

Index of Lab/Tutorial 4

Page	Title
1	Notes for exercises in Session 4
2	Logistic regression menu (Minitab)
3	Logistic regression results (Minitab)

NOTES FOR EXERCISES IN SESSION 4

- VER:15, 16.1; add2:7; (add4:1,3; add2:10)

Outline of lab session:

- review of VER model-building exercise (# 15): Q1-6,
- brief demonstration of Minitab facilities for logistic regression (using `mice` and `nocardia` datasets),
- further work on VER exercises: 15.Q7-11, 16.1,
- perhaps also work on extra model building exercise (Exercise 7 for Session 2),
- perhaps also work on extra “Exercises from Session 4: Logistic regression”,
 - # 3: logistic two-factor design,
 - # 1: logistic regression with single quantitative predictor, comparing Minitab and Stata outputs.
- any questions on linear and logistic regression models (conceptual and/or practical).

Solution files: at VHM 802, Exercises and Datasets page.

LOGISTIC REGRESSION MENUS (MINITAB)

- simple Binary 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* (lecture 5aL),
- Results: need to request expanded 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 likelihood-ratio (LR) tests,
- Options — useful features:
 - * choice of residuals (deviance or Pearson; lecture 5aL),
 - * choice of # groups for Hosmer-Lemeshow (H-L) test (lecture 5aL),
- 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: covariances¹ 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 fit²:
 - * total number of pairs of obs. with different outcomes = # 0's \times # 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 18:

- 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.