

Cvičenie 6 – Špeciálne spojité rozdelenia
– Centrálna limitná veta

Rovnomerné (uniform) $R(a,b)$, $Ro(a,b)$, $U(a,b)$

$$f(x) = \begin{cases} 1/(b-a) & \text{na intervale } (a,b) \\ 0 & \text{inde .} \end{cases}$$

$$F(x) = \begin{cases} (x-a)/(b-a) & \text{na intervale } (a,b) \\ 0 & \text{na } (-\infty, a] \\ 1 & \text{na } [b, \infty) \end{cases}$$

Exponenciálne $\text{Exp}(\lambda)$

$$f(x) = \begin{cases} \lambda * e^{-\lambda x} & \text{na } (0, \infty) \\ 0 & \text{na } (-\infty, 0] \end{cases}$$

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{na } (0, \infty) \\ 0 & \text{na } (-\infty, 0] \end{cases}$$

Stredná hodnota je $1/\lambda$.

Všimnime si , že F rastie od 0 po 1 na nekonečnom intervale, teda nie je tu žiaden bod „b“, kde by sa dosiahla 1-ka a ďalej by už F pokračovala konštantne na tejto hodnote.

https://en.wikipedia.org/wiki/Exponential_distribution

Normálne¹ (Gaussovo) rozdelenie

$N(0,1)$, $N(\mu, \sigma^2)$

Parametre Gaussovho rozdelenia sú stredná hodnota a rozptyl.

Základný tvar pre strednú hodnotu 0 a rozptyl 1 sa nazýva **normované normálne rozdelenie**

$N(0,1)$

$$f_n(x) = (2\pi)^{(-1/2)} * \exp(-x^2/2) \quad \text{na } (-\infty, \infty)$$

$$F_n(x) = \text{normcdf}(x) \quad \text{na } (-\infty, \infty)$$

Pre všeobecné parametre

$N(\mu, \sigma^2)$

$$F(x) = \text{normcdf}(x, \mu, \sigma) = \frac{1}{2} \left[1 + \text{erf}\left(\frac{x - \mu}{\sigma\sqrt{2}}\right) \right] \quad \text{na } (-\infty, \infty)$$

Platí (!!!) nasledujúci „konverzný“ vzťah:

$$F(x) = F_n((x - \mu)/\sigma)$$

Poznámky:

1. Distribučná funkcia F je integrálom z f , avšak analytický zápis F si vyžaduje použitie špeciálnych funkcií nad rámec „klasického zoznamu“ elementárnych funkcií. Na jednoduchších kalkulačkách preto hodnoty F vyčíslit nedokážeme. Dnes našťastie trochu lepší softvér už má distribučnú funkciu pre normálne rozdelenie priamo zabudovanú, môžeme rovno dosadzovať.

→ V prostrediach Matlab, Octave je to funkcia `normcdf(xu, mu, sigma)`, prípadne sa dá zostaviť s použitím funkcie `erf` (error function).

Rovnako dnes už je k dispozícii množstvo online kalkulačiek, stačí pogúliť.

2. V prípade, ak tieto moderné možnosti nemáte k dispozícii (napr. na skúške), je riešením použitie tabuliek. Tie sa dajú nájsť na webe a aj na našich stránkach by niečo malo visieť. V nich ale nájdete iba hodnoty pre normované $N(0,1)$, takže k hodnotám $N(\mu, \sigma^2)$ sa treba dopracovať pomocou vyššie uvedeného konverzného vzorca.

https://en.wikipedia.org/wiki/Normal_distribution

¹ Môžete sa občas stretnúť s pojmom „normálové rozdelenie“ (žiaľ, aj v niektorých učebniciach). Tento názov je nesprávny a nezmyselný (doslova znamená *kolmicové* rozdelenie) – kto ho použije, neprejde cez skúšku z tohto predmetu.

Centrálna limitná veta

https://en.wikipedia.org/wiki/Central_limit_theorem

Riešené príklady

Príklad 1:

Spojité náhodná veličina má rovnomerné rozdelenie $R(4,9)$.

a) Nájdite funkciu hustoty a distribučnú funkciu, nakreslite ich grafy.

