Ściąga na egzamin z Rachunku Prawdopodobieństwa i Elementów Statystyki Matematycznej - ©Piotr Kowalski, 29 stycznia 2020, wersja 0.3, MIT License

Zdarzenie elementarne

Każdy możliwy wynik eksperymentu losowego nazywamy **zdarzeniem elementarnym** ω, a zbiór wszystkich możliwych wyników eksperymentu (wszystkich zdarzeń elementarnych) nazywamy **zbiorem zdarzeń elementarnych** i oznaczamy $\Omega, (\omega \in \Omega)$.

Aksjomaty prawdopodobieństwa

Dla danego zbioru zdarzeń elementarnych Ω oraz σ-ciała zdarzeń losowych $\mathcal{F}$, **prawdopodobieństwem** nazywamy funkcję $P : \mathcal{F} \to \mathcal{R}$ spełniającą:

1. Dla dowolnego zdarzenia losowego $A \in \mathcal{F}$, $P(A) \geqslant 0$.
2. $P(\Omega) = 1$.
3. Dla dowolnego nieskończonego ciągu zdarzeń losowych $A_1, A_2, \ldots, \forall_{n\in\mathcal{N}} A_n \in \mathcal{F}$, parami rozłącznych, mamy $P(\bigcup_{n=1}^{\infty}) = \sum_{n=1}^{\infty} P(A_n)$.

Dla dowolnych zdarzeń A, B mamy $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Prawdopodobieństwo warunkowe

Prawdopodobieństwo A pod warunkiem że zaszło zdarzenie B: $P(A|B) = \frac{P(A \cup B)}{P(B)}$.

Jeżeli $P(A_1 \cap \ldots \cap A_n) > 0$, to $P(A_1 \cap \ldots \cap A_n) = P(A_1) \prod_{i=2}^{n} P(A_i | A_1 \cap \ldots \cap A_{i-1})$.

Prawdopodobieństwo zupełne

Ciąg zdarzeń nazywamy zupełnym, jeśli:

1. $\bigcup_i A_i = \Omega$,
2. $\forall_{i \neq j} A_i \cap A_j = \emptyset$,
3. $\forall_i P(A_i) > 0$.

Twierdzenie

Jeśli zdarzenia tworzą układ zupełny, to dla dowolnego zdarzenia B mamy $P(B) = \sum_i P(B|A_i)P(A_i)$

Reguła Bayesa

Twierdzenie Niech A_i tworzą układ zupełny. Wtedy dla dowolnego zdarzenia losowego $B, P(B) > 0$ i dowolnego j zachodzi $P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_i P(B|A_i)P(A_i)}$

Niezależność zdarzeń

Definicja Zdarzenia są wzajemnie niezależne gdy $P(A \cap B) = P(A) \cdot P(B)$

Jeżeli zdarzenia A i B są niezależne, to niezależne są również zdarzenia A i $\overline{B}$, $\overline{A}$ i B, $\overline{A}$ i $\overline{B}$.

Zdarzenia są wzajemnie niezależne jeśli $P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$.

Jeśli $A_1 \ldots$ są zdarzeniami wzajemnie niezależnymi, to $(\bigcup_{i=1}^{n} A_i) = 1 - \prod_{i=1}^{n}(1 - P(A_i))$

Łączenie prawdopodobieństw

szeregowe: $P(A_s) = \prod_{i=1}^{n} p_i$

równoległe: $P(A_r) = 1 - \prod_{i=1}^{n}(1 - p_i)$

Dystrybuanta

$F(x) = P(X \leqslant x) = P(\{\omega \in \Omega : X(\omega) \leqslant x\})$

Własności:

$\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$

niemalejąca, prawostronnie ciągła

$P(a < X \leqslant b) = F(b) - F(a)$

Zmienne losowe dyskretne

$p(a) = P(X = a)$ - funkcja prawdopodobieństwa, własności:

$p(x) \geqslant 0$, $\sum_{x\in X} p(x) = 1$

Dystrybuanta dyskretna zmiennej X o nośniku χ: $F(x) = \sum_{\{x_i\in\chi : x_i \leqslant x\}} p(x_i)$.

