Prawdopodobieństwo warunkowe

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \tag{1}$$

$$P(E_1 \cap \cdots \cap E_n) = \prod_{i=1}^{n} P(X_i|X_1, \ldots, X_{i-1}) \tag{2}$$

$$P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i) \text{ dla } \bigcap_{i=1}^{n} F_i = \Omega \tag{3}$$

	zmienne dyskretne	zmienne ciągłe
definicja	$P(x) = P(X = x)$	$f(x) = F'(x)$
obliczanie prawdopodobieństwa	$P(X \in A) = \sum_{x \in A} P(x)$	$P(X \in A) = \int_A f(x)dx$
skumulowana f. rozkładu	$F(x) = P(X \leqslant x) = \sum_{y \leqslant x} P(y)$	$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(y)dy$
całkowite prawdopodobieństwo	$\sum_x P(x) = 1$	$\int_{-\infty}^{\infty} f(x)dx = 1$
wartość oczekiwana	$EX = \sum_x xP(x)$	$EX = \int xf(x)dx$
wariancja	$VarX = \sigma^2 = E[(X - \mu)^2]$	$VarX = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx$

Nierówność Czebyszewa

$$P(|X - \mu| > \epsilon) \leqslant (\frac{\sigma}{\epsilon})^2 \tag{4}$$

Rozkłady dyskretne

Rozkład	P(x)	EX	VarX	
Bernoulli(p)	$P(x) = \begin{cases} p, & \text{for } x = 1 \\ q = (1-p), & \text{for } x = 0 \end{cases}$	$EX = p$	$VarX = pq$	próba
Binomial(n, p)	$P(x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, \ldots$	$EX = np$	$VarX = npq$	liczba sukcesów z n prób
Geometric(p)	$P(x) = (1-p)^{x-1}p$ for $x = 1, 2, \ldots$ $P(X > k) = (1-p)^k$	$EX = \frac{1}{p}$	$VarX = \frac{1-p}{p^2}$	liczba prób do sukcesu
Poiss(λ)	$P(x) = e^{-\lambda} \frac{\lambda^x}{x!}$ for $x = 0, 1, \ldots$	$EX = \lambda$	$VarX = \lambda$	rozkład zdarzeń rzadkich

Rozkłady ciągłe

Rozkład	f(x), F(x)	EX	VarX	
Unif(a,b)	$f(x) = \frac{1}{b-a}$ for $a \leqslant x \leqslant b$ $F(x) = \begin{cases} 0, & \text{for } x < a \\ \frac{x-a}{b-a}, & \text{for } a \leqslant x < b \\ 1, & \text{for } x \geqslant b \end{cases}$	$EX = \frac{a+b}{2}$	$VarX = \frac{(b-a)^2}{12}$	
Exp(λ)	$f(x) = \lambda e^{-\lambda x}$ for $x \geqslant 0$ $F(x) = 1 - e^{-\lambda x}$	$EX = \frac{1}{\lambda}$	$VarX = \frac{1}{\lambda^2}$	modelowanie czasu, brak pamięci
Gamma(α, λ)	$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$ $F(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-\lambda t} dt$	$EX = \frac{\alpha}{\lambda}$	$VarX = \frac{\alpha}{\lambda^2}$	łączny czas α niezależnych zdarzeń $\sim Exp(\lambda)$
N(μ, σ)	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$ $F(x) = \Phi(x)$ dla N(0,1)	$EX = \mu$	$VarX = \sigma^2$	

$$Bin(n,p) \approx Poiss(\lambda) \tag{5}$$

$$P(T \leqslant t) = P(X \geqslant \alpha) \tag{6}$$
$$T \sim Gamma(\alpha, \lambda), X \sim Poiss(\lambda t)$$

$$Binomial(n,p) = N(np, \sqrt{np(1-p})) \tag{7}$$
$$X_i \sim Bernoulli(p), S_n = \sum_{i=1}^{n} X_i, 0.05 \leqslant p \leqslant -.095$$

