MV011 Statistika I

6. Transformace náhodné veličiny

Jan Koláček (kolacek@math.muni.cz)

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Brno

Motivační příklad

Příklad 1

Do výroby pracovních nástrojů vstupuje tyč délky $X\,cm$*. Ta je dále strojově opracována tak, že se její délka zdvojnásobí a je přišroubována k dalšímu dílu délky* $10\,cm$*. Jaká bude celková délka pracovního nástroje?*

Zápis: $X \sim f_X$, $\boxed{Y = 2X + 10}$ $\Rightarrow Y \sim ?$, víme $E(Y) = E(2X + 10) = 2E(X) + 10$
např. $X \sim N(50, 1)$, $Y = 2X + 10 \Rightarrow Y \sim ?$; $E(Y) = 2 \cdot 50 + 10 = 110$, $D(Y) = 2^2 \cdot D(X) = 4$

Další otázky:

Z ... délka druhého dílu, tj. $X \sim f_X$, $Z \sim f_Z$, $\boxed{Y = 2X + Z}$, $\Rightarrow Y \sim ?$
A ... koeficient prodloužení tyče, tj. $X \sim f_X$, $Z \sim f_Z$, $A \sim f_A$, $\boxed{Y = AX + Z}$
Obecně:

$$X \sim f_X,\ Y = h(X)\ \Rightarrow f_Y = ?$$

nebo

$$\mathbf{X} = (X_1, \dots, X_n)' \sim f_{\mathbf{X}},\ Y = h(\mathbf{X})\ \Rightarrow f_Y = ?$$

Obecná transformace

Věta 1

Nechť $X \sim f_X(x)$ *a transformace* $y = h(x)$ *je* **vzájemně jednoznačná** *(prostá a na), tj. když* **existuje derivace** $\frac{d}{dy}h^{-1}(y)$ *a* **je spojitá**. *Pak platí*

$$f_Y(y) = f_X\left(h^{-1}(y)\right)\left|\frac{dh^{-1}(y)}{dy}\right|. \tag{1}$$

Příklad 2 (Lognormální rozdělení, Lognormal distribution)

Náhodná veličina $X \sim N(0,1)$. *Vypočtěte hustotu náhodné veličiny* $Y = e^X$.

$X \sim N(0,1) \Rightarrow f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\ x \in \mathbb{R}$

$y = \underbrace{e^x}_{h(x)} \Rightarrow x = \underbrace{\ln(y)}_{h^{-1}(y)},\ y > 0$; derivace: $\frac{dh^{-1}(y)}{dy} = \frac{d\ln(y)}{dy} = \frac{1}{y}$

$f_Y(y) = f_X\left(h^{-1}(y)\right)\left|\frac{dh^{-1}(y)}{dy}\right| = \frac{1}{\sqrt{2\pi}}e^{-\frac{\ln^2 y}{2}} \cdot \frac{1}{y},\ y > 0$

Lineární transformace

Příklad 3 (Lineární transformace)

Nechť náhodná veličina X je absolutně spojitá s hustotou $f_X(x)$. Nalezněte hustotu transformované náhodné veličiny

$$Y = a + bX, \text{ kde } a, b \in \mathbb{R},\ b \neq 0.$$

Dva způsoby řešení:

- Dosazení do vzorce (1)

  K transformaci $y = a + bx$ existuje inverzní transformace $h^{-1}(y) = \frac{y-a}{b}$, která má derivaci $\frac{dh^{-1}(y)}{dy} = \frac{1}{b}$, takže

$$f_Y(y) = f_X\left(h^{-1}(y)\right)\left|\frac{dh^{-1}(y)}{dy}\right| = f_X\left(\frac{y-a}{b}\right)\frac{1}{|b|}$$

Lineární transformace

- Výpočet přes distribuční funkci

$$
\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= P(a + bX \le y) = \begin{cases} P\left(X \le \frac{y-a}{b}\right) = F_X\left(\frac{y-a}{b}\right) & \text{pro } b > 0 \\ P\left(X \ge \frac{y-a}{b}\right) = 1 - F_X\left(\frac{y-a}{b}\right) & \text{pro } b < 0 \end{cases}
\end{aligned}
$$

