Opakování pravděpodobnostních pojmů II.

Distribuční funkce diskrétní „jednorozměrné“ náhodné proměnné

Mějme náhodnou proměnnou ξ, která je definována na některé části celočíselné osy $D \subset Z_1$, potom pod její distribuční funkcí budeme rozumět $F_\xi(x) = P(\xi < x); \quad x \in R_1$. Obvykle se distribuční funkce dodefinovává pro celou R_1 a to tak, že pro definiční obor $D \subset Z_1$ náhodné proměnné ξ a mimo něj:

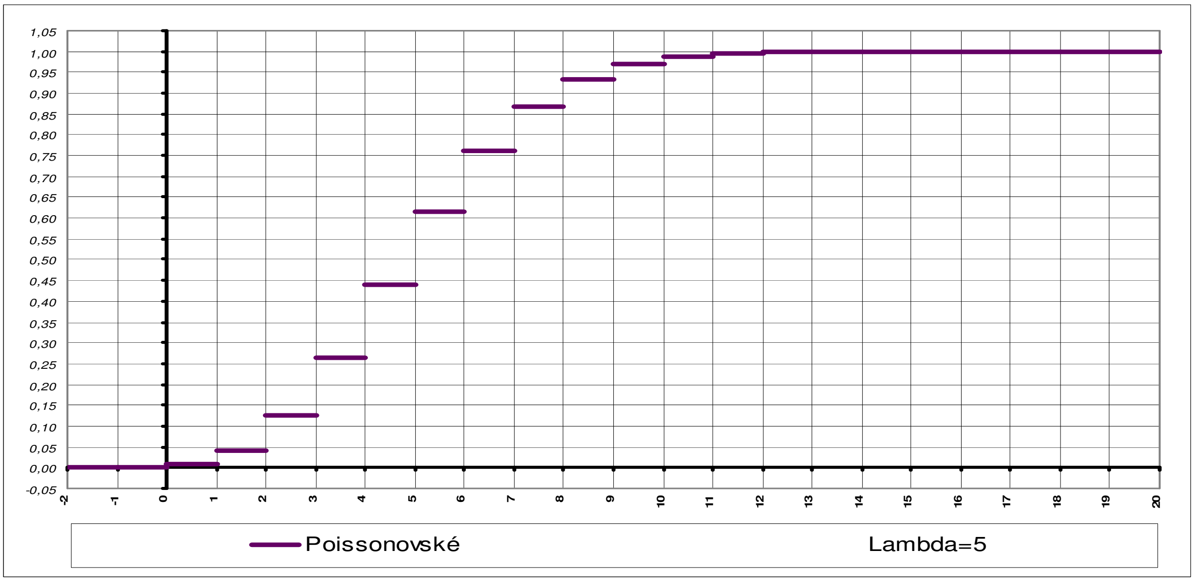
1. Pro $x \in D$, platí $F_\xi(x) = P(\xi < x)$
2. Pro x takové, že $\forall y \in D \Rightarrow x < y$, $F_\xi(x) = 0$
3. Pro x takové, že $\forall y \in D \Rightarrow x > y$, $F_\xi(x) = 1$
4. Protože D není souvislý (ve smyslu a v R_1) pak pro bod $x : x \notin D; \inf(D) < x < \sup(D)$ je „dodefinováno“ $F_\xi(x) = \sup\{F_\xi(z) : z \in D; z < x\}$. Poznámka, striktně řečeno nejedná se o dodefinování.

