

Applied Statistical Modelling (M348) content listing

Unit 1: <i>An introduction to statistical modelling and R</i>	Revises simple linear regression and is also an introduction to the statistical software used in M348, namely R.
Unit 2: <i>Multiple linear regression</i>	Description of the multiple linear regression model, diagnostics including leverage and Cook's distance, transformations, measuring model fit, stepwise regression.
Unit 3: <i>Regression with a categorical explanatory variable</i>	Regression with a factor, indicator variables, analysis of variance and the ANOVA table, contrasts.
Unit 4: <i>Multiple regression with both covariates and factors</i>	One covariate and one factor: parallel slopes and non-parallel slopes. regression with two factors, interactions, regression with any number of factors and covariates.
Unit 5: <i>Case study</i>	An example of statistical modelling in practice. Preparing data for analysis, missing data, conveying results, replication.
Unit 6: <i>Regression for a binary response variable</i>	The logistic function, logistic regression, assessing model fit for logistic regression models, diagnostic plots for logistic regression models
Unit 7: <i>Generalised linear models</i>	Poisson regression, the generalized linear model, inference for GLMs, GLM applications.
Unit 8: <i>Log-linear models for contingency tables</i>	Two-way contingency tables, contingency tables with more than two variables, links between logistic and loglinear models.
Strand study <i>Introduction to Econometrics (Strand A Econometrics)</i> Or <i>Cluster Analysis (Strand B Data Science)</i>	<i>Implications of when the classical linear regression model assumption do not hold, simultaneous equation models, panel data models.</i> Or Dissimilarity, assessing clusters, hierarchical clustering, partitional clustering, density-based clustering.
Strand study <i>Introduction to Time Series Econometrics (Strand A Econometrics)</i> Or <i>Big Data (Strand B Data Science)</i>	<i>Stock and flow variables, persistence and momentum, random walks, testing for stationarity, spurious regression and cointegration.</i> Or Characterising big data, handling big data, algorithms and models, data analysis implications, consent, anonymisation and good practice.
Final unit: Review unit: <i>Data analysis in practice</i>	The general linear model, embedding the general linear model in generalised linear modelling, choosing the right GLM, outliers.