Course 2 · Week 4 — Measurement, change, survival, reporting

Cheatsheet — biostats_courses

R. Heller

Don't dichotomise

Cutting a continuous variable at the median loses ≈ 37% of the information for a linear association. Keep predictors continuous; model non-linearity with a spline if needed.

Change scores vs ANCOVA

- Change scores suffer from **regression to the mean**.
- ANCOVA (adjust for baseline) is the efficient analysis in an RCT.
- In observational studies with unequal baselines, both approaches answer subtly different questions — state which.

Agreement & reliability

Statistic	Use	R
Cohen's κ	two raters, categorical	`irr::kappa2(cbind(r1, r2))`
Weighted κ	ordinal categorical	`irr::kappa2(..., weight = "squared")`
ICC(2,1) / ICC(3,1)	continuous	`psych::ICC(df)`
Bland-Altman	two methods, continuous	plot $\overline{xy}$ vs $y - x$

```r
mean_diff <- mean(y - x)
loa <- mean_diff + c(-1.96, 1.96) * sd(y - x)
ggplot(df, aes((x + y)/2, y - x)) + geom_point() +
  geom_hline(yintercept = c(mean_diff, loa), linetype = 2)
```

Survival primer

Concept	R
Kaplan-Meier	`survfit(Surv(time, event) ~ g)` + `ggsurvfit::ggsurvfit`
Log-rank test	`survdiff(Surv(time, event) ~ g)`
Cox PH model	`coxph(Surv(time, event) ~ x1 + x2)`
PH check	`cox.zph(fit)`

- Report a hazard ratio with 95% CI.
- PH violated? ⟶ time-varying coefficient or stratify.
- Interpret HR only after checking for non-informative censoring.

Decision curves, NRI, IDI

- **Decision curve**: net benefit vs threshold probability — dominates "treat all" and "treat none" when useful.
- **NRI**: how many events move up / non-events move down in risk.
- **IDI**: mean change in predicted probability by class.
- Report decision curve first; NRI/IDI as secondary.

Explanation vs prediction (Shmueli)

Explanatory	Predictive
Goal: inference about β	Goal: minimise out-of-sample loss
Tools: ANOVA, diagnostics, intervals	Tools: CV, regularisation, ensembles
Metric: interval coverage, bias	Metric: RMSE, AUC, Brier on hold-out

Reporting guidelines

Design	Guideline
Randomised trial	CONSORT
Observational	STROBE
Diagnostic-accuracy	STARD
Prediction model (incl. AI)	TRIPOD / TRIPOD-AI
Systematic review	PRISMA
Animal research	ARRIVE

Decision rule for Week 4

- Continuous predictor + continuous outcome ⟶ no median splits.
- Two-group with repeated measurement ⟶ ANCOVA.
- New rater / device ⟶ Bland-Altman + ICC.
- Time-to-event ⟶ KM + Cox, always check PH.
- Prediction model ⟶ TRIPOD checklist at submission.

Common pitfalls

- Reporting κ on near-constant outcomes (high % agreement, low κ).
- Quoting median survival that is never reached.
- Forgetting to mention PH assumption checking.
- Claiming a "prediction model" built on the same data used for evaluation.

Further reading

- Harrell, *BBR*, ch. 17–18.
- Royston & Altman, *Prognosis research*.