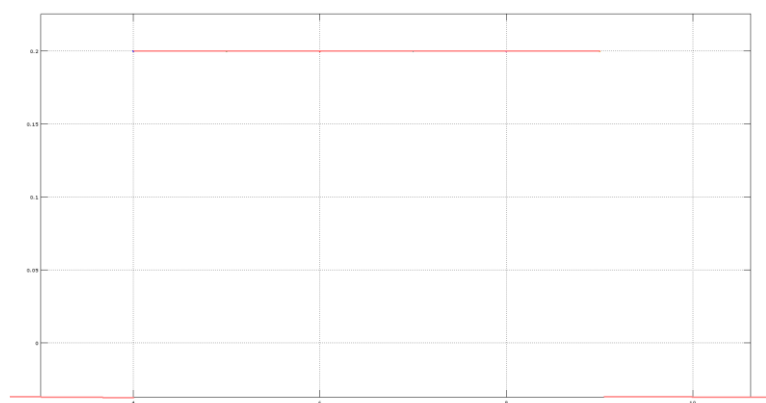
Vypočítajte strednú hodnotu a rozptyl.

b) Vypočítajte pravdepodobnosť $P(5.7 < x < 7.2)$ a $P(7.7 < x < 10.1)$

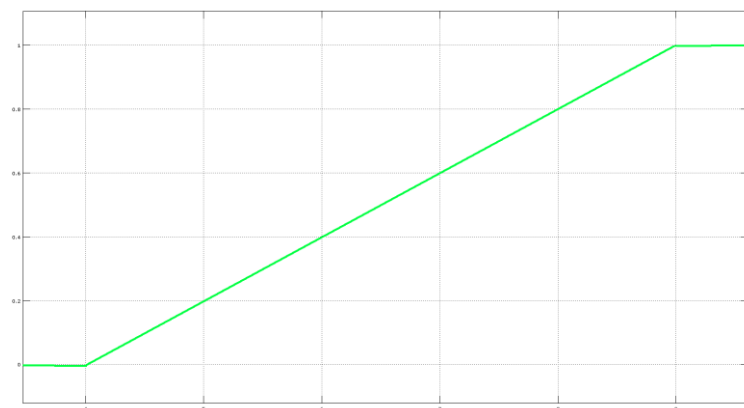
Riešenie:

a) Náhodná veličina je podľa zadania nenulová na intervale $(4,9)$.

$$f(x) = 1/(9-4) = 1/5 \quad \text{na } (4,9)$$



$$F(x) = (x-4)/5 \quad \text{na } (4,9)$$



Strednú hodnotu a rozptyl vyrátame podľa dostupných vzorcov pre rovnomerné rozdelenie.

$$E(X) = (4+9)/2 = 6.5,$$

$$D(X) = (9-4)^2 / 12 = 2.0833333$$

$$b) \quad P(5.7 < x < 7.2) = F(7.2) - F(5.7) = (7.2 - 5.7) / 5 = 0.3$$

$$P(7.7 < x < 10.1) = F(10.1) - F(7.7) = 1 - (7.7-4)/5 = 0.26$$

Príklad 2:

Stredná životnosť vysávača modelu XYZ je 1000 hodín. Samotná životnosť ako náhodná veličina má exponenciálne rozdelenie.

- a) Aká je pravdepodobnosť, že vysávač vydrží najviac 800, 1000, 1200 hodín?
- b) Aká je pravdepodobnosť, že vysávač vydrží aspoň 800, 1000, 1200 hodín?
- c) Aká je pravdepodobnosť, že vysávač vydrží najviac 800 až 1200 hodín?
- d) Pravdepodobnosť, že vydrží aspoň x hodín, je 0.5 (0.95, 0.99). Zistite hodnotu x .