Vlastnosti (shodné se spojitým případem)

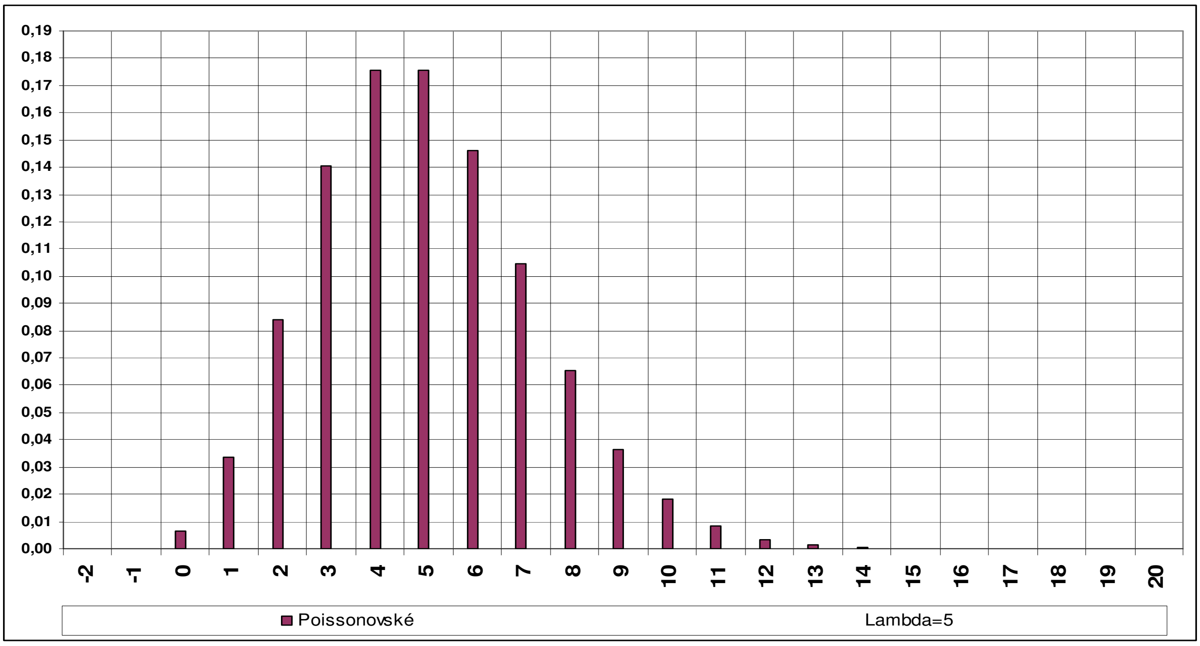
1. $F_\xi(x)$ je neklesající.
2. $F_\xi(x)$ je zleva spojitá.
3. $\lim\limits_{x \to -\infty} F_\xi(x) = 0$.
4. $\lim\limits_{x \to +\infty} F_\xi(x) = 1$.
5. $F_\xi(x)$ je shora i zdola omezená. Tato vlastnost je nadbytečná, lze ji dokázat z předchozích čtyř.

Každá distribuční funkce má uvedené vlastnosti 1-4. Každá funkce mající uvedené vlastnosti 1-4 je distribuční nějaké náhodné proměnné.

Příklady:



Poissonovské rozdělení – distribuční funkce $F(x) = \sum_{i<x} e^{-\lambda} \frac{\lambda^i}{i!}$.



Poissonovské rozdělení – hustota = pravděpodobnost $P\{\ \xi = i\ \} = f(x) = \begin{matrix} e^{-\lambda}\dfrac{\lambda^i}{i!} & \Leftrightarrow x = i \\ 0 & \Leftrightarrow x \neq i \end{matrix}$.

Pokud je, v některých případech, potřebné lze pro popis diskrétních pravděpodobností použít aparátu „$\delta -$funkce“ = „$\delta -$impulsu“ $f(x) = P\{\xi = i\}\delta(x-i)$. Potom: $F(x) = \int\limits_{-\infty}^{x} P\{\xi = i\}\delta(y-i)dy = \sum\limits_{i<x} P\{\xi = i\}$.

Diskrétní rozdělení nad souborem elementárních jevů bez struktury *R₁*

V takovém případě je nosičem „diskrétních“ pravděpodobností nějaká „abeceda elementárních jevů“ D bez další struktury R_1 (tj. vzdálenosti, vztahy býti větší, ….). Taková „abeceda“ obvykle bývá konečná, obecně může být i nekonečná (=spočetná = s mohutností množiny přirozených čísel). $\forall d \in D; P(d) \geq 0;\quad \sum\limits_{d\in D} P(d) = 1;\quad \forall A \subset D; P(A) = \sum\limits_{d\in A} P(d);\quad P(\emptyset) = 0$.