Przykład:

$$F(x) = \begin{cases} 0, x < 1 \\ 0.4, 1 \leqslant x < 2 \\ 0.9, 2 \leqslant x \end{cases}$$

Zmienne losowe ciągłe

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

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

$f(x) = F'(x)$

$P(a \leqslant X \leqslant b) = F(b) - F(a)$

$P(X = a) = 0$ dla dowolnego $a \in R$

Własności funkcji gęstości:

$\forall_{x\in R} f(x) \geqslant 0$, $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$

Funkcje zmiennych losowych

Dyskretne:

$P(Y = y_i) = \sum_{\{x_j\in\chi : g(x_j) = y_i\}} P(X = x_j)$

Ciągłe: gęstość liniowej funkcji zm.los.

$\forall_{a\neq 0} b \in R f_{aX+b}(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$

gęstość kwadratu zm.los.

$$f_{X^2}(y) = \begin{cases} 0, x \leqslant 0 \\ \frac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right], x > 0 \end{cases}$$

Jeśli g jest ściśle monotoniczna i różniczkowalna, to $Y = g(X)$: $f_Y(y) = f_X(g^{-1}(y)) \cdot \frac{1}{|g'(g^{-1}(y))|}$

Wartość oczekiwana

Dyskretne: $E[X] = \sum_{x_i\in X} x_i \cdot P(X = x_i)$

Ciągłe: $E[X] = \int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x$

Dla typu mieszanego o $F(x) = pF_d(x) + (1 - p)F_c(x)$, $E[X] = pE[X_d] + (1 - p)E[X_c]$.

Dla funkcji zmiennej losowej $Y = g(X)$:

$$E[g(X)] = \begin{cases} \sum_i g(x_i)P(X = x_i), \text{ jeśli } X \text{ jest zm.los. dyskretną} \\ \int_{-\infty}^{+\infty} g(x)f_X(x)\mathrm{d}x, \text{ jeśli } X \text{ jest zm.los. ciągłą} \end{cases}$$

Jeśli istnieje wartość oczekiwana $E[X]$, to $E[aX + b] = aE[X] + b$

Momenty

Momentem rzędu n-tego względem $c \in R$ zmiennej losowej X nazywamy: $E[(X - c)^n]$

Momenty zwykłe: $c = 0$

Pierwszy moment: $E[X]$

Jeśli istnieje n-ty moment zwykły, to isnieją wszystkie momenty rzędu mniejszego od n

Momenty centralne: $c = E[X]$

Wariancja: $V(X) = \sigma_x^2 = \sigma^2 = E[(X-\mu)^2] = E[X^2] - (E[X])^2$, gdzie $\mu = E[X]$

$V(aX + b) = V(aX) = a^2 V(X)$

Odchylenie standardowe: $\sigma = \sqrt{V(X)}$

Zmienna X jest standaryzowana jeśli $E[X] = 0$ i $V(X) = 1$

Standaryzacja: $X^* = \frac{X-\mu}{\sigma}$

Współczynnik skośności: $\gamma_1 = E\left[\left(\frac{X-\mu}{\sigma}\right)\right] = \frac{E[(X-\mu)^3]}{(E[(X-\mu)^2])^{3/2}}$

Kurtoza: $\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$

Kwantyle, kwartyle

Kwantyl: $\forall_{p\in(0,1)} x_p = Q(p) = F^{-1}(p) = \inf\{x \in R : p \leqslant F(x)\}$

Mediana: kwantyl $x_0.5$ rzędu 0.5.

Dolny kwartyl Q_1 - kwantyl rzędu 0.25, górny kwartyl Q_3 - kwantyl rzędu 0.75

Rozkład międzykwartylowy $IQR = Q_3 - Q_1$

Doświadczenie Bernoulliego

Doświadczenie kończące się sukcesem z prawdopodobieństwem p lub porażką z prawdopodobieństwem 1-p.

Ciąg n doświadczeń z prawd. sukcesu p oznaczamy $b(n, p)$.

Prawdopodobieństwo uzyskania ciągu składającego się z k sukcesów, przy założeniu niezależności: $p^k(1 - p)^{n-k}$.

Prawdopodobieństwo uzyskania k sukcesów w n niezależnych doświadczeniach z $p \in [0, 1]$: $b(k; n, p) = \binom{n}{k} p^k (1 - p)^{n-k}$.

Poisson (fr. Ryba)

Dla $n \geqslant 25$ i $\lambda = n \cdot p \leqslant 10$ możemy przybliżyć rozkładem Poißona: $b(k; n, p) \approx e^{-np} \frac{(np)^k}{k!}$, na przykład: $\sum_{k=0}^{14} b(k; 500, 0.02) \approx F(14; 500 \cdot 0.02)$, gdzie F jest dystrybuantą rozkładu Ryby, dostępna w tablicach.

