MS Module 5: Hypothesis testing of proportions (overview)

(The attached PDF file has better formatting.)

Reading: §9.3: Tests Concerning a Population Proportion

- Large sample tests use the central limit theorem and an approximate normal distribution.
- Small sample tests use the binomial distribution.

If the null hypothesis is not true, the Z statistic is still normally distributed but its mean and variance are not 0 and 1. The β values (probabilities of Type II errors) depend on the type of null hypothesis (one-tailed or two-tailed) and the value of p_0 and p' (the proportions in the null hypothesis and the assumed alternative).

For two-sided null hypotheses, we use the absolute value of $p' - p_0$. For a one-side (single tailed) null hypothesis, check whether p' is larger of smaller than p_0 before plugging values into the formula. If p' is in the null hypothesis, computing β makes no sense.

Module 5, continued: Hypothesis testing – *p* values

Reading: §9.4: *p* values
Reading: §9.5: Comments on Selecting a Test Procedure, subsections

- statistical versus practical significance
- power and uniformly most powerful tests (through the end of example 9.22, including Figure 9.10)

The other subsections in §9.5 ("best tests for simply hypotheses" and "likelihood ratio tests") are not on the syllbus for this course.

Final exam problems often ask for the *p* values (by interpolation in the tables provided). The exam problems are multiple choice questions, and the choices are sufficiently different that simple interpolation is sufficient.

The final exam does not test the proof of the Neyman–Pearson theorem. You may skip Example 9.23; the mathematics in this example is not tested on the final exam.

Regulators and journals set critical *z* values: new drugs must meet strict thresholds to be deemed safe and journal submissions must meet thresholds to be published. These critical *z* values and critical *t* values (such as 5% or 1%) are arbitrary. They help standardize regulatory tests and papers in research journals, but they are not appropriate for business decisions.

Actuaries are more concerned with *p* values and economic costs of using new research. A *p* value of 6% when the costs of using the new information are small is better than a *p* value of 4% when the costs are high.

The textbook refers to these costs as statistical versus practical significance. These costs differ for each use of the statistical analysis; no general formulas can be given.

The textbook shows how *p* values are computed. Final exam problems ask for *p* values, which you derive by interpolation on the exam. (In practice, computer programs give the *p* values for most distributions).