

Matematika III – 8. týden

Základní typy a vlastnosti náhodných veličin

Jan Slovák

Masarykova univerzita
Fakulta informatiky

5.11. 2013

Obsah přednášky

- 1 Literatura
- 2 Pravděpodobnostní funkce a hustoty
- 3 Náhodné vektory
- 4 Číselné charakteristiky veličin

Plán přednášky

- 1 Literatura
- 2 Pravděpodobnostní funkce a hustoty
- 3 Náhodné vektory
- 4 Číselné charakteristiky veličin

Kde je dobré číst?

- vlastní poznámky, texty současného přednášejícího, GOOGLE, atd.
- Karel Zvára, Josef Štěpán, Pravděpodobnost a matematická pravděpodobnost statistika, Matfyzpress, 2006, 230pp.
- J. Slovák, M. Panák, M. Bulant, Matematika drsně a svižně, Muni Press, Brno 2013, v+773 s., elektronická edice www.math.muni.cz/Matematika_drsne_svizne
- Marie Budíková, Štěpán Mikoláš, Pavel Osecký, Teorie pravděpodobnosti a matematická statistika (sbírka příkladů), Masarykova univerzita, 3. vydání, 2004, 117 stran, ISBN 80-210-3313-4.
- Marie Budíková, Tomáš Lerch, Štěpán Mikoláš, Základní statistické metody, Masarykova univerzita, 2005, 170 stran, ISBN 80-210-3886-1.
- Riley, K.F., Hobson, M.P., Bence, S.J. Mathematical Methods for Physics and Engineering, second edition, Cambridge University Press, Cambridge 2004, ISBN 0 521 89067 5, xviii

Plán přednášky

- 1 Literatura
- 2 Pravděpodobnostní funkce a hustoty
- 3 Náhodné vektory
- 4 Číselné charakteristiky veličin

Diskrétní náhodné veličiny

Jestliže náhodná veličina X na pravděpodobnostním prostoru (Ω, \mathcal{A}, P) nabývá jen konečně nebo spočetně mnoha hodnot $x_1, x_2, \dots \in \mathbb{R}$, pak existuje **pravděpodobnostní funkce** $f(x)$ taková, že

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \\ 0 & \text{jinak.} \end{cases}$$

Spojité náhodné veličiny

Hustota $f(x)$ **pravděpodobnosti** pro náhodnou veličinu X je funkce splňující pro $-\infty \leq a \leq b \leq \infty$

$$P(a < X < b) = \int_a^b f(x) dx. \quad (*)$$

Náhodná veličina X , pro kterou existuje její **hustota pravděpodobnosti** splňující (*), se nazývá **spojitá**.

Degenerované rozdělení $Dg(\mu)$.

Toto rozdělení odpovídá konstantní hodnotě $X = \mu$. Distribuční funkce F_X a pravděpodobnostní funkce f_X jsou tedy rovny

$$F_X(t) = \begin{cases} 0 & t \leq \mu \\ 1 & t > \mu \end{cases} \quad f_X(t) = \begin{cases} 1 & t = \mu \\ 0 & \text{jinak} \end{cases}.$$

Degenerované rozdělení $Dg(\mu)$.

Toto rozdělení odpovídá konstantní hodnotě $X = \mu$. Distribuční funkce F_X a pravděpodobnostní funkce f_X jsou tedy rovny

$$F_X(t) = \begin{cases} 0 & t \leq \mu \\ 1 & t > \mu \end{cases} \quad f_X(t) = \begin{cases} 1 & t = \mu \\ 0 & \text{jinak} \end{cases}.$$

Alternativní rozdělení $A(p)$ popisuje pokus s pouze dvěma možnými výsledky, kterým budeme říkat zdar a nezdar. Náhodné veličině X pro určitost přiřadíme hodnotu 0 pro nezdar a 1 pro zdar. Pokud má zdar pravděpodobnost p , pak nezdar musí mít pravděpodobnost $1 - p$. Jsou tedy distribuční a pravděpodobnostní funkce tvaru:

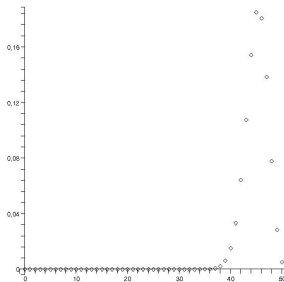
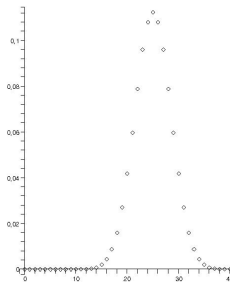
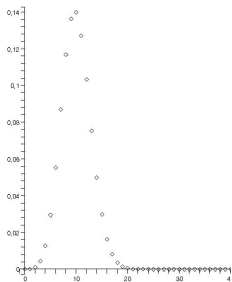
$$F_X(t) = \begin{cases} 0 & t \leq 0 \\ 1 - p & 0 < t \leq 1 \\ 1 & t > 1 \end{cases} \quad f_X(t) = \begin{cases} p & t = 1 \\ 1 - p & t = 0 \\ 0 & \text{jinak} \end{cases}.$$

Binomické rozdělení $Bi(n, p)$

odpovídá n -krát nezávisle opakovanému pokusu popsanému alternativním rozdělením, přičemž naše náhodná veličina měří počet zdarů. Je tedy zjevné, že pravděpodobnostní funkce bude mít nenulové hodnoty právě v celých číslech $0, \dots, n$ odpovídajícím celkovému počtu úspěchů v pokusech (a nezáleží nám na pořadí). Je tedy

$$f_X(t) = \begin{cases} \binom{n}{t} p^t (1-p)^{1-t} & t \in \{0, 1, \dots, n\} \\ 0 & \text{jinak} \end{cases}.$$

Na obrázku jsou pravděpodobnostní funkce pro $Bi(50, 0.2)$, $Bi(50, 0.5)$ a $Bi(50, 0.9)$. Rozdělení pravděpodobnosti dobře odpovídá intuici, že nejvíce výsledků bude blízko u hodnoty np :



S binomickým rozdělením se setkáváme velice často v praktických úlohách. Jednou z nich je popis náhodné veličiny, která popisuje počet X předmětů v jedné zvolené přihrádce z n možných, do nichž jsme náhodně rozdělili r předmětů.

Umístění kteréhokoliv předmětu do pevně zvolené přihrádky má pravděpodobnost $1/n$ (každá z nich je stejně pravděpodobná).

Zjevně tedy bude pro jakýkoliv počet $k = 0, \dots, r$

$$P(X = k) = \binom{r}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{r-k} = \binom{r}{k} \frac{(n-1)^{r-k}}{n^r},$$

jde proto o rozložení X typu $\text{Bi}(r, 1/n)$.

Jestliže nám bude vzrůstat počet přihrádek n společně s počtem předmětů r_n tak, že v průměru nám na každou přihrádku bude připadat (přibližně) stejný počet prvků λ , můžeme dobře vyjádřit chování našeho rozdělení veličin X_n při limitním přechodu $n \rightarrow \infty$:

Jestliže nám bude vzrůstat počet přihrádek n společně s počtem předmětů r_n tak, že v průměru nám na každou přihrádku bude připadat (přibližně) stejný počet prvků λ , můžeme dobře vyjádřit chování našeho rozdělení veličin X_n při limitním přechodu $n \rightarrow \infty$:

$$\begin{aligned} \lim_{n \rightarrow \infty} P(X_n = k) &= \lim_{n \rightarrow \infty} \binom{r_n}{k} \frac{(n-1)^{r_n-k}}{n^{r_n}} \\ &= \lim_{n \rightarrow \infty} \frac{r_n(r_n-1)\dots(r_n-k+1)}{(n-1)^k} \frac{1}{k!} \left(1 - \frac{1}{n}\right)^{r_n} \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 + \frac{-\frac{r_n}{n}}{r_n}\right)^{r_n} \\ &= \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

protože obecně funkce $(1 + x/n)^n$ konvergují stejnoměrně k funkci e^x na každém omezeném intervalu v \mathbb{R} .

