1 Náhodné veličiny

1.1 Základní pojmy

Při studiu reálných dějů se neomezujeme pouze na jejich sledování a popis. Bývá zapotřebí stanovit zákonitosti dějů a využít je k určení jejich dalšího průběhu. K tomu se často používají matematické modely, mezi něž patří také modely pravděpodobnostní a statistické.

Není-li výsledek nějakého pokusu nebo děje jednoznačně určen podmínkami, za nichž se odehrává, můžeme různé možné výsledky považovat za *elementární jevy*. Označujeme je symbolem ω s případnými indexy. Množinu všech elementárních jevů označujeme Ω a nazýváme ji *prostor elementárních jevů*. Většinou se nezajímáme o jednotlivé elementární jevy, ale o určité jejich množiny. Z matematického hlediska se ukázalo výhodné zajímat se o takový systém množin elementárních jevů, který tvoří σ-algebru. Nechť je tedy na prostoru Ω dána nějaká σ-algebra $\mathcal{A}$ jeho podmnožin. Tyto podmnožiny se nazývají *náhodné jevy*. Jednotlivým množinám patřícím do $\mathcal{A}$ (a jenom jim) se pak připisuje pravděpodobnost pomocí nějaké pravděpodobnostní míry P. Trojice $(\Omega, \mathcal{A}, \mathsf{P})$ se nazývá *pravděpodobnostní prostor*.

V matematické statistice se stává, že na $\mathcal{A}$ je dán celý systém pravděpodobnostních měr $\{\mathsf{P}_\alpha\}$. Tím je dán celý systém modelů náhodného děje. Na základě analýzy reálných dat je pak třeba rozhodnout, které z těchto modelů se dají použít a které je třeba zamítnout.

Nechť $\mathbb{R}$ je reálná přímka a $\mathcal{B}$ systém jejích borelovských množin. Nechť $X(\omega)$ je měřitelná funkce z $(\Omega, \mathcal{A}, \mathsf{P})$ do $(\mathbb{R}, \mathcal{B})$. Pak se $X(\omega)$ nazývá *náhodná veličina* a značí se stručně X. Každé borelovské množině $B \in \mathcal{B}$ lze přiřadit její vzor $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$ a pravděpodobnostní míru $Q(B) = \mathsf{P}\{X^{-1}(B)\}$. Míra Q se nazývá *indukovaná míra* (je indukovaná funkcí X) nebo také *zákon rozdělení náhodné veličiny* X. Většinou se říká stručně jen *rozdělení*. Ve statistice se často píše $\mathcal{L}(X)$ místo Q. Pak je vidět, které veličině rozdělení přísluší. To je výhodné tam, kde se vyšetřuje několik náhodných veličin. Písmeno $\mathcal{L}$ vzniklo z anglického slova law (zákon).

Položíme-li specielně $B = (-\infty, x)$, dostaneme

$$F(x) = \mathsf{P}\{X < x\}.$$

Funkce F se nazývá *distribuční funkce*. Z teorie míry je známo, že mezi Q a F je vzájemně jednoznačný vztah. V teorii pravděpodobnosti se dokazuje, že každá distribuční funkce je neklesající, zleva spojitá a platí pro ni $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$. Distribuční funkce může mít nejvýše spočetně mnoho bodů nespojitosti. V matematické statistice mají nejdůležitější uplatnění následující dva typy distribučních funkcí:

a) Distribuční funkce F je funkce skoků. V tomto případě říkáme, že jde o *diskrétní rozdělení*. Jak jsme připomněli, bodů nespojitosti může být nejvýš spočetně mnoho, a tak je můžeme označit $x_1, x_2, \ldots$. Nechť p_k je velikost skoku funkce F v bodě x_k. Lze zjistit, že platí $p_k = \mathsf{P}\{X = x_k\}$.

b) Existuje taková funkce $f(x)$, že platí

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t. \tag{1.1}$$

Pak jde o *spojité rozdělení*. Jak známo, F má vyjádření (1.1) tehdy a jen tehdy, je-li absolutně spojitá. Protože F je neklesající, musí platit $f(x) \geq 0$ skoro všude. Je-li třeba, upraví se $f(x)$ tak, aby byla všude nezáporná. Proto dále pod pojmem hustota budeme rozumět jen nějakou její nezápornou verzi.

Mnohé vlastnosti rozdělení a jeho charakteristik plynou z věty o přenosu integrace, kterou nyní připomeneme.

