
- ▶ vyberme náhodně $n = 60$ dvacetiletých mužů, označme jako Y počet kuřáků mezi nimi, je tedy $Y \sim \text{bi}(60, 0,35)$
- ▶ střední hodnota (očekávaný počet), rozptyl

$$\mu_Y = 60 \cdot 0,35 = 21 \quad \sigma_Y^2 = 60 \cdot 0,35 \cdot 0,65 = 13,65 \doteq (3,7)^2$$

- ▶ ukázky pravděpodobností možných hodnot

k	15	17	19	21	23	25
$P(Y = k)$	0,029	0,062	0,095	0,107	0,091	0,059

- ▶ psti počítány pomocí `[dbinom(0:60,60,0.35)]`

příklad: kuřáci

- ▶ mezi dvacetiletými muži je (řekněme) 35 % kuřáků ($\pi = 0,35$)
- ▶ je-li dvacetiletých 70 tisíc ($m = 70000$), pak je kuřáků asi $m\pi = 70\,000 \cdot 0,35 = 24\,500$, ale nevíme, kteří to jsou
- ▶ vyberme náhodně $n = 60$ dvacetiletých mužů, označme jako Y počet kuřáků mezi nimi, je tedy $Y \sim \text{bi}(60, 0,35)$
- ▶ střední hodnota (očekávaný počet), rozptyl

$$\mu_Y = 60 \cdot 0,35 = 21 \quad \sigma_Y^2 = 60 \cdot 0,35 \cdot 0,65 = 13,65 \doteq (3,7)^2$$

- ▶ ukázky pravděpodobností možných hodnot

k	15	17	19	21	23	25
$P(Y = k)$	0,029	0,062	0,095	0,107	0,091	0,059

- ▶ psti počítány pomocí `[dbinom(0:60,60,0.35)]`

Poissonovo rozdělení

[Poisson distribution]

- ▶ $X \sim \text{Po}(\lambda)$ ($\lambda > 0$)
- ▶ zákon vzácných (řídkých) jevů
- ▶ kolikrát nastal jev během jednotkového časového intervalu, na jednotkové ploše, v jednotkovém objemu ...

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

- ▶ $\mu_X = \lambda, \sigma_X^2 = \lambda$
- ▶ pro velké n a malé π lze rozdělení $\text{bi}(n, \pi)$ aproximovat pomocí rozdělení $\text{Po}(n\pi)$
- ▶ např. počet kolonií na Petriho misce

Poissonovo rozdělení

[Poisson distribution]

- ▶ $X \sim \text{Po}(\lambda)$ ($\lambda > 0$)
- ▶ zákon vzácných (řídkých) jevů
- ▶ kolikrát nastal jev během jednotkového časového intervalu, na jednotkové ploše, v jednotkovém objemu ...
- ▶
$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots$$
- ▶ $\mu_X = \lambda, \sigma_X^2 = \lambda$
- ▶ pro velké n a malé π lze rozdělení $\text{bi}(n, \pi)$ aproximovat pomocí rozdělení $\text{Po}(n\pi)$
- ▶ např. počet kolonií na Petriho misce

Poissonovo rozdělení

[Poisson distribution]

- ▶ $X \sim \text{Po}(\lambda)$ ($\lambda > 0$)
- ▶ zákon vzácných (řídkých) jevů
- ▶ kolikrát nastal jev během jednotkového časového intervalu, na jednotkové ploše, v jednotkovém objemu ...

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

- ▶ $\mu_X = \lambda, \sigma_X^2 = \lambda$
- ▶ pro velké n a malé π lze rozdělení $\text{bi}(n, \pi)$ aproximovat pomocí rozdělení $\text{Po}(n\pi)$
- ▶ např. počet kolonií na Petriho misce

Poissonovo rozdělení

[Poisson distribution]

- ▶ $X \sim \text{Po}(\lambda)$ ($\lambda > 0$)
- ▶ zákon vzácných (řídkých) jevů
- ▶ kolikrát nastal jev během jednotkového časového intervalu, na jednotkové ploše, v jednotkovém objemu ...