To dává **Poissonovo rozdělení** $Po(\lambda)$.

Poissonovo rozdělení $Po(\lambda)$ popisuje např. události, které se vyskytují náhodně v čase a přitom pravděpodobnost výskytu v následujícím časovém intervalu o jednotkové délce nezávisí na předchozí historii a je rovna stále stejné hodnotě λ .

V praxi jsou takové procesy spojeny např. s poruchovostí strojů a zařízení.

Poissonovo rozdělení $Po(\lambda)$ popisuje např. události, které se vyskytují náhodně v čase a přitom pravděpodobnost výskytu v následujícím časovém intervalu o jednotkové délce nezávisí na předchozí historii a je rovna stále stejné hodnotě λ .

V praxi jsou takové procesy spojeny např. s poruchovostí strojů a zařízení.

Theorem (Poissonova věta)

Jsou-li $X_n \sim \text{Bi}(n, p_n)$ a $\lim_{n \rightarrow \infty} n \cdot p_n = \lambda$ je konečná, pak

$$\lim_{n \rightarrow \infty} P(X_n = k) = P(X = k),$$

kde $X \sim \text{Po}(\lambda)$.

Příklady spojitých rozdělání

Nejjednodušší je tzv. **rovnoměrné rozdělání**. Jestliže chceme, aby pravděpodobnost každé hodnoty v předem daném intervalu $(a, b) \subset \mathbb{R}$ byla stejná, pak hustota f_X našeho rozdělání náhodné veličiny X má být konstantní. Pak ovšem jsou pro libovolná reálná čísla $-\infty < a < b < \infty$ jen jediné možné hodnoty

$$f_X(t) = \begin{cases} 0 & t \leq a \\ \frac{1}{b-a} & t \in (a, b) \\ 0 & t \geq b, \end{cases} \quad F_X(t) = \begin{cases} 0 & t \leq a \\ \frac{t-a}{b-a} & t \in (a, b) \\ 1 & t \geq b. \end{cases}$$

Exponenciální rozdělení $\text{ex}(\lambda)$

je dalším rozdělením, které je snadno určeno požadovanými vlastnostmi náhodné veličiny. Předpokládejme, že sledujeme výskyt náhodného jevu tak, že výskyty v nepřekrývajících se intervalech jsou nezávislé. Je-li tedy $P(t)$ pravděpodobnost, že jev **nenastane** během intervalu délky t , pak nutně $P(t + s) = P(t)P(s)$ pro všechna $t, s > 0$. Předpokládejme navíc diferencovatelnost funkce P a $P(0) = 1$. Pak $\ln P(t + s) = \ln P(t) + \ln P(s)$.

Exponenciální rozdělení $\text{ex}(\lambda)$

je dalším rozdělením, které je snadno určeno požadovanými vlastnostmi náhodné veličiny. Předpokládejme, že sledujeme výskyt náhodného jevu tak, že výskyty v nepřekrývajících se intervalech jsou nezávislé. Je-li tedy $P(t)$ pravděpodobnost, že jev **nenastane** během intervalu délky t , pak nutně $P(t+s) = P(t)P(s)$ pro všechna $t, s > 0$. Předpokládejme navíc diferencovatelnost funkce P a $P(0) = 1$. Pak $\ln P(t+s) = \ln P(t) + \ln P(s)$. Limitním přechodem:

$$\lim_{s \rightarrow 0^+} \frac{\ln P(t+s) - \ln P(t)}{s} = (\ln P)'(0) = -\lambda.$$

Odtud vyplývá diferenciální rovnice

$$(\ln P(t))' = -\lambda.$$

Odtud dostáváme $\ln P(t) = -\lambda t + C$ a počáteční podmínka dává jediné řešení

$$P(t) = e^{-\lambda t}.$$

Všimněme si, že z definice našich objektů vyplývá, že $\lambda > 0$.

