PHYS 509 – Theory of Measurements

Enhanced Formula Sheet with Intuitive Explanations

Contents

Week 1 – Probability Basics

Sample space and events: A sample space S is the set of all possible outcomes. Events are subsets of S.

Probability axioms (Kolmogorov):

- $P(E) \geq 0$
- $P(S) = 1$
- If E_i are disjoint, $P\left(\bigcup_i E_i\right) = \sum_i P(E_i)$

These rules define how probability behaves and allow everything else to be derived.

Conditional probability:

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

This tells us how likely A is if we know B happened — it updates probabilities with new information.

Law of total probability:

$$P(A) = \sum_i P(A|B_i)P(B_i)$$

We can compute $P(A)$ by breaking it into contributions from a complete set of mutually exclusive events B_i.

Bayes' theorem:

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

Reverses conditional probabilities — crucial for inference from data.

Random variables and PDFs: A random variable X assigns a number to each outcome. PDF $p(x)$: probability density for continuous variables. CDF $P(X \leq x) = \int_{-\infty}^{x} p(x')dx'$

Mean and variance:

$$E[X] = \int x\, p(x)\, dx, \quad \mathrm{Var}(X) = E[(X - E[X])^2]$$

Mean is the "balance point," variance measures spread.

Week 2 – Common Distributions

Binomial distribution:

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

For n independent Bernoulli trials. Mean $= np$, variance $= np(1-p)$.

Poisson distribution:

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

Models rare events. Mean and variance both $= \lambda$.

Gaussian (normal) distribution:

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

Arises from the central limit theorem. Symmetric, fully described by mean μ and variance σ^2.

Exponential distribution:

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

Describes waiting times between random events.

Week 3 – Transformations and Error Propagation

Variable transformations: If $y = g(x)$ and $p_X(x)$ known:

$$p_Y(y) = p_X(x) \left| \frac{dx}{dy} \right|$$

Error propagation (linear approximation): For $y = f(x_1, \ldots, x_n)$:

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_{x_i}^2 + 2 \sum_{i<j} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \mathrm{Cov}(x_i, x_j)$$

Small uncertainties in inputs propagate linearly to outputs.

Covariance and correlation:

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

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

Covariance shows how two variables change together; correlation normalizes it to ± 1.

Week 4 – Joint and Conditional Distributions

Joint PDF:

$$p(x, y)$$

Describes probability density over two variables simultaneously.

Marginal distributions:

$$p(x) = \int p(x, y)\, dy$$

Conditional distributions:

$$p(x|y) = \frac{p(x, y)}{p(y)}$$

Independence: X and Y independent $\iff p(x, y) = p(x)p(y)$.

Week 5 – Bayesian Inference

Posterior:

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)}$$

Updates our belief about parameter θ given data D.

Evidence (marginal likelihood):

$$p(D) = \int p(D|\theta)p(\theta)\, d\theta$$

Acts as a normalization factor.

Maximum a posteriori (MAP):

$$\hat{\theta}_{MAP} = \arg\max_{\theta} p(\theta|D)$$

Maximum likelihood (ML):

$$\hat{\theta}_{ML} = \arg\max_{\theta} p(D|\theta)$$

MAP incorporates prior knowledge; ML relies only on data.

Week 6 – Parameter Estimation and Likelihood

Likelihood function:

$$L(\theta) = p(D|\theta)$$

Not a PDF over data but a function of θ given fixed data.

Log-likelihood:

$$\ln L(\theta) = \sum_i \ln p(x_i|\theta)$$

Often maximized instead of L for numerical stability.

Fisher information:

$$I(\theta) = E\left[\left(\frac{\partial \ln L}{\partial \theta}\right)^2\right]$$

Cramér–Rao bound:

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{I(\theta)}$$

Sets a lower limit on the variance of unbiased estimators.

Week 7 – Confidence Intervals and Hypothesis Testing

Confidence interval: An interval that will contain the true parameter a certain fraction of the time if the experiment is repeated.

Test statistic: A quantity computed from data to decide between hypotheses.

p-value: Probability of obtaining a result as extreme or more extreme than observed under the null hypothesis.

Significance level α: Reject H_0 if $p < \alpha$.

Week 8 – Chi-Squared and Goodness of Fit

Chi-squared statistic:

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{\sigma_i^2}$$

Measures how well observed data O_i matches expected E_i.

Reduced chi-squared:

$$\chi^2_\nu = \frac{\chi^2}{\nu}$$

Should be ~ 1 for a good fit.

Week 9 – Linear Regression

Best-fit slope and intercept:

$$m = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}, \quad b = \bar{y} - m\bar{x}$$

Find the straight line that minimizes squared errors.

Uncertainties:

$$\sigma_m^2 = \frac{\sigma^2}{\sum(x_i - \bar{x})^2}, \quad \sigma_b^2 = \sigma^2\left[\frac{1}{N} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2}\right]$$

Week 10 – Covariance Matrices and Transformations

Covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \mathrm{Cov}(x, y) \\ \mathrm{Cov}(y, x) & \sigma_y^2 \end{pmatrix}$$

Encodes uncertainties and correlations in multiple variables.

Linear transformations: If $\mathbf{y} = A\mathbf{x}$:

$$\Sigma_y = A\Sigma_x A^T$$

Week 11 – Multivariate Gaussian and Measurement Ellipses

Multivariate normal:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^k|\Sigma|}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mu)^T\Sigma^{-1}(\mathbf{x} - \mu)\right]$$

Uncertainty ellipses: Contours of constant χ^2 form ellipses:

$$(\mathbf{x} - \mu)^T\Sigma^{-1}(\mathbf{x} - \mu) = c$$

c chosen from χ^2 distribution for desired confidence.

Week 12 – Advanced Topics and Review

Tchebyshev's inequality:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Holds for any distribution with finite variance — shows most probability is near the mean.

Central Limit Theorem: The sum of many independent random variables tends to a normal distribution, regardless of their original distribution.

Bayesian vs Frequentist: Bayesian treats parameters as random variables with priors; frequentist treats them as fixed but unknown.