Věta 1.1 (o přenosu integrace) *Budiž t měřitelné zobrazení z $(\Omega, \mathcal{A}, \mathsf{P})$ do nějakého měřitelného prostoru $(\Lambda, \mathcal{D})$. Budiž g měřitelná funkce na $(\Lambda, \mathcal{D})$. Nechť Q je míra na $\mathcal{D}$ indukovaná zobrazením t, tj. $Q(D) = \mathsf{P}\{t^{-1}(D)\}$ pro $D \in \mathcal{D}$. Potom platí*

$$\int_{\Omega} g[t(\omega)]\,\mathrm{d}\mathsf{P}(\omega) = \int_{\Lambda} g(t)\,\mathrm{d}Q(t), \tag{1.2}$$

jakmile jeden z integrálů existuje.

Důkaz. Viz Halmos (1950), věta 3 v § 39. Věta má však natolik fundamentální význam, že hlavní kroky důkazu uvedeme.

1) Zvolme nejprve $g = I_B$, kde I_B je charakteristická funkce množiny B. Nechť $B \in \mathcal{D}$. Připomeňme, že $I_B[t(\omega)] = 1$ pro $t(\omega) \in B$, tj. pro $\omega \in t^{-1}(B)$. Proto levá strana vzorce (1.2) je

$$\int_{\Omega} I_B[t(\omega)]\,\mathrm{d}\mathsf{P}(\omega) = \int_{t^{-1}(B)} \mathrm{d}\mathsf{P}(\omega) = \mathsf{P}[t^{-1}(B)],$$

zatímco pravá strana vzorce (1.2) je

$$\int_{\Lambda} I_B(t)\,\mathrm{d}Q(t) = \int_{B} \mathrm{d}Q(t) = Q(B) = \mathsf{P}[t^{-1}(B)].$$

Tím je pro měřitelnou charakteristickou funkci množiny vzorec (1.2) dokázán.

2) Je-li g jednoduchá měřitelná funkce, tvrzení plyne z toho, že integrál je lineární funkcionál.

3) Nechť g je nezáporná měřitelná funkce. Pak existuje posloupnost g_k nezáporných jednoduchých měřitelných funkcí taková, že $g_k(t) \uparrow g(t)$ pro každé $t \in \Lambda$. Dále se užije Leviho věta o monotónní konvergenci.

4) Je-li g reálná měřitelná funkce, užije se rozklad $g = g^+ - g^-$.

5) Je-li g komplexní měřitelná funkce, užije se její rozklad na reálnou a imaginární část. $\square$

Ve speciálním případě, kdy $(\Lambda, \mathcal{D}) = (\mathbb{R}, \mathcal{B})$ a zobrazení t je dáno náhodnou veličinou X, dostává vzorec (1.2) tvar

$$\int_\Omega g[X(\omega)]\,\mathrm{dP}(\omega) = \int_{\mathbb{R}} g(t)\,\mathrm{d}Q(t), \tag{1.3}$$

kde míra Q odpovídá distribuční funkci F. Platí-li (1.1), máme

$$\int_{\mathbb{R}} g(t)\,\mathrm{d}Q(t) = \int_{\mathbb{R}} g(t)\,\mathrm{d}F(t) = \int_{\mathbb{R}} g(t)f(t)\,\mathrm{d}t. \tag{1.4}$$

Existuje-li hustota, pak kombinací vzorců (1.3) a (1.4) dostaneme, že platí

$$\mathsf{P}\{X \in B\} = \int_B f(x)\,\mathrm{d}x, \qquad B \in \mathcal{B}.$$

Je-li F funkce skoků, která má v bodech $x_1, x_2, \ldots$ skoky o velikostech $p_1, p_2, \ldots$, pak platí

$$\mathsf{P}\{X \in B\} = \sum_{\{i:x_i \in B\}} p_i.$$

Ve statistice se často používají integrály náhodných veličin. Používá se pro ně symbol E (expectation). Říkáme, že $\mathsf{E}X = \int_\Omega X(\omega)\,\mathrm{dP}(\omega)$ je *střední hodnota* náhodné veličiny X, pokud uvedený integrál existuje. Protože obvykle bývá známa jenom distribuční funkce, většinou se provádí konkrétní výpočet podle vzorce $\mathsf{E}X = \int_{\mathbb{R}} x\,\mathrm{d}F(x)$. Je-li a nějaká konstanta, pak $\mathsf{E}a = a$. Poznamenejme, že se velmi často vyskytuje situace, kdy integrujeme přes celý prostor. Pro stručnost se tento prostor vynechává, a tak v dalším např. místo $\int_\Omega X(\omega)\,\mathrm{dP}(\omega)$ budeme psát jenom $\int X(\omega)\,\mathrm{dP}(\omega)$.

