24 Glossary

R. Heller · Date: 2026-05-11 · Last changed: 2026-10-30

APPENDIX · GLOSSARY

Plain-English definitions of the terms this curriculum uses the most.

Each entry links to the first lab in which the term appears.

24.1 A

α (alpha). The probability of a type I error — rejecting a true null hypothesis. Conventional values are 0.05 and 0.01. Hypothesis testing p values type i ii errors.

ANCOVA. Analysis of covariance: a linear model combining categorical predictors and continuous covariates, commonly used to adjust RCT analyses for baseline. Ancova in rcts baseline adjustment.

ANOVA. Analysis of variance: a linear model with categorical predictors. One way anova and contrasts emmeans.

Assumption. A condition the chosen test needs to be interpretable — normality, equal variance, independence, linearity. Two sample and paired t tests.

24.2 B

Bias. Systematic deviation of an estimator from the quantity being estimated. Populations samples and the central limit theorem.

Bootstrap. Resampling with replacement to approximate the sampling distribution of a statistic. Bootstrap and permutation tests.

Brier score. Mean squared error between predicted probabilities and observed outcomes. Calibration discrimination roc auc brier score.

24.3 C

CI (confidence interval). A range of plausible values for a parameter; in repeated sampling, 95% CIs cover the true value 95% of the time. Bootstrap and permutation tests.

Cohen's d. Standardised difference in means; a common effect size for two-group comparisons. Two sample and paired t tests.

Competing risk. An event whose occurrence precludes or alters the probability of the event of interest. Competing risks and multistate models.

Confounding. A third variable distorting the association between exposure and outcome. Multiple regression confounding interaction centring.

Cox model. A proportional-hazards regression for time-to-event outcomes. Survival primer km log rank cox ph.

CV (cross-validation). Splitting data into folds to estimate generalisation error. Cross validation nested cv bootstrap 632.

24.4 D

DAG (directed acyclic graph). A graphical representation of causal assumptions. Dags with dagitty and ggdag.

24.5 E

Effect size. A standardised measure of the magnitude of an effect, independent of sample size. Two sample and paired t tests.

24.6 F

FDR (false discovery rate). The expected proportion of false positives among rejected nulls. Fdr knockoffs replication crisis.

Fisher's exact test. Exact test of independence for 2×2 tables, appropriate when expected cell counts are small. Two proportions chi square risk and odds.

24.7 G

GAM (generalised additive model). A regression with smooth non-linear terms, fitted via penalised splines. Gams with mgcv.

GEE (generalised estimating equations). A marginal-model approach for clustered or repeated data. Glmms and gee.

GLM (generalised linear model). A regression with a link function and exponential-family error. Logistic regression binomial glm.

24.8 H

Hazard ratio. Ratio of hazard rates between two groups in a survival model. Survival primer km log rank cox ph.

24.9 I

ICC (intraclass correlation). Proportion of variance attributable to clustering. Kappa icc bland altman.

Interaction. The effect of one predictor depends on the value of another. Multiple regression confounding interaction centring.

IPTW (inverse-probability-of-treatment weighting). Propensity-score method that reweights observations to emulate a trial. Propensity scores and iptw.

24.10 K

Kaplan-Meier. Non-parametric estimator of the survival function. Survival primer km log rank cox ph.

Kruskal-Wallis. Non-parametric one-way ANOVA on ranks. Non parametric tests.

24.11 L

Lasso. L1-regularised regression; produces sparse coefficient estimates. Regularisation ridge lasso elastic net.

Likelihood. The probability of the observed data under a model, viewed as a function of the parameters. Maximum likelihood estimation.

Linear model. A model of the form $y = X\beta + \varepsilon$. Simple linear regression.

Logistic regression. GLM with a logit link for binary outcomes. Logistic regression binomial glm.

24.12 M

MAR (missing at random). Missingness depends on observed variables only. Mcar mar mnar.

MCAR (missing completely at random). Missingness is independent of all variables. Mcar mar mnar.

MCMC. Markov-chain Monte Carlo; the Bayesian posterior-sampling workhorse. Brms stan loo hierarchical models.

MDE (minimum detectable effect). Smallest effect your study has power to detect. Power closed form and simulation.

Meta-analysis. Combining effect estimates across studies. Meta analysis basics.

Mixed model. Regression combining fixed and random effects. Linear mixed models with lme4.

MNAR (missing not at random). Missingness depends on unobserved values. Mcar mar mnar.

Multiple imputation. Imputing missing values several times and pooling the analyses. Multiple imputation with mice.

24.13 N

Non-parametric. A test or estimator that makes weak distributional assumptions. Non parametric tests.

24.14 O

Odds ratio. Ratio of odds between two groups; the natural scale for logistic regression. Logistic regression binomial glm.

Outlier. An observation far from the bulk of the data. Diagnostics residuals qq leverage cook s distance vif.

Overdispersion. Variance exceeding the model's nominal variance; common in Poisson regression. Poisson and negative binomial regression.

24.15 P

p-value. Probability of data as or more extreme than observed, assuming the null. Hypothesis testing p values type i ii errors.

Paired test. A test comparing matched observations rather than independent samples. Two sample and paired t tests.

PCA (principal components analysis). Linear dimension reduction by orthogonal projection onto directions of maximum variance. Pca factor analysis cca lda.

Permutation test. Inference by shuffling labels to build a null distribution. Bootstrap and permutation tests.

Poisson regression. GLM with a log link for counts. Poisson and negative binomial regression.

Power. Probability of detecting an effect if it exists; 1 − β. Sample size power and quarto reporting.

Pre-registration. A timestamped record of the research plan before data analysis. Pre registration and statistical analysis plans.

Propensity score. Probability of treatment given covariates. Propensity scores and iptw.

Pseudoreplication. Treating correlated observations as independent replicates. Bench and translational design.

24.16 R

Random effect. A model coefficient treated as a draw from a distribution. Linear mixed models with lme4.

Randomisation. Allocating units to arms by a chance mechanism. Randomised controlled trials.

Regression to the mean. Tendency of extreme values to be closer to the mean on remeasurement. Dichotomisation change scores regression to the mean.

Reliability. Consistency of repeated measurements. Kappa icc bland altman.

Resampling. Bootstrap, permutation, and cross-validation collectively. Bootstrap and permutation tests.

Residual. Observed minus predicted. Diagnostics residuals qq leverage cook s distance vif.

Risk ratio. Ratio of event probabilities between two groups. Two proportions chi square risk and odds.

Robust SE. A standard error computed without assuming homoscedasticity. Robust regression weighted ls hc sandwich standard errors.

ROC / AUC. Receiver operating characteristic; area under it measures discrimination. Calibration discrimination roc auc brier score.

24.17 S

SAP (statistical analysis plan). The formal plan for a trial's analysis, written before data are seen. Pre registration and statistical analysis plans.

SE (standard error). Standard deviation of a sampling distribution. Populations samples and the central limit theorem.

SEM (standard error of the mean). Standard error of a sample mean. Populations samples and the central limit theorem.

Shapiro-Wilk. A test of normality, powerful in small samples. Two sample and paired t tests.

Spearman correlation. Rank-based correlation. Pearson and spearman correlation.

Survival. Time-to-event analysis. Survival primer km log rank cox ph.

24.18 T

Type I / II error. False-positive and false-negative errors of a test. Hypothesis testing p values type i ii errors.

24.19 V

VIF (variance inflation factor). Measure of collinearity among regression predictors. Diagnostics residuals qq leverage cook s distance vif.

24.20 W

Wilcoxon test. Non-parametric test for paired (signed-rank) or unpaired (rank-sum) data. Non parametric tests.