Riešenie:

Životnosť vysávača je veličina s rozdelením $\text{Exp}(1/1000)$. Vieme totiž, že $EX) = \mu = 1000$, a platí $\mu = 1/\lambda$, odkiaľ $\lambda = 1/\mu$. Odpovede na prvé tri otázky v zadaní získame jednoduchým dosadením do vzorca

$$F(x) = 1 - e^{-x/1000}$$

- a) $P(\text{najviac } 800) = F(800) - F(0) = 0.55067 - 0 = 0.55067$
 $P(\text{najviac } 1000) = F(1000) - F(0) = 0.63212 - 0 = 0.63212$
 $P(\text{najviac } 1200) = F(1200) - F(0) = 0.63212 - 0 = 0.69881$

Všimnime si, že odpoveď pri „najviac tisíc“ prekvapivo (?) nie je 0.5. Je to preto, lebo stredná hodnota a medián sú dve rôzne veci.

- b) $P(\text{aspoň } x) = F(\infty) - F(x) = 1 - (1 - e^{-x/1000}) = e^{-x/1000}$

$$\begin{aligned} P(\text{aspoň } 800) &= 0.44933 \\ P(\text{aspoň } 1000) &= 0.36788 \\ P(\text{aspoň } 1200) &= 0.30119 \end{aligned}$$

V akom vzťahu sú výsledky častí a), b) ?

- c) $P(800 \text{ až } 1200) = F(1200) - F(800) = 0.14813$

- d) Hoci to nie je povedané rovno, ide tak trochu o hľadanie kvantilov. Keď raz zaznie na prednáškach tento pojem, už budete vedieť, o čom to je.

Hodnota $F(x)$ hovorí, aká je pravdepodobnosť, že vysávač vydrží najviac x hodín.

$$F(x) = 1 - e^{-\lambda x}$$

My však chceme mať v tom výroku aspoň x hodín. Musíme teda vychádzať z takéhoto vzorca:

$$y = 1 - F(x) = e^{-\lambda x}$$

Teraz vyjadríme y (pozor, výsledok nie je klasická kvantilová funkcia):

$$\begin{aligned} e^{-\lambda x} &= y \\ -\lambda x &= \ln(y) \\ x &= -\ln(y) / \lambda \end{aligned}$$

Je teda 0.5 (50-percentná) pravdepodobnosť, že vysávač vydrží aspoň
 $-\ln(0.5) / (1/1000) = \ln(2) * 1000 = 693.15$ hodín.

Je 0.95 (95-percentná) pravdepodobnosť, že vysávač vydrží aspoň $-\ln(0.95) * 1000 = 51.293$ hodín.

Je 0.99 (99-percentná) pravdepodobnosť, že vysávač vydrží aspoň $-\ln(0.99) * 1000 = 10.05$ hodín.

Príklad 3:

Životnosť počítača (model XY) v hodinách je daná rozdelením $\text{Exp}(1/5000)$.

a) Aká je pravdepodobnosť, že z 10 kúpených počítačov aspoň 8 sa dožije 2 rokov, ak priemerne pracujú 5 hodín denne (aj cez víkendy a sviatky, iba 29.2. sú vypnuté celý deň).

b) Aká je pravdepodobnosť, že z 10 počítačov sa aspoň 1 dožije 7 rokov?

Riešenie:

Životnosť počítača je veličina s rozdelením $\text{Exp}(1/5000)$.

Platí teda

$$\begin{aligned} F(x) &= 1 - e^{-x/5000} && (\text{dožije sa najviac } x \text{ hodín}) \\ 1 - F(x) &= e^{-x/5000} && (\text{dožije sa aspoň } x \text{ hodín}) \end{aligned}$$

Rok má 365 dní, ročne teda počítač pracuje 1825 hodín.

a) Počítač jednotliviec sa dožije 2 rokov (tj. 2 a viac) s pravdepodobnosťou

$$p = 1 - F(2 \cdot 1825) = e^{-3650/5000} = 0.48191$$

Ak 2 rokov sa dožije aspoň 8 počítačov, to je 8, 9 alebo 10.