Často se setkáváme s nezápornými náhodnými veličinami. Ty popisují třebas dobu čekání, dobu životnosti apod. Při výpočtu jejich střední hodnoty bývá užitečná následující věta.

Věta 1.2 *Nechť X je nezáporná náhodná veličina s distribuční funkcí F s konečnou střední hodnotou. Pak $\mathsf{E}X = \int_0^\infty [1 - F(x)]\,\mathrm{d}x$.*

Důkaz. Nechť $T > 0$ je bod spojitosti distribuční funkce F. Integrací per partes dostaneme

$$\int_0^T x\,\mathrm{d}F(x) = TF(T) - \int_0^T F(x)\,\mathrm{d}x = -T[1-F(T)] + \int_0^T [1-F(x)]\,\mathrm{d}x. \quad (1.5)$$

Přitom

$$T[1-F(T)] = T\int_T^\infty \mathrm{d}F(x) \le \int_T^\infty x\,\mathrm{d}F(x).$$

Vzhledem k předpokladu, že střední hodnota je konečná, platí $\int_T^\infty x\,\mathrm{d}F(x) \to 0$ pro $T \to \infty$, takže tvrzení věty dostaneme z (1.5) limitním přechodem. □

Označme $\mu'_k = \mathsf{E}X^k$, $k = 1, 2, \dots$. Číslo μ'_k se nazývá *obecný moment* k*-tého řádu*. Existuje-li moment μ'_1 a je-li konečný, pak definujeme $\mu_k = \mathsf{E}(X - \mathsf{E}X)^k$, $k = 0, 1, \dots$. Číslo μ_k se nazývá *centrální moment* k-tého řádu. Nejdůležitější centrální moment je μ_2. Nazývá se *rozptyl* a obvykle se označuje symbolem σ^2. Přitom $\sigma \ge 0$ je *směrodatná odchylka*. Někdy místo σ^2 píšeme $\mathsf{var}\,X$ a čteme *variance* veličiny X. Úpravou dostaneme, že $\sigma^2 = \mathsf{E}X^2 - (\mathsf{E}X)^2$. Je-li $\mathsf{E}X^2 < \infty$, pak také $|\mathsf{E}X| < \infty$, neboť

$$|\mathsf{E}X| = |\mathsf{E}(1.X)| \le \sqrt{\mathsf{E}1^2.\mathsf{E}X^2} = \sqrt{\mathsf{E}X^2} < \infty.$$

Nechť a a b jsou reálná čísla. Existuje-li $\mathsf{E}X$, pak $\mathsf{E}(a + bX) = a + b\mathsf{E}X$. Je-li $\mathsf{E}X^2 < \infty$, pak $\mathsf{var}(a + bX) = b^2\mathsf{var}\,X$. Máme-li náhodné veličiny $X_1, \dots, X_n$ s konečnými středními hodnotami, pak pro libovolná reálná čísla $c_1, \dots, c_n$ platí $\mathsf{E}\sum_{i=1}^n c_i X_i = \sum_{i=1}^n c_i \mathsf{E}X_i$.

Nechť $\mathsf{E}X^4 < \infty$. Pak také $\mathsf{E}X^2 < \infty$. Předpokládejme, že $\sigma > 0$. Pak definujeme *šikmost* $\alpha_3 = \mu_3/\sigma^3$ a *špičatost* $\alpha_4 = \mu_4/\sigma^4$. Při výpočtu šikmosti a špičatosti se často používá vzorců

$$\begin{aligned}\mu_3 &= \mathsf{E}X^3 - 3\mathsf{E}X^2\mathsf{E}X + 2(\mathsf{E}X)^3,\\ \mu_4 &= \mathsf{E}X^4 - 4\mathsf{E}X^3\mathsf{E}X + 6\mathsf{E}X^2(\mathsf{E}X)^2 - 3(\mathsf{E}X)^4.\end{aligned}$$