Hustotu pak dostaneme jako derivaci distribuční funkce

$$
\begin{aligned}
f_Y(y) = \frac{dF_Y(y)}{dy} &= \begin{cases} F'_X\left(\frac{y-a}{b}\right)\frac{1}{b} = \ \ f_X\left(\frac{y-a}{b}\right)\frac{1}{b} & \text{pro } b > 0 \\ -F'_X\left(\frac{y-a}{b}\right)\frac{1}{b} = -f_X\left(\frac{y-a}{b}\right)\frac{1}{b} & \text{pro } b < 0 \end{cases} \\
&= f_X\left(\frac{y-a}{b}\right)\frac{1}{|b|}
\end{aligned}
$$

Lineární transformace

Věta 2 (Lineární transformace normálního rozdělení)

Mějme náhodnou veličinu s normálním rozdělením $X \sim N(\mu, \sigma^2)$. Dále nechť $a, b \in \mathbb{R}$, $b \neq 0$ jsou reálné konstanty. Potom náhodná veličina, která je lineární transformací původní, má opět normální rozdělení, a to

$$Y = a + bX \ \sim \ N(a + b\mu, b^2\sigma^2).$$

Speciálně náhodná veličina

$$U = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

má **standardizované normální rozdělení**.

Příklad

Příklad 4 (Standardizace normálního rozdělení)

Při prodeji vánočních kaprů má hmotnost kapra v jedné z kádí přibližně normální rozdělení s parametry $\mu = 2,3$ a $\sigma^2 = 0,3^2$.

- **a)** *Jaký podíl kaprů přesáhne svou hmotností 2,6 kg?*
- **b)** *Jaký podíl kaprů má hmotnost mezi 2,1 kg a 2,6 kg?*
- **c)** *Jak volit hmotnostní hranici, aby podíl kaprů přesahujících tuto hranici byl 10 %?*

X . . . hmotnost kapra $\Rightarrow$ $\boxed{X \sim N(2,3; 0,3^2)}$

a) Jaký podíl kaprů přesáhne svou hmotností 2,6 kg?

$$P(X > 2,6) = 1 - P(X \leq 2,6) = 1 - P\left(\frac{X-\mu}{\sigma} \leq \frac{2,6-2,3}{0,3}\right)$$
$$= 1 - P(U \leq 1) = 1 - \Phi(1) = 1 - 0,84 = \boxed{0,16}$$

Příklad

b) Jaký podíl kaprů má hmotnost mezi 2,1 kg a 2,6 kg?

$$\begin{aligned} P(2,1 < X \leq 2,6) &= P\left(\frac{2,1-2,3}{0,3} < \frac{X-\mu}{\sigma} \leq \frac{2,6-2,3}{0,3}\right) \\ &= P\left(-\frac{2}{3} < U \leq 1\right) = \Phi(1) - \Phi(-2/3) \\ &= \Phi(1) - (1 - \Phi(2/3)) = 0,84 + 0,74 - 1 = \boxed{0,58} \end{aligned}$$

c) Jak volit hmotnostní hranici, aby podíl kaprů přesahujících tuto hranici byl 10 %?

$$0,1 = P(X > c) = 1 - P\left(\frac{X-\mu}{\sigma} \leq \frac{c-2,3}{0,3}\right) = 1 - \Phi\left(\frac{c-2,3}{0,3}\right)$$

$$\Phi\left(\frac{c-2,3}{0,3}\right) = 0,9$$

$$\frac{c-2,3}{0,3} = u_{0,9} = 1,28 \Rightarrow c = 0,3 \cdot 1,28 + 2,3 = \boxed{2,684}$$

Střední hodnota transformované n. v.

Věta 3

Nechť $h(x)$ *je borelovská funkce. Potom střední hodnota transformované náhodné veličiny* $Y = h(X)$ *existuje právě když existuje a je konečný integrál*

$\int\limits_{-\infty}^{\infty} h(x)dF(x) < \infty$. *V tomto případě platí* $\boxed{EY = Eh(X) = \int\limits_{-\infty}^{\infty} h(x)dF(x)}$.

