DWUPUNKTOWY	$B(p)$	$: \{0,1\}$	$P(X=1)=p$	$EX=p$	$D^2X=p(1-p)$
DWUMIANOWY	$B(n,p)$	$: \{0,\dots,n\}$	$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$	$EX=np$	$D^2X=np(1-p)$
JEDNOSTAJNY	$Unif[a,b]$	$: [a,b]$	$f(x)=\frac{1}{b-a}\cdot \mathbb{1}\{x\in[a,b]\}$	$EX=(a+b)/2$	$D^2X=(b-a)^2/12$
NORMALNY	$N(\mu,\sigma^2)$	$: \mathbb{R}$	$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{\frac{-(x-\mu)^2}{2\sigma^2}}$	$EX=\mu$	$D^2X=\sigma^2$

$\Theta(x)=\begin{cases}1, & x\geq 0\\ 0, & x<0\end{cases}$ $\quad (\log_b x)'=\frac{1}{x\ln b}$

$\mathbb{1}\{x\}=\begin{cases}1, & x\\ 0, & wpw\end{cases}$ $\quad (a^x)'=a^x\ln a$

$\log x^y = y\log x$ $\quad (e^{-x})'=-e^{-x}$

$\log_a a = 1$ $\quad f(g(x))'=f'(g(x))g'(x)$

$\log 1 = 0$ $\quad \log(x\cdot y)=\log x+\log y$

$a^{\log_a b}=b$ $\quad \log(\frac{x}{y})=\log x-\log y$

$\underbrace{P(X=x)}_{\text{R. BRZEGOWY}}=\sum_y P(X=x,Y=y)$ $\quad \left(f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dx\right)$

$\underbrace{P(X=x,Y=y)}_{\text{R. ŁĄCZNY}}=\underbrace{P(X=x|Y=y)}_{\text{R. WARUNKOWY}}P(Y=y)$

BAYES :) $P(Y=y|X=x)=\frac{P(X=x|Y=y)P(Y=y)}{P(X=x)}$

$EX=\sum_x xP(X=x)$

$E[g(X)]=\sum_x g(x)P(X=x)$

$E[X|Y=y]=\sum_x xP(X=x|Y=y)$

$E[E[X|Y]]=\sum_y E[X|Y=y]P(Y=y)$

$\cancel{E[EX]}$ $E[E[X|Y]]=EX$

$\cancel{E[aX]=aEX}$ $E[aX+Y]=aEX+EY$

$D^2[aX+b]=a^2D^2X$

$D^2X=E[(X-EX)^2]$

$D^2X=EX^2-(EX)^2$

$E[(X-a)^2]=E[(X-EX)^2]+(EX-a)^2$

DLA NIEZAL. X i Y

$P(Y=y|X=x)=P(Y=y)$ $\quad E[XY]=EX\cdot EY$

$P(X=x,Y=y)=P(X=x)P(Y=y)$ $\quad D^2[X+Y]=D^2X+D^2Y$

$\hat{y}=h(x)$ $\quad \underbrace{\ell(y,\hat{y})}_{\text{f-a STRATY}}$ $\quad \underbrace{L(h)}_{\text{RYZYKO (śr. wart. straty na populacji)}}$ $\quad \underbrace{\Delta L(h)}_{\text{NADWYŻKA RYZYKA (ŻAL)}}=L(h)-L^*(h)\geq 0$

$\underbrace{L(\hat{y}|x)}_{\text{RYZYKO WARUNKOWE}}=E[\ell(Y,\hat{y})|X=x]=\sum_{y\in Y}\ell(y,\hat{y})P(y|x)$

$L(h)=E[\ell(Y,\hat{y})]=E[L(h(X)|X)]=\sum_{x\in X}L(h(x)|x)P(x)$

$h^*=\underset{h}{\operatorname{argmin}}\,L(h)$ $\quad \underbrace{L^*=L(h^*)}_{\text{RYZYKO BAYESOWSKIE}}$

$h^*(x)=\underset{\hat{y}\in Y}{\operatorname{argmin}}\,L(\hat{y}|x)$

$\Delta L(h)=L(h)-L(h^*)=\underbrace{L(h)-L(h^*_H)}_{\text{B. ESTYMACJI}}+\underbrace{L(h^*_H)-L(h^*)}_{\text{B. APROKSYM.}}$

KLASYF. BINARNA $\ell(y,\hat{y})=\mathbb{1}\{y\neq\hat{y}\}$

$L(h(x)|x)=\ell(0,h(x))P(y=0|x)+\ell(1,h(x))P(y=1|x)=\begin{cases}P(y=1|x), & h(x)=0\\ P(y=0|x), & h(x)=1\end{cases}$

$L(h)=E[\mathbb{1}\{Y\neq h(X)\}]=\sum_{x\in X,\,y\in\{0,1\},\,y\neq h(x)}P(x,y)=P(h(X)\neq Y)$

$h^*(x)=\begin{cases}1, & P(Y=1|x)>P(Y=0|x) \\ 0, & P(Y=1|x)<P(Y=0|x) \\ \text{DOWOLNIE}, & P(Y=1|x)=P(Y=0|x)\end{cases}$ $\quad \begin{matrix}L(1|x)<L(0|x)\\ L(1|x)>L(0|x)\\ \Leftrightarrow L(1|x)=L(0|x)\end{matrix}$