Rozkłady zmiennych dyskretnych

Bernoulli

1. Eksperyment składa się z n mniejszych doświadczeń, zwanych próbami, gdzie n jest ustalone i znane przed doświadczeniem. 2. Każda próba kończy się sukcesem lub porażką. 3. Próby są niezależne.

X ma rozkład Bernoulliego jeśli

$$p(x) = P(X = x) = \begin{cases} 1 - p, x = 0 \\ p, x = 1 \end{cases}$$

$E[X] = p$, $V(X) = p(1 - p)$

Dwumianowy, czyli dalej Bernoulli

Zmienną losową X równą liczbie sukcesów w n niezależnych doświadczeniach Bernoulliego nazywamy zmienną o rozkładzie dwumianowym z param. (n, p), oznaczamy $X \sim b(n, p)$.

$P(X = k) = b(k; n, p) = \binom{n}{k} p^k (1 - p)^{n-k}$

$E[X] = np$, $V(X) = np(1 - p)$

Hipergeometryczny

1. Losujemy ze zbioru N elementów. 2. Każdemu z N obiektów można przypisać sukces lub porażka (mamy M sukcesów). 3. Wybieramy n obiektów bez zwracania tak, że wybór każdego n elementowego podzbiotu ma to samo prawdopodobieństwo.

Liczba sukcesów X ma rozkład hipergeometryczny $X \sim HG(n, M, N)$, $P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$, $E[X] =$

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

Ujemny dwumianowy (Pascala)

1. Niezależne próby Bernoulliego. 2. Eksperyment kończy się w chwili uzyskania r-tego sukcesu, r ustalone.

X - Liczba porażek do uzyskania r-tego sukcesu ma rozkład ujemny dwumianowy $X \sim NB(r,p)$, $nb(x;r,p) = \binom{x+r-1}{r-1}p^r(1-p)^x$, $E[X] = \frac{r(1-p)}{p}$, $V(X)]\frac{r(1-p)}{p^2}$

Geometryczny

X równa liczbie porażek w ciągu niezależnych doświadczeń Bernoulliego, do uzyskania pierwszego sukcesu, ma rozkład geometryczny.

Dla $Y = X+1$, $E[Y] = \frac{1}{p}$, $V(Y) = \frac{1-p}{p^2}$

Poisson, The Sequel

1. Istnieje $\lambda > 0$, że w dowolnie krótkim przedziale czasowym Δt prawdop. zaobserwowania dokładnie jednego zdarzenia wynosi $\lambda \cdot \Delta t + \circ(\Delta t)$. 2.Prawdop. zaobserwowania w Δt więcej niż jednego zdarzenia wynosi $\circ(\Delta t)$. 3. Liczba zaobserwowanych zd.los. w Δt jest niezależna od liczby wcześniej zaobserwowanych zdarzeń.

Jeśli 1-3 spełnione, to $P_k(t)$, że liczba zdarzeń zaobserwowanych do chwili t jest równa k wynosi $P_k(t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$, $E[X] = \lambda = V(X)$

Rozkłady zmiennych ciągłych

Jednostajny

X ma rozkład jednostajny na $[a,b]$, $X \sim U(a,b)$ jeśli $f(x) = \frac{1}{b-a}$ dla $x \in [a,b]$ i 0 dla pozostałych. Dla $x \in [a,b)$, $F(x) = \frac{x-a}{b-a}$.

$E[X] = \frac{a+b}{2}$, $V(X) = \frac{(b-a)^2}{12}$

Normalny

$X \sim N(\mu,\sigma^2)$, $f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$

$F(x) = P(X \leqslant x) = \Phi(\frac{x-\mu}{\sigma})$, gdzie Φ patrz poniżej, $E[X] = \mu$, $V(X) = \sigma^2$

Jeśli X ma normalny, to $Y = aX+b \sim N(a\mu+b, a^2\sigma^2)$

Standardowy normalny

$U \sim N(0,1)$, gęstość: $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$, $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\mathrm{d}t$.

Wykładniczy

Własność braku pamięci: $P(X > s+t | X > t) = P(X > s)$

$X \sim \exp(\lambda)$, $f(x) = \lambda e^{-\lambda x}$, $x \geqslant 0$ i 0 gdy $x < 0$. $F(x) = 1 - e^{-\lambda x}$, $x \geqslant 0$.

Gamma

Gamma Eulera: $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x$, $\alpha > 0$.