- *Nechť* $X \sim (M, p)$ *je diskrétního typu, pak platí* $Y \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \Leftrightarrow \sum\limits_{x \in M} h(x)p(x)$ *absolutně konverguje. V tomto případě* $\boxed{EY = Eh(X) = \sum\limits_{x \in M} h(x)p(x)}$.
- *Nechť* $X \sim f(x)$ *je absolutně spojitého typu. Potom* EY *existuje právě když je funkce* $h(x)f(x)$ *integrovatelná vzhledem k Lebesgueově míře a přitom platí* $\boxed{EY = Eh(X) = \int\limits_{-\infty}^{\infty} h(x)f(x)dx}$, *tj.* $EY = Eh(X) \in \mathcal{L}_1(\Omega, \mathcal{A}, P) \Leftrightarrow h(x)f(x)$ *je integrovatelná vzhledem k Lebesgueově míře.*

Př. $Y = X^2 \Rightarrow EY = EX^2 = \int x^2 f(x)dx$ nebo $EY = EX^2 = \sum x^2 p(x)$

Příklad

Příklad 5

Náhodná veličina X má binomické rozdělení $X \sim Bi(n,\theta)$. Vypočtěte střední hodnotu náhodné veličiny $Y = e^{2X}$.

$$EY = E(e^{2X}) = \sum_{x=0}^{n} e^{2x} \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} \left(\theta e^2\right)^x (1-\theta)^{n-x}$$

$$\overset{\text{binom. věta}}{=} \left(\theta e^2 + 1 - \theta\right)^n$$

Momentová vytvořující funkce

Definice 4

Nechť X je náhodná veličina definovaná na $(\Omega, \mathcal{A}, P)$. Pak funkce $m : \mathbb{R} \to \mathbb{R}$ daná vztahem $m(t) = Ee^{tX}$, $t \in \mathbb{R}$, se nazývá **momentovou vytvořující funkcí náhodné veličiny** X (moment-generating function).

Definice 5

Nechť X je náhodná veličina definovaná na $(\Omega, \mathcal{A}, P)$. Pak funkce $\psi : \mathbb{R} \to \mathbb{C}$ daná vztahem $\psi(t) = Ee^{itX}$, $t \in \mathbb{R}$, se nazývá **charakteristickou funkcí náhodné veličiny** X (characteristic function). Tj. $\psi(t) = m(it)$.

Věta 6

Za předpokladu, že existují příslušné momenty náhodné veličiny X, tak existují i příslušné derivace momentové vytvořující funkce a platí

$$m^{(k)}(0) = EX^k.$$

Příklad

Příklad 6

Náhodná veličina X má binomické rozdělení $X \sim Bi(n,\theta)$. Vypočtěte střední hodnotu náhodné veličiny X pomocí momentové vytvořující funkce.

Z Příkladu 5 máme

$$m(t) = \left(\theta e^t + 1 - \theta\right)^n .$$

Podle předchozí věty je $EX = m'(0)$.
Derivujeme

$$m'(t) = n\left(\theta e^t + 1 - \theta\right)^{n-1} \theta e^t.$$

Takže

$$EX = m'(0) = n\left(\theta e^0 + 1 - \theta\right)^{n-1} \theta e^0 = n\theta.$$

Příklad 7

Náhodná veličina X má geometrické rozdělení $X \sim Ge(\theta)$. Vypočtěte střední hodnotu náhodné veličiny X.

Pravděpodobnostní funkce geometrického rozdělení je tvaru

$$p(x) = \begin{cases} (1-\theta)^x\theta & x = 0, 1, 2, \ldots, \quad \theta \in (0,1) \\ 0 & jinak \end{cases}$$

Z definice

$$EX = \sum_{x=0}^{\infty} xp(x) = \sum_{x=0}^{\infty} x(1-\theta)^x\theta = \ldots ? \text{ (viz tabule)}.$$

Momentová vytvořující funkce

$$m(t) = \sum_{x=0}^{\infty} e^{tx}(1-\theta)^x\theta = \theta \sum_{x=0}^{\infty} [e^t(1-\theta)]^x = \frac{\theta}{1-e^t(1-\theta)}.$$

