Kurs Prawdopodobieństwo

Wzory

Elementy kombinatoryki

	Kolejność ma znaczenie	Kolejność nie ma znaczenia
Te same elementy nie mogą się powtarzać	Reguła mnożenia $\sqcup \quad \sqcup \quad \sqcup$	Wzór z dwumianem Newtona $\binom{n}{k}$
Te same elementy mogą się powtarzać		Wzór: $\binom{n+k-1}{k}$

„Klasyczna” definicja prawdopodobieństwa

$$P(A)=\frac{\overline{\overline{A}}}{\overline{\overline{\Omega}}}$$

gdzie:

$\overline{\overline{A}}$ - liczba zdarzeń sprzyjających A

$\overline{\overline{\Omega}}$ - liczba wszystkich zdarzeń

eTrapez Usługi Edukacyjne E-learning Krystian Karczyński
www.etrapez.pl
Tel. 603 088 274

Prawdopodobieństwo – definicja Kołmogorowa

Ω - zbiór wszystkich zdarzeń elementarnych

S – „sigma-ciało" na zbiorze Ω, czyli zbiór jego podzbiorów spełniający warunki:

1) $\phi \in S$
2) $A \in S \Rightarrow A' \in S$
3) $A_1, A_2, A_3, \ldots \in S \Rightarrow A_1 \cup A_2 \cup A_3 \cup \ldots = \bigcup_{n=1}^{\infty} A_n \in S$

P – funkcja o argumentach ze zbioru S i wartościach będących liczbami rzeczywistymi, spełniająca warunki („aksjomaty"):

1. $\forall_{A \in S} P(A) \geq 0$
2. $P(\Omega) = 1$
3. $P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$ - dla zdarzeń parami rozłącznych, tzn. $(A_i \cap A_j = \phi \;\; dla \;\; i \neq j)$

Wartości funkcji $P(A)$ możemy nazywać „prawdopodobieństwem"

Własności prawdopodobieństwa

1. $P(A) \in < 0,1 >$
2. $P(\phi) = 0$
3. $A \subseteq B \Rightarrow P(A) \leq P(B)$
4. $P(A') = 1 - P(A)$
5. $A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)$
6. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Niezależność zdarzeń

Zdarzenia A i B są niezależne, gdy:
$P(A \cap B) = P(A) \cdot P(B)$

Prawdopodobieństwo warunkowe

Prawdopodobieństwo warunkowe zajścia zdarzenia A pod warunkiem zajścia zdarzenia B oznaczamy jako $P(A|B)$ i liczymy ze wzoru:

$$P(\mathrm{A}|\mathrm{B})=\frac{P(\mathrm{A}\cap\mathrm{B})}{P(\mathrm{B})}$$

Prawdopodobieństwo całkowite i wzór Bayesa

Zakładając, że $\mathrm{B}_i\cap\mathrm{B}_j=\varnothing\ (dla\ i\neq j)\ i\ P(\mathrm{B}_1)+P(\mathrm{B}_2)+\ldots+P(\mathrm{B}_n)=1$:

Prawdopodobieństwo całkowite

$$P(\mathrm{A})=P(\mathrm{B}_1)P(\mathrm{A}|\mathrm{B}_1)+P(\mathrm{B}_2)P(\mathrm{A}|\mathrm{B}_2)+\ldots+P(\mathrm{B}_n)P(\mathrm{A}|\mathrm{B}_n)$$

Wzór Bayesa

$$P(\mathrm{B}_i|\mathrm{A})=\frac{P(\mathrm{B}_i)P(\mathrm{A}|\mathrm{B}_i)}{P(\mathrm{A})}$$

Schemat Bernoulliego

Prawdopodobieństwo zajścia k „sukcesów" w n niezależnych i identycznych doświadczeniach, z których każde może zakończyć się tylko na dwa sposoby (z prawdopodobieństwami p dla „sukcesu" i q dla „porażki") wynosi:

$$P(S=k)=\binom{n}{k}p^kq^{n-k}$$

Zmienne losowe

Dyskretne zmienne losowe

Rozkład

x_i	...	...	...
p_i	...	...	...

