PHYS509 Formula Sheet - Theory of Measurements

Tobias Faehndrich

September 2025

1 Lecture 1: Foundations of Probability Theory

1.1 Basics

- $P(A \cap B)$ = probability of both A and B occurring
- $P(A \cup B)$ = probability of either A or B occurring (or both)
- $P(A^c)$ = probability of A not occurring
- Disjoint events: $A \cap B = \emptyset$
- Events: subsets of sample space S
- Sample space: set of all possible outcomes
- Complement: $A^c = S \setminus A$
- Independent events: $P(A \cap B) = P(A)P(B)$

1.2 Kolmogorov's Axioms

For probability measure P on sample space S:

1. $P(S) = 1$
2. $P(\emptyset) = 0$
3. For disjoint events $E_1, E_2, \ldots$:

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

1.3 Key Consequences

$$P(E^c) = 1 - P(E) \tag{1.1}$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{1.2}$$
$$P(B) \leq P(A) \text{ if } B \subset A \tag{1.3}$$

1.4 Uniform Probability

For finite sample space with N equally likely outcomes:

$$P(E) = \frac{|E|}{|S|} = \frac{\text{favorable outcomes}}{\text{total outcomes}}$$

1.5 Conditional Probability

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \quad \text{for } P(F) > 0$$

1.6 Set Operations

$$A \cup B = \{x : x \in A \text{ or } x \in B\} \tag{1.4}$$
$$A \cap B = \{x : x \in A \text{ and } x \in B\} \tag{1.5}$$
$$A^c = \{x \in S : x \notin A\} \tag{1.6}$$

1.7 De Morgan's Laws

$$(A \cup B)^c = A^c \cap B^c \tag{1.7}$$
$$(A \cap B)^c = A^c \cup B^c \tag{1.8}$$

2 Lecture 2: Bayesian Inference and Random Variables

2.1 Bayes' Theorem

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

2.2 Law of Total Probability

For partition $\{A_1, A_2, \ldots, A_n\}$ of S:

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

2.3 Independence

Events A and B are independent if:

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

Equivalently: $P(A|B) = P(A)$ (when $P(B) > 0$)

2.4 Random Variables

A random variable X is a function $X : S \to \mathbb{R}$

2.5 Probability Mass Function (PMF)

For discrete random variable X:

$$p_X(k) = P(X = k)$$

Properties:

$$p_X(k) \geq 0 \text{ for all } k \tag{2.1}$$
$$\sum_k p_X(k) = 1 \tag{2.2}$$

2.6 Cumulative Distribution Function (CDF)

$$F_X(x) = P(X \leq x)$$

3 Lecture 3: Bayesian Reasoning and Probability Distributions

3.1 Continuous Random Variables

3.1.1 Probability Density Function (PDF)

$$f_X(x) \geq 0, \quad \int_{-\infty}^{\infty} f_X(x)dx = 1$$

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$

3.1.2 Relationship to CDF

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

$$f_X(x) = \frac{dF_X(x)}{dx}$$

3.2 Expectation Value

$$\text{Discrete: } E[X] = \sum_k k \cdot p_X(k) \tag{3.1}$$
$$\text{Continuous: } E[X] = \int_{-\infty}^{\infty} x f_X(x)dx \tag{3.2}$$

3.3 Variance

$$\mathrm{Var}(X) = E[(X - E[X])^2] \tag{3.3}$$
$$= E[X^2] - (E[X])^2 \tag{3.4}$$

3.4 Standard Deviation

$$\sigma_X = \sqrt{\mathrm{Var}(X)}$$

3.5 Important Distributions

3.5.1 Uniform Distribution

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

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

3.5.2 Exponential Distribution

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

$$E[X] = \frac{1}{\lambda}, \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

3.5.3 Normal (Gaussian) Distribution

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

$$E[X] = \mu, \quad \mathrm{Var}(X) = \sigma^2$$

3.6 Characteristic Function

$$\varphi_X(t) = E[e^{itX}] = \begin{cases} \int_{-\infty}^{\infty} e^{itx} f_X(x)\, dx & \text{for continuous } X \\ \sum_k e^{itx_k} p_X(x_k) & \text{for discrete } X \end{cases}$$

Properties:

- $\varphi_X(0) = 1$
- For independent X, Y: $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$
- Moments: $E[X^n] = \frac{1}{i^n}\varphi_X^{(n)}(0)$