Rozkład łączny

$$F_{(X,Y)}(x,y) = P(X \leqslant x \cap Y \leqslant y)$$
$$f_{(X,Y)}(x,y) = \frac{\delta^2}{\delta x \delta y} F_{(X,Y)}(x,y)$$

	zmienne dyskretne	zmienne ciągłe
rozkłady brzegowe	$P(x) = \sum_y P(x,y)$ $P(y) = \sum_x P(x,y)$	$f(x) = \int_Y f(x,y)dy$ $f(y) = \int_X f(x,y)dx$
niezależność	$P(x,y) = P(x)P(y)$	$f(x,y) = f(x)f(y)$
obliczanie prawdopodobieństwa	$P((X,Y) \in A) = \sum_{(x,y)\in A} P(x,y)$	$\int\int_{(x,y)\in A} f(x,y)dxdy$

Centralne Twierdzenie Graniczne

$$Z_n = \frac{S_n - E(S_n)}{Std(S_n)} = \frac{S_n - n\mu}{\sqrt{(n)}\sigma} \to N(0,1) \text{ for } n \text{ to } \infty \tag{8}$$

$$S_n = \sum_{i=1}^{n} X_i, E(X_i) = \mu, Std(X_i) = \sigma$$

Estymatory - Monte Carlo

dla X, $p = P(X \in A)$

$$\hat{p} = \hat{P}(X \in A) = \frac{\#(X_i \in A)}{n}$$

$$E\hat{p} = \frac{1}{n}(np) = p$$

$$Std\hat{p} = \frac{1}{n}\sqrt{np(1-p)} = \sqrt{(}\frac{p(1-p)}{n})$$

Dokładność

$$P(|\hat{p} - p| > \epsilon) = P(\frac{|\hat{p} - p|}{\sqrt{\frac{p(1-p)}{n}}} > \frac{\epsilon}{\sqrt{\frac{p(1-p)}{n}}}) \approx 2\Phi(\frac{-\epsilon\sqrt{n}}{\sqrt{p(1-p)}})$$

Estymacja średniej $\bar{X}$ z $X_1, \ldots, X_n$ ze wspólnym μ i σ

$$E\bar{X} = \frac{1}{n}(EX_1 + \cdots + EX_n) = \frac{1}{n}n\mu = \mu$$

$$Var\bar{X} = \frac{1}{n^2}(Var(X_1) + \cdots + Var(X_n)) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

Estymator wariancji

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Procesy Markowa

$$P = \begin{bmatrix} p_{11} & p_{12} & \ldots & p_{1n} \\ p_{21} & p_{22} & \ldots & p_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ p_{n1} & p_{n2} & \ldots & p_{nn} \end{bmatrix}$$

Rozkład w czasie h: $P_h = P_0 * P^h$

Rozkład stacjonarny: $\pi P = \pi$, $\sum \pi_i = 1$

Estymacja: **metoda momentów**

k-ty moment z populacji	$\mu_k = E(X^k)$
k-ty centralny moment z populacji	$\mu'_k = E((X - \mu_1)^k)$
k-ty moment z próby	$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$
k-ty centralny moment z próby	$m'_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k$

$\mu_1 = EX, \mu_2 = VarX, \mu_k = m_k$

Estymacja: **metoda największej wiarygodności**

$$P(X = (X_1, \ldots, X_n)) = P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i)$$

Szukamy ekstremum: $\frac{\delta P}{\delta \theta}(x) = 0$, używając logarytmu.

Przedziały ufności

$$\hat{\theta} \pm z_{\frac{\alpha}{2}}\sigma(\hat{\theta})$$

$$\bar{X} \pm z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$\bar{X} - \bar{Y} = z_{\frac{\alpha}{2}} \pm \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$$

Rozkład t-studenta

$t = \frac{\hat{\theta}=\theta}{s(\hat{\theta})} \leftarrow$ zastępujemy $Std(\hat{\theta})$ przez $S(\hat{\theta})$, n-1 stopni swobody

$$\bar{X} \pm t_{\frac{\alpha}{2}}^{(n-1)} \frac{S}{\sqrt{n}}$$