$\sum p_i = 1$

Dystrybuanta

$F(x) = P(X < x)$

Wartość oczekiwana

$EX = \sum x_i p_i$

Mediana $x_{0,5}$**,** Me

Wartość zmiennej losowej, dla której skumulowane prawdopodobieństwa „przekraczają” $\frac{1}{2}$

Dominanta, moda D

Wartość zmiennej losowej osiągana z największym prawdopodobieństwem

Kwantyl rzędu p x_p

Wartość zmiennej losowej, dla której skumulowane prawdopodobieństwa „przekraczają” p

Wariancja $D^2(X)$**,** σ^2

$D^2(X) = \sum (x_i - EX)^2 p_i$, $D^2(X) = EX^n - (EX)^n$

Odchylenie standardowe $D(X)$**,** σ

$D(X) = \sqrt{D^2(X)}$

Współczynnik zmienności V

$V = \frac{D(X)}{E(X)}$

Moment zwykły n-tego rzędu EX^n, α_n

$$EX^n = \sum x_i^n p_i$$

Moment centralny n-tego rzędu μ_n

$$\mu_n = \sum (x_i - EX)^n p_i$$

Współczynnik asymetrii

$$\gamma_1 = \frac{\mu_3}{(D(X))^3}$$

Współczynnik koncentracji

$$K = \frac{\mu_4}{(D(X))^4}$$

Przykłady rozkładów dyskretnych zmiennych losowych

Rozkład Bernoulliego

W rozkładzie Bernoulliego prawdopodobieństwa określane są ze wzoru:

$$P(X = k) = \binom{n}{k} p^k q^{n-k}$$

$$EX = np$$

$$D^2(X) = npq$$

Rozkład Poissona

W rozkładzie Poissona prawdopodobieństwa określane są ze wzoru:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$EX = \lambda$$

$$D^2(X) = \lambda$$

Dla dużych n i małych p rozkład Bernoulliego można zastępować rozkładem Poissona

Rozkład hipergeometryczny

W rozkładzie hipergeometrycznym prawdopodobieństwa określane są ze wzoru:

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

Gdzie N to ilość wszystkich elementów w populacji, M to ilość wszystkich elementów w populacji o określonej cesze, n to ilość elementów w próbce, k to ilość elementów w próbce o określonej cesze

$$EX=\frac{M\cdot n}{N}$$

$$D^2X=n\cdot\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$$

Dla dużych N i M, oraz $\frac{M}{N}\to p$ rozkład Bernoulliego można zastępować rozkładem hipergeometrycznym.

Ciągłe zmienne losowe

Funkcja gęstości

$f(x)$

$$P(a \le X \le b) = \int_a^b f(x)dx$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Dystrybuanta

$$F(x) = P(X < x) = \int_{-\infty}^{x} f(t)dt$$

Wartość oczekiwana

$$EX = \int_{-\infty}^{\infty} xf(x)dx$$

Mediana $x_{0,5}$**,** Me

Wartość $x_{0,5}$, dla której $F(x_{0,5}) = 0,5$

Dominanta, moda D

Maksimum globalne funkcji gęstości $f(x)$

Kwantyl rzędu p x_p

Wartość x_p, dla której $F(x_p) = p$

Wariancja $D^2(X)$**,** σ^2

$$D^2(X) = \int_{-\infty}^{\infty} (x - EX)^2 f(x)dx, D^2(X) = EX^n - (EX)^n$$

Odchylenie standardowe $D(X)$**,** σ

$$D(X) = \sqrt{D^2(X)}$$

Współczynnik zmienności V

$$V = \frac{D(X)}{E(X)}$$

Moment zwykły n-tego rzędu EX^n, α_n

$$EX^n = EX = \int_{-\infty}^{\infty} x^n f(x) dx$$

Moment centralny n-tego rzędu μ_n

$$\mu_n = \int_{-\infty}^{\infty} (x - EX)^n f(x) dx$$

Współczynnik asymetrii

$$\gamma_1 = \frac{\mu_3}{(D(X))^3}$$

Współczynnik koncentracji

$$K = \frac{\mu_4}{(D(X))^4}$$

Przykłady rozkładów ciągłych zmiennych losowych

Rozkład normalny

W rozkładzie normalnym prawdopodobieństwa określane są z funkcji gęstości o wzorze:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(x-m)^2}{2\sigma^2}}$$

$$EX = m$$

$$D^2(X) = \sigma^2$$

Standaryzacja rozkładu normalnego:

$$Z = \frac{X - m}{\sigma}$$

Rozkład jednostajny

W rozkładzie jednostajnym prawdopodobieństwa określane są z funkcji gęstości o wzorze:

$$f\left(x\right)=\begin{cases}\dfrac{1}{b-a} & w\ przedziale\ \ x\in\left\langle a,b\right\rangle\\ 0 & dla\ pozostalych\ \ x\end{cases}$$