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

- ▶ $\mu_X = \lambda, \sigma_X^2 = \lambda$
- ▶ pro velké n a malé π lze rozdělení $\text{bi}(n, \pi)$ aproximovat pomocí rozdělení $\text{Po}(n\pi)$
- ▶ např. počet kolonií na Petriho misce

Poissonovo rozdělení

[Poisson distribution]

- ▶ $X \sim \text{Po}(\lambda)$ ($\lambda > 0$)
- ▶ zákon vzácných (řidkých) jevů
- ▶ kolikrát nastal jev během jednotkového časového intervalu, na jednotkové ploše, v jednotkovém objemu ...

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

- ▶ $\mu_X = \lambda, \sigma_X^2 = \lambda$
- ▶ pro velké n a malé π lze rozdělení $\text{bi}(n, \pi)$ aproximovat pomocí rozdělení $\text{Po}(n\pi)$
- ▶ např. počet kolonií na Petriho misce

Poissonovo rozdělení

[Poisson distribution]

- ▶ $X \sim \text{Po}(\lambda)$ ($\lambda > 0$)
- ▶ zákon vzácných (řídkých) jevů
- ▶ kolikrát nastal jev během jednotkového časového intervalu, na jednotkové ploše, v jednotkovém objemu ...
- ▶
$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots$$
- ▶ $\mu_X = \lambda, \sigma_X^2 = \lambda$
- ▶ pro velké n a malé π lze rozdělení $\text{bi}(n, \pi)$ aproximovat pomocí rozdělení $\text{Po}(n\pi)$
- ▶ např. počet kolonií na Petriho misce

Poissonovo rozdělení

[Poisson distribution]

- ▶ $X \sim \text{Po}(\lambda)$ ($\lambda > 0$)
- ▶ zákon vzácných (řídkých) jevů
- ▶ kolikrát nastal jev během jednotkového časového intervalu, na jednotkové ploše, v jednotkovém objemu ...
- ▶
$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots$$
- ▶ $\mu_X = \lambda, \sigma_X^2 = \lambda$
- ▶ pro velké n a malé π lze rozdělení $\text{bi}(n, \pi)$ aproximovat pomocí rozdělení $\text{Po}(n\pi)$
- ▶ např. počet kolonií na Petriho misce

příklad

s jakou pstí udělá 5 z 55 stejně připravených studentů zkoušku na výbornou, je-li pst jedničky 0,1?

- binomické rozdělení $Y \sim \text{bi}(55, 0,1)$ [dbinom(5,55,0.1)]

$$P(Y = 5) = \binom{55}{5} \cdot 0,1^5 \cdot 0,9^{50} = 0,179$$

- aproximace Poissonovým rozdělením
 $Y \sim \text{Po}(55 \cdot 0,1) = \text{Po}(5,5)$ [dpois(5, 5.5)]

$$P(Y = 5) = \frac{5,5^5}{5!} e^{-5,5} = 0,171$$

příklad

s jakou pstí udělá 5 z 55 stejně připravených studentů zkoušku na výbornou, je-li pst jedničky 0,1?

- ▶ binomické rozdělení $Y \sim \text{bi}(55, 0,1)$ [dbinom(5,55,0.1)]

$$P(Y = 5) = \binom{55}{5} \cdot 0,1^5 \cdot 0,9^{50} = 0,179$$

- ▶ aproximace Poissonovým rozdělením
 $Y \sim \text{Po}(55 \cdot 0,1) = \text{Po}(5,5)$ [dpois(5, 5.5)]

$$P(Y = 5) = \frac{5,5^5}{5!} e^{-5,5} = 0,171$$

normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

[normal (Gaussian) distribution]