ALTERNATYWNIE

$\eta(x)=P(Y=1|X)$ $\quad 1-\eta(x)=P(Y=0|X)$

$L(h(x)|x)=\begin{cases}\eta(x), & \eta(x)\geq 1/2\\ 1-\eta(x), & \eta(x)<1/2\end{cases}=\min\{\eta(x),1-\eta(x)\}$

DLA $a\in[0,1]$ $\min\{a,1-a\}=\frac{1}{2}-|a-\frac{1}{2}|$

$L(h)=\sum_x \min\{\eta(x),1-\eta(x)\}P(x)=E[\min\{\eta(x),1-\eta(x)\}]$

$h^*(x)=\begin{cases}1, & \eta(x)>\frac{1}{2}\\ 0, & \eta(x)<\frac{1}{2}\\ \text{ind}, & \eta(x)=1/2\end{cases}$ $\quad \Delta L(h)=E[\mathbb{1}\{h(X)\neq h^*(X)\}|2\eta(X)-1|]$

BEZWZGLĘDNY

$\ell(y,\hat{y})=|y-\hat{y}|$

$h^*(x)=med[Y|X=x]$

KWADRATOWY

$\ell(y,\hat{y})=(y-\hat{y})^2$

$h^*(x)=E[Y|X=x]=\mu$

$\Delta L(h)=E[(h^*(X)-h(X))^2]$

LOGISTYCZNY

$f^*(x)=\ln\frac{\eta(x)}{1-\eta(x)}$

WYKŁADNICZY

$\ell(y,\hat{y})=e^{-\bar{y}\hat{y}}$

$\bar{y}=2y-1\in\{-1,1\}$

LOGARYTMICZNA

$l(k,\hat{y}) = -\ln \hat{y}_k \quad \hat{y} = (\hat{y}_1, \dots, \hat{y}_k) \in \Delta^k$

$\rightarrow$ BINARNA $\quad l(y,\hat{y}) = \begin{cases} -\ln \hat{y}, & y=1 \\ -\ln(1-\hat{y}), & y=0 \end{cases}$

$L(\hat{y}|x) = -\sum_{k=1}^{k} \eta_k(x) \ln \hat{y}_k$

$\eta(x) = (\eta_1(x), \dots, \eta_k(x))$

$h^*(x) = E[Y|X=x] \xrightarrow{\text{BINARNA}} \eta(x)$

LOGISTYCZNA

$\eta(x) = \sigma(f^*(x)) = \frac{1}{1+e^{-f^*(x)}} \Rightarrow$

$\Rightarrow f^*(x) = \ln \frac{\eta(x)}{1-\eta(x)}$

KLASYF. Z KOSZTAMI

$l(y,\hat{y}) = \begin{cases} 0, & y=\hat{y} \\ c_0, & y=0 \wedge \hat{y}=1 \\ c_1, & y=1 \wedge \hat{y}=0 \end{cases}$

$h^*(x) = \begin{cases} 1 & \eta(x) \geq \frac{c_0}{c_0+c_1} \\ 0 & \eta(x) < \frac{c_0}{c_0+c_1} \end{cases} = H\left(\eta(x) - \frac{c_0}{c_0+c_1}\right)$

ZAWIASOWY

$l(y,\hat{y}) = \max(0, 1-\tilde{y}\hat{y}), \quad \tilde{y} = 2y-1$

$L(\hat{y}|x) = \begin{cases} (1-\eta(x))(1+\hat{y}), & \hat{y}>1 \\ 1+(1-2\eta(x))\hat{y}, & \hat{y} \in [-1,1] \\ \eta(x)(1-\hat{y}), & \hat{y}<-1 \end{cases}$

$f^*(x) = \begin{cases} 1, & \eta(x) \geq 1/2 \\ -1, & \eta(x) < 1/2 \end{cases} = \operatorname{sgn}(\eta(x) - 1/2)$

STRATEGIA

ZNAJDŹ h^*/L DLA KLASYF

1. WYZNACZ $P(y|x)$ i podstaw do h^*
2. WYZNACZ $\eta(0), \eta(1), h^*(0), h^*(1), h^*(x), L(h(x)|x), L(h)$
3. RYZYKO WARUNKOWE ZE WZORU $L(h(x)|x)$, PODSTAWIĆ DO $L(1|x) < L(0|x)$

WYMIAR VC

1. POKAŻ, ŻE ISTNIEJE JAKIKOLWIEK ZBIÓR $\{x_1, \dots x_n\}$, GDZIE DA SIĘ POETYKIETOWAĆ DANE NA WSZYSTKIE 2^n SPOSOBÓW $\longrightarrow V_n \geq n$
2. ŻADEN ZBIÓR O ROZM. $n+1$ NIE MOŻE ZOSTAĆ PRZEZ H POETYKIETOWANY NA WSZYSTKIE 2^{n+1} SPOSOBÓW $(V_n \leq n)$
3. WYNIKA, ŻE $V_n = n$

DLA $\hat{h}$ ZERM

$\underbrace{L(\hat{h}) - \min_{h \in H} L(h)}_{\text{BŁĄD ESTYMACJI}} = O\left(\sqrt{\frac{\ln|H|}{n}}\right)$

DLA EXPON. WEIGHTS

$\frac{\hat{L}_n - \min_{h \in H} L_n(h)}{n} = O\left(\sqrt{\frac{\ln|H|}{n}}\right)$

N. MARKOVA

$P(x \geq a) \leq \frac{Ex}{a}$

N. CZEBYSZEWA

$P(|X - EX| \geq \varepsilon) \leq \frac{D^2 X}{\varepsilon^2}$