

NÁHODNÉ VELIČINY V SIMULAČNÍCH METODÁCH

Milan Guštar

Abstract

This paper describes basic types of random variables, forms of their representation and the approximation of real-life random variables.

1. Náhodné veličiny

Přiřazením číselných hodnot elementárním jevům nebo výsledkům realizací pokusů dostáváme náhodné veličiny. Náhodné veličiny lze definovat jako funkce ξ na množině elementárních jevů Ω . Čísla $\xi(\omega)$ pro $\omega \in \Omega$ jsou pak realizacemi náhodné veličiny ξ . Náhodné veličiny mohou být **diskrétní** nebo **spojité**. Příkladem diskrétní náhodné veličiny jsou čísla získaná házením hrací kostkou. Ta mohou nabývat pouze hodnot z množiny $\{1,2,3,4,5,6\}$.

$$f(x) = \begin{cases} \frac{1}{6} & \text{pro } x \in \{1,2,3,4,5,6\} \\ 0 & \text{jinde} \end{cases}$$

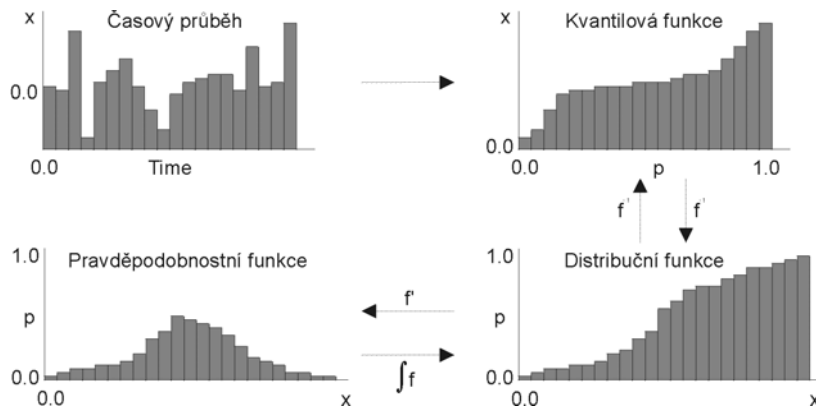
Příkladem spojitého rozdělení je rozdělení rovnoměrné. Veličina s rovnoměrným rozdělením může nabývat libovolné hodnoty z intervalu (a,b) , přičemž všechny možné výsledky jsou rovnocenné:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{pro } a < x < b \\ 0 & \text{jinde} \end{cases}$$

Některé náhodné veličiny mohou být kombinací veličin diskrétních a spojitých.

2. Reprezentace náhodných veličin

Pravděpodobnostní chování náhodných veličin lze popsat mnoha způsoby. Jedním z nejobvyklejších je popis **pravděpodobnostní funkcí** nebo **funkcí hustoty pravděpodobnosti**, jejíž tvar podává obraz o důležitých vlastnostech rozdělení. Další obvyklou formou je **distribuční funkce**. Oba způsoby popisu charakterizují rozdělení náhodných veličin úplně, tedy např. pokud mají dvě veličiny stejné distribuční funkce, mají i stejná rozdělení a naopak. Další ekvivalentní formou popisu je např. kvantilová funkce nebo charakteristická funkce (momentová vytvořující funkce). Vztahy mezi některými formami reprezentace jsou znázorněny na obrázku 1. Ve všech uvedených případech je k popisu náhodných veličin obvykle použito nějaké funkce vyjádřené v uzavřeném tvaru a jejích parametrů. Proto je tento způsob popisu možné označit za **parametrický**. V případě **neparametrického** popisu je namísto parametrů rozdělení definováno např. polohami a rozsahy tříd a četnostmi v jednotlivých třídách v případě po částech rovnoměrného nebo obecného diskrétního rozdělení apod. Takto vyjádřená rozdělení jsou označována za neparametrická [2].

