Solution to Exercise 10.26

⋆10.26 Consider the ML estimator $\hat{\theta}$ from the previous exercise. Explain how you could obtain an asymptotic confidence interval for θ in three different ways. The first should be based on inverting a Wald test in the θ parametrization, the second should be based on inverting a Wald test in the ϕ parametrization, and the third should be based on inverting an LR test.

Generate 100 observations from the exponential distribution with $\theta = 0.5$, find the ML estimate based on these artificial data, and calculate 95% confidence intervals for θ using the three methods just proposed. **Hint:** To generate the data, use uniformly distributed random numbers and the inverse of the exponential CDF.

From the results of the previous exercise, it is easy to see that a Wald statistic for the hypothesis $\theta = \theta_0$ is

$$\frac{n(\hat{\theta} - \theta_0)^2}{\hat{\theta}^2}.$$

If c_α denotes the $1 - \alpha$ quantile of the $\chi^2(1)$ distribution, we can find the limits of a $1 - \alpha$ confidence interval by solving the equation

$$\frac{n(\hat{\theta} - \theta_0)^2}{\hat{\theta}^2} = c_\alpha, \tag{S10.67}$$

as we did in Section 5.2. There are two solutions for θ_0, and these are the upper and lower limits of the confidence interval. Equation (S10.67) can be rewritten as

$$\theta_0^2 - 2\hat{\theta}\theta_0 + \hat{\theta}^2(1 - c_\alpha/n) = 0. \tag{S10.68}$$

By the standard formula for the roots of a quadratic equation, the solutions to equation (S10.68) are

$$\theta_l = \hat{\theta} - n^{-1/2}\hat{\theta}c_\alpha^{1/2} \quad \text{and} \quad \theta_u = \hat{\theta} + n^{-1/2}\hat{\theta}c_\alpha^{1/2}.$$

Notice that θ_l and θ_u are, respectively, $c_\alpha^{1/2}$ standard errors below and $c_\alpha^{1/2}$ standard errors above the estimate $\hat{\theta}$, where the standard error is the square root of

$$\widehat{\mathrm{Var}}_{\mathrm{H}}(\hat{\theta}) = \frac{1}{n}\hat{\theta}^2.$$

This is the right distance to go, because the relationship between the $\chi^2(1)$ and standard normal distributions implies that $c_\alpha^{1/2}$ is the $1-\alpha/2$ critical value of the standard normal distribution. Thus our first confidence interval is

$$\big[\hat{\theta} - n^{-1/2}\hat{\theta}c_\alpha^{1/2},\ \hat{\theta} + n^{-1/2}\hat{\theta}c_\alpha^{1/2}\big]. \tag{S10.69}$$

Of course, this interval does not make sense if the lower limit is negative, which happens whenever $n < c_\alpha$. For reasonable values of α, however, this can happen only if n is extremely small.

The second confidence interval requires us to invert a Wald test in the ϕ parametrization. From the fact that

$$\widehat{\mathrm{Var}}_{\mathrm{H}}(\hat\phi) = \left(\frac{2n\hat\phi}{\hat\phi^3} - \frac{n}{\hat\phi^2}\right)^{-1} = \tfrac{1}{n}\hat\phi^2,$$

we can see that the Wald statistic has precisely the same form as it does in the θ parametrization, and so the confidence interval for ϕ must be

$$\big[\hat\phi - n^{-1/2}\hat\phi c_\alpha^{1/2},\ \ \hat\phi + n^{-1/2}\hat\phi c_\alpha^{1/2}\big]. \tag{S10.70}$$

Taking the inverse of each of the limits and interchanging them then yields a confidence interval for θ:

$$\big[(\hat\phi + n^{-1/2}\hat\phi c_\alpha^{1/2})^{-1},\ \ (\hat\phi - n^{-1/2}\hat\phi c_\alpha^{1/2})^{-1}\big]. \tag{S10.71}$$

Of course, we can only do this if the lower limit of the interval (S10.70) is positive, which it must be whenever $c_\alpha < n$.

The third confidence interval requires us to invert a likelihood ratio test statistic. When we evaluate the loglikelihood function (10.04) at $\hat\theta$, it simplifies to

$$\ell(\boldsymbol{y}, \theta) = n \log\hat\theta - n.$$

Thus the LR statistic for $\theta = \theta_0$ is

$$2n(\log\hat\theta - \log\theta_0 + \theta_0/\hat\theta - 1).$$

We must solve the equation

$$2n(\log(\hat\theta/\theta_0) + \theta_0/\hat\theta - 1) = c_\alpha \tag{S10.72}$$

to find the two ends of the confidence interval. It appears that we have to solve equation (S10.72) numerically. Since it is an equation in only one unknown, and we know that one solution must be below $\hat\theta$ and one must be above $\hat\theta$, it should not be difficult to do so.

In the second part of the question, readers are asked to draw a single sample and construct the three confidence intervals at the .95 level. The answer they obtain depends on the particular set of random numbers that they use, so there is no single "correct" answer, although there could be many incorrect ones. The file **ar1.data** contains the particular set of 100 observations that we used. Students could be asked to use these data instead of generating their own data.

The data in the file **ar1.data** were generated by using the fact that the CDF of the exponential distribution is $F(y_t, \theta) = 1 - \exp(-\theta y_t)$. We know that $F(y_t, \theta)$ must follow the uniform U(0, 1) distribution. Therefore, if u_t denotes a random variable that follows this distribution, we have

$$u_t = 1 - \exp(-\theta y_t).$$

Solving this equation, we find that

$$y_t = -\frac{1}{\theta}\log(1 - u_t). \tag{S10.73}$$

To generate the sample, we simply generated 100 realizations of u_t from the U(0, 1) distribution and then used equation (S10.73) to transform these into the y_t. Because u_t has the same distribution as $1 - u_t$, we can simply replace $\log(1 - u_t)$ by $\log u_t$ in (S10.73).

Our estimate of θ is $\hat\theta = 0.53175$. Since the .95 quantile of the $\chi^2(1)$ distribution is 3.841459, $n^{-1/2}c_\alpha^{1/2} = 0.19600$, and the first interval (S10.69) is

$$\big[0.53175(1 - 0.196),\ \ 0.53175(1 + 0.196)\big] = \big[0.42753,\ \ 0.63597\big]. \tag{S10.74}$$

The estimate of ϕ is $\hat\phi = 1/0.53175 = 1.88058$. The confidence interval (S10.70) for ϕ is then

$$\big[1.88058(1 - 0.196),\ \ 1.88058(1 + 0.196)\big] = \big[1.51199,\ \ 2.24918\big].$$

Thus the corresponding confidence interval for θ is

$$\big[0.44461,\ \ 0.66138\big]. \tag{S10.75}$$

Both limits of this interval are larger than the corresponding limits of the interval (S10.74).

The third interval is obtained by solving equation (S10.72), with $\hat\theta = 0.53175$, $n = 100$, and $c_\alpha = 3.841459$. The resulting confidence interval is

$$\big[0.43422,\ \ 0.64289\big],$$

which is somewhat closer to the interval (S10.74) than it is to the interval (S10.75). However, like the latter, it is not symmetric around $\hat\theta$.