Momenty – modifikace vyjádření pro diskrétní náhodné proměnné

Střední hodnota: $E\{\xi\} = \int\limits_{-\infty}^{+\infty} s f_\xi(s) ds = \int\limits_{-\infty}^{+\infty} s dF_\xi(s) = \sum\limits_{i\in D} iP\{\xi = i\}$**,** pokud existuje.

Námět: Nalezněte diskrétní rozdělení ke kterému neexistuje střední hodnota (návod, využijte vlastností harmonické řady).

Obecné momenty: $E\{\xi^k\} = \sum\limits_{i\in D} i^k P\{\xi = i\}$**.**

Centrální momenty: $E\{(\xi - E\{\xi\})^k\} = \sum\limits_{i\in D} (i - E\{\xi\})^k P\{\xi = i\}$**.** Speciální postavení má druhý centrální moment = rozptyl = variance:

$$Var(\xi) = \sigma^2(\xi) = E\{(\xi - E\{\xi\})^2\} = \sum_{i\in D} (i - E\{\xi\})^2 P\{\xi = i\}.$$

Kvantily: p-kvantilem je řešení x_p rovnice: $F_\xi(x_p) = p$. Jak je zřejmé z průběhu distribuční funkce diskrétní náhodné proměnné p-kvantily nebudou až na zřídké výjimky existovat. Diskutujte možnosti alternativní definice $P(\xi \le x_p) \ge p; P(\xi \ge x_p) \ge 1-p$.

Vytvořující funkce pravděpodobností

Dále budeme používat označení $P\{\xi = i\} = p_i; i \in D; p_i = 0 \Leftrightarrow (i \notin D) \wedge (i \ge 0)$ a to pro takové diskrétní náhodné proměnné, pro které $P\{\xi = i\} > 0 \Rightarrow (i \ge 0) \wedge (i \in Z)$.

Pod vytvořující funkcí pravděpodobností budeme rozumět funkci $G_\xi(z) = \sum_{i=0}^{+\infty} p_i z^i$, jinak napsáno $G_\xi(z) = E\{z^\xi\}$. Pak platí: $p_0 = G_\xi(0)$ a $p_k = \frac{d^k}{dz^k}\frac{G_\xi(z)}{k!}\Big|_{z=0}$. Dokažte!

Příklady

Rozdělení	Pravděpodobnost	Vytvořující funkce pravděpodobností
Binomické	$p_i = \binom{n}{i} p^i (1-p)^{n-i}$	$G(z) = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} z^i = \sum_{i=0}^{n} \binom{n}{i} (pz)^i (1-p)^{n-i} = (pz + (1-p))^n$
Poissonovo	$p_i = e^{-\lambda} \frac{\lambda^i}{i!}$	$G(z) = \sum_{i=0}^{+\infty} e^{-\lambda} \frac{\lambda^i}{i!} z^i = \sum_{i=0}^{+\infty} e^{-\lambda} \frac{(\lambda z)^i}{i!} = e^{\lambda(z-1)}$
Geometrické	$p_i = (1-q)q^i$	$G(z) = \sum_{i=0}^{+\infty} (1-q) q^i z^i = \sum_{i=0}^{+\infty} (1-q)(qz)^i = \frac{1-q}{1-qz}$
….		
Negativně binomické	$p_i = \binom{i+r-1}{r-1} p^r (1-p)^i; r \ge 1$	$G(z) = \sum_{i=0}^{+\infty} \binom{i+r-1}{r-1} p^r (1-p)^i z^i = \left(\frac{pz}{1-(1-p)z}\right)^r$