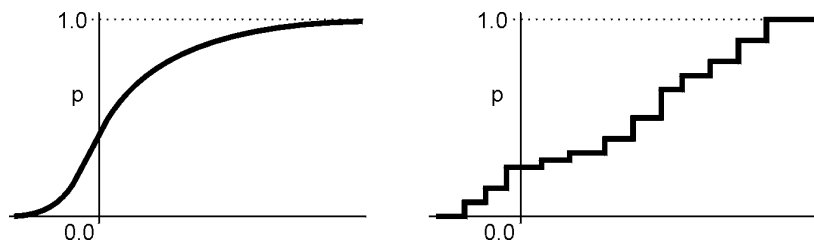


Obr.1: Reprezentace náhodných veličin

Distribuční funkce $F: \mathbb{R} \rightarrow [0;1]$ je pravděpodobnostní charakteristika náhodné veličiny, definována vztahem:

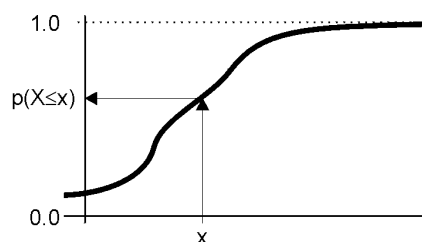
$$F(x) = P[X \leq x] \text{ nebo } F(x) = P[X < x]$$

V případě spojitého rozdělení je distribuční funkce absolutně spojitá, v případě diskrétního rozdělení má schodovitý průběh, viz obrázek 2.



Obr.2: Distribuční funkce spojitého a diskrétního rozdělení

Distribuční funkce udává pravděpodobnost s jakou náhodná veličina X nepřekročí zadanou hodnotu x , viz obrázek 3. V této souvislosti se lze setkat s termínem **pravděpodobnost nepřekročení**.



Obr.3: Distribuční funkce a pravděpodobnost nepřekročení.

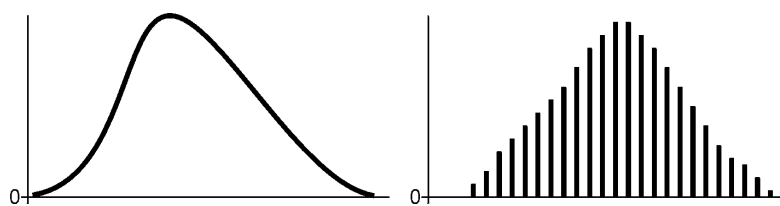
Spojitou náhodnou veličinu X lze charakterizovat **hustotou pravděpodobnosti** $f(x)$, definovanou:

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P[x \leq X \leq x + \Delta x]}{\Delta x}$$

Pro diskrétní náhodné veličiny hustotě pravděpodobnosti odpovídá **pravděpodobnostní funkce** nebo **frekvenční funkce**, definovaná:

$$f(x) = \sum_i p_i \delta(x - x_i) \quad p_i = P[X = x_i]$$

kde $\delta(x)$ je impulsní funkce, mající pro $x = 0$ hodnotu 1 a pro ostatní x hodnotu 0. Diskrétní náhodnou veličinu lze tedy definovat výčtem hodnot x_i , kterých může nabývat a pravděpodobnostmi p_i s jakými jednotlivé hodnoty nabývá, viz obrázek 4.



Obr.4: Hustota pravděpodobnosti a pravděpodobnostní funkce

Pro všechna $x_1 < x_2$ platí:

$$P[x_1 \leq X \leq x_2] = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx$$

To znamená, že pravděpodobnost, že spojitá náhodná veličina X nabývá hodnot z intervalu $x_1 \leq X < x_2$ je rovna velikosti plochy pod křivkou funkce hustoty pravděpodobnosti $f(x)$ mezi $x = x_1$ a $x = x_2$, viz obrázek 5. V případě diskretních náhodných veličin udává hodnota pravděpodobnostní funkce $f(x)$ pravděpodobnost s jakou veličina nabývá hodnoty x .

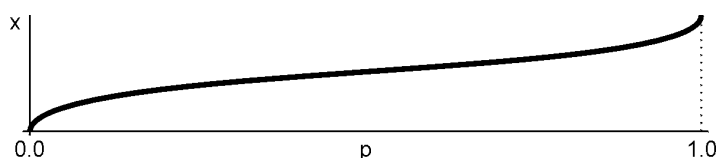
Obr.5: Hustota pravděpodobnosti a pravděpodobnost $P[x_1 \leq X < x_2]$