Příklad

Derivujeme

$$m'(t) = \frac{\theta(1-\theta)e^t}{(1-e^t(1-\theta))^2}.$$

Takže

$$EX = m'(0) = \frac{1-\theta}{\theta}.$$

Transformace náhodného vektoru

Obecně: $Y = h(X_1, \dots, X_n) \Rightarrow Y \sim ?$
Konkrétně: $Y = X_1 + X_2 \Rightarrow Y \sim ?$

Věta 7

Jestliže náhodné veličiny spojitého typu $X_1 \sim f_{X_1}$ a $X_2 \sim f_{X_2}$ jsou **nezávislé**, *pak náhodná veličina $Y = X_1 + X_2$ má hustotu*

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(y - x_2) f_{X_2}(x_2) dx_2 = \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(y - x_1) dx_1$$

Hustotu $f_Y(y)$ potom nazýváme **konvolucí** *(***convolution***) hustot f_{X_1} a f_{X_2} a značíme $\boxed{f_Y(y) = f_{X_1} * f_{X_2}}$.*

Transformace normálního rozdělení

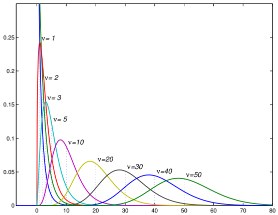
Definice 8 (χ^2 rozdělení, Chi-square distribution)

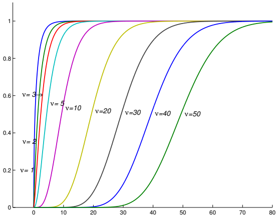
Řekneme, že náhodná veličina X má χ^2 **rozdělení** s $\nu > 0$ stupni volnosti, pokud její hustota má tvar

$$f_X(x) = \begin{cases} \frac{1}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{1}{2}x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

a budeme psát

$$\boxed{X \sim \chi^2(\nu)}.$$





Transformace normálního rozdělení

Věta 9 (Součet n nezávislých χ^2 veličin)

Nechť $U_1, \ldots, U_n$ jsou **nezávislé** *náhodné veličiny se standardizovaným normálním rozdělením, t.j.*

$$U_i \sim N(0,1) \qquad \textit{pro} \qquad i = 1, \ldots, n.$$

Pak náhodná veličina

$$K = \sum_{i=1}^{n} U_i^2 \sim \chi^2(n)$$

má χ^2 rozdělení o n stupních volnosti.

Transformace normálního rozdělení

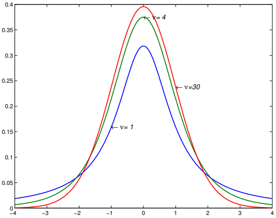
Definice 10 (Studentovo rozdělení, Student's distribution)

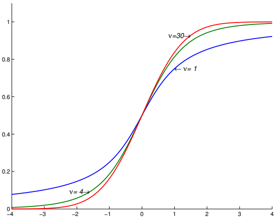
Řekneme, že náhodná veličina X má **Studentovo t rozdělení o $\nu > 0$ stupních volnosti**, pokud její hustota je tvaru

$$f_X(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} \nu^{-\frac{1}{2}} \left(\frac{x^2}{\nu} + 1\right)^{-\frac{\nu+1}{2}} \qquad \text{pro } x \in \mathbb{R}$$

Pak píšeme

$$\boxed{X \sim t(\nu)}.$$





Student

- **William Sealy Gosset** (13.6.1876 – 16.10.1937) vystudoval Winchester College a poté matematiku a chemii na New College v Oxfordu
- hlavní sládek v pivovaru Arthur Guinness & Son v Dublinu
- zkoumal možnosti, jak statisticky testovat kvalitu surovin – zejména ječmene a chmele
- 1906 – 1907 pracoval v laboratoři K. Pearsona
- vypracoval vlastní t-test pro malou velikost statistického souboru
- nesměl publikovat pod vlastním jménem, používal pseudonym **Student**

„Posílám Vám kopii Studentových tabulek, protože jste zřejmě jediný člověk, který je kdy použije.“

