

Generalized Linear Models

Lecture 13. Summary

Generalized linear models

- ① Framework and general theory
- ② Models with continuous response
- ③ Models with binary response
- ④ Count data models
- ⑤ Tweedie models. Compound Poisson models
- ⑥ Models with censoring and truncation

Framework and general theory

- Form of a generalized linear model. Link and response function. Exponential family of distributions: likelihood function, mean and variance
- Estimation of the model parameters. Score function. Empirical and theoretical Fisher information matrix. Solving the maximum likelihood equations. Newton-Raphson and Fisher scoring methods
- Testing the hypothesis about the model. Likelihood ratio statistic, score statistic, Wald's statistic
- Goodness of fit of a model. Deviance
- Diagnostics of a GLM. Generalized residuals.

Models with continuous response

- Normally distributed response. Estimating the model. Diagnostics. Transformation of variables: Box-Cox family of transformations
- Gamma-distributed response (models with constant coefficient of variation). Estimating the model. Deviance. Link functions. Canonical link
- Models with log-normal and inverse Gaussian response
- Parametrization of distributions through the parameter of exponential family (θ)

Models with binary response

- Grouped and ungrouped data. Fitting the model. Goodness of fit (Hosmer-Lemeshow test, Pearson χ^2 -test)
- Bernoulli and (scaled) binomial as members of exponential family
- Estimation of parameters. Separation problem and options to deal with it
- Theory of latent variable. Different choices for tolerance distribution. Different link functions. Canonical link. Interpretation of models

Count data models

- Poisson models. Apparent overdispersion. Quasi-likelihood. Grouped and ungrouped data. Interpretation of models
- Negative binomial models. Log-likelihood of the model. Link functions. Interpretation of models
- Models with zero-modifications. Derivation of the log-likelihood. Interpretation of models
 - Zero-inflated models (ZIP, ZINB)
 - Zero-altered / hurdle models (ZAP, ZANB)
 - Zero-truncated models (ZTP, ZTNB)
- Generalized Poisson models. ZIGP, ZTGP models

Compound Poisson models

- Compound Poisson models (CPo)
- Zero-inflated and zero-altered compound Poisson models (ZICPo, ZACPo)
- Poisson-Gamma compound distribution: Tweedie model Tw_p ($1 < p < 2$)

Models with censoring and truncation

- Censored and truncated count data models. Derivation of likelihood
- Survival models:
 - Survival/survivor function. Hazard function. Cumulative hazard
 - Distributions used. Parametric and semiparametric regression models (AFT and PH)
 - Log-likelihood with censored data
 - Proportional hazards model as Poisson model
- Censored/truncated continuous response:
 - Tobit model: idea, log-likelihood
 - Two-part models. Sample selection model