Odtud dostáváme $\ln P(t) = -\lambda t + C$ a počáteční podmínka dává jediné řešení

$$P(t) = e^{-\lambda t}.$$

Všimněme si, že z definice našich objektů vyplývá, že $\lambda > 0$. Uvažme náhodnou veličinu X udávající okamžik, kdy náš jev poprvé **nastane**. Zřejmě tedy je distribuční funkce rozdělení pro X dána

$$F_X(t) = 1 - P(t) = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \leq 0. \end{cases}$$

Je vidět, že je to rostoucí funkce s hodnotami mezi nulou a jedničkou a správnými limitami v $\pm\infty$.

Hustotu tohoto rozdělení dostaneme derivováním distribuční funkce, tj.

$$f_X = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & t \leq 0. \end{cases}$$

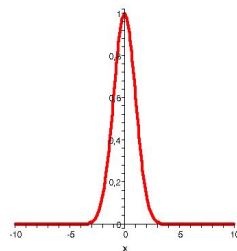
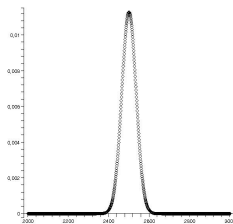
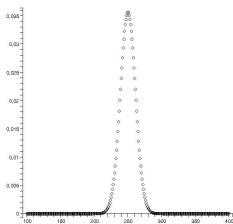
Normální rozdělení

je ze všech nejdůležitější.

Normální rozdělení

je ze všech nejdůležitější.

Jestliže v binomiálním rozdělení zachováme konstantní úspěšnost p , ale budeme přidávat počet pokusů n , bude pravděpodobnostní funkce kupodivu pořád mít podobný tvar (i když jiné rozměry). Na obrázku při rostoucím n se budou vynesené bodové hodnoty slévat do křivky, pro hodnoty $Bi(500, 0.5)$ a $Bi(5000, 0.5)$ je výsledek vidět na obrázku níže. Třetí křivka na obrázku je grafem funkce $f(x) = e^{-x^2/2}$.



Hledáme-li podobné spojité rozdělení, potřebovali bychom spočítat $\int_a^b e^{-x^2/2} dx$ což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Hledáme-li podobné spojité rozdělení, potřebovali bychom spočítat $\int_a^b e^{-x^2/2} dx$ což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Rozdělení s touto hustotou se nazývá **normální rozdělení** $N(0, 1)$. Příslušnou distribuční funkci

$$F_X(x) = \int_{-\infty}^x e^{-x^2/2} dx$$

nelze vyjádřit pomocí elementárních funkcí, přesto se s ní numericky běžně počítá (pomocí tabulek nebo softwarových aplikací).

Hledáme-li podobné spojité rozdělení, potřebovali bychom spočítat $\int_a^b e^{-x^2/2} dx$ což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Rozdělení s touto hustotou se nazývá **normální rozdělení** $N(0, 1)$. Příslušnou distribuční funkci

$$F_X(x) = \int_{-\infty}^x e^{-x^2/2} dx$$

nelze vyjádřit pomocí elementárních funkcí, přesto se s ní numericky běžně počítá (pomocí tabulek nebo softwarových aplikací). Hustotě f_X se také často říká **Gaussova křivka**.

Plán přednášky

- 1 Literatura
- 2 Pravděpodobnostní funkce a hustoty
- 3 Náhodné vektory**
- 4 Číselné charakteristiky veličin

Obdobně definujeme distribuční funkce a hustotu a pravděpodobnostní funkci pro spojité a diskrétní náhodné vektory. Hovoříme také o **simultánních (sdružených) pravděpodobnostních funkcích a hustotách**.