Další užití vytvořující funkce pravděpodobností

$$G_\xi(1) = \sum_{i=0}^{+\infty} p_i = 1, \quad \frac{d}{dz} G_\xi(z)\Big|_{z=1} = \frac{d}{dz} \sum_{i=0}^{+\infty} p_i z^i \Big|_{z=1} = \sum_{i=0}^{+\infty} i p_i z^{i-1} \Big|_{z=1} = \sum_{i=0}^{+\infty} i p_i = E\{\xi\},$$

$$\frac{d^2}{dz^2} G_\xi(z)\Big|_{z=1} = \frac{d^2}{dz^2} \sum_{i=0}^{+\infty} p_i z^i \Big|_{z=1} = \sum_{i=0}^{+\infty} i(i-1) p_i z^{i-2} \Big|_{z=1} = \sum_{i=0}^{+\infty} i(i-1) p_i = E\{\xi^2\} - E\{\xi\}$$

Námět: Odvoďte výraz pro $Var(\xi) = \sigma^2(\xi) = E\{(\xi - E\{\xi\})^2\} = E\{\xi^2\} - (E\{\xi\})^2$ pomocí vytvořující funkce pravděpodobností.

Tvrzení: Mějme náhodnou proměnnou η, která je tvořena součtem nezávislých náhodných proměnných $\eta = \sum_{i=1}^{n} \xi_i$, potom $G_\eta(z) = \prod_{i=1}^{n} G_{\xi_i}(z)$.

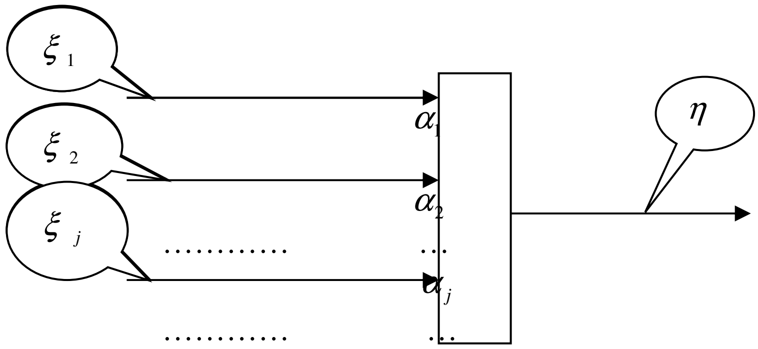
Důkaz: $G_\eta(z)=E\{z^\eta\}=E\left\{z^{\sum_{i=1}^{n}\xi_i}\right\}=\prod_{i=1}^{n}E\{z^{\xi_i}\}=\prod_{i=1}^{n}G_{\xi_i}(z)$.

Tvrzení: Mějme směs rozdělení diskrétních náhodných proměnných $\xi_1,\xi_2\ldots\xi_j,\ldots$

$$P_i = P(\eta = i) = \sum_{j=1}^{+\infty}\alpha_j p_{\xi_j i};\; 1 = \sum_{j=1}^{+\infty}\alpha_j; \alpha_j \geq 0 ,$$

kde $p_{\xi_j i}$ je pravděpodobnost toho, že náhodná proměnná ξ_j nabývá hodnoty i. Potom náhodná proměnná η jež je „výstupem“ takové směsi má vytvořující funkci: $G_\eta(z)=\sum_{j=1}^{+\infty}\alpha_j G_{\xi_j}(z)$.

Důkaz: Směs rozdělení si lze představit jako přepínač, který na výstup přepíná náhodně jednu ze vstupních náhodných proměnných a pravděpodobnosti toho, že na výstup bude přepnuta *j-tá* vstupní náhodná proměnná jsou rovny „vahám“ směsi α_j. Jev přepnutí je nezávislý na jevech vstupních hodnot jednotlivých náhodných proměnných.



$$G_\eta(z)=\sum_i\sum_{j=1}^{+\infty}\alpha_j p_{\xi_j i}z^i=\sum_{j=1}^{+\infty}\alpha_j\sum_i p_{\xi_j i}z^i=\sum_{j=1}^{+\infty}\alpha_j G_{\xi_j}(z).$$

