

Formulary for Statistical Data Analysis

https://github.com/Grufoony/Physics_Unibo

Contents

1 Probability theory	5
1.1 Combinatorics	5
1.2 Probability	5
1.3 Random variables and distributions	6
1.4 Important distributions	7
1.4.1 Binomial distribution	7
1.4.2 Poisson distribution	8
1.4.3 Uniform distribution	8
1.4.4 Exponential distribution	8
1.4.5 Gaussian distribution	9
1.4.6 Multivariate Gaussian distribution	9
1.4.7 Chi-square distribution	9
1.4.8 Cauchy (Breit-Wigner) distribution	9
1.4.9 Student's t distribution	10

Chapter 1

Probability theory

1.1 Combinatorics

Permutations without repetitions:

$$P_n = n! \quad (1.1)$$

Permutations with repetitions:

$$P_n^r = \frac{n!}{\prod k_i!} \quad (1.2)$$

Dispositions without repetitions:

$$D_{n,k} = \frac{n!}{(n-k)!} \quad (1.3)$$

Dispositions with repetitions:

$$D_{n,k}^r = n^k \quad (1.4)$$

Combinations without repetitions:

$$C_{n,k} = \binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (1.5)$$

Combinations with repetitions:

$$C_{n,k}^r = \binom{n+k-1}{k} \quad (1.6)$$

1.2 Probability

Conditional probability:

$$P(A|B) = \frac{P(A \cap B)}{p(B)} \quad (1.7)$$

Probability of intersection for independent events:

$$P(A \cap B) = P(A)P(B) \quad (1.8)$$

Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1.9)$$

Law of total probability:

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_i P(B|A_i)P(A_i)} \quad (1.10)$$

Bayes' theorem in Bayesian thinking:

$$P(H|\vec{x}) = \frac{P(\vec{x}|H)\pi(H)}{\int P(\vec{x}|H)\pi(H)dH} \quad (1.11)$$

1.3 Random variables and distributions

Marginal pdf:

$$f_i(x_i) = \int f(\vec{x}) \prod_j dx_j \quad (1.12)$$

Conditional pdfs:

$$f(y|x) = \frac{f(x,y)}{f_x(x)} f(x|y) = \frac{f(x,y)}{f_y(y)} \quad (1.13)$$

Bayes' theorem for distributions

$$f(x|y) = \frac{f(y|x)f_x(x)}{f_y(y)} \quad (1.14)$$

Condition for independent variables:

$$f(x,y) = f_x(x)f_y(y) \quad (1.15)$$

Distribution of a function of a random variable in 1-D:

$$g(a) = f(x(a)) \left| \frac{dx}{da} \right| \quad (1.16)$$

Distribution of a function of a random variable in N-D:

$$g(\vec{y}) = |J|f(\vec{x}) \quad (1.17)$$

Expectation value:

$$E[x] = \int xf(x)dx = \mu_x \quad (1.18)$$

Variance:

$$V[x] = E[x^2] - E[x]^2 \quad (1.19)$$

Covariance:

$$cov[x, y] = E[xy] - E[x]E[y] = E[(x - \mu_x)(y - \mu_y)] \quad (1.20)$$

Correlation coefficient:

$$\rho_{xy} = \frac{cov[xy]}{\sigma_x \sigma_y} \quad (1.21)$$

Correlation matrix:

$$V_{ij} = cov[x_i, x_j] \quad (1.22)$$

Variance of a function of random variables:

$$\sigma_y^2 \approx \sum_{i,j} \left[\frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j} \right]_{\vec{x}=\vec{\mu}} V_{ij} \quad (1.23)$$

Variance of a vector function of random variables:

$$U_{kl} \approx \sum_{i,j} \left[\frac{\partial y_k}{\partial x_i} \frac{\partial y_k}{\partial x_j} \right]_{\vec{x}=\vec{\mu}} V_{ij} \quad (1.24)$$

Error propagation for sum of uncorrelated variables:

$$\sigma_y^2 = \sigma_1^2 + \sigma_2^2 + 2cov[x_1, x_2] \quad (1.25)$$

Error propagation for product of uncorrelated variables:

$$\frac{\sigma_y^2}{y^2} = \frac{\sigma_1^2}{x_1^2} + \frac{\sigma_2^2}{x_2^2} + 2 \frac{cov[x_1, x_2]}{x_1 x_2} \quad (1.26)$$

