MATH-449 - Biostatistics
EPFL, Spring 2025
Problem Set 10

1. In this exercise we use data from a study of an intrauterine device (IUD) for prevention [A. V. Peterson, 1975]. One hundred women participated in the study and were followed for a maximum of 300 days. For each woman, the number of days from the insertion of the device to the first of the following events were recorded:

   (i) The IUD is expelled

   (ii) The IUD is removed (for personal or medical reasons)

   (iii) The woman is lost to follow-up (censoring)

   The recorded times of these events are given in the table below:

IUD expelled	IUD removed	Censoring
2	14	86
8	21	203
10	27	207
25	40	
28	42	
28	92	
32	110	
63	147	
86	148	
158	165	
	166	
	178	
	183	
	272	
	272	
	288	
	288	
	288	
	297	

   (a) Describe the competing risks in this situation.

   (b) Compute and plot the Nelson–Aalen estimates $\widehat{A}_{0h}(t)$ for the cumulative cause-specific hazard rates for causes $h = 1, 2$ and interpret the plots. (Use `R`)

   (c) Calculate and plot the cumulative incidence functions $\widehat{P}_{0h}(0, t)$ for causes $h = 1, 2$ and interpret the plots.

   (d) Discuss the interpretation of the estimates in (b) and (c).

2. Let $N(t)$ be a counting process that satisfies the multiplicative intensity model such that the intensity $\lambda(t) = \alpha(t)Z(t)$, where $\alpha(t)$ is a hazard function and $Z(t) = \sum_{i=1}^{n} I(\tilde{T}_i \geq t)$ is a (left-continuous and adapted) predictable process that does not depend on $\alpha(t)$. Consider the null hypothesis

$$\alpha(t) = \alpha_0(t) \text{ for all } t \in [0, \tau],$$

   where $\alpha_0(t)$ is a known function. You will now derive a test for this null hypothesis.

   (a) Let $J(t) = I(Z(t) > 0)$ and $\hat{H}(t) = \int_0^t \frac{J(s)}{Z(s)} dN(s)$, and define $H_0^*(t) = \int_0^t J(s)\alpha_0(s)ds$. Show that $\hat{H} - H_0^*$ is a mean zero martingale when the null hypothesis is true.

(b) Find an expression for the predictable variation process of $\hat{H} - H_0^*$ when the null hypothesis is true. *Hint:* The expression will include $J(t), Z(t)$ and $\alpha_o(t)$.

(c) Consider

$$Q(\tau) = \int_0^\tau L(s)\big(d\hat{H}(s) - dH_0^*(s)\big),$$

where $L(t)$ is a nonnegative predictable process that takes the value 0 whenever $Z(t) = 0$. Show that $Q(\tau)$ is a mean zero martingale when the null hypothesis is true (when considered as a process in τ), and explain why it is reasonable to use $Q(\tau)$ as a test statistic.

(d) Show that

$$\langle Q\rangle(\tau) = \int_0^\tau \frac{L(s)^2}{Z(s)}\alpha_0(s)ds$$

when the null hypothesis holds true, and explain why this is predictable variation process is an unbiased estimator for the variance of $Q(\tau)$ under the null hypothesis.

(e) Give a brief argument why the test statistic $Q(\tau)/\sqrt{\langle Q\rangle(\tau)}$ is approximately standard normally distributed when the null hypothesis holds true.

(f) One possible choice of weight process is $L(t) = Z(t)$. The test that is obtained by using this weight process is called the one-sample log-rank test. Show that, for the one-sample log-rank test, $Q(\tau) = N(\tau) - E(\tau)$, where $E(\tau) = \int_0^\tau Z(s)\alpha_0(s)ds$. Explain why $E(\tau)$ can be interpreted as (an estimate of) the expected number of events under the null hypothesis.

(g) Show that the standardized version of the one-sample log-rank statistic takes the form $\frac{N(\tau)-E(\tau)}{\sqrt{E(\tau)}}$.

3. Read[1] the paper by Greenland and colleagues [Greenland et al., 2016] who described 25 common misinterpretations of p-values. Now read the abstract of a study conducted by [Sharma et al., 2017], which compared 3 groups of men with low testosterone levels:

- Group 1 received testosterone replacement therapy (TRT) resulting in normalization of testosterone levels
- Group 2 received TRT not resulting in normalization of testosterone levels
- Group 3 did not receive TRT

Consider the conclusion of the abstract. Does the text reflect any of the following misinterpretations discussed in Greenland et al?

(a) Misinterpretation #1

(b) Misinterpretation #4

(c) Misinterpretation #12

(d) Misinterpretation #13

Why or why not? (1-2 sentences for each answer)

4. Consider an empirical study that reported a 95% confidence interval (CI). The principal investigator is confused about how to interpret this interval. State whether the following propositions are true or false.

(a) The probability that the 95% CI from our study includes the true parameter is 95%.

(b) If we repeat our study in many random samples from the same population, the 95% CI will include the true parameter in 95% of the samples.

[1]The required readings have been uploaded to Moodle alongside the current Exercise Sheet.

(c) If we published a hundred studies, at the end of our career we expect that, the 95% CI included the true parameter in 95% of our studies.

(d) The p-value for the null hypothesis is the probability that the test hypothesis is true.

(e) The p-value for the null hypothesis is the probability that chance alone produced the observed association.

(f) The p-value for the null hypothesis is the probability of obtaining an estimate at least as far from the null as the estimate we have obtained.

References

Jr A. V. Peterson. Nonparametric Estimation in the Competing Risks Problem. 1975. URL `https://purl.stanford.edu/dh962hh2024`.

Sander Greenland, Stephen J. Senn, Kenneth J. Rothman, John B. Carlin, Charles Poole, Steven N. Goodman, and Douglas G. Altman. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*, 31(4):337–350, April 2016. ISSN 1573-7284. doi: 10.1007/s10654-016-0149-3. URL `https://doi.org/10.1007/s10654-016-0149-3`.

Rishi Sharma, Olurinde A. Oni, Kamal Gupta, Mukut Sharma, Ram Sharma, Vikas Singh, Deepak Parashara, Surineni Kamalakar, Buddhadeb Dawn, Guoqing Chen, John A. Ambrose, and Rajat S. Barua. Normalization of Testosterone Levels After Testosterone Replacement Therapy Is Associated With Decreased Incidence of Atrial Fibrillation. *Journal of the American Heart Association: Cardiovascular and Cerebrovascular Disease*, 6 (5):e004880, May 2017. ISSN 2047-9980. doi: 10.1161/JAHA.116.004880. URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5524065/`.