Własności: 1. $\Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha)$ 2. $\forall_{n\in N}\Gamma(n) = (n-1)!$ 3. $\forall_{x\in(0,1)}\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin(x\pi)}$

$X \sim \gamma(\alpha,\beta)$, $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$ dla $x \geqslant 0$ i 0 w p.p.

$aX \sim \gamma(\alpha, \frac{\beta}{a})$, $E[X] = \frac{\alpha}{\beta}$, $V(X) = \frac{\alpha}{\beta^2}$

Zmienne losowe dwuwymiarowe

dystrybuanta: $F(x,y) = P(X \leqslant x, Y \leqslant y)$

Własności: 1. $\lim_{x\to-\infty}F(x,y) = 0$, $\lim_{y\to-\infty}F(x,y) = 0$, $\lim_{x,y\to\infty}F(x,y) = 1$. 2. Niemalejąca ze względu na każdą ze zmiennych. 3. Prawostronnie ciągła ze względu na każdą ze zmiennych. 4. $F(x_1,y_1)+F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1) \geqslant 0$

Dyskretne

funkcja prawdop. $p(x,y) = P(X=x, Y=y)$

Własności: 1. $p(x,y) \geqslant 0$ 2. $\sum_x\sum_y p(x,y) = 1$

Rozkłady brzegowe X: $p_X(x) = P(X=x) = \sum_{y\in Y}p(x,y)$, y analogicznie

Zmienne X i Y są niezależne, gdy $p(x,y) = p_X(x)\cdot p_Y(y)$ **Ciągłe**

f gęstość, jeśli: $P\{(X,Y) \in A\} = \iint_A f(x,y)\mathrm{d}xy$

$f(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\mathrm{d}u\mathrm{d}v$

$f(x,y) \geqslant 0$, $\iint_{R^2} f(x,y)\mathrm{d}x\mathrm{d}y = 1$

Gęstości brzegowe X: $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y$, y analogicznie

Zmienne X i Y są niezależne, gdy $F(x,y) = F_X(x)\cdot F_Y(y)$, co jest równoważne $f(x,y) = f_X(x)\cdot f_Y(y)$

Funkcje zmiennych losowych

Dyskretne:

Jeśli $X_1,\ldots,X_n$ - niezależnie zmienne o jednakowych rozkładach Bernoulliego, to $S_n = \sum_{i=1}^n X_i \sim b(n,p)$

Jeśli $X_i \sim b(n_i,p)$, to $\sum_{i=1}^n \sim b(\sum_{i=1}^n n_i, p)$.

Jeśli $X_i \sim P(\lambda_i)$, to $\sim_{i=1}^n X_i \sim P(\sum_{i=1}^n \lambda_i)$

Ciągłe:

Niech $Z = g(X,Y)$. Dystrybuanta $F_Z(z) = P(Z \leqslant z) = \iint_{\{(x,y):g(x,y)\leqslant z\}} f(x,y)\mathrm{d}x\mathrm{d}y$

Gęstość sumy $Z = X+Y$: $f_Z(z) = \int_{-\infty}^{\infty} f(x, z-x)\mathrm{d}x$, gdy X i Y niezależne, to $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)\mathrm{d}x$

Jeśli $X_i \sim N(\mu_i,\sigma_i^2)$ (niezależne), to $\sum_{i=1}^n \sim N(\sum_{i=1}^n, \sum_{i=1}^n \sigma_i^2)$

Jeśli $X_i \sim \gamma(\alpha_i,\beta)$, to $\sum_{i=1}^n X_i \sim \gamma(\sum_i \alpha_i, \beta)$

$E[h(X,Y)] = \sum_x\sum_y h(x,y)p(x,y)$ dla dyskretnych, $= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f(x,y)dxdy$

$E[aX+bY] = aE[X]+bE[Y]$, $V(aX+bY) = a^2V(X)+b^2V(Y)+2ab\mathrm{Cov}(X,Y)$

$\mathrm{Cov}(X,Y) = E(X,Y) - E(X)E(Y)$

Współczynnik korelacji: $\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_x\cdot\sigma_y}$

Regresja liniowa

Jeśli (X,Y) dwuwymiarowa zmienna losowa, dla której istnieją σ_X^2, σ_Y^2 oraz Cov, to $E\{[Y-(aX+b)]^2\}$ osiąga minimum gdy $a = \rho_{XY}\frac{\sigma_Y}{\sigma_X}$ oraz $b = m_Y - am_X$ tzn. funkcja regresji zmiennej Y przy $X = x$ jest równa $\frac{y-m_Y}{\sigma_Y} = \rho_{XY}\frac{x-m_X}{\sigma_X}$