Pro dvě proměnné (vektor (X, Y) náhodných veličin):

$$f(x, y) = \begin{cases} P(X = x_i \wedge Y = y_j) & x = x_i \wedge y = y_j \\ 0 & \text{jinak.} \end{cases}$$

u diskrétních a pro všechny $a, b \in \mathbb{R}$ pro spojité:

$$F(b, a) = P(-\infty < X < b, \infty < Y < a) = \int_{-\infty}^a \int_{-\infty}^b f(x, y) dx dy.$$

Marginální rozložení

pro jednu z proměnných obdržíme tak, že přes ostatní počítáme nebo zintegrujeme. Např. u diskrétních vektorových veličin (X, Y) tvoří jevy $(X = x_i, Y = y_j)$ pro všechny možné hodnoty x_i a y_j s nenulovými pravděpodobnostmi pro X a Y úplný systém jevů pro vektor (X, Y) a dostáváme vztah:

$$P(X = x_i) = \sum_{j=1}^{\infty} P(X = x_i, Y = y_j)$$

mezi **marginálním rozdělením pravděpodobnosti** náhodné veličiny X a **sduženým rozdělením pravděpodobnosti** náhodného vektoru (X, Y) .

Náhodné veličiny X a Y jsou **stochasticky nezávislé**, jestliže jejich sdružená distribuční funkce splňuje

$$F(x, y) = G(x) \cdot H(y),$$

kde F a G jsou distribuční funkce veličin X a Y .

Plán přednášky

- 1 Literatura
- 2 Pravděpodobnostní funkce a hustoty
- 3 Náhodné vektory
- 4 Číselné charakteristiky veličin

Uvažme náhodnou veličinu Y s rozdělením pravděpodobnosti $P(Y = k) = 1/n$, $k = 1, 2, \dots, n$, tj. pro $n = 6$ popisujeme hod férovou kostkou. Průměr $\frac{1}{2}(n + 1)$ bude patrně dobře vystihovat „očekávanou průměrnou hodnotu“.

Uvažme náhodnou veličinu Y s rozdělením pravděpodobnosti $P(Y = k) = 1/n$, $k = 1, 2, \dots, n$, tj. pro $n = 6$ popisujeme hod férovou kostkou. Průměr $\frac{1}{2}(n + 1)$ bude patrně bude dobře vystihovat „očekávanou průměrnou hodnotu“.

Při binomickém rozdělení $\text{Bi}(n, p)$ můžeme stejně dobře použít „váženého průměru“, který vyjádří „průměrné očekávání“:

$$\begin{aligned} & \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^n \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} (1-p)^{n-k} \\ &= np \sum_{j=0}^{n-1} \frac{(n-1)!}{(n-j-1)!j!} p^j (1-p)^{n-j-1} = np, \end{aligned}$$

(protože součet pravděpodobností v každém binomickém rozdělení je jedna).

Střední hodnota

Nechť X je náhodná veličina s diskrétním rozdělením. Jestliže řada $\sum_{k=1}^{\infty} x_i P(X = x_i)$ konverguje absolutně (zejména tedy pro všechny X s konečně mnoha možnými hodnotami x_i), pak její součet $E X$ nazýváme **střední hodnotou** X .

Střední hodnota

Nechť X je náhodná veličina s diskrétním rozdělením. Jestliže řada $\sum_{k=1}^{\infty} x_i P(X = x_i)$ konverguje absolutně (zejména tedy pro všechny X s konečně mnoha možnými hodnotami x_i), pak její součet EX nazýváme **střední hodnotou** X .

Je-li X náhodná veličina se spojitým rozdělením s hustotou $f(x)$ a nevlastní integrál $\int_{-\infty}^{\infty} xf(x)dx$ konverguje absolutně, pak jeho hodnota EX se nazývá **střední hodnota** X .

Je tedy $EX = np$, je-li $X \sim \text{Bi}(n, p)$, zatímco pro rovnoměrné rozdělení na intervalu (a, b) dostaneme dle očekávání

$$EX = \int_a^b \frac{1}{b-a} dx = \frac{1}{2} \frac{b^2 - a^2}{b-a} = \frac{1}{2}(a+b).$$

Střední hodnota

Nechť X je náhodná veličina s diskrétním rozdělením. Jestliže řada $\sum_{k=1}^{\infty} x_i P(X = x_i)$ konverguje absolutně (zejména tedy pro všechny X s konečně mnoha možnými hodnotami x_i), pak její součet EX nazýváme **střední hodnotou** X .

