Màster universitari en Estadística i Investigació Operativa

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH — UNIVERSITAT DE BARCELONA

Lifetime Data Analysis, Course 2021/22

Exercises Topics 5 and 6

Delivery of exercises: December 30, 2021.

Note: You can do the exercises in groups of two students.

Exercise 1 (3 points)

The data frame `tongue` of the R package `KMsurv` (https://CRAN.R-project.org/package=KMsurv) contains the survival times (in weeks) of 80 patients with oral cancer. The objective of this exercise is to study the possible relation of this cancer with the tumour DNA profile, which is either aneuploid[1] or diploid[2].

(a) Use the function `survplot` of the `rms` package to draw both survival curves. What do you observe?

(b) Use the logrank test to test the hypothesis that survival is not related to the tumour type. Which are the null and the alternative hypotheses? Comment on the result.

(c) Fit the log-logistic regression model with the single covariate 'Tumour type' and use now this model to test the same hypothesis as in **b**. Which are the null and the alternative hypotheses in the framework of this model? Comment on the result.

(d) Estimate the odds ratio and the acceleration factor and give an interpretation of both measures.

(e) Assess the goodness of fit of the model graphically using the residuals of the log-logistic regression model.

Exercise 2 (1.5 points)

Following, you have to generate survival times from a log-normal distribution and to check whether the Weibull or the log-logistic distribution fit better to the data. For this purpose, do the following:

(a) Use function `rlnorm` to generate 300 survival times (random variable T) from a lognormal distribution with parameters $\mu = 2$ and $\sigma = 1$.

(b) Generate 300 censoring times (random variable C) from an exponential distribution with mean 20.

(c) Create the variables $Y = \min(T, C)$ and $\delta = \mathbf{1}_{\{T \leq C\}}$. Which is the proportion of right-censored survival times?

(d) Draw the cumulative hazard plots for the Weibull and the log-logistic distribution proposed on Slide 48/64 of Topic 5. Which parametric model seems to fit the data better?

[1] https://en.wikipedia.org/wiki/Aneuploidy

[2] https://en.wikipedia.org/wiki/Ploidy#Diploid

Exercise 3 (1.5 points)
We assume that the model assumptions of the Cox proportional hazards model hold for a continuous survival time T and several covariates $\boldsymbol{Z} = (Z_1, \ldots, Z_p)'$, i.e.,

$$\lambda(t; \boldsymbol{z}) = \lambda_0(t) \exp(\boldsymbol{\beta}'\boldsymbol{z}),$$

but that the survival times are grouped in the intervals $[0 = a_0, a_1), \ldots, [a_{g-1}, a_g)$. We define the corresponding hazards as conditional probabilities:

$$\lambda_j(\boldsymbol{z}) = P(T < a_j | T \geq a_{j-1}; \boldsymbol{z}), \quad j = 1, \ldots, g.$$

Show that the following relation holds:

$$\log\left(1 - \lambda_j(\boldsymbol{z})\right) = \log\left(1 - \lambda_j(\boldsymbol{0})\right) \exp(\boldsymbol{\beta}'\boldsymbol{z}).$$

Exercise 4 (4 points)
The data frame `hodg` of the `KMsurv` package contains the times until relapse or death of 43 lymphoma patients that underwent a bone marrow transplant. For information on the data set, execute the following R commands:

```
> library(KMsurv)
> ? hodg
```

(a) Convert the variables `gtype` and `dtype` into factors.
(b) Draw the survival functions corresponding to the four combinations of graft and disease types measuring time until relapse or death in years. Comment on the graph.
(c) Fit the proportional hazards model that includes graft type, disease type, the interaction of both, and the Karnofsky index. Interpret the model fit.
(d) Estimate the hazard ratios associated to the graft type (comparing autologous to allogenic transplantations) and interpret both values.
(e) Check the proportional hazards assumption.
(f) Concerning the estimation of the four model parameters, are there any influential observations?

Note: For information on autologous and allogenic transplantations, see, e.g., `https://en.wikipedia.org/wiki/Allotransplantation` and `https://en.wikipedia.org/wiki/Autotransplantation`, respectively.