Lab 2, Significance Testing

Dave Harrington

May 2018

The lab files come in two versions: a PDF with the problem statements and an Rmd file that produces the PDF. In most cases, you can work in the Rmd file and enter in your solutions. For the purely algebraic questions, you may either use LaTeX commands to enter your solutions in the Rmd file or write out the answers on paper.

The solution files (a PDF with the solutions, and the Rmd file to produce them) are contained in a separate folder under each lab. Your learning experience with the labs will be more effective if you do not look at the solutions until after you finish working on the questions.

Problem 1: Tests for differences at a single time point

This problem uses the Cox and Oakes leukemia data mentioned in the slides for Units 2 and 3, which is in the `eventtimedata` package as `cox.oakes.leukemia`. You may use the code shown in the slides as needed.

a) In the Cox and Oakes leukemia data, find a 95% confidence interval for the difference in survival curves at time point $t^\star = 10$ weeks.

b) Does the confidence interval imply that the values of the survival functions at $t^\star = 10$ are significantly different at significance level $\alpha = 0.05$?

Problem 1 Solution.

a)

b)

Problem 2: The numerator of the log-rank statistic

Show that the $(o - e)$ terms in the numerator of the log-rank statistic can be written as

$$\frac{r_{0j} r_{1j}}{r_j} (\widehat{\lambda}_{1j} - \widehat{\lambda}_{0j}).$$

Problem 2 Solution.

Problem 3: Prognosis in lymphoma

The Unit 3 lecture slides used the lymphoma prognosis data to illustrate a four-group log-rank test, testing for survival differences in lymphoma by stage. This example examines a simpler two-group test but with an interesting complication. The data are in the dataset `lymphoma.prognosis` in the `eventtimedata` package.

a) Estimate the survival probability by status of bulky disease. Bulky disease is coded in the numeric variable `BULK`, with `1` denoting `not present` and `2` denoting `present`. Be careful about the coding of the status variable `SURVIVAL`. See the slides or the code chunk below for how to use it appropriately.

   Do the data appear to satisfy the proportional hazards assumption?

b) Using a log-rank statistic, test for significant differences in survival in patients with bulky disease versus those who do not have bulky disease. State precisely the null and alternative hypotheses that are being tested.

c) The validity of the p-value from a log-rank test does not require that the data satisfy proportional hazards, but the test does lose power for some settings in which the hazards are not proportional. Comment on the effect of absence of proportional hazards on the outcome of the test.

d) In the Fleming-Harrington tests, setting the parameters $\rho = 1, \gamma = 0$ produces a generalized Wilcoxon test which emphasizes early differences. Re-do part b) with such a test. How does it change the outcome? Explain why.

```
library(survival)
library(eventtimedata)
data("lymphoma.prognosis")

#adjusting coding of status variable
died = lymphoma.prognosis$SURVIVAL - 1
died[died == 2] = 0  #recoding those lost to follow-up as censored

#examine survival probability by status of bulky disease

#conduct log-rank test

#conduct generalized wilcoxon test
```

Problem 3 Solution.

a)

b)

c)

d)

Problem 4: Prognosis in lymphoma, continued

It is common in cancer studies to stratify treatment assignment by age, since older patients often have a much different prognosis than younger ones. Stratification in an analysis can also be useful in studying prognostic factors.

a) In the lymphoma data, assess the association between bulk disease and survival again, but this time by using a stratified log-rank test, stratifying on the binary variable `AGE60`. Is the result the same as in the earlier, unstratified log-rank test?

b) Try to explain the differences between the stratified and unstratified tests by exploring the data. This question is a bit open-ended, and so requires some thought. Are the assumptions for a stratified test met in this example?

c) Suppose you had only the variables for survival, censoring, bulky disease, and the binary variable `AGE60`. What do you think is the best way to analyze the relationship between survival and bulky disease, accounting for the possible confounder `AGE60`? Assume you cannot use a regression model.

```
library(survival)
library(eventtimedata)
data("lymphoma.prognosis")

#adjusting coding of status variable in lymphoma.prognosis dataframe
lymphoma.prognosis$died = lymphoma.prognosis$SURVIVAL - 1
lymphoma.prognosis$died[lymphoma.prognosis$died == 2] = 0

#conduct stratified log-rank test


#subset patients
younger.patients = subset(lymphoma.prognosis, AGE60 == 1)
older.patients = subset(lymphoma.prognosis, AGE60 == 2)

#explore data
```

Problem 4 Solution.

a)

b)

c)

Problem 5: UMARU Impact Study (UIS)

Unit 4, the unit on proportional hazards regression, uses data from the University of Massachusetts AIDS Research Unit (UMARU) IMPACT Study. The data are stored as the dataset `uis` in the package `eventtimedata`.

The study was a 5-year collaborative project comprised of two concurrent randomized trials of residential treatment for drug abuse that enrolled a total of 628 patients.

- *Program A:* Randomized 444 subjects to a 3- or 6-month program of health education and relapse prevention. Clients were taught to recognize "high-risk" situations that are triggers to relapse, and taught skills to cope with these situations without using drugs.
- *Program B:* Randomized 184 participants to a 6- or 12-month program with highly structured lifestyle in a communal living setting.

The main outcome in the study was relapse into drug use during the follow-up period.

Variables in `uis`

`id`	Subject ID (1-628)
`age`	Age in years
`beck`	Beck depression score
`hercoc`	Heroine or cocaine use prior to entry
`ivhx`	IV drug use at admission
`ndrugtx`	Number of previous drug treatments
`race`	Race (`white`, `other`)
`treat`	Treatment assignment (`short`, `long`)
`site`	Treatment program
`los`	Length of stay in treatment (days)
`time`	Time to return to drug use (days)
`censor`	Indicator of drug use relapse (1 = `yes`, 0 = `censored`)

This problem examines possible differences in outcome by intervention group. Unit 4 examines those differences after adjusting for some of the covariates.

a) Plot the survival curves for the times to relapse by treatment group. The variable `treat` is a factor variable and is stored with value 1 for `short` and 2 for `long`.

b) Do the curves appear to come from distributions with proportional hazards? Plot the cumulative hazard functions for the two groups to confirm your answer.

c) Test for significant differences between the curves using the log-rank statistic, using a two-sided test with significance level $\alpha = 0.05$.

d) State precisely the null and alternative hypotheses being tested in part c). Does the significant p-value from the log-rank test reflect a systematic difference between the curves?

e) The short and long treatment groups both consist of a mixture of treatment programs. What might be a better analysis approach than the simple, unstratified log-rank test?

f) What do you notice about the time point at which the survival curves begin to diverge? Speculate on the possible reason for the pattern in the curves.

Problem 5 Solution.

a)

b)

c)

d)

e)

f)