1 Linear regression

Simple Linear regression

- State the model and the assumptions for a simple linear regression. Interpret the parameters.
- Describe the principle of Least Squares, set the Normal Equations and find their solution.
- Derive the distributions of the least squares estimates, their means and variances.
- State the distribution of a predicted new observation.
- Derive confidence intervals for the estimates.
- Describe the difference between a confidence interval and a prediction interval.
- Derive the procedure for hypothesis testing on parameters (t-tests).
- Construction and use of p-values.

Multiple linear regression

- State the model and the assumptions for a multiple linear regression. Interpret the parameters.
- Set the system of Normal Equations. State the solution of the system.
- Derive the distributions of the least squares estimates, their means and (co)variances.
- t-test for multiple regression.
- Variability decomposition and the construction of ANOVA tables.
- Global F-test and Partial F-test. Their construction and use.
- Discuss how categorical predictors/covariates can be parametrized. Parameter interpretations for categorical covariates.
- Interactions.
- Construct $R^2$, $R^2_{adj}$, AIC, BIC and describe their use.
- Discuss the problem of collinearity among the independent variables.
- Discuss the situations where the assumption of normal distribution is crucial and where it is not (estimates? tests? confidence intervals? prediction intervals? . . . )
- Outlier detection. Construct different types of residuals (standardised, studentized), their properties, and how they can be used.
- Potentially influential observations. Define the Leverage and describe its use.
- Influence analysis: what is an influential observation? Cook's distance and DFBETAs.
2 Logistic regression

1. The Binomial distribution, odds and odds ratios.

2. State the model and the assumptions for simple and multiple logistic regression. Interpret the parameters.

3. Maximum Likelihood procedure for a general logistic regression model; Newton-Raphson algorithm and asymptotics for the maximum likelihood estimates.

4. Confidence intervals for log odds, odds ratios and probabilities.

5. The Wald test.

6. Use of the deviance for hypothesis testing and the Likelihood-Ratio test.

7. Comparison of non-nested models.

8. Define some different types of residuals and influence measures, their properties, and how they can be used.


3 Other types of regression


2. Construction of the Poisson distribution and rare events; Poisson regression.

3. Negative binomial distribution and negative binomial regression; relation with the Poisson distribution.

4. Quantile regression. Not this year