Je-li X náhodná veličina se spojitým rozdělením s hustotou $f(x)$ a nevlastní integrál $\int_{-\infty}^{\infty} xf(x)dx$ konverguje absolutně, pak jeho hodnota EX se nazývá **střední hodnota** X .

Je tedy $EX = np$, je-li $X \sim \text{Bi}(n, p)$, zatímco pro rovnoměrné rozdělení na intervalu (a, b) dostaneme dle očekávání

$$EX = \int_a^b \frac{1}{b-a} dx = \frac{1}{2} \frac{b^2 - a^2}{b-a} = \frac{1}{2}(a+b).$$

Střední hodnotou náhodného vektoru (X_1, \dots, X_n) rozumíme vektor středních hodnot (EX_1, \dots, EX_n) .

Vlastnosti střední hodnoty

Theorem (Afinní chování střední hodnoty)

Uvažme náhodné veličiny X, Y , skaláry $a, b \in \mathbb{R}$, náhodný vektor $W = (X_1, \dots, X_n)$ a čtvercovou skalární matici B s n řádky.

- Pro konstantní náhodnou veličinu $X = a \in \mathbb{R}$ je $E a = a$.
- $E(a + bX) = a + bE X$.
- $E(X + Y) = E X + E Y$.
- $E(a + BX) = a + B(E X)$.

kvantilová funkce

Je-li $F(x)$ distribuční funkce náhodné veličiny X , pak

$$F^{-1}(u) = \inf\{x \in \mathbb{R}; F(x) \geq u\}, \quad 0 < u < 1$$

je kvantilová funkce náhodné veličiny X .

Hodnota $F^{-1}(\alpha)$ se nazývá α -kvantil.

Tzv. kritické hodnoty pro veličinu X jsou pak $F^{-1}(1 - \alpha)$.

Rozptyl

Další charakteristika popisuje, jak moc se dá čekat, že se hodnoty náhodné veličiny „hemží“ kolem nějaké hodnoty.

Definition

Nechť X je náhodná veličina s konečnou střední hodnotou. Pak definujeme **rozptyl** veličiny X výrazem

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

pokud taková konečná hodnota existuje.

Rozptyl

Další charakteristika popisuje, jak moc se dá čekat, že se hodnoty náhodné veličiny „hemží“ kolem nějaké hodnoty.

Definition

Nechť X je náhodná veličina s konečnou střední hodnotou. Pak definujeme **rozptyl** veličiny X výrazem

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

pokud taková konečná hodnota existuje.

Odmocnina z rozptylu $\sqrt{\text{var } X}$ se nazývá **směrodatná odchylka** náhodné veličiny X .

Jde o zjevnou obdobu definice kvadrátu vzdálenosti vektorů nebo funkcí. Zachycujeme tak „očekávanou vzdálenost“ hodnot X od její střední hodnoty.

Theorem

Jestliže má náhodná veličina X konečný rozptyl, pro libovolné skaláry $a, b \in \mathbb{R}$ platí

- $\text{var } X = E X^2 - (E X)^2$
- $\text{var}(a + bX) = b^2 \text{var } X$
- $\sqrt{\text{var}(a + bX)} = |b| \sqrt{\text{var } X}$.

Theorem

Jestliže má náhodná veličina X konečný rozptyl, pro libovolné skaláry $a, b \in \mathbb{R}$ platí

- $\text{var } X = E X^2 - (E X)^2$
- $\text{var}(a + bX) = b^2 \text{var } X$
- $\sqrt{\text{var}(a + bX)} = |b| \sqrt{\text{var } X}$.

Občas přiřazujeme k X **normovanou** veličinu Z ,

$$Z = \frac{X - E X}{\sqrt{\text{var } X}},$$

kteřá má zjevně nulovou střední hodnotu a jednotkový rozptyl.

Theorem

Má-li X rozptyl a $\epsilon > 0$ je libovolné, pak platí

$$P(|X - EX| \geq \epsilon) \leq \frac{\text{var } X}{\epsilon^2}.$$