- ▶ $\mu_X = \mu, \sigma_X^2 = \sigma^2$
- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
- ▶ maximální hodnota hustoty úměrná $1/\sigma$
- ▶ $N(0, 1)$ (normované normální rozdělení):
 $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ (hustota),
 $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$ (distr. fce)
- ▶ $X \sim N(\mu, \sigma^2)$, pak $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

- ▶ model vzniku: součet velkého počtu nepatrných příspěvků
- ▶ velmi často modeluje znaky v poměrovém měřítku

normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

[normal (Gaussian) distribution]

- ▶ $\mu_X = \mu, \sigma_X^2 = \sigma^2$
- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
 - ▶ maximální hodnota hustoty úměrná $1/\sigma$
 - ▶ $N(0, 1)$ (normované normální rozdělení):
 $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ (hustota),
 $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$ (distr. fce)
- ▶ $X \sim N(\mu, \sigma^2)$, pak $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

- ▶ model vzniku: součet velkého počtu nepatrných příspěvků
- ▶ velmi často modeluje znaky v poměrovém měřítku

normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

[normal (Gaussian) distribution]

- ▶ $\mu_X = \mu, \sigma_X^2 = \sigma^2$
- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
- ▶ maximální hodnota hustoty úměrná $1/\sigma$
- ▶ $N(0, 1)$ (normované normální rozdělení):
 $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ (hustota),
 $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$ (distr. fce)
- ▶ $X \sim N(\mu, \sigma^2)$, pak $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

- ▶ model vzniku: součet velkého počtu nepatrných příspěvků
- ▶ velmi často modeluje znaky v poměrovém měřítku

normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

[normal (Gaussian) distribution]

- ▶ $\mu_X = \mu, \sigma_X^2 = \sigma^2$
- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
- ▶ maximální hodnota hustoty úměrná $1/\sigma$
- ▶ $N(0, 1)$ (normované normální rozdělení):
 $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ (hustota),
 $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$ (distr. fce)
- ▶ $X \sim N(\mu, \sigma^2)$, pak $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

- ▶ model vzniku: součet velkého počtu nepatrných příspěvků
- ▶ velmi často modeluje znaky v poměrovém měřítku

normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

[normal (Gaussian) distribution]

- ▶ $\mu_X = \mu, \sigma_X^2 = \sigma^2$
- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
- ▶ maximální hodnota hustoty úměrná $1/\sigma$
- ▶ $N(0, 1)$ (normované normální rozdělení):
 $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ (hustota),
 $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$ (distr. fce)
- ▶ $X \sim N(\mu, \sigma^2)$, pak $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

- ▶ model vzniku: součet velkého počtu nepatrných příspěvků
- ▶ velmi často modeluje znaky v poměrovém měřítku

normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

[normal (Gaussian) distribution]

- ▶ $\mu_X = \mu, \sigma_X^2 = \sigma^2$
- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
- ▶ maximální hodnota hustoty úměrná $1/\sigma$
- ▶ $N(0, 1)$ (normované normální rozdělení):
 $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ (hustota),
 $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$ (distr. fce)
- ▶ $X \sim N(\mu, \sigma^2)$, pak $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

- ▶ model vzniku: součet velkého počtu nepatrných příspěvků
- ▶ velmi často modeluje znaky v poměrovém měřítku

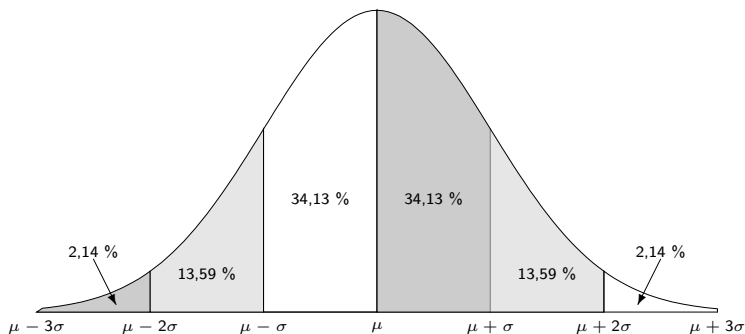
normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