W. Gosset v dopise R. A. Fisherovi

Transformace normálního rozdělení

Věta 11 (Podíl standardizovaného normálního a χ^2)

Nechť náhodné veličiny $U \sim N(0,1)$ *a* $K \sim \chi^2(\nu)$ *jsou* **nezávislé***. Pak náhodná veličina*

$$T = \frac{U}{\sqrt{K/\nu}} \ \sim \ t(\nu)$$

má Studentovo t*-rozdělení o* ν *stupních volnosti.*

Transformace normálního rozdělení

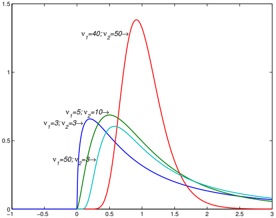
Definice 12 (Fisherovo–Snedecorovo F rozdělení)

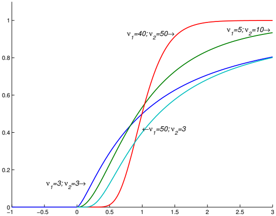
Řekneme, že náhodná veličina X má **Fisherovo–Snedecorovo F rozdělení o $\nu_1 > 0$ a $\nu_2 > 0$ stupních volnosti**, pokud její hustota je tvaru

$$f_X(x) = \begin{cases} \dfrac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\dfrac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} y^{\frac{\nu_1}{2}-1} \left(\dfrac{\nu_1}{\nu_2}y + 1\right)^{-\frac{\nu_1+\nu_2}{2}} & y \geq 0, \\ 0 & y < 0. \end{cases}$$

Pak píšeme

$$\boxed{X \sim F(\nu_1, \nu_2)}.$$





Transformace normálního rozdělení

Věta 13 (Podíl dvou nezávislých χ^2)

Nechť K_1 a K_2 jsou **nezávislé** *náhodné veličiny a*

$$K_i \sim \chi^2(\nu_i), \qquad i = 1, 2.$$

Pak náhodná veličina

$$F = \frac{K_1/\nu_1}{K_2/\nu_2} \sim F(\nu_1, \nu_2)$$

má Fisherovo–Snedecorovo F rozdělení o ν_1 a ν_2 stupních volnosti.

Motivační příklad

Příklad 8

Nechť náhodná veličina X_1 *značí výsledek hodu kostkou. Popište rozdělení této veličiny.*
Nechť X_2 *značí výsledek hodu druhou kostkou. Popište rozdělení veličiny* $X_1 + X_2$.
Dále popište rozdělení veličiny $X_1 + X_2 + X_3$.

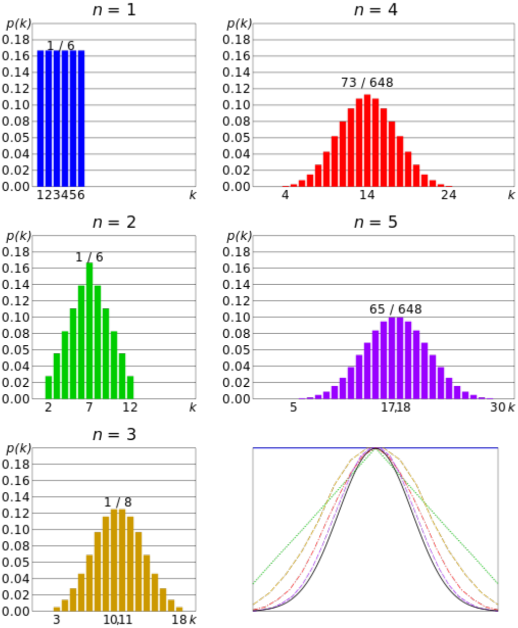
$$\vdots$$

X_1	1	2	3	4	5	6
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

X_1+X_2	2	3	4	5	6	7	8	9	10	11	12
$p(x)$	$\frac{1}{6^2}$	$\frac{2}{6^2}$	$\frac{3}{6^2}$	$\frac{4}{6^2}$	$\frac{5}{6^2}$	$\frac{6}{6^2}$	$\frac{5}{6^2}$	$\frac{4}{6^2}$	$\frac{3}{6^2}$	$\frac{2}{6^2}$	$\frac{1}{6^2}$

