Index of Lab/Tutorial 4

Page	Title
1	Logistic regression menu (Minitab)
2	Logistic regression results (Minitab)

Notes for Exercises in Session 4

- add4:1; VER:16.1,2; add4:3 (add2:8,10)

Outline of lab session:

- brief demonstration of Minitab facilities for logistic regression (using `mice` and `nocardia` datasets),
- (upon request only) further demonstration/discussion of Minitab/Stata analysis for linear (regression) models,
- individual work/discussions on the exercises (using Minitab and/or Stata),
  - skip logistic regression parts not covered yet: model goodness-of-fit and diagnostics,
- any questions on linear and logistic regression models (conceptual and/or practical).

Logistic Regression Menus (Minitab)

- simple Fitted Line plot menu (one continuous predictor),
- main analysis menu: `Stat-Regression-Binary Logistic...`,
- Model setup:
  - choice between binary and grouped data formats,
  - similar layout as for linear regression; note that derived terms (interactions, polynomials) are defined in the Model submenu,
- Graphs: diagnostic plots similar to linear models based on *covariate patterns* (next lecture),
- Results: need to request extended tables in order to get Wald tests and CIs, otherwise only tests based on deviance
  - deviance $\sim -2 \ln L$, with a reference level ("saturated model"),
  - tests based on deviance $\sim$ likehood-ratio (LR) tests,
- Options — useful features:
  - choice of residuals (deviance or Pearson; next lecture),
  - choice of # groups for Hosmer-Lemeshow (H-L) test (next lecture),
- Storage — useful variables:
  - diagnostics/residuals + fitted prob.,
  - log-likelihood value for model,
  - variance/covariance matrix of estimates,
    - diagonal: variances (squared SE's) of estimates,
    - off-diagonal: covariances[1] between estimates.

[1] Correlations can be computed as: covariance/product of SE's.

Logistic Regression Results (Minitab)

Notes on interpretation of output:

- Wald test for single parameters/coefficients, but also LR test for all terms in the Deviance table,
- goodness-of-fit tests:
  - Pearson and Deviance (LR) tests against model with one parameter per covariate pattern,
  - table of observed and expected counts for H-L test,
- measures of association or model fit[2]:
  - total number of pairs of obs. with different outcomes = # 0's × # 1's,
  - concordant pairs (c): $\hat{p}_1 > \hat{p}_0$, ($\hat{p}_1 \sim$ "event")
  - discordant pairs (d): $\hat{p}_1 < \hat{p}_0$, ($\hat{p}_0 \sim$ "non-event")
  - tied pairs: $\hat{p}_1 = \hat{p}_0$,
  - summary measures $\sim$ single value for concordance (higher is better), e.g. G-K gamma $= (c-d)/(c+d)$.

Features missing for logistic regression in Minitab 17:

- descriptive tools for assessing linearity,
- fit statistics ($-2 \ln L$, pseudo R^2, BIC) and ROC curves.

[2] Not covered in VER; Stata `somersd` command gives Somer's D & Kendall's tau.