[normal (Gaussian) distribution]

- ▶ $\mu_X = \mu, \sigma_X^2 = \sigma^2$
- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
- ▶ maximální hodnota hustoty úměrná $1/\sigma$
- ▶ $N(0, 1)$ (normované normální rozdělení):
 $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ (hustota),
 $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$ (distr. fce)
- ▶ $X \sim N(\mu, \sigma^2)$, pak $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$

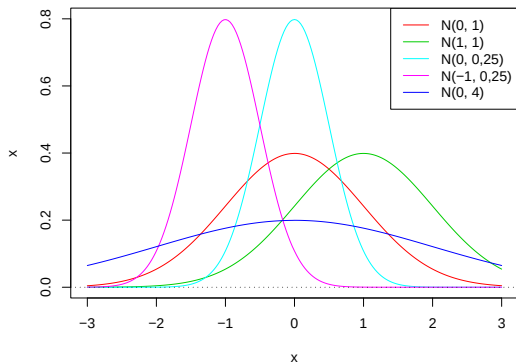
$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

- ▶ model vzniku: součet velkého počtu nepatrných příspěvků
- ▶ velmi často modeluje znaky v poměrovém měřítku

hustota $N(\mu, \sigma^2)$ $[dnorm(x, mu, sigma)]$ 

normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

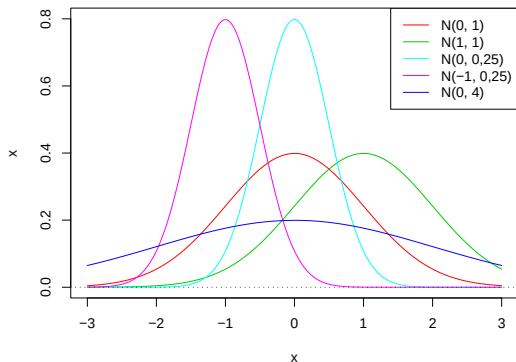
význam parametrů



- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
- ▶ maximální hodnota hustoty přibližně $0,4/\sigma$
- ▶ model vzniku: součet velkého počtu nepatrných příspěvků

normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

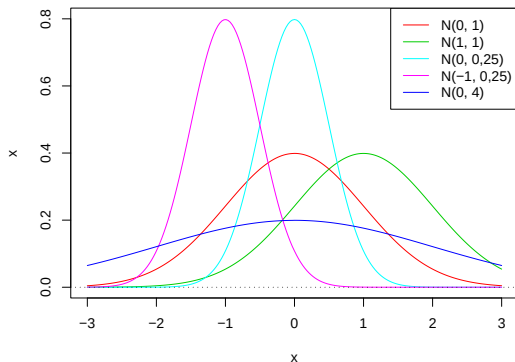
význam parametrů



- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
- ▶ maximální hodnota hustoty přibližně $0,4/\sigma$
- ▶ model vzniku: součet velkého počtu nepatrných příspěvků

normální (Gaussovo) rozdělení $N(\mu, \sigma^2)$

význam parametrů



- ▶ spojité rozdělení, symetrické okolo střední hodnoty μ
- ▶ maximální hodnota hustoty přibližně $0,4/\sigma$
- ▶ model vzniku: součet velkého počtu nepatrných příspěvků

výpočet pravděpodobnosti, že $a < X < b$
 použije distribuční funkci $N(0, 1)$

$$P(a < X < b) = F_X(b) - F_X(a) \quad \text{platí obecně pro spoj. rozdělení.}$$

$$X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

$$P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P\left(Z \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

$$P(a < X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

$$[\text{pnorm}((b-\mu)/\text{sigma}) - \text{pnorm}((a-\mu)/\text{sigma})]$$

v programu R je distribuční funkce $N(\mu, \sigma^2)$ s obecnými parametry:

$$[\text{pnorm}(b, \mu, \text{sigma}) - \text{pnorm}(a, \mu, \text{sigma})]$$

příklad

- ▶ jaký díl populace desetiletých hochů má výšku od 135 do 140 cm, když pro výšku desetiletých platí $X \sim N(136,1, 6,4^2)$
- ▶ předpokládáme zaokrouhlování na celá čísla při měření

$$\begin{aligned} P(134,5 < X < 140,5) &= \Phi\left(\frac{140,5 - 136,1}{6,4}\right) - \Phi\left(\frac{134,5 - 136,1}{6,4}\right) \\ &= 0,754 - 0,401 = 0,353 \end{aligned}$$

`[pnorm((140.5-136.1)/6.4)-pnorm((134.5-136.1)/6.4)]`

- ▶ pomocí distr. fce s obecnými parametry
`[pnorm(140.5,136.1,6.4)-pnorm(134.5,136.1,6.4)]`

příklad

- ▶ jaký díl populace desetiletých hochů má výšku od 135 do 140 cm, když pro výšku desetiletých platí $X \sim N(136,1, 6,4^2)$
- ▶ předpokládáme zaokrouhlování na celá čísla při měření

$$\begin{aligned} P(134,5 < X < 140,5) &= \Phi\left(\frac{140,5 - 136,1}{6,4}\right) - \Phi\left(\frac{134,5 - 136,1}{6,4}\right) \\ &= 0,754 - 0,401 = 0,353 \end{aligned}$$

`[pnorm((140.5-136.1)/6.4)-pnorm((134.5-136.1)/6.4)]`

- ▶ pomocí distr. fce s obecnými parametry

`[pnorm(140.5,136.1,6.4)-pnorm(134.5,136.1,6.4)]`

příklad

- ▶ jaký díl populace desetiletých hochů má výšku od 135 do 140 cm, když pro výšku desetiletých platí $X \sim N(136,1, 6,4^2)$
- ▶ předpokládáme zaokrouhlování na celá čísla při měření

$$\begin{aligned} P(134,5 < X < 140,5) &= \Phi\left(\frac{140,5 - 136,1}{6,4}\right) - \Phi\left(\frac{134,5 - 136,1}{6,4}\right) \\ &= 0,754 - 0,401 = 0,353 \end{aligned}$$

`[pnorm((140.5-136.1)/6.4)-pnorm((134.5-136.1)/6.4)]`

- ▶ pomocí distr. fce s obecnými parametry
`[pnorm(140.5,136.1,6.4)-pnorm(134.5,136.1,6.4)]`

kritické hodnoty normálního a Studentova t -rozdělení

[Student distribution]

- **normální rozdělení** $N(0, 1)$ [qnorm(1-alpha)]

$$Z \sim N(0, 1) : P(Z > z(\alpha)) = \alpha$$

ze symetrie platí $P(|Z| > z(\alpha/2)) = \alpha$

- **Studentovo t -rozdělení** t_k
(podobné normálnímu, protože místo σ používá jeho odhad, má větší rozptyl)

$$T \sim t_k : P(|T| > t_k(\alpha)) = \alpha$$

- jsou to spíše kritické hodnoty $|T|$ [qt(1-alpha/2,k)]

kritické hodnoty normálního a Studentova t -rozdělení

[Student distribution]

- **normální rozdělení** $N(0, 1)$ [qnorm(1-alpha)]

$$Z \sim N(0, 1) : \quad P(Z > z(\alpha)) = \alpha$$

ze symetrie platí $P(|Z| > z(\alpha/2)) = \alpha$

- **Studentovo t -rozdělení** t_k
(podobné normálnímu, protože místo σ používá jeho odhad, má větší rozptyl)

$$T \sim t_k : \quad P(|T| > t_k(\alpha)) = \alpha$$

- jsou to spíše kritické hodnoty $|T|$ [qt(1-alpha/2,k)]

kritické hodnoty normálního a Studentova t -rozdělení

[Student distribution]