$X_1 + X_2 + X_3$:

3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$\frac{1}{6^3}$	$\frac{3}{6^3}$	$\frac{6}{6^3}$	$\frac{10}{6^3}$	$\frac{15}{6^3}$	$\frac{21}{6^3}$	$\frac{25}{6^3}$	$\frac{27}{6^3}$	$\frac{27}{6^3}$	$\frac{25}{6^3}$	$\frac{21}{6^3}$	$\frac{15}{6^3}$	$\frac{10}{6^3}$	$\frac{6}{6^3}$	$\frac{3}{6^3}$	$\frac{1}{6^3}$


n = 1
p(k)
0.18
0.16
0.14
0.12
0.10
0.08
0.05
0.04
0.02
0.00
1 / 6
123456
k
n = 4
73 / 648
4
14
24
n = 2
1 / 6
2
7
12
n = 5
65 / 648
5
17,18
30
n = 3
1 / 8
3
10,11
18

Centrální limitní věta

Značení

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Věta 14 (Lindebergova–Lévyho CLV, Central Limit Theorem)

Nechť $\{X_n\}_{n=1}^{\infty}$ *je posloupnost* **nezávislých** *náhodných veličin se stejným rozdělením se střední hodnotou* μ *a nenulovým rozptylem* σ^2*. Potom náhodné veličiny*

$$U_{\overline{X}_n} = \frac{(\overline{X}_n - \mu)\sqrt{n}}{\sigma}$$

mají asymptoticky **standardizované normální rozdělení** $N(0,1)$*, což budeme značit*

$$U_{\overline{X}_n} \overset{A}{\sim} N(0,1).$$

Centrální limitní věta

Příklad 9

Zatížení letadla s 64 místy nemá překročit $6\,000\ kg$*. Jaká je pravděpodobnost, že při plném obsazení bude tato hodnota překročena, má-li hmotnost cestujícího střední hodnotu* $90\ kg$ *a směrodatnou odchylku* $10\ kg$*?*

X_i ... hmotnost i-tého cestujícího, $E(X_i) = 90$, $D(X_i) = 100$, $i = 1, \dots, 64$

$Y = X_1 + \cdots + X_{64} = \sum_{i=1}^{64} X_i$, $\boxed{P(Y > 6\,000) = ?}$

CLV $\Rightarrow \sum_{i=1}^{n} X_i \approx N(nE(X_i), nD(X_i)) \Rightarrow \dfrac{\sum_{i=1}^{n} X_i - nE(X_i)}{\sqrt{nD(X_i)}} \approx N(0,1)$

Proto

$$P(Y > 6\,000) = 1 - P\left(\sum_{i=1}^{64} X_i \leq 6\,000\right) = 1 - P\left(\frac{Y - 64 \cdot 90}{\sqrt{64 \cdot 100}} \leq \frac{6\,000 - 64 \cdot 90}{\sqrt{64 \cdot 100}}\right)$$
$$= 1 - \Phi(3) = 1 - 0{,}9985 = 0{,}0015$$

Centrální limitní věta

Příklad 10

Předpokládejme, že žák má při písemce stejnou šanci dostat kteroukoli ze známek 1–5*. Jaká je pravděpodobnost, že průměr známek ve třídě se* 40 *žáky bude lepší než* 2,5*?*

X_i ... známka i-tého žáka, $E(X_i) = \frac{1}{5}(1+2+3+4+5) = 3$,
$D(X_i) = \frac{1}{5}(2^2+1+0+1+2^2) = 2$, $i = 1, \ldots, 40$
$Y = \frac{1}{40}\sum_{i=1}^{40} X_i$, $\boxed{P(Y < 2{,}5) = ?}$

CLV $\Rightarrow \sum_{i=1}^{n} X_i \approx N(nE(X_i), nD(X_i)) \Rightarrow \dfrac{\frac{1}{n}\sum_{i=1}^{n} X_i - E(X_i)}{\sqrt{\frac{D(X_i)}{n}}} \approx N(0,1)$

