Andrzej Giniewicz

strona prywatna

Report 2, Periodic inspections and interval censored data

Step 0: read assignment description

Read the description for all steps, so you know what you will be doing later. The bigger picture might influence separate steps!

Step 1: choose your group

You should prepare your report in 2 person groups. One 3 person group per class is allowed in case of odd number of students attending.

Step 2: create a generator

Create a generator for periodically inspected lightbulb.

You should assume, that the life-times of lightbulbs is a vector of iid random variables from exponential distribution with failure rate λ.

When the lightbulb fails, you replace it with a new one. Unfortunately, it is not easy to see the bulb failed, so you are making periodic inspections. The times between inspections is a vector of iid random variables from exponential distribution with rate ν. **Change of a failed lightbulb can occur only at the moment of inspection!** It means, that you will have interval censored data (note, that the last lightbulb in any time window can be interval or right censored).

Your generator should take 3 parameters: λ, ν and T_0 — the time horizon of simulation. Return a list of:

1. interval censored lifetimes of lightbulbs (for analysis),
2. true moments of inspection (a Poisson process, for plotting),
3. true moments of light failures (for plotting, note it is not a Poisson process — why?).

Step 3: make sure your generator works, analyse the process

Make an analysis to answer the following questions (in each case, answer depending on the 3 parameters of generator):

1. what is the percentage of time when there is no light,

2. what is the average number of lightbulb replacements,
3. propose an additional question and answer it.

Step 4: naive estimator

The naive estimator ignores the censoring. Look only at the finite right sides of intervals (assume failures occured during inspections). Estimate the failure rate using an average of such right sides of intervals. Answer the questions:

1. how the mean of such estimator depends on true failure rate and inspection rate,
2. how the variance of such estimator depends on true failure rate and inspection rate,
3. how the bias of such estimator depends on true failure rate and inspection rate,
4. how the mean square error of such estimator depends on true failure rate and inspection rate.

Step 5: taking into account interval censoring

Create your own implementation of Turnbull estimator of survival function. Assume linear interpolation where possible and exponential if needed to receive single representative of Turnbull equivalcence class. Propose estimate of failure rate using the mean calculated from survival function estimator. Answer the questions:

1. how the mean of such estimator depends on true failure rate and inspection rate,
2. how the variance of such estimator depends on true failure rate and inspection rate,
3. how the bias of such estimator depends on true failure rate and inspection rate,
4. how the mean square error of such estimator depends on true failure rate and inspection rate.

Step 6: Prepare the report

The report should contain:

1. Short introduction.
2. Description of generator.
3. Implementation of generator.
4. Two plots of sample realisations of the process.
5. Analysis and answer to step 3 questions.
6. Description of naive estimator.
7. Implementation of naive estimator.
8. Analysis and answer to step 4 questions.
9. Description of Turnbull estimator.
10. Implementation of Turnbull estimator.
11. Analysis and answer to step 5 questions.
12. Include conclusions for your work.
13. Avoid spelling mistakes.
14. Take care for aesthetics of report.
15. Use KnitR, ggplot and LintR.

Also **(possible negative points!)**:

1. Correct structure, with introduction, analysis and (most important) conclusions.
2. No single unnecessary table or figure, which is not referenced in analysis or conclusions („In the figure 1 we see a histogram” does not count — draw some conclusions about the variables you have, do not describe what you see).
3. Correctly label all figures, add legends where needed, make sure they are publication quality.
4. Make sure it compiles without errors.
5. Prefer writing less but well, than a lot and sloppy (look out for less and sloppy if you don't want to score a lot of negative points).

Finally **(possible 10% extra points!)**:

1. Do not limit yourself to the checklist above, amaze me!

Step 7: send your report (deadline: 31.05.2020, 23:59)

Send:

1. the report in .Rnw file,
2. the report in .pdf file,
3. list of R packages to install.

With packages installed, I should be able to copy the files into single place (data and .Rnw) and press compile PDF to obtain same ,pdf file as you.

Do not get late! If you get late $x \in \mathbb{R}_+$ days, we multiply your score by

$$\max\left\{0, 1 - \frac{\lceil x \rceil}{30}\right\}.$$