- **normální rozdělení** $N(0, 1)$ [qnorm(1-alpha)]

$$Z \sim N(0, 1) : \quad P(Z > z(\alpha)) = \alpha$$

ze symetrie platí $P(|Z| > z(\alpha/2)) = \alpha$

- **Studentovo t -rozdělení** t_k
(podobné normálnímu, protože místo σ používá jeho odhad, má větší rozptyl)

$$T \sim t_k : \quad P(|T| > t_k(\alpha)) = \alpha$$

- jsou to spíše kritické hodnoty $|T|$ [qt(1-alpha/2,k)]

některé kritické hodnoty

α	0,50	0,25	0,10	0,05	0,01
$z(\alpha/2)$	0,674	1,150	1,645	1,960	2,576
$t_{100}(\alpha)$	0,677	1,157	1,660	1,984	2,626
$t_{20}(\alpha)$	0,687	1,185	1,725	2,086	2,845
$t_5(\alpha)$	0,727	1,301	2,015	2,571	4,032

- ▶ $T \sim t_k$ má jediný parametr k (počet stupňů volnosti)
- ▶ s rostoucím k se chování blíží normálnímu rozdělení
- ▶ pro $T \sim t_5$ je 95 % hodnot v intervalu $(-2,571; 2,571)$
- ▶ pro $Z \sim N(0, 1)$ je 95 % hodnot v intervalu $(-1,960; 1,960)$

některé kritické hodnoty

α	0,50	0,25	0,10	0,05	0,01
$z(\alpha/2)$	0,674	1,150	1,645	1,960	2,576
$t_{100}(\alpha)$	0,677	1,157	1,660	1,984	2,626
$t_{20}(\alpha)$	0,687	1,185	1,725	2,086	2,845
$t_5(\alpha)$	0,727	1,301	2,015	2,571	4,032

- ▶ $T \sim t_k$ má jediný parametr k (počet stupňů volnosti)
- ▶ s rostoucím k se chování blíží normálnímu rozdělení
- ▶ pro $T \sim t_5$ je 95 % hodnot v intervalu $(-2,571; 2,571)$
- ▶ pro $Z \sim N(0, 1)$ je 95 % hodnot v intervalu $(-1,960; 1,960)$

některé kritické hodnoty

α	0,50	0,25	0,10	0,05	0,01
$z(\alpha/2)$	0,674	1,150	1,645	1,960	2,576
$t_{100}(\alpha)$	0,677	1,157	1,660	1,984	2,626
$t_{20}(\alpha)$	0,687	1,185	1,725	2,086	2,845
$t_5(\alpha)$	0,727	1,301	2,015	2,571	4,032

- ▶ $T \sim t_k$ má jediný parametr k (počet stupňů volnosti)
- ▶ s rostoucím k se chování blíží normálnímu rozdělení
- ▶ pro $T \sim t_5$ je 95 % hodnot v intervalu $(-2,571; 2,571)$
- ▶ pro $Z \sim N(0, 1)$ je 95 % hodnot v intervalu $(-1,960; 1,960)$

některé kritické hodnoty

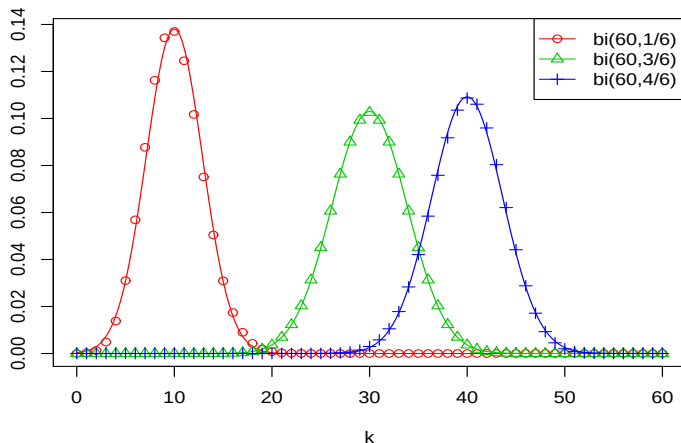
α	0,50	0,25	0,10	0,05	0,01
$z(\alpha/2)$	0,674	1,150	1,645	1,960	2,576
$t_{100}(\alpha)$	0,677	1,157	1,660	1,984	2,626
$t_{20}(\alpha)$	0,687	1,185	1,725	2,086	2,845
$t_5(\alpha)$	0,727	1,301	2,015	2,571	4,032