Inverzní distribuční funkce, zvaná také **kvantilová funkce** $Q: [0;1] \rightarrow \square \mathbf{R}$, definovaná vztahem

$$Q(p) = F_x^{-1}(x)$$

je další formou reprezentace náhodných veličin. Pomocí kvantilové funkce lze pro danou pravděpodobnost p najít odpovídající hodnotu p -kvantilu x_p náhodné veličiny X , tj. platí

$$Q(p) = x_p \text{ a } F(x_p) = p.$$

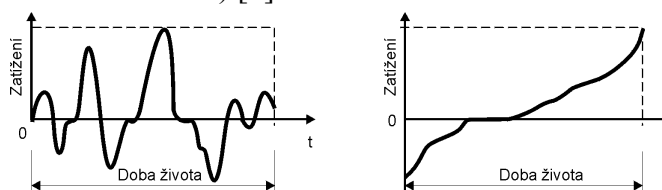


Obr.6: Kvantilová funkce

Inverzní funkci lze vytvořit pouze ke spojitě, rostoucí distribuční funkci, tedy pouze pro absolutně spojitá rozdělení. V případě diskretních a smíšených rozdělení inverzní funkce F^{-1} neexistuje. Pro tyto případy lze definici kvantilové funkce zobecnit:

$$Q(p) = \inf \{x | F(x) \geq p\}$$

V metodě SBRA jsou kvantilové funkce pro zatížení označovány jako **křivky trvání zatížení** (Load Duration Curve – LDC) [3].

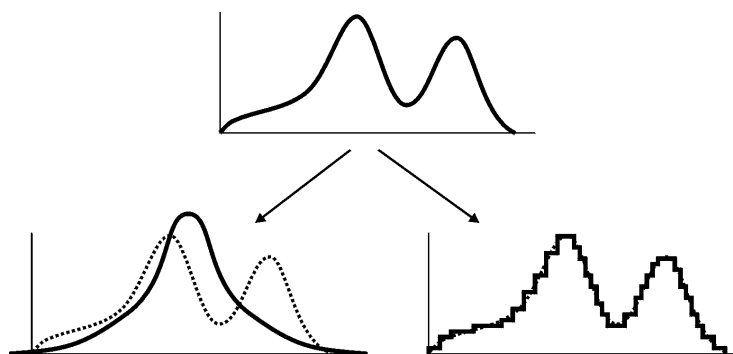


Obr.7: Historie a křivka trvání zatížení

Křivka trvání vznikne z časového průběhu veličiny setříděním hodnot podle velikosti, (viz obrázek 7) a reprezentuje trvání jednotlivých úrovní zatížení. Křivka trvání je jako kvantilová funkce ekvivalentem distribuční funkce nebo funkce hustoty pravděpodobnosti a plně charakterizuje stochastické vlastnosti veličiny.

3. Aproximace reálných náhodných veličin

Rozdělení některých náhodných veličin jsou získávána na základě teoretických znalostí o těchto veličinách. Pro většinu fyzikálních veličin nejsou typy a parametry rozdělení dostatečně přesně známy a proto jsou k jejich určení používána data získaná na základě měření, testů a zkoušek. Tříděním těchto dat je vytvářen histogram (sloupcový graf četností), reprezentující empirickou pravděpodobnostní funkci nebo funkci hustoty pravděpodobnosti. Toto empirické rozdělení bývá aproximováno vhodným rozdělením teoretickým [2], jehož typ bývá zvolen buď na základě znalostí teoretických principů vedoucích k vytvoření posloupnosti nebo na základě vlastností vzorku dat. V situacích, kdy není k dispozici dostatečně rozsáhlý vzorek, je třeba typ rozdělení odhadnout. Při volbě typu rozdělení hraje roli i jeho předpokládané budoucí použití. Proto bývá často použito normálního rozdělení. Jeho vlastnosti jsou dobře známy, mnoho výpočtů lze provést analyticky, mnohdy lze předpokládat možnost aplikace centrální limitní věty a tedy i očekávat přibližnou normalitu hledaného rozdělení. Rozdělení, která mají funkci hustoty pravděpodobnosti nepravidelného tvaru a pro jejichž aproximaci nelze najít vhodná teoretická rozdělení, lze dobře aproximovat některým z neparametrických typů rozdělení [1], viz obrázek 8.



Obr.8: Aproximace rozdělení

Jelikož u reálných náhodných veličin lze předpokládat, že nabývají pouze hodnot z omezeného intervalu $[a,b]$, lze pro jejich aproximaci použít omezená (useknutá) rozdělení. To je výhodné při implementaci generátoru náhodných čísel při použití těchto rozdělení v rámci simulačních metod.

Oznámení

Příspěvek byl vypracován v rámci výzkumu podporovaného GAČR (projekt č. 103/01/1410).

Literatura

- [1] HURT, J., *Simulační metody*. SPN Praha, 1982.
- [2] SILVERMAN, B. W., *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, 1986.
- [3] MAREK, P., GUŠTAR, M., ANAGNOS, T., *Simulation-Based Reliability Assessment for Structural Engineers*, CRC Press, Inc., Boca Raton, Florida, 1995.