3.7 Moments

3.7.1 Raw Moments

The n-th raw moment:

$$\mu'_n = E[X^n] = \begin{cases} \sum_k x_k^n p_X(x_k) & \text{discrete} \\ \int_{-\infty}^{\infty} x^n f_X(x)\, dx & \text{continuous} \end{cases}$$

3.7.2 Calculating Moments from Characteristic Function

$$E[X^n] = \frac{\varphi_X^{(n)}(0)}{i^n}$$

3.7.3 Central Moments

The n-th central moment:

$$\mu_n = E[(X-\mu)^n]$$

- $\mu_0 = 1$ (by definition)
- $\mu_1 = 0$ (by definition)
- $\mu_2 = \mathrm{Var}(X) = \sigma^2$
- μ_3 related to skewness
- μ_4 related to kurtosis

3.7.4 Standardized Moments

$$\text{Skewness: } \gamma_1 = \frac{\mu_3}{\sigma^3} = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] \tag{3.5}$$

$$\text{Kurtosis: } \gamma_2 = \frac{\mu_4}{\sigma^4} = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] \tag{3.6}$$

3.7.5 Moment Generating Function (MGF)

$$M_X(t) = E[e^{tX}] = \sum_{n=0}^{\infty} \frac{t^n}{n!} E[X^n]$$

Property: $M_X^{(n)}(0) = E[X^n]$

Note: MGF uses real t, while characteristic function uses imaginary it.

4 Lecture 4: Joint Distributions, Correlations, and Variable Transformations

4.1 Joint Probability

4.1.1 Discrete Case

$$p_{X,Y}(x,y) = P(X=x, Y=y)$$

4.1.2 Continuous Case

$$f_{X,Y}(x,y) \geq 0, \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$$

4.2 Marginal Distributions

$$p_X(x) = \sum_y p_{X,Y}(x,y) \tag{4.1}$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy \tag{4.2}$$

4.3 Independence

Random variables X and Y are independent if:

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

4.4 Convolution Formula

For independent random variables X and Y, the distribution of their sum $Z = X+Y$ is:

$$P(Z=k) = \begin{cases} \sum_{r=0}^{k} P(X=r)\cdot P(Y=k-r) & \text{discrete} \\ \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)\,dx & \text{continuous} \\ \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)\,dy & \text{also continuous} \end{cases}$$

Note: Convolution is much easier using characteristic functions:

$$\varphi_{X+Y}(t) = \varphi_X(t)\cdot\varphi_Y(t)$$

4.5 Covariance

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

4.6 Correlation Coefficient

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

Property: $-1 \leq \rho(X,Y) \leq 1$

4.7 Covariance Matrix

$$V = \begin{pmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(Y,X) & \mathrm{Var}(Y) \end{pmatrix}$$

5 Lecture 5: Propagation of Uncertainty in Measurements

5.1 Linear Transformation

For $z = ax + b$:

$$\sigma_z^2 = a^2\sigma_x^2$$

5.2 General Function of One Variable

For $z = f(x)$ and small σ_x:

$$\sigma_z^2 \approx \left(\frac{df}{dx}\bigg|_{x=\mu_x} \right)^2 \sigma_x^2$$

5.3 Function of Multiple Variables

For $z = f(x, y)$:

$$\sigma_z^2 \approx \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 + 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\mathrm{Cov}(x, y)$$

5.4 Common Cases

$$z = x + y: \quad \sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2\mathrm{Cov}(x, y) \tag{5.1}$$

$$z = x - y: \quad \sigma_z^2 = \sigma_x^2 + \sigma_y^2 - 2\mathrm{Cov}(x, y) \tag{5.2}$$

$$z = xy: \quad \frac{\sigma_z^2}{z^2} = \frac{\sigma_x^2}{x^2} + \frac{\sigma_y^2}{y^2} + \frac{2\mathrm{Cov}(x, y)}{xy} \tag{5.3}$$

$$z = \frac{x}{y}: \quad \frac{\sigma_z^2}{z^2} = \frac{\sigma_x^2}{x^2} + \frac{\sigma_y^2}{y^2} - \frac{2\mathrm{Cov}(x, y)}{xy} \tag{5.4}$$

6 Lecture 6: Covariance Transformations and the Binomial Distribution

6.1 Linear Transformation of Covariance Matrix

For $\vec{z} = A\vec{x}$:

$$V_z = AV_xA^T$$

6.2 General Transformation

For $\vec{z} = f(\vec{x})$, the Jacobian matrix J has elements:

$$J_{ij} = \frac{\partial z_i}{\partial x_j}$$

Then: $V_z \approx JV_xJ^T$

6.3 Binomial Distribution

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

$$E[X] = np, \quad \mathrm{Var}(X) = np(1 - p)$$

6.4 Poisson Distribution

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$E[X] = \lambda, \quad \mathrm{Var}(X) = \lambda$$

