UNIVERSIDADE NOVA DE LISBOA
Faculdade de Ciências e Tecnologia
Departamento de Informática

MIEI
EPGCI

Information Theory

Marginal Probability

$$P(x) = \sum_{y \in A_Y} P(x,y)$$

Product Rule

$$P(x,y) = p(x \mid y) p(y)$$

$$P(x,y) = p(y \mid x) p(x)$$

Bayes Theorem

$$P(y \mid x, \boldsymbol{H}) = \frac{P(x \mid y, \boldsymbol{H}) P(y \mid \boldsymbol{H})}{P(x \mid \boldsymbol{H})}$$

$$P(y \mid x, \boldsymbol{H}) = \frac{P(x \mid y, \boldsymbol{H}) P(y \mid \boldsymbol{H})}{\sum_{y'} P(x \mid y', \boldsymbol{H}) P(y' \mid \boldsymbol{H})}$$

Binomial Distribution

Lets f be the probability of one outcome of a random experiment. Let r be a random variable that represents the number of times the outcome occurs in N independent experiments

$$P(r \mid f, N) = \binom{N}{r} f^r (1-f)^{N-r} \qquad E(r) = Nf \qquad Var(r) = Nf(1-f)$$

Poisson Distribution

Is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time

$$P(r \mid \lambda) = e^{-\lambda} \frac{\lambda^r}{r!} \qquad E(r) = \lambda \qquad Var(r) = \lambda$$

For large values of λ the Poisson distribution is well approximated, for values of r around λ, by a Gaussian distribution with mean λ and variance λ

The probability density of the normal distribution is

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Stirling Approximations

Factorial

$$x! \approx x^x e^{-x} \sqrt{2\pi x}$$

$$\ln x! \approx x \ln x - x + \tfrac{1}{2} \ln 2\pi x$$

$$\binom{N}{r} \approx 2^{N H_2(r/N)}$$

$$\ln \binom{N}{r} \approx N H_2(r/N)$$

$$\ln \binom{N}{r} \approx N H_2(r/N) - \tfrac{1}{2} \log \left[2\pi N \frac{N-r}{N} \frac{r}{N} \right]$$

Logarithm rules

Rule name	Rule
Logarithm product rule	$\log_b(x \cdot y) = \log_b(x) + \log_b(y)$
Logarithm quotient rule	$\log_b(x / y) = \log_b(x) - \log_b(y)$
Logarithm power rule	$\log_b(x^y) = y \cdot \log_b(x)$
Logarithm base switch rule	$\log_b(c) = 1 / \log_c(b)$
Logarithm base change rule	$\log_b(x) = \log_c(x) / \log_c(b)$

Hamming Code H(7, 4)

$$t_5 = s_1 \oplus s_2 \oplus s_3$$
$$t_6 = s_2 \oplus s_3 \oplus s_4$$
$$t_7 = s_1 \oplus s_3 \oplus s_4$$

Entropy

$$H(x) = \sum_{x \in A_X} P(x) \log_2 \frac{1}{P(x)} = -\sum_{x \in A_X} P(x) \log_2 P(x)$$

$$H(\boldsymbol{p}) = H(p_1, 1 - p_1) + (1 - p_1) H\left(\frac{p_2}{1-p_1}, \frac{p_3}{1-p_1}, \ldots, \frac{p_I}{1-p_1}\right)$$

Gibs inequality

$$D_{KL}(P \| Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$

$$D_{KL}(P \| Q) \geq 0$$

$$D_{KL}(P \| Q) = 0 \quad \text{only if} \quad P = Q$$

Jensen inequality

$$\varepsilon[f(x)] \geq f(\varepsilon[x])$$

Raw bit Content

$$H_0(X) = \log_2 |A_X|$$

Kraft inequality

$$\sum_{i=1}^{I} 2^{-l_i} \leq 1 \quad , I = |A_X|$$

Essential bit content of X

$$H_\partial(X) = \log_2 |S_\partial|$$

Expected codewords length

$$L(C, X) = H(X) + D_{KL}(\boldsymbol{p} \| \boldsymbol{q})$$