- ▶ $T \sim t_k$ má jediný parametr k (počet stupňů volnosti)
- ▶ s rostoucím k se chování blíží normálnímu rozdělení
- ▶ pro $T \sim t_5$ je 95 % hodnot v intervalu $(-2,571; 2,571)$
- ▶ pro $Z \sim N(0, 1)$ je 95 % hodnot v intervalu $(-1,960; 1,960)$

aproximace binomického rozdělení normálním

se stejnou střední hodnotou a stejným rozptylem

rozdělení $bi(n, \pi)$ lze aproximovat pomocí $N(n\pi, n\pi(1 - \pi))$



další rozdělení související s normálním

[F-distribution, chi-square distribution]

- ▶ V má rozdělení (musí být $P(V > 0) = 1$!!)

logaritmicko-normální, platí-li $\ln V \sim N(\mu, \sigma^2)$

- ▶ **Fisherovo F -rozdělení** $F_{k,m}$ [qf(1-alpha,k,m)]

$$F \sim F_{k,m} : P(F > F_{k,m}(\alpha)) = \alpha$$

- ▶ **rozdělení chí-kvadrát** χ_k^2 [qchisq(1-alpha,k)]

$$X^2 \sim \chi_k^2 : P(X^2 > \chi_k^2(\alpha)) = \alpha$$

- ▶ speciálně platí:

- ▶ $\chi_1^2(0,05) = 3,841 = 1,960^2$
- ▶ $\chi_1^2(\alpha) = z(\alpha/2)^2$
- ▶ $F_{1,m}(\alpha) = (t_m(\alpha))^2$

další rozdělení související s normálním

[F-distribution, chi-square distribution]

- ▶ V má rozdělení (musí být $P(V > 0) = 1$!!)

logaritmicko-normální, platí-li $\ln V \sim N(\mu, \sigma^2)$

- ▶ Fisherovo F -rozdělení $F_{k,m}$ [qf(1-alpha,k,m)]

$$F \sim F_{k,m} : P(F > F_{k,m}(\alpha)) = \alpha$$

- ▶ rozdělení chí-kvadrát χ_k^2 [qchisq(1-alpha,k)]

$$X^2 \sim \chi_k^2 : P(X^2 > \chi_k^2(\alpha)) = \alpha$$

- ▶ speciálně platí:

- ▶ $\chi_1^2(0,05) = 3,841 = 1,960^2$
- ▶ $\chi_1^2(\alpha) = z(\alpha/2)^2$
- ▶ $F_{1,m}(\alpha) = (t_m(\alpha))^2$

další rozdělení související s normálním

[F-distribution, chi-square distribution]

- ▶ V má rozdělení (musí být $P(V > 0) = 1$!!)

logaritmicko-normální, platí-li $\ln V \sim N(\mu, \sigma^2)$

- ▶ **Fisherovo F -rozdělení** $F_{k,m}$ [qf(1-alpha,k,m)]

$$F \sim F_{k,m} : P(F > F_{k,m}(\alpha)) = \alpha$$

- ▶ **rozdělení chí-kvadrát** χ_k^2 [qchisq(1-alpha,k)]

$$X^2 \sim \chi_k^2 : P(X^2 > \chi_k^2(\alpha)) = \alpha$$

- ▶ speciálně platí:

$$\chi_1^2(0,05) = 3,841 = 1,960^2$$

$$\chi_1^2(\alpha) = z(\alpha/2)^2$$

$$F_{1,m}(\alpha) = (t_m(\alpha))^2$$

další rozdělení související s normálním

[F-distribution, chi-square distribution]

