

Assignment Chapter 2 for Statistics 415, Spring 2005

- 2.2** Start with equation (2.1) and derive the expression for the exponential distribution quantile $t_p = -\theta \log(1 - p)$.
- 2.3** The following table gives survival times for 81 patients that had been diagnosed with melanoma. The observations marked with a $>$ represent patients who were lost to follow-up (e.g., some patients dropped out of the study because they had moved or for some other reasons) or who were still alive at the time the data were analyzed.

136	58	>55	>181	21	23	>190	65
234	>194	14	90	20	130	>213	>215
124	>108	54	98	>193	138	141	110
>67	50	26	103	59	>134	>147	>162
65	40	34	57	>81	>152	>125	>151
34	158	27	>148	27	>132	>140	32
>130	38	85	>129	>100	19	118	53
>140	66	46	37	>50	>114	>124	26
102	>93	>80	60	>86	>21	>44	23
80	>73	19	38	31	25	>76	>13
>16							

- a. Attempt to find an adequate distributional model to describe these data. Suppose that the purpose of the analysis is to present a survival curve to doctors who need to advice their patients. List the steps that you followed in your analysis and present some of the graphical displays that you used along the way.
 - b. Describe what assumptions one needs to make about the censored observations. Discuss the reasonableness of those assumptions for this application.
- 2.4** Nelson (1982) gives failure and censoring times for a turbine fans. The data were collected in order to estimate the failure-time distribution of the fans so that appropriate maintenance policies could be developed. The data are available as a SPLIDA object `Fan.1d`. Attempt to find an adequate distributional model to describe these data. List the steps that you followed in your analysis and present some of the graphical displays that you used along the way.
- 2.5** Provide an intuitive explanation for the reason that likelihood ratio based confidence intervals would be expected to provide a better approximation to the nominal coverage probability than the simpler normal-approximation intervals.
- 2.9** Refer to Section 2.4.3 of the text.
- a. Show how equations (2.14) and (2.15) lead to equation (2.16).
 - b. Explain the implication of this result. As part of your explanation, note which of the quantities in these equations are fixed and which are random.
 - c. Explain the proper interpretation of a confidence interval (or a confidence interval procedure).
- 2.10** Give an intuitive explanation for why the relative likelihood should play such an important role in statistical inference.
- 2.11** Coverage probability is defined as the probability that a confidence interval procedure will contain the quantity that it is supposed to contain. Refer to Example 2.11 in the text. Explain why it is that the coverage probability for the confidence interval for λ is the same as the coverage probability for the confidence interval for θ .
- 2.12** Equation (2.19) gives the correct likelihood assuming that the remission times were rounded to the nearest week.
- a. How should the “correct” likelihood term be written if the reported value of 29 indicated that the patient came out of remission in the 29th week? Draw a simple picture to illustrate.
 - b. For which recording method is using the density approximation $f(29; \theta)$ more accurate? Explain.
 - c. Explain the justification for the approximation in (2.21). Draw a simple picture to illustrate.
 - d. Suggest a “refined” density approximation that might be used in cases where the density approximation cannot be used. Draw a simple picture to illustrate.
- 2.13** Consider the melanoma data given in Exercise 2.3.
- a. Compute $\hat{\theta}$ the ML estimate of the mean of the exponential distribution. Compare with output from an appropriate computer program.

- b. Compute and estimate of the standard error of $\hat{\theta}$.
- c. Compute and compare 95% confidence intervals using the normal approximation methods given in Section 2.43 and the chisquare method [use, specifically, equations (2.17) and (2.12)] and the chisquare method in (2.26). Can you determine which approximation is being used by the computer?