Diskrétne rozdelenia pravdepodobnosti

Diskrétne rozdelenie pravdepodobnosti poznáme vtedy, ak pre všetky možné hodnoty náhodnej premennej $\{x_1, x_2, \dots x_n, \dots\}$ určíme pravdepodobnosti $P(X = x_i) = p(x_i)$, pričom $\sum_i p(x_i) = 1$. Ak je množina všetkých možných hodnôt konečná, rozdelenie pravdepodobnosti môžeme vyjadriť vo forme tabuľky, kde prvý riadok sú hodnoty tejto premennej a v druhom riadku ich odpovedajúce pravdepodobnosti.

Rovnomerné (diskrétne) rozdelenie

Náhodná premenná X má rovnomerné rozdelenie s parametrom n, keď pre jej pravdepodobnostnú funkciu $p(x)$ platí:

$$p(x) = P(X = x) = \begin{cases} \frac{1}{n} & \text{pre} \quad x = 1, 2, \dots, n \\ 0 & \text{inde} \end{cases}$$

Túto náhodnú premennú priraďujeme náhodnému pokusu, ktorý má konečný počet n rovnako možných výsledkov.

Uvedené rozdelenie pravdepodobnosti $p(x)$ možno definovať aj pomocou nasledovnej tabuľky:

x	1	2	…	n
$p(x)$	$1/n$	$1/n$	...	$1/n$

Stredná hodnota a rozptyl náhodnej premennej X sa rovná:

$$E(X)=\sum_{i=1}^{n} x_i \cdot p(x_i)=\frac{n+1}{2}$$

$$D(X)=\sum_{i=1}^{n}\left(x_i-E(X)\right)^2 \cdot p(x_i)=E(X^2)-\left[E(X)\right]^2=\frac{n^2-1}{12}$$

Alternatívne rozdelenie

Najjednoduchším typom diskrétneho rozdelenia pravdepodobnosti je alternatívne rozdelenie. Alternatívne rozdelenie sa používa v prípade sledovania výskytu určitej náhodnej udalosti.

Kvantifikáciu nečíselných výsledkov pokusu prevedieme tým spôsobom, že náhodnej premennej priradíme hodnotu 1 s pravdepodobnosťou p, ak pri náhodnom pokuse nastala sledovaná udalosť A a hodnotu 0 s pravdepodobnosťou $1-p$, ak sledovaná udalosť nenastala.

Rozdelenie pravdepodobnosti $A(p)$ je potom dané tabuľkou hodnôt $p(x)$:

x	0	1
$p(x)$	1-p	p

Pre strednú hodnotu a rozptyl náhodnej premennej s alternatívnym rozdelením platí:

$$E(X)=\sum_{i=1}^{2} x_i\, p(x_i)=0\cdot(1-p)+1\cdot p=p$$

$$D(X)=\sum_{i=1}^{2}\left(x_i-E(X)\right)^2 \cdot p(x_i)=p(1-p)$$

Binomické rozdelenie

Uvažujme sériu n nezávislých pokusov, za tých istých podmienok. Pri každom z nich môže sledovaná udalosť A nastať s rovnakou pravdepodobnosťou p a nenastať s pravdepodobnosťou $1-p$. Pravdepodobnosť udalosti A, pri ľubovoľnom pokuse, nezávisí od výsledku predchádzajúcich pokusov. Náhodná premenná X, ktorá udáva, koľkokrát udalosť A pri n-násobnom opakovaní pokusu nastala, môže nadobúdať niektorú hodnotu z množiny $\{0,1,2,\dots,n\}$. Rozdelenie pravdepodobnosti tejto premennej X sa nazýva binomické rozdelenie pravdepodobnosti, označujeme ho $X \sim Bi(n,p)$ a je určené pravdepodobnostnou funkciou $p(x)$:

$$p(x) = P(X = x) = \binom{n}{x} \cdot p^x (1-p)^{n-x} \qquad x = 0,1,2,\dots,n$$

Stredná hodnota a rozptyl náhodnej premennej $X \sim Bi(n,p)$ sa rovná:

$$E(X) = \sum_{i=1}^{n} x_i\, p(x_i) = np$$

$$D(X) = \sum_{i=0}^{n} \left(x_i - E(X)\right)^2 p(x_i) = np(1-p)$$

Rozdelenie súčtu nezávislých náhodných premenných

Nech náhodné premenné $X_1,\dots,X_k$ sú vzájomne nezávislé, pričom premenná X_j má rozdelenie typu $B_i(n_j, p)$ ($j = 1,2,\dots,k$), potom náhodná premenná $X = \sum_{j=1}^{k} X_j$ má rozdelenie $B_i\left(\sum_{j=1}^{k} n_j, p\right)$.

Poznámka. Ak náhodné premenné X_i ($i = 1,2,\dots,n$) majú alternatívne rozdelenie $A(p)$ a sú vzájomne nezávislé, potom náhodná premenná $X = \sum_{i=1}^{n} X_i$ má binomické rozdelenie $B_i(n, p)$. Alternatívne rozdelenie je vlastne špeciálnym prípadom binomického rozdelenia pravdepodobnosti s parametrom $n = 1$.

Príklady rozdelenia pravdepodobnosti $B_i(n, p)$

Náhodná premenná *X* má binomické rozdelenie, ak hodnoty predstavujú napr.

- počet chybných výrobkov zistených vo výbere *n* výrobkov vyrábaných určitým strojom za stálych podmienok,
- počet chybných výrobkov zistených vo výbere *n* výrobkov z danej dávky, pri ktorom výrobok do nej vraciame
- počet prípadov, v ktorých sa prejavila účinnosť podaného prípravku skúšaného na *n* objektoch
- počet prípadov, v ktorých doba výrobnej operácie prekročila určenú hodnotu, za predpokladu, že bola evidovaná doba *n* výrobných operácií.