Jak uvidíme dále, u některých rozdělení neexistuje ani první moment (např. u *Cauchyova rozdělení*). U jiných rozdělení naopak existují a jsou konečné momenty μ'_k pro všechna $k = 1, 2, \dots$. Příkladem je třeba *normální rozdělení*, ale i mnohá další. Pak se může stát, že touto posloupností momentů je takové rozdělení jednoznačně určeno (např. právě normální rozdělení). Říkáme, že jde o *determinované rozdělení* (v Hamburgerově smyslu). Naproti tomu *logaritmicko-normální rozdělení* s hustotou

$$f_1(x) = \frac{1}{\sqrt{2\pi}x}\exp\left\{-\frac{1}{2}\ln^2 x\right\}, \qquad x > 0,$$

a rozdělení s hustotou

$$f_2(x) = [1 + a\sin(2\pi\ln x)]f_1(x), \qquad x > 0,$$

pro libovolné $a \in [-1, 1]$ mají stejné momenty μ'_k, $k = 1, 2, \dots$. Takové rozdělení, které není svými momenty jednoznačně určeno, se nazývá *nedeterminované*.

Nejznámější kritérium determinovanosti je obsaženo v následující větě.

Věta 1.3 *Jestliže řada* $\sum_{k=1}^{\infty} \mu'_k t^k/k!$ *pro některé* $t > 0$ *absolutně konverguje, pak posloupnost momentů* μ'_k, $k = 1, 2, \ldots$, *jednoznačně určuje distribuční funkci.*

Důkaz viz Rao (1978). □

V mnoha případech je nutné použít jemnější kritéria (viz Lin, Huang 1997). Rozdělení náhodné veličiny X je determinované, platí-li

$$\sum_{n=1}^{\infty} (\mathsf{E}X^{2n})^{-\frac{1}{2n}} = \infty$$

(*Carlemanovo kritérium*). Rozdělení náhodné veličiny X s hustotou g je nedeterminované, platí-li

$$\int_{-\infty}^{\infty} \frac{-\ln g(y)}{1+y^2}\,\mathrm{d}y < \infty$$

(*Kreinovo kritérium*).

Povšimněme si, že ve všech těchto případech jsou udány pouze postačující podmínky.

K popisu rozdělení se používá i velké množství dalších charakteristik. Jednou z nich je tzv. *medián* $\tilde{x}$. Je to (libovolné) číslo, pro které platí $F(\tilde{x}) \leq \frac{1}{2}$, $F(\tilde{x}^+) \geq \frac{1}{2}$. Medián obecně těmito podmínkami není jednoznačně určen. Nechť $|\mathsf{E}X| < \infty$. Pak se dá dokázat, že výraz $\mathsf{E}|X - c|$ nabývá své minimální hodnoty při $c = \tilde{x}$.

Další důležitá charakteristika se nazývá *modus* a značí se $\hat{x}$. Je-li diskrétní rozdělení soustředěno v bodech $x_1, x_2, \ldots$, je $\hat{x}$ ten bod, pro který platí $\mathsf{P}(X = \hat{x}) \geq \mathsf{P}(X = x_i)$, $i = 1, 2, \ldots$. Je-li rozdělení spojité s hustotou $f(x)$, pak modus je ta hodnota $\hat{x}$, pro kterou platí $f(\hat{x}) \geq f(x)$, $-\infty < x < \infty$. Ani modus nemusí být určen jednoznačně.

Nechť F je nějaká distribuční funkce. Zaveďme funkci F^{-1} předpisem

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}, \qquad 0 < u < 1.$$

Pak se F^{-1} nazývá *kvantilová funkce* odpovídající distribuční funkci F. Hodnotám $F^{-1}(u)$ se říká *kvantily*. Je-li F rostoucí funkce, pak F^{-1} je obyčejná inverzní funkce k F.

Věta 1.4 *Nechť náhodná veličina* X *má rostoucí spojitou distribuční funkci* F. *Pak pro náhodnou veličinu* $U = F(X)$ *platí*

$$\mathsf{P}(U < u) = \begin{cases} 0 & \textit{pro} \quad u \leq 0, \\ u & \textit{pro} \quad 0 < u < 1, \\ 1 & \textit{pro} \quad 1 \leq u. \end{cases} \tag{1.6}$$

Důkaz. Je-li $u \leq 0$, nebo $1 < u$, je tvrzení zřejmé. Nechť $0 < u < 1$. Pak máme

$$\mathsf{P}(U < u) = \mathsf{P}[F(X) < u] = \mathsf{P}[X < F^{-1}(u)] = F[F^{-1}(u)] = u.$$