Sep 30, 2015

Contents

I Introduction

I.1 Motivating Example: PTLD following renal (kidney) transplant

1. Note: In this course we are concerned with the statistical foundations (theory, methods) needed for statistical inference.

   - Our goal is to learn the best way to answer a particular scientific question using statistical analysis.
   - A major aspect of the course is learning the varied ways in which we define "best" (i.e., learning about optimality criteria).
   - We also learn approaches that are likely to behave well with respect to the specific optimality criteria we use for any particular problem.

   We illustrate the "components" of a statistical problem with an example.

2. Example: People who have kidney failure require dialysis and eventually kidney transplantation.

   - After transplant, patients take immunosuppressive drugs to prevent rejection.
   - Some transplant patients develop a cancer-like syndrome: post transplant lymphoproliferative disorder (PTLD).
   - Hypothesize that prior infection with Epstein-Barr virus (EBV) is a risk factor: immunotherapy allows reactivation of a dormant virus.
   - Hypothesize that risk of developing PTLD is highest soon after transplant, and then wanes.
   - Want to use (observational) registry data to quantify any such patterns in the incidence of PTLD.

3. Note: In using statistical analyses to answer this question, we must

   - choose a probability model that describes variability of response,
   - identify the ultimate statistical goal,
   - identify a framework for judging precision of inference,
   - choose/identify a sampling framework for obtaining data,
   - choose a summary measure of the distribution within groups as a target of inference,
   - choose a method of estimating the summary measure,
   - choose a method of estimating the precision of estimates,
   - (perhaps) choose a contrast to compare some summary measure across groups as a target of inference,
   - (perhaps) choose a method of estimating the contrast, and

- (perhaps) choose a method of estimating the precision of the contrast.

4. Note: The probability model has to account for the fact that there is variability in the time to development of PTLD in subjects, which variability includes the distinct possibility that some subjects never develop PTLD. Probability models used in statistical analyses are often characterized as being

   - parametric,
   - semi-parametric, and
   - non-parametric (distribution-free).

   In this example, a typical choice of probability model might be the semi-parametric proportional hazards model for the time to development of PTLD, though I tend to really consider a more distribution-free model.

5. Note: The "statistical goal" of a particular application is often categorized as

   - clustering observations,
   - clustering variables,
   - estimation of summary measures of the distribution of "response" within homogeneous groups,
   - using some contrast to compare distributions of "response" across different groups, and/or
   - predicting "response".

   We note that the summary measure or contrast of summary measures is sometimes referred to as a "targeted parameter of interest". A distinction will have to be made between this use of the word "parameter" and its use within a "parametric probability model".
   Our example can be best characterized as focused on contrasts comparing the distribution of time to PTLD across groups.

6. Note: The "framework for judging precision inference" is most often categorized as

   - frequentist, or
   - Bayesian.

   In the example used here, statistical analysis was reported using frequentist criteria including consistent estimates, confidence intervals, and p values.

7. Note: The sampling framework for collection of data might involve

   - an observational "convenience" sample,
   - a designed (possibly stratified) random sampling scheme from observational data (e.g., surveys, case-control, cohort), or

- an interventional experiment (e.g., randomized clinical trial).

In the example used here, data was abstracted from a (convenience) observational registry of patients on dialysis (USRDS).

8. Note: Inference may be ultimately based on some distributional "summary measure" defined within homogeneous groups having similar probability distributions of the "response". Commonly used summary measures include:

   - mean,
   - geometric mean,
   - median (or other quantile),
   - proportion or odds of exceeding some specified threshold, and
   - (average) hazard.

   In the example used here, the presumption was that the inference would be used on a weighted average of the hazard function that could be estimated from the data with the greatest precision when a proportional hazards relationship was presumed.

9. Note: Methods used for estimation of a summary measure might be based on estimating equations derived from (among others)

   - least squares,
   - method of moments,
   - maximum likelihood,
   - maximum partial likelihood, or
   - other methods motivated by likelihood methods including
     - maximum partial likelihood,
     - quasi likelihood,
     - generalized estimating equations,
     - profile likelihood, or
     - marginal likelihood

   Choosing among these methods might be based on various optimality criteria such as

   - frequentist accuracy or central tendency (e.g. unbiased or median unbiased in small samples, asymptotic unbiased or consistent in large samples),
   - frequentist precision (e.g. minimum variance, minimum mean squared error, asymptotic efficiency, most powerful),
   - Bayesian accuracy (e.g., posterior mean, posterior median, posterior mode),
   - Bayesian precision (e.g. credible interval based on highest posterior density), or
   - robustness to incorrect model presumptions.

In the example used here, a presumption of a proportional hazards model might lead to use of the Cox proportional hazards (PH) model.

10. Note: Methods used for estimating precision of a summary measure might be broadly categorized as

    - model based (e.g., using likelihood methods in "regular" statistical problems),
    - method of moments based (e.g., using empirical estimates of variance from contributions to an estimating equation), or
    - based on resampling methods (e.g., bootstrapping, empirical likelihood, permutation).

    In the example used here, model based estimates of standard errors were derived under the partial maximum likelihood theory of the PH model.

11. Note: In statistical problems comparing distributions across groups, contrasts of summary measures might include

    - differences or weighted averages of stratified differences,
    - ratios or weighted geometric means of stratified ratios,
    - probability that a randomly chosen individual from one group has a measurement exceeding that of a randomly chosen individual from another group, or
    - maximal difference between each group's cumulative distribution function.

    In the example used here, the hazard ratio estimated from the PH model can be viewed as a weighted geometric mean of hazard ratios over time.

12. Note: Methods of estimating contrasts can be broadly characterized as

    - contrasts that represent a fundamental "parameter" of a regression model,
    - contrasts that represent weighted averages of regression "parameters", or
    - contrasts computed across predictions derived from "highly predictive models".

    In the example used here, the regression parameters associated with EBV and time model the desired contrast.

13. Note: Methods used for estimating precision of a contrast might again be broadly categorized as

    - model based (e.g., using likelihood methods in "regular" statistical problems),
    - method of moments based (e.g., using empirical estimates of variance from contributions to an estimating equation), or
    - based on resampling methods (e.g., bootstrapping, empirical likelihood, permutation).

    In the example used here, model based estimates of standard errors were derived under the partial maximum likelihood theory of the PH model.