Tvrzení: Nechť nezávislé náhodné veličiny $\xi_1,\xi_2,\ldots,\xi_j,\ldots$ mají stejné rozdělení pravděpodobnosti s vytvořující funkcí $G(z)$ a nechť náhodná veličina N má rozdělení pravděpodobnosti na množině přirozených čísel $D=\{1,2,\ldots,i,\ldots\}$ s pravděpodobnostmi q_i. ξ_j, N jsou nezávislé $\forall j=1,\ldots$. Potom náhodná veličina $\eta=\sum_{j=1}^{N}\xi_j$ (součet náhodného počtu náhodných proměnných) má vytvořující funkci $G_\eta(z)=G_N(G(z))$.

Důkaz:

Pro každé pevné n má náhodná proměnná $\eta_n=\sum_{j=1}^{n}\xi_j$ vytvořující funkci $G_n(z)=G^n(z)$, jedná se o součet n stejně rozdělených a nezávislých náhodných proměnných. Náhodná proměnná η je vlastně směsí náhodných proměnných $\eta_n=\sum_{j=1}^{n}\xi_j$ s vahami (pravděpodobnostmi) q_n, proto $G_\eta(z)=\sum_{n=1}^{+\infty}q_n G^n(z)=\sum_{n=1}^{+\infty}q_n(G(z))^n=G_N(G(z))$.

Námět: Odvoďte podobné tvrzení pro „spojité“ náhodné proměnné $\xi_1, \xi_2 ..., \xi_j, ...$ s využitím aparátu charakteristických funkcí. Inspirace: Alfréd Rényi; Teorie pravděpodobnosti, ACADEMIA, Praha 1972, str. 262-314.

Některá diskrétní rozdělení pravděpodobností:

Rozdělení	**Pravděpodobnost**	**Střední hodnota**	**Rozptyl**
Rovnoměrné	$p_i = \frac{1}{N}$, $i=1,\dots,N$	$\frac{N+1}{2}$	$\frac{N^2-1}{12}$
Alternativní	$p_1 = p; p_0 = 1-p$	p	$p(1-p)$
Alternativní (Rademacher)	$p_{-1} = 1/2; p_{+1} = 1/2$	0	1
Binomické	$p_i = \binom{n}{i} p^i (1-p)^{n-i}$ $i=0,\dots,n$	np	$np(1-p)$
Poissonovo	$p_i = e^{-\lambda}\frac{\lambda^i}{i!}$, $i=0,1,\dots$	λ	λ
Geometrické	$p_i = (1-q)q^i$, $i=0,1,\dots$	$\frac{q}{1-q}$	$\frac{q}{(1-q)^2}$
Negativně binomické	$p_i = \binom{i+r-1}{r-1} p^r (1-p)^i; r \geq 1$, $i=0,1,\dots$ někdy bývají uváděny hodnoty $r, r+1, r+2,\dots$ v tom případě budou střední hodnota a rozptyl vypadat jinak – odvoďte.	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$
Hypergeometrické	$p_i = \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$, $i=max(0,M-N+n),\dots, min(M,n)$	$n\frac{M}{N}$	$n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$
Logaritmické	$\frac{-p^i}{i\,lg(1-p)}$, $i=1,2,\dots$	$\frac{-1}{lg(1-p)}\frac{p}{1-p}$	$\frac{-p(p+lg(1-p))}{(1-p)^2 lg^2(1-p)}$

Doporučená a zdrojová literatura:

Jiří Reif	Metody matematické statistiky, ZČU v Plzni 2004
Jaroslav Hátle, Jiří Likeš	Základy počtu pravděpodobnosti a matematické statistiky. SNTL Praha 1974.
Alfréd Rényi	Teorie pravděpodobnosti, ACADEMIA, Praha 1972
C. Radhakrishna Rao	Lineární metody statistické indukce a jejich aplikace, ACADEMIA, Praha 1978
http://en.wikipedia.org/wiki/Probability_distribution	