6.5 Central Limit Theorem

For large n, the sum of independent random variables approaches a normal distribution:

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} N(0, 1)$$

7 Key Constants and Identities

7.1 Combinatorics

$$\binom{n}{k} = \frac{n!}{k!(n - k)!}$$

7.2 Gaussian Integrals

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

7.3 Standard Normal Distribution

$\Phi(z) = P(Z \leq z)$ where $Z \sim N(0, 1)$

8 Lecture 8: The Gaussian Distribution and CLT

8.1 Gaussian Properties

- **PDF**: $G(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- **Expectation**: $E[x] = \mu$
- **Variance**: $V(x) = \sigma^2$
- **Kurtosis**: $E\left[\left(\frac{x-\mu}{\sigma}\right)^4\right] = 3$ (Excess kurtosis = 0)

8.2 Central Limit Theorem (CLT)

For n independent random variables x_i with mean μ and variance σ^2, the sample mean $\bar{x} = \frac{1}{n}\sum x_i$ approaches a normal distribution as $n \to \infty$:

$$\bar{x} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

8.3 Error Function

The cumulative distribution of a Gaussian can be expressed using the error function:

$$\text{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-y^2} dy \tag{8.1}$$

$$\text{erfc}(t) = 1 - \text{erf}(t) = \frac{2}{\sqrt{\pi}} \int_t^\infty e^{-y^2} dy \tag{8.2}$$

8.4 Gaussian Confidence Intervals

The probability of a result falling within k standard deviations of the mean:

- $1\sigma \implies 68.27\%$
- $2\sigma \implies 95.45\%$
- $3\sigma \implies 99.73\%$

9 Lecture 9: Estimators and Least Squares

9.1 Properties of Estimators

An estimator $\hat{a}$ for a parameter a should be:

- **Consistent**: $\lim_{n\to\infty} \hat{a} = a$
- **Unbiased**: $E[\hat{a}] = a$
- **Efficient**: Smallest variance among all unbiased estimators.

9.2 Least Squares Estimation (χ^2)

The best estimate for parameters $\vec{\theta}$ in a model $f(x, \vec{\theta})$ is found by minimizing the chi-squared statistic.

$$\chi^2(\vec{\theta}) = \sum_{i=1}^{n} \frac{(y_i - f(x_i, \vec{\theta}))^2}{\sigma_i^2}$$

The minimum is found by solving the system of equations $\frac{\partial \chi^2}{\partial \theta_j} = 0$ for each parameter θ_j.

9.3 Linear Least Squares

For a model linear in its parameters, $\vec{f} = A\vec{\theta}$:

- **Parameter estimates**:
$$\hat{\vec{\theta}} = (A^T V^{-1} A)^{-1} A^T V^{-1} \vec{y}$$
- **Covariance of parameters**:
$$V(\hat{\vec{\theta}}) = (A^T V^{-1} A)^{-1}$$

9.4 Goodness of Fit

The reduced chi-squared, where 'dof' is the degrees of freedom (n data points - m parameters), indicates the quality of the fit.

$$\frac{\chi^2_{\text{min}}}{\text{dof}} \approx 1 \implies \text{Good Fit}$$

10 Lecture 10: Advanced χ^2 and Parameter Errors

10.1 General Chi-Squared for Correlated Data

For correlated measurements with a covariance matrix V:

$$\chi^2 = (\vec{y} - \vec{\mu})^T V^{-1} (\vec{y} - \vec{\mu}) = \sum_{i,j} (y_i - \mu_i) V_{ij}^{-1} (y_j - \mu_j)$$

10.2 Parameter Errors from χ^2 Curvature

The inverse covariance matrix of the parameter estimates $\hat{\vec{\theta}}$ is determined by the Hessian (second derivative) matrix of the χ^2 function at its minimum.

10.2.1 General (Non-Linear) Case

$$(V(\hat{\vec{\theta}}))^{-1}_{jk} = \frac{1}{2} \frac{\partial^2 \chi^2}{\partial \theta_j \partial \theta_k} \bigg|_{\vec{\theta} = \hat{\vec{\theta}}}$$

The diagonal elements of $V(\hat{\vec{\theta}})$ give the variances, $\sigma^2_{\theta_j}$, for each parameter.