Pravdepodobnosť toho je

$$C(10,10) \cdot p^{10} + C(10,9) \cdot p^9 \cdot (1-p) + C(10,8) \cdot p^8 \cdot (1-p)^2 = 0.043074$$

b) Počítač jednotliviec sa dožije 7 rokov (tj. 7 a viac) s pravdepodobnosťou

$$p = 1 - F(7 \cdot 1825) = 0.077692$$

Pre nás však je zaujímavá pravdepodobnosť, že sa nedožije 7 rokov, a to je

$$1-p = 0.92231$$

Z 10 sa dožije 7 rokov aspoň jeden znamená, že sa nestane, aby sa nedožil ani jeden.

$$1-(1-p)^{10} = 0.55458$$

Príklad 4:

Pre dané Gaussovo rozdelenie zistite pravdepodobnosť :

- a) $P(1 < x < 2)$ pre $N(0, 1)$
- b) $P(-1 < x < 0.5)$ pre $N(0, 1)$
- c) $P(1 < x < 2)$ pre $N(1, 2.25)$
- d) $P(1 < x < 2)$ pre $N(4, 2.25)$

Riešenie:

Ak máte k dispozícii dobrý softvér, výsledky sú len otázkou dosadenia. Tu si ukážeme postup pomocou tabuliek.

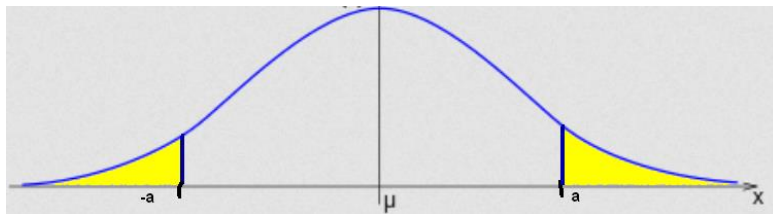
$$a) P(1 < x < 2) = F_n(2) - F_n(1) = 0.9772 - 0.8413 = 0.1359 .$$

$$b) P(-1 < x < 0.5) = F_n(0.5) - F_n(-1) = \dots$$

Zlá správa: V tabuľkách nemáme k dispozícii hodnoty pre záporné x .

Dobrá správa: Nepotrebujeme ich. Platí totiž $F(-a) = 1 - F(a)$. Dôvod vidno z obrázku –

$F(-a)$ je žltá plocha vľavo, $1 - F(a)$ je žltá plocha vpravo. (Majú byť rovnaké, obrázok je jemne nepresný)



$$\dots = F_n(0.5) - (1 - F_n(1)) = 0.6915 - (1 - 0.8413) = 0.5328 .$$

c) Tu musíme použiť „konverzný vzorec“.

$$\begin{aligned} P(1 < x < 2) &= F(2) - F(1) = F_n((2-1) / 1.5) - F_n((1-1) / 1.5) = F_n(2/3) - F_n(0) = \\ &= 0.7454 - 0.5 = 0.2454 \end{aligned}$$

d) Tu musíme použiť „konverzný vzorec“ a zároveň fintu z bodu b).

$$\begin{aligned} P(1 < x < 2) &= F(2) - F(1) = F_n((2-4) / 1.5) - F_n((1-4) / 1.5) = F_n(-4/3) - F_n(-2) = \\ &= F_n(2) - F_n(4/3) = 0.9772 - 0.9088 = 0.0684 \end{aligned}$$

Príklad 5:

Výška mužov (v cm) na Slovensku má (zhruba) normálne rozdelenie $N(179, 64)$.

- a) Aká je pravdepodobnosť, že náhodne vybraný muž bude mať výšku medzi 155-203?
- b) Aká je pravdepodobnosť, že z náhodne vybraných 5 mužov budú mať aspoň dvaja aspoň 185 cm?

Riešenie:

$$\begin{aligned} \text{a) } P(155 < x < 203) &= F(203) - F(155) = F_n((203-179)/8) - F_n((155-179)/8) = \\ &= F_n(3) - F_n(-3) = 2 \cdot 0.9987 - 1 = 0.9974 \end{aligned}$$

- b) Pravdepodobnosť, že aspoň jeden bude mať aspoň 185:

$$p = 1 - F(185) = 1 - F_n((185-179)/8) = 1 - F_n(0.75) = 1 - 0.7734 = 0.2266$$

Aspoň dvaja z 5 znamená 2,3,4, alebo 5. Otočme to – ide tu o doplnok k situácii, keď najviac jeden (tj. 0 alebo 1) má 185+.

$$P_d = P(0 \text{ alebo } 1 \text{ sú } 185+) = (1-p)^5 + 5 \cdot (1-p)^4 \cdot p = 0.68202$$

A teraz späť, zaujíma nás pravdepodobnosť $1 - P_d = 0.31798$.