Rozkład wykładniczy

W rozkładzie wykładniczym prawdopodobieństwa określane są z funkcji gęstości o wzorze:

$$f\left(x\right)=\begin{cases}\dfrac{1}{\lambda}e^{-\frac{x}{\lambda}} & dla\ \ x\geq 0\\ 0 & dla\ \ x<0\end{cases}$$

$$EX=\lambda$$

$$D^2\left(X\right)=\lambda$$

Zmienne losowe dwuwymiarowe

Dyskretne zmienne losowe dwuwymiarowe

Rozkład

$X_{x_i} \diagdown Y_{y_j}$	y_1	y_2	$\cdots$	y_j	
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\sum p_{1.}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\sum p_{2.}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\sum p_{i.}$
	$\sum p_{.1}$	$\sum p_{.2}$	$\cdots$	$\sum p_{.j}$	1

Rozkłady brzegowe

$$\sum_i p_{i.} \text{ , } \sum_j p_{.j}$$

Prawdopodobieństwo warunkowe

$$P\left(X = x_i \middle| Y = y_j\right) = \frac{p_{ij}}{p_{.j}} \text{ , } P\left(Y = y_j \middle| X = x_i\right) = \frac{p_{ij}}{p_{i.}}$$

Niezależność zmiennych losowych

Dwie zmienne losowe X i Y nazywamy niezależnymi, gdy:

$$\underset{i,j}{\forall} P\left(X = x_i, Y = y_j\right) = P\left(X = x_i\right) \cdot P\left(Y = y_j\right)$$

Dystrybuanta

$$F(x, y) = P(X < x, Y < y) = \sum_{x_i < x} \sum_{y_j < y} p_{ij}$$

Wartości oczekiwane

Wartości oczekiwane EX , EY liczymy z rozkładów brzegowych

Wariancje

Wariancje $D^2(X)$ $D^2(Y)$, liczymy z rozkładów brzegowych.

eTrapez Usługi Edukacyjne E-learning Krystian Karczyński
www.etrapez.pl
Tel. 603 088 274

Kowariancja

$$C(X,Y)=\sum_i\sum_j\left(x_i-E(X)\right)\left(y_j-E(Y)\right)p_{ij}$$

Współczynnik korelacji

$$\rho=\frac{C(X,Y)}{D(X)\cdot D(Y)}$$

Jeśli $\rho=0$ zmienne losowe nazywamy „nieskorelowanymi". Nie oznacza to jednak, że są niezależne. Jeśli jednak zmienne losowe są niezależne, to na pewno $\rho=0$.

Ciągłe zmienne losowe dwuwymiarowe

Funkcja gęstości

$$f(x,y)$$

$$P(a \le X \le b,\ c \le Y \le d) = \int_a^b \int_c^d f(x,y)\,dydx$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dydx = 1$$

Rozkłady brzegowe

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)\,dy, \quad f_2(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

Rozkłady warunkowe

$$f(X|Y) = \frac{f(x,y)}{f_2(y)}, \quad f(Y|X) = \frac{f(x,y)}{f_1(x)}$$

Niezależność zmiennych losowych

Dwie zmienne losowe X i Y nazywamy niezależnymi, gdy dla dowolnych x i y:

$$f(x,y) = f_1(x) \cdot f_2(y)$$

Dystrybuanta

$$F(x,y) = P(X < x, Y < y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\,dudv$$

Wartości oczekiwane

$$E(X) = \int_{-\infty}^{\infty} x f_1(x)\,dx$$

$$E(Y) = \int_{-\infty}^{\infty} y f_2(y)\,dy$$

Wariancje

$$D^2(X)=\int_{-\infty}^{\infty}(x-EX)^2\, f_1(x)dx$$

$$D^2(Y)=\int_{-\infty}^{\infty}(y-EY)^2\, f_2(y)dx$$

Kowariancja

$$C(X,Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-E(X))(y-E(Y))\, f(x,y)dxdy$$

Współczynnik korelacji

$$\rho=\frac{C(X,Y)}{D(X)\cdot D(Y)}$$

Jeśli $\rho=0$ zmienne losowe nazywamy „nieskorelowanymi”. Nie oznacza to jednak, że są niezależne. Jeśli jednak zmienne losowe są niezależne, to na pewno $\rho=0$.