Proto

$$P(Y < 2{,}5) = P\left(\frac{Y-3}{\sqrt{\frac{2}{40}}} \le \frac{2{,}5-3}{\sqrt{\frac{2}{40}}}\right) = \Phi(-\sqrt{5})$$
$$= 1 - \Phi(\sqrt{5}) = 1 - 0{,}98713 = 0{,}013$$

Centrální limitní věta

Věta 15 (Integrální věta Moivre–Laplaceova)

Nechť náhodná veličina Y_n *udává* **počet úspěchů** *v posloupnosti délky* **n** *nezávislých alternativních pokusů s pravděpodobností úspěchu* θ. *Pak náhodné veličiny*

$$\frac{Y_n - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{A}{\sim} N(0,1).$$

Příklad

Příklad 11 (Anketa)

Při anketě rozdáme 160 dotazníků. Pravděpodobnost, že se nám vrátí dotazník vyplněný, je $0{,}7$*. Jaká je pravděpodobnost, že se nám vrátí alespoň* 100 *vyplněných dotazníků?*

X_i ... i-tý dotazník se vrátí vyplněný, $X_i \in \{0,1\}$, $X_i \sim A(0{,}7)$, $i = 1, \dots, 160$

Y ... počet vyplněných dotazníků ze 160 rozdaných, $Y \sim Bi(160; 0{,}7)$, $Y = \sum_{i=1}^{160} X_i$

$E(Y) = 160 \cdot 0{,}7 = 112$, $D(Y) = 160 \cdot 0{,}7 \cdot 0{,}3 = 33{,}6$

$$P(Y \geq 100) = \binom{160}{100} 0{,}7^{100} 0{,}3^{60} + \binom{160}{101} 0{,}7^{101} 0{,}3^{59} + \dots + \binom{160}{160} 0{,}7^{160} 0{,}3^{0} = ?$$

Oprava na spojitost a užití **CLV**:

$$P(Y \geq 100) = P(Y > 99) \doteq P(Y \geq 99{,}5) = 1 - P\left(\frac{Y - 112}{\sqrt{33{,}6}} \leq \frac{99{,}5 - 112}{\sqrt{33{,}6}}\right)$$

$$= 1 - \Phi\left(\frac{-12{,}5}{\sqrt{33{,}6}}\right) = \Phi(2{,}0702) = 0{,}98$$

Příklad

Příklad 12 (Anketa)

Kolik lístků musíme v předchozím příkladě rozdat, aby pravděpodobnost, že se jich vrátí minimálně 100 *vyplněných byla alespoň* $0,99$*?*

Y_n ... počet vyplněných dotazníků z n rozdaných, $Y_n \sim Bi(n;0,7)$
$E(Y_n) = n \cdot 0,7$, $D(Y_n) = n \cdot 0,7 \cdot 0,3$

$P(Y_n \geq 100) = P(Y_n > 99) \doteq P(Y_n \geq 99,5)$
Řešíme nerovnici

$$\begin{aligned} P(Y_n \geq 99,5) &\geq 0,99 \\ 1 - P\left(\frac{Y_n - n\cdot 0,7}{\sqrt{n\cdot 0,7\cdot 0,3}} \leq \frac{99,5 - n\cdot 0,7}{\sqrt{n\cdot 0,7\cdot 0,3}}\right) &\geq 0,99 \\ \Phi\left(\frac{99,5 - n\cdot 0,7}{\sqrt{n\cdot 0,7\cdot 0,3}}\right) &\leq 0,01 \end{aligned}$$

$\Rightarrow \frac{99,5-n\cdot 0,7}{\sqrt{n\cdot 0,7\cdot 0,3}} \leq u_{0,01} = -u_{0,99} = -2,326$
$\Leftrightarrow 0,7\cdot n - 2,326\cdot\sqrt{0,21}\sqrt{n} - 99,5 \geq 0 \Rightarrow \sqrt{n}_{1,2} = \frac{1,07\pm 16,72}{1,4} \Rightarrow n \geq 161,5$

Příklad

Příklad 13 (Oslava)

Kupujeme chlebíčky na oslavu, které se zúčastní 100 lidí. Označíme X_i počet snědených chlebíčků i-tého účastníka oslavy a ze zkušenosti víme, že $E(X_i) = 3$ a $D(X_i) = 3$. Kolik musíme koupit chlebíčků, aby s pravděpodobností $0,95$ nedošly?