Relatívne binomické rozdelenie

Nech náhodná premenná X má rozdelenie $B_i(n,p)$, pričom parameter p je pravdepodobnosť sledovanej udalosti A. Náhodná premenná $\frac{X}{n}$ má relatívne binomické rozdelenie pravdepodobnosti a predstavuje relatívnu početnosť udalosti A.

Stredná hodnota a rozptyl tejto premennej sa rovná

$$E\left(\frac{X}{n}\right) = p$$

$$D\left(\frac{X}{n}\right) = \frac{p \cdot (1-p)}{n}.$$

Poznámka. Stredná hodnota relatívnej početnosti náhodnej udalosti A sa teda rovná jej pravdepodobnosti. Uvedená náhodná premenná tvorí základ tzv. p-regulačných diagramov, ktoré sa používajú v štatistickom riadení kvality.

Poissonovo rozdelenie pravdepodobnosti

Ak vychádzame zo situácie opísanej v binomickom rozdelení a predpokladáme, že pravdepodobnosť náhodnej udalosti A je malá ($P(A) \leq 0{,}1$) a $n > 30$, používame namiesto binomického rozdelenia Poissonovo rozdelenie s parametrom λ, ktoré označujeme $Po(\lambda)$. Toto rozdelenie za uvedených podmienok dobre aproximuje binomické rozdelenie.

Pravdepodobnostná funkcia $p(x)$ Poissonovho rozdelenie je definovaná vzťahom

$$p(x) = P(X = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}, x = 0,1,\dots,n,\dots$$

a vyjadruje pravdepodobnosť, že v sérii veľkého počtu nezávislých pokusov *n*, udalosť *A* nastane práve *x*-krát, za predpokladu, že pravdepodobnosť v jednom pokuse $P(A)=\frac{\lambda}{n}$ je veľmi malá. Hodnoty funkcie $p(x)$ sú pre praktické účely tabelované.

Stredná hodnota a rozptyl náhodnej premennej *X* sa rovná

$$E(X)=\sum_{x=0}^{\infty} x\cdot e^{-\lambda}\frac{\lambda^x}{x!}=\lambda$$

$$D(X)=\sum_{x=0}^{\infty}(x-\lambda)^2 e^{-\lambda}\frac{\lambda^x}{x!}=\lambda$$

Teda parameter λ Poissonovho rozdelenia predstavuje strednú hodnotu príslušnej náhodnej premennej *X*.

Pre rozdelenie $P_o(\lambda)$, $\lambda>0$ vždy platí, že $E(X)=D(X)$.

Poznámka. Rovnosť medzi uvedenými charakteristikami sa často využíva v praxi, pri riešení otázky, či je správna hypotéza (tvrdenie) o tom, že náhodná premenná má Poissonovo rozdelenie pravdepodobnosti. Niekedy sa toto rozdelenie nazýva rozdelenie výskytu zriedkavých udalostí.

Príklady Poissonovho rozdelenia pravdepodobnosti

Náhodná premenná X má Poissonovo rozdelenie, ak hodnoty predstavujú napríklad

- počet ťažkých dopravných nehôd v určitom meste za deň,
- počet smrteľných pracovných úrazov v nejakom rezorte za rok,
- počet telefónnych výziev na určitom prístroji v danom časovom intervale,
- počet chýb na zvare pri nedeštruktívnej skúške,
- počet bublín na ploche danej veľkosti na odliatku,
- počet trhlín na 1km kábla a pod.

Poznámka. Každá udalosť, ktorá sa vyskytuje na jednotke plochy (času, objemu a pod.) sa zväčša dobre aproximuje rozdelením $P_o(\lambda)$

Poznámka. Pre dostatočne veľké n a dostatočne malé p môžeme binomické rozdelenie aproximovať rozdelením Poissonovým. Pre $n > 30$ a $p < 0{,}1$ sa dopustíme chyby menšej ako 10^{-2}.

Hypergeometrické rozdelenie

Náhodná premenná X má hypergeometrické rozdelenie, ak jej funkcia rozdelenia pravdepodobnosti má tvar

$$p(x) = P(X = x) = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}, x = \max\{0, M - N + n\}, \ldots, \min\{M, n\}$$

kde N, M, n sú prirodzené čísla, pre ktoré platí $1 \le n < N$, $1 \le M < N$. Označujeme ho $X \sim H(N, M, n)$.

Hypergeometrické rozdelenie predstavuje nasledujúca modelová situácia: Máme konečnú množinu *N* jednotiek, z ktorých *M* má určitú vlastnosť V. Z tejto množiny vyberieme náhodne, naraz alebo postupne, bez vracania, *n* jednotiek. Náhodná premenná X, ktorá predstavuje počet *x* jednotiek z *n* výbraných jednotiek, ktoré majú vlastnosť V, má hypergeometrické rozdelenie.

Pre strednú hodnotu tohto rozdelenia a rozptyl náhodnej premennej platí

$$E(X) = n\frac{M}{N}$$

$$D(X) = n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$$

Poznámka. Dá sa dokázať, že pre malé hodnoty $\frac{n}{N} < 0{,}1$ možno hypergeometrické rozdelenie aproximovať binomickým rozdelením s parametrami *n* a $p = \frac{M}{N}$.

Poznámka. Z porovnania charakteristík binomického a hypergeometrického rozdelenia môžme usúdiť, že nastáva ich zhoda v prípade, že počet prvkov N v súbore je oproti počtu vybraných prvkov *n* veľmi veľký . Potom podiel $\frac{N-n}{N-1}$ sa približne rovná 1.