- ▶ V má rozdělení (musí být $P(V > 0) = 1$!!)

logaritmicko-normální, platí-li $\ln V \sim N(\mu, \sigma^2)$

- ▶ **Fisherovo F -rozdělení** $F_{k,m}$ [qf(1-alpha,k,m)]

$$F \sim F_{k,m} : P(F > F_{k,m}(\alpha)) = \alpha$$

- ▶ **rozdělení chí-kvadrát** χ_k^2 [qchisq(1-alpha,k)]

$$X^2 \sim \chi_k^2 : P(X^2 > \chi_k^2(\alpha)) = \alpha$$

- ▶ speciálně platí:

- ▶ $\chi_1^2(0,05) = 3,841 = 1,960^2$
- ▶ $\chi_1^2(\alpha) = z(\alpha/2)^2$
- ▶ $F_{1,m}(\alpha) = (t_m(\alpha))^2$

další rozdělení související s normálním

[F-distribution, chi-square distribution]

- ▶ V má rozdělení (musí být $P(V > 0) = 1$!!)

logaritmicko-normální, platí-li $\ln V \sim N(\mu, \sigma^2)$

- ▶ **Fisherovo F -rozdělení** $F_{k,m}$ [qf(1-alpha,k,m)]

$$F \sim F_{k,m} : P(F > F_{k,m}(\alpha)) = \alpha$$

- ▶ **rozdělení chí-kvadrát** χ_k^2 [qchisq(1-alpha,k)]

$$X^2 \sim \chi_k^2 : P(X^2 > \chi_k^2(\alpha)) = \alpha$$

- ▶ speciálně platí:

- ▶ $\chi_1^2(0,05) = 3,841 = 1,960^2$
- ▶ $\chi_1^2(\alpha) = z(\alpha/2)^2$
- ▶ $F_{1,m}(\alpha) = (t_m(\alpha))^2$

další rozdělení související s normálním

[F-distribution, chi-square distribution]

- ▶ V má rozdělení (musí být $P(V > 0) = 1$!!)

logaritmicko-normální, platí-li $\ln V \sim N(\mu, \sigma^2)$

- ▶ **Fisherovo F -rozdělení** $F_{k,m}$ [qf(1-alpha,k,m)]

$$F \sim F_{k,m} : P(F > F_{k,m}(\alpha)) = \alpha$$

- ▶ **rozdělení chí-kvadrát** χ_k^2 [qchisq(1-alpha,k)]

$$X^2 \sim \chi_k^2 : P(X^2 > \chi_k^2(\alpha)) = \alpha$$

- ▶ speciálně platí:

- ▶ $\chi_1^2(0,05) = 3,841 = 1,960^2$
- ▶ $\chi_1^2(\alpha) = z(\alpha/2)^2$
- ▶ $F_{1,m}(\alpha) = (t_m(\alpha))^2$

další rozdělení související s normálním

[F-distribution, chi-square distribution]

- ▶ V má rozdělení (musí být $P(V > 0) = 1$!!)

logaritmicko-normální, platí-li $\ln V \sim N(\mu, \sigma^2)$

- ▶ Fisherovo F -rozdělení $F_{k,m}$ [qf(1-alpha,k,m)]

$$F \sim F_{k,m} : P(F > F_{k,m}(\alpha)) = \alpha$$

- ▶ rozdělení chí-kvadrát χ_k^2 [qchisq(1-alpha,k)]

$$X^2 \sim \chi_k^2 : P(X^2 > \chi_k^2(\alpha)) = \alpha$$

- ▶ speciálně platí:

- ▶ $\chi_1^2(0,05) = 3,841 = 1,960^2$
- ▶ $\chi_1^2(\alpha) = z(\alpha/2)^2$
- ▶ $F_{1,m}(\alpha) = (t_m(\alpha))^2$