X_i ... počet snědených chlebíčků i-tého účastníka oslavy, $E(X_i) = 3$, $D(X_i) = 3$, $i = 1, \dots, 100$

$Y = X_1 + \cdots + X_{100} = \sum_{i=1}^{100} X_i$ ... počet všech snědených chlebíčků

Hledáme c tak, aby $P(Y \leq c) = 0,95$.

$$0,95 = P(Y \leq c) = P\left(\frac{Y - 100 \cdot 3}{\sqrt{100 \cdot 3}} \leq \frac{c - 100 \cdot 3}{\sqrt{100 \cdot 3}}\right) = \Phi\left(\frac{c - 300}{\sqrt{300}}\right)$$

$$\Phi\left(\frac{c - 300}{\sqrt{300}}\right) = 0,95$$

$$\frac{c - 300}{\sqrt{300}} = u_{0,95} = 1,645 \Rightarrow c = 300 + 1,645 \cdot \sqrt{300} = 328,49.$$

Příklad

Příklad 14 (Pojišťovna)

Pojišťovna má $1\,000$ klientů stejné věkové skupiny. Pravděpodobnost úmrtí klienta této skupiny v daném roce je $0,01$. Každý klient zaplatí pojistné $1\,200$ Kč ročně. Jaká je pravděpodobnost, že pojišťovna nebude mít daném roce zisk, když v případě úmrtí klienta vyplatí jeho rodině $80\,000$ Kč?

X_i … i-tý klient zemře, $X_i \in \{0, 1\}$, $X_i \sim A(0,01)$, $i = 1, \dots, 1\,000$

Y … počet úmrtí v daném roce, $Y \sim Bi(1\,000; 0,01)$, $Y = \sum\limits_{i=1}^{1000} X_i$

$E(Y) = 1000 \cdot 0,01 = 10$, $D(Y) = 1000 \cdot 0,01 \cdot 0,99 = 9,9$

Oprava na spojitost a užití **CLV**:

$$\begin{aligned} P(80000 \cdot Y \geq 1000 \cdot 1200) &= P(Y \geq 15) \doteq P(Y > 14,5) \\ &= 1 - P\left(U \leq \frac{14,5 - 10}{\sqrt{9,9}}\right) \\ &= 1 - \Phi\left(\frac{4,5}{\sqrt{9,9}}\right) = \Phi(1,43) = 0,0763. \end{aligned}$$

Příklad

Příklad 15 (Síťový disk)

Počítačový správce poskytuje 100 uživatelům neomezené místo na síťovém disku. Označme X_i počet MB obsazených i-tým uživatelem. Z předchozích zkušeností víme, že $EX_i = 1\,200$ a $DX_i = 160\,000$. Jakou kapacitu musí mít síťový disk, aby byla překročena s pravděpodobností $0,01$?

X_i ... počet MB obsazených i-tým uživatelem, $EX_i = 1\,200$, $DX_i = 160\,000$.

$Y = X_1 + \cdots + X_{100} = \sum\limits_{i=1}^{100} X_i$ ... obsazené místo na disku

Hledáme $\boxed{c}$ tak, aby $\boxed{P(Y \leq c) = 0,99}$.

$$0,99 = P(Y \leq c) = P\left(\frac{Y - 100 \cdot 1200}{\sqrt{100 \cdot 160000}} \leq \frac{c - 100 \cdot 1\,200}{\sqrt{100 \cdot 160\,000}}\right) = \Phi\left(\frac{c - 120\,000}{4\,000}\right)$$

$$\Phi\left(\frac{c - 120\,000}{4\,000}\right) = 0,99$$

$$\frac{c - 120\,000}{4\,000} = u_{0,99} = 2,326 \Rightarrow \boxed{c} = 120\,000 + 2,326 \cdot 4\,000 = \boxed{129\,304\ MB}.$$