STAT 687 - Homework # 2

Due 0800, 21 July 2014

From the text problem # 3.5

The supplier of an electromechanical control for a household appliance ran an accelerated life test on sample controls. In the test, 25 controls were put on test and run until 30 thousand cycles had been accumulated. Failures occurred at 5, 21,and 28 thousand cycles. The other 22 controls did not fail by the end of the test.

(a) Compute and plot a nonparametric estimate for $F(t)$.

(b) Compute an approximate 95% confidence interval for the probability that an electromechanical device from the same production process, tested in the same way, would fail before 30 thousand cycles. Use the conservative binomial approach.

(c) Compute an approximate 95% confidence interval for the probability that an electromechanical device from the same production process, tested in the same way, would fail before 30 thousand cycles. Use the normal-approximation method based on $Z_{\hat{F}(30)} \sim NOR(0, 1)$

(d) Explain why, in this situation, the approach in part (b) would be preferred to the approach in part (c)

(e) The appliance manufacturer is really interested in the probability of the number of days to failure for its product. Use-rate differs from household to household, but the average rate is 2.3 cycles per day. What can the manufacturer say about the proportion of devices that would fail in 10 years of operation (the expected technological life of the product)?

(f) Refer to part (e). Describe an appropriate model to use when use-rate varies in the population of units. To simplify, start by assuming that there are only two differnt use-rates. Discuss, using appropriate expressions

From the text problem # 3.7

Consider the Plant 1 heat exchanger data in Figure 3.1

(a) Write the likelihood for these data in terms of π_1, π_2, π_3, and π_4

(b) Write the likelihood for these data in terms of p_1, p_2, p_3, and p_4

From the text problem # 3.18

Explain why the nonparametric estimate of $F(t)$ is a set of points for the heat exchanger data in Example 3.6 but a step-function for the shock absorber data in Example 3.9

From the text problem # 3.23

Consider the relationship $S(t_i) = \exp[-H(t_i)]$, where $H(t)$ is the cumulative hazard function. Note that a nonparametric estimator (based on the product limit estimator) of $H(t)$ without assuming a distributional form is

$$\hat{H}(t_i) = -\sum_{j=1}^{i} \log(1 - \hat{p}_j) \approx \sum_{j=1}^{i} \hat{p}_j = \sum_{j=1}^{i} \frac{d_j}{n_j} = \hat{\hat{H}}(t_i)$$

$\hat{\hat{H}}(t_i)$ is known as the Nelson-Aalen Estimator of $H(t_i)$. Thus $\hat{\hat{F}}(t_i) = 1 - \exp[-\hat{\hat{t}}_i]$ is another nonparametric estimator for $F(ti)$.

(a) Give conditions to assure a good agreement between $\hat{H}(t_i)$ and $\hat{\hat{H}}(t_i)$ and thus between $\hat{F}(t_i)$ and $\hat{\hat{F}}(t_i)$

(b) Use the delta method to compute approximate expressions for $Var[\hat{H}(t_i)]$ and $Var[hat\hat{H}(t_i)]$. Comment on the expression(s) you get.

(c) Compute the Nelson-Aalen estimate of $F(t)$ and compare with the estimate computed in Exercise 3.20. Describe similarities and differences.

(d) Show that $\hat{\hat{H}}(t_i) < \hat{H}(t_i)$ and that $\hat{\hat{F}}(t_i) < \hat{F}(t_i)$.

From the text problem # 4.5

Let $T \sim WEIB(\mu, \sigma), \eta = \exp(\mu)$, and $\beta = 1/\sigma$.

(a) For $m > 0$, show that $E(T^m) = \eta^m \Gamma(1 + m/\beta)$, where $\Gamma(x)$ is the gamma function.

(b) Use the result in (a) to show that $E(T) = \eta\Gamma(1 + 1/\beta)$ and $Var(T) = \eta^2[\Gamma(1 + 2/\beta) - \Gamma^2(1 + 1/\beta)]$.

From the text problem # 4.7

Consider the Weibull $h(t)$. Note that when $\beta = 1, h(t)$ is constant and that when $\beta = 2, h(t)$ increases linearly.

Show that:

(a) If $0 < \beta < 1$, then $h(t)$ is decreasing in t.

(b) If $1 < \beta < 2$, then $h(t)$ is concave increasing.

(c) If $\beta > 2$, then $h(t)$ is convex increasing.

From the text problem # 4.15

The coefficient of variation, γ_2, is a useful scale-free measure of relative variability for a random variable.

(a) Derive an expression for the coefficient of variation for the Weibull distribution

(b) Compute, γ_2 for all combinations of $\beta = .5, 1, 3, 5$ and $\eta = 50, 100$. Also, plot the Weibull pdfs for the same combinations of parameters.

(c) Explain the effect that changes in η and β have on the shape of the Weibull density and the effect that they have on γ_2.

From the text problem # 5.2

In some applications a sample of failure times comes from a mixture of subpopulations.

(a) Write down the expression for the cdf $F(t)$ for a mixture of two exponential distributions with means_ $\theta_1 = 1$ and $\theta_2 = 10$ (subpopulations 1 and 2, respectively) with ζ _being the proportion from subpopulation 1.

(b) For $\zeta = 0, .1, .5, .9, 1$, compute the mixture $F(t)$ for a number of values of t ranging between 0 and 30. Plot these distributions on one graph.

(c) Plot $\log(t)$ versus $\log[-\log[1 - F(t)]]$ for each $F(t)$ computed in part (b). Comment on the shapes of the mixtures of exponential distributions, relative to a pure exponential distribution or a Weibull distribution.

(d) Plot the hazard function $h(t)$ of the mixture distributions in part (b).

(e) Qualitatively, what do the Weibull plots in part (c) suggest about the hazard function of a mixture of two exponential distributions?

From the text problem # 5.4

Refer to Exercise 5.1. Show that a mixture of two exponential distributions with different θ values will always have a decreasing hazard function.

From the text problem # 5.12

Let $T_{(1)}$ denote the minimum of m independent Weibull random variables with parameters $\mu_i, i = 1, ..., m$, and constant σ. Show that $T_{(1)}$ has a Weibull distribution.