Príklad 6:

Pokračujeme v zadaní príkladu 5.

- a) Aká je pravdepodobnosť, že súčet výšok náhodne vybraných 5 mužov bude aspoň 900 cm?
- b) Aká je pravdepodobnosť, že priemer výšok náhodne vybraných 4 mužov bude najviac 178 cm?

Riešenie:

Na riešenie týchto úloh použijeme teóriu z prednášok (z kapitoly o centrálnej limitnej vete). Zároveň pri riešení si už uľahčíme prácu a použijeme na výpočet softvér.

- a) Ak výška náhodného jednotlivca v populácii je náhodná veličina s $N(179, 64)$, potom súčet výšok piatich náhodne vybraných šťastlivcov má rozdelenie $N(179 \cdot 5, 64 \cdot 5)$.

Odpoveď na otázku je potom $P = 1 - F(900) = 1 - \text{normcdf}(900, 179 \cdot 5, 64 \cdot 5) = 0.38993$

- b) Priemer výšok 4 náhodne vylosovaných výhercov má rozdelenie $N(179, 64/4)$.

Odpoveď na otázku je $P = F(178) = \text{normcdf}(178, 179, 64/4) = 0.40129$

Neriešené príklady

1. Spojitá náhodná veličina má rovnomerné rozdelenie $X \sim R(7,12)$.
 - a) Nájdite charakteristiky X , funkciu hustoty a distribučnú funkciu.
 - b) Vypočítajte pravdepodobnosti $P(9 < x < 10.5)$, $P(6 < x < 13.5)$

2. Bezporuchovosť auta je veličina s rozdelením $\text{Exp}(10^{-4})$. Meria sa v hodinách, predpokladáme 2 hodiny jazdy denne.
 - a) Aká je pravdepodobnosť, že auto vydrží bez poruchy aspoň 5 rokov? (0.69420)
 - b) Aká je pravdepodobnosť, že sa prvá porucha objaví počas 7. roku? (0.045430)
 - c) Aká je pravdepodobnosť, že auto nevydrží bez poruchy viac ako 10 rokov? (0.51809)

3. Pravdepodobnosť, že študent (nie na FEI) príde na cvičenie s meškaním x minút, je veličina s rozdelením $\text{Exp}(0.1)$. V krúžku je 10 študentov.
 - a) Aká je pravdepodobnosť, že do 10 minút po začiatku budú prítomní všetci? (0.010186)
 - b) Aká je pravdepodobnosť, že aspoň jeden z tých desiatich sa nedostaví vôbec (=meškanie 50 minút a viac) (0.065373)
 - c)* Pravdepodobnosť, že v čase t bude prítomná aspoň polovica študentov, je 99%. Vypočítajte t (s presnosťou v celých minútach). (16)

4. Normálne rozdelenie:
 - a) Zistite $P(2 < x < 7)$ pre $N(-1, 4)$. (0.0668)
 - b) Zistite $P(2 < x < 7)$ pre $N(3, 9)$. (0.53750)
 - c) Zistite $P(2 < x < 7)$ pre $N(10, 6.25)$. (0.11438)

5. Úspešnosť študenta na skúške (počet bodov zo 100) má Gaussovo rozdelenie $N(60, 25)$
 - a) Aká je pravdepodobnosť, že študent spraví skúšku (aspoň 56 bodov)? (0.7881)
 - b) Aká je pravdepodobnosť, že z vybraných 5 študentov prejdú všetci? (0.30402)
 - c) Aká je pravdepodobnosť, že vybraných 5 študentov dosiahne spolu aspoň 320 bodov? (0.0368)
 - d) Aká je pravdepodobnosť, že 50 študentov dosiahne v priemere aspoň 61 bodov? (0.0787)