CPH 931 — Fall 2009 — Dr. Charnigo

Written Assignment 4

Written Assignment 4 is due on Friday 06 November at the end of class.

All three exercises pertain to a data set on kidney infection that can be acquired from {http://www.mcw.edu/FileLibrary/Groups/Biostatistics/Publicfiles/DataFromSection/DataFromSectionTXT/Data_from_section_1.4.txt}.

Descriptions of the variables are provided on the website. For convenience, let T denote the time to infection and X the catheter placement.

[30] 1. Consider the Weibull model

$$\log T_i = \alpha + \beta x_i + \sigma W_i,$$

with W_i having the distribution shown on p. 4 of Lecture 8.

[10] a. Which of the following would indicate a greater hazard of kidney infection with surgical catheter placement: $\beta < 0$ or $\beta > 0$? Please give a brief explanation for your answer.

[10] b. Fit the Weibull model. Create a plot of estimated survival functions (or estimated cumulative distribution functions) based on the Weibull model. More specifically, display the estimated survival functions for the following two groups of patients: those with surgical catheter placement and those with percutaneous catheter placement.

[10] c. For patients with percutaneous catheter placement, find the estimated median time to kidney infection. Do the same for patients with surgical catheter placement. Can the relationship between these two estimated median times be described in terms of $\widehat{\beta}$?

Hint: You can answer this with pencil and paper. For each group, set the estimated survival function equal to 0.50 and solve for t.

[30] 2. Consider the log logistic model

$$\log T_i = \alpha + \beta x_i + \sigma W_i,$$

with W_i having the distribution shown on p. 8 of Lecture 8.

[10] a. Which of the following would indicate a greater hazard of kidney infection with surgical catheter placement: $\beta < 0$ or $\beta > 0$? Please give a brief explanation for your answer.

[10] b. Fit the log logistic model. Create a plot of estimated survival functions (or estimated cumulative distribution functions) based on the log logistic model. More specifically, display the estimated survival functions for the following two groups of patients: those with surgical catheter placement and those with percutaneous catheter placement. Compare the results to those obtained in exercise 1b.

[10] c. For patients with percutaneous catheter placement, find the estimated median time to kidney infection. Do the same for patients with surgical catheter placement. Can the relationship between these two estimated median times be described in terms of $\widehat{\beta}$? Compare the results to those obtained in exercise 1c.

[40] 3. Consider the familiar proportional hazards model from CPH 930,

$$\log h_{x_i}(t) = \alpha(t) + \beta x_i,$$

where $h_{x_i}(t)$ is the hazard of kidney infection at time t.

[10] a. Which of the following would indicate a greater hazard of kidney infection with surgical catheter placement: $\beta < 0$ or $\beta > 0$? Please give a brief explanation for your answer.

[10] b. Fit the proportional hazards model. Create a plot of estimated survival functions (or estimated cumulative distribution functions) based on the proportional hazards model. More specifically, display the estimated survival functions for the following two groups of patients: those with surgical catheter placement and those with percutaneous catheter placement. Compare the results to those obtained in exercises 1b and 2b.

[10] c. Test the proportional hazards assumption.

[10] d. For this data set, with which of the three models (Weibull, log logistic, proportional hazards) do you feel comfortable? With which, if any, do you feel uncomfortable? Please give a brief explanation for your answer.

CPH 931 — Fall 2009 — Dr. Charnigo

Written Assignment 4

Written Assignment 4 is due on Friday 06 November at the end of class.

All three exercises pertain to a data set on kidney infection that can be acquired from {http://www.mcw.edu/FileLibrary/Groups/Biostatistics/Publicfiles/DataFromSection/DataFromSectionTXT/Data_from_section_1.4.txt}.

Descriptions of the variables are provided on the website. For convenience, let T denote the time to infection and X the catheter placement.

[30] 1. Consider the Weibull model

$$\log T_i = \alpha + \beta x_i + \sigma W_i,$$

with W_i having the distribution shown on p. 4 of Lecture 8.

[10] a. Which of the following would indicate a greater hazard of kidney infection with surgical catheter placement: $\beta < 0$ or $\beta > 0$? Please give a brief explanation for your answer.

[10] b. Fit the Weibull model. Create a plot of estimated survival functions (or estimated cumulative distribution functions) based on the Weibull model. More specifically, display the estimated survival functions for the following two groups of patients: those with surgical catheter placement and those with percutaneous catheter placement.

[10] c. For patients with percutaneous catheter placement, find the estimated median time to kidney infection. Do the same for patients with surgical catheter placement. Can the relationship between these two estimated median times be described in terms of $\widehat{\beta}$?

Hint: You can answer this with pencil and paper. For each group, set the estimated survival function equal to 0.50 and solve for t.

[30] 2. Consider the log logistic model

$$\log T_i = \alpha + \beta x_i + \sigma W_i,$$

with W_i having the distribution shown on p. 8 of Lecture 8.

[10] a. Which of the following would indicate a greater hazard of kidney infection with surgical catheter placement: $\beta < 0$ or $\beta > 0$? Please give a brief explanation for your answer.

[10] b. Fit the log logistic model. Create a plot of estimated survival functions (or estimated cumulative distribution functions) based on the log logistic model. More specifically, display the estimated survival functions for the following two groups of patients: those with surgical catheter placement and those with percutaneous catheter placement. Compare the results to those obtained in exercise 1b.

[10] c. For patients with percutaneous catheter placement, find the estimated median time to kidney infection. Do the same for patients with surgical catheter placement. Can the relationship between these two estimated median times be described in terms of $\widehat{\beta}$? Compare the results to those obtained in exercise 1c.

[40] 3. Consider the familiar proportional hazards model from CPH 930,

$$\log h_{x_i}(t) = \alpha(t) + \beta x_i,$$

where $h_{x_i}(t)$ is the hazard of kidney infection at time t.

[10] a. Which of the following would indicate a greater hazard of kidney infection with surgical catheter placement: $\beta < 0$ or $\beta > 0$? Please give a brief explanation for your answer.

[10] b. Fit the familiar proportional hazards model. Create a plot of estimated survival functions (or estimated cumulative distribution functions) based on the familiar proportional hazards model. More specifically, display the estimated survival functions for the following two groups of patients: those with surgical catheter placement and those with percutaneous catheter placement. Compare the results to those obtained in exercises 1b and 2b.

[10] c. Now fit an extended proportional hazards model in which catheter placement is permitted to interact with time,

$$\log h_{x_i}(t) = \alpha(t) + \beta x_i + \gamma x_i \log t.$$

Provide a formula for the estimated hazard ratio (hazard for a patient with percutaneous catheter placement divided by hazard for a patient with surgical catheter placement) at time t. Evaluate your formula at $t = 1$ month, $t = 6$ months, and $t = 12$ months.

[10] d. Continuing from part c, test the null hypothesis that catheter placement does not interact with time. Does the outcome of the hypothesis test make you more comfortable or less comfortable with applying the familiar proportional hazards model to this data set?