

## STA 4504 First Exam Topic List

What is the difference between ordinal and nominal categorical variables?

In what situations would we use a binomial model for categorical data?

How would we draw inferences for the binomial parameter  $\pi$  based on observed data? What does it mean for an estimator to be the maximum likelihood estimator?

What type of model is used for binary data when the number of observed responses is random?

When data are cross-classified in a two-way contingency table, how do we form relative risks, odds and odds ratios? Be able to estimate these from data. Why do we express results in terms of odds and odds ratios? How are a relative risk and an odds ratio interpreted? Be able to find a CI for an odds ratio?

When it comes to describing a study, what is the meaning of the terms: experimental, observational, retrospective, prospective, case-control, cohort, clinical trial and cross-sectional?

What does independence mean in a contingency table and why is this condition important?

Be able to find the Pearson Chi-square and the likelihood ratio tests for testing independence?

Be able to implement Fisher's exact test. What is the rationale for its null distribution?

What is the meaning of the terms: partial tables, marginal table, conditional odds ratios, marginal odds ratio, homogeneous association, and interaction. Be able to estimate conditional and marginal odds ratios. What is Simpson's paradox and what does it tell us about interpreting marginal tables?

How do we test for homogeneity of the odds ratios in partial tables, how do we estimate the common odds ratio and what hypotheses are tested by the Cochran-Mantel-Haenszel test?

What are the three components in a Generalized Linear Model? What are the identity, logit and log links?

With binomial data and an identity or logit link, how do we fit the data and what are the inference questions we want to ask about the model?

What are the three tests used to determine whether the predictor ( $x$ ) is useful in predicting the response? What are the Deviance and Pearson Chi-Square criteria for determining whether the model fits the data adequately?

How do we form a GLM for count data using the Poisson model and a log link? Be able to perform inferences on a GLM for Poisson data?

What modifications are made when we seek to model a rate for Poisson data? How does this model differ from the usual Poisson model setting?

What does overdispersion mean? Why and how do we perform inferences on a GLM using a negative binomial model?