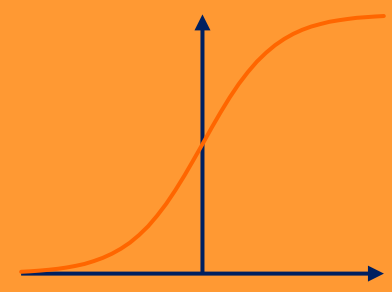
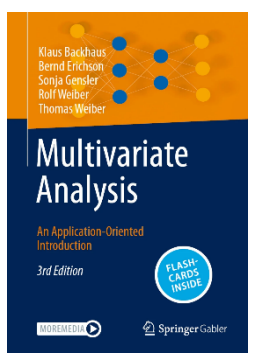
Logistic Regression



Main research question	Which of two or more alternatives will occur? What factors influence an event and what is their effect?
Example	Which individual characteristics influence the choice between certain car brands?
Type of analysis	Structure-testing method
Measurement level	
dependent variable	Input: Categorical (binary or multinomial logistic regression) Output: probabilities
independent variables	Metric or categorical
Recommendations	▪ Start your analysis with only two categories, which require the estimation of only one function (binary logistic regression). ▪ As logistic regression yields probabilities, a nonlinear functional form becomes necessary. A linear alternative for problems with a categorical dependent variable is discriminant analysis, and for problems with two categories, linear regression can be used. These methods are easier to handle but do not provide probabilities. ▪ An advantage of logistic regression is that it requires fewer assumptions (compared to discriminant analysis and regression analysis) concerning the data and is thus more robust. A disadvantage of logistic regression is that it requires a larger sample size.
Keywords	Akaike information criterion, Bayesian information criterion, Chi-square statistic, Classification table, Confusion matrix, Cox & Snell R2, Cutoff value, Deviance, hit rate, Leave-one-out method, Leverage effect, Likelihood ratio statistic, Linear probability model, Link function, Logistic function, Logit choice model, Maximum likelihood method, McFadden's R2, Nagelkerke's R2, Null model, Odds ratio, Pearson residuals, Predictive accuracy, Pseudo-R2 statistics, ROC curve, Sensitivity, Specificity, Wald test



Backhaus, Klaus; Erichson, Bernd; Gensler, Sonja; Weiber, Rolf; Weiber, Thomas (2025)
Multivariate Analysis – An Application-oriented Introduction, Springer: Berlin

www.multivariate-methods.info