Characteristic function:

$$\phi_x(k) = E[e^{ikx}] = \int_{-\infty}^{\infty} e^{ikx} f(x) dx \quad (1.27)$$

Moments of Characteristic function:

$$\frac{d^m}{dk^m} \phi_z(k) = i^m \mu'_m \quad (1.28)$$

1.4 Important distributions

1.4.1 Binomial distribution

Characteristic function:

$$\phi_p(k) = p \left[(e^{ik} - 1) + 1 \right]^N \quad (1.29)$$

Distribution:

$$f(n; N, p) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} \quad (1.30)$$

Expectation value:

$$E[n] = Np \quad (1.31)$$

Variance:

$$V[n] = Np(1-p) \quad (1.32)$$

1.4.2 Poisson distribution

Characteristic function:

$$\phi_\nu(k) = e^{\nu(e^{ik}-1)} \quad (1.33)$$

Distribution:

$$f(n; \nu) = \frac{\nu^n}{n!} e^\nu \quad (1.34)$$

Expectation value:

$$E[n] = \nu \quad (1.35)$$

Variance:

$$V[n] = \nu \quad (1.36)$$

1.4.3 Uniform distribution

Characteristic function:

$$\phi_{\alpha,\beta}(k) = \frac{e^{i\beta k} - e^{i\alpha k}}{(\beta - \alpha)ik} \quad (1.37)$$

Distribution:

$$f(x; \alpha, \beta) = \frac{1}{\beta - \alpha} \text{ for } \alpha \leq x \leq \beta \quad (1.38)$$

Expectation value:

$$E[x] = \frac{\alpha + \beta}{2} \quad (1.39)$$

Variance:

$$V[x] = \frac{(\beta - \alpha)^2}{12} \quad (1.40)$$

1.4.4 Exponential distribution

Characteristic function:

$$\phi_\xi(k) = \frac{1}{1 - ik\xi} \quad (1.41)$$

Distribution:

$$f(x; \xi) = \frac{1}{\xi} e^{-x/\xi} \text{ for } x \geq 0 \quad (1.42)$$

Expectation value:

$$E[x] = \xi \quad (1.43)$$

Variance:

$$V[x] = \xi^2 \quad (1.44)$$

1.4.5 Gaussian distribution

Characteristic function:

$$\phi_{\mu,\sigma}(k) = e^{i\mu k - \frac{1}{2}\sigma^2 k^2} \quad (1.45)$$

Distribution:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1.46)$$

Expectation value:

$$E[x] = \mu \quad (1.47)$$

Variance:

$$V[x] = \sigma^2 \quad (1.48)$$

1.4.6 Multivariate Gaussian distribution

Distribution:

$$f(\vec{x}; \vec{\mu}, V) = \frac{1}{(2\pi)^{n/2}|V|^{1/2}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu})^t V^{-1} (\vec{x} - \vec{\mu})\right) \quad (1.49)$$

1.4.7 Chi-square distribution

Characteristic function:

$$\phi_n(k) = (1 - 2ik)^{-\frac{n}{2}} \quad (1.50)$$

Distribution:

$$f(z; n) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1} e^{-z/2} \quad (1.51)$$

Expectation value:

$$E[z] = n \quad (1.52)$$

Variance:

$$V[z] = 2n \quad (1.53)$$

1.4.8 Cauchy (Breit-Wigner) distribution

Characteristic function:

$$\phi_{x_0, \Gamma}(k) = e^{-ikx_0 - |k|\frac{\Gamma}{2}} \quad (1.54)$$

Distribution:

$$f(x; \Gamma, x_0) = \frac{1}{\pi} \frac{\Gamma/2}{\Gamma^2/4 + (x - x_0)^2} \quad (1.55)$$

1.4.9 Student's t distribution

Distribution:

$$f(x; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu}\right)^{-(\frac{\nu+1}{2})} \quad (1.56)$$

Expectation value:

$$E[x] = 0 \quad (1.57)$$

Variance:

$$V[x] = \frac{\nu}{\nu - 2} \quad (1.58)$$