In lecture 16, we looked at precipitation amounts in Madison County (at Morrisville station). We found that the Weibull distribution had a good fit to the monthly precipitation amounts.

We found that the MLEs for the Weibull distribution were

$$\hat{a} = 2.1871$$
$$\hat{\sigma} = 3.9683$$

and

$$-\mathcal{L}(\{\hat{a}, \hat{\sigma}\}|\mathbf{x}) = 2166.496$$

is the realized negative log-likelihood. Note this means that the log-likelihood is

$$\mathcal{L}(\{\hat{a}, \hat{\sigma}\}|\mathbf{x}) = -2166.496,$$

and the usual likelihood is

$$L(\{\hat{a}, \hat{\sigma}\}|\mathbf{x}) = e^{[\mathcal{L}(\{\hat{a},\hat{\sigma}\}|\mathbf{x})]} \approx= e^{-2166.496},$$

which `R` cannot differentiate from 0.

1. Someone asked "why Weibull?" in class. That is, why wouldn't we use another right-skewed distribution like the Gamma (see Lecture 15), or the Log-Normal (see Lecture 17).

   (a) Compute the MLEs for these data using a Gamma distribution.

   $$\alpha = 4.174581$$
   $$\beta = 1.189099$$

   (b) Compute the MLEs for these data using the Log-Normal distribution.

   $$\sigma = 1.131261$$
   $$\mu = 0.5333417$$

   (c) Compute the likelihood ratio to compare the Weibull and the Gamma distribution.

   $$\text{Likelihood Ratio} = 2.161318e - 07$$

   Which has a better fit according to the likelihood ratio?
   Since the overall result is less than one, the denominator is greater than the numerator. Therefore, the gamma (denominator) has a greater likelihood ratio and is a better fit.

   $$Q = \frac{L(\{\hat{a}, \hat{\sigma}\}|\mathbf{x})}{L(\{\hat{\alpha}, \hat{\beta}\}|\mathbf{x})} = e^{\left[\mathcal{L}(\{\hat{a},\hat{\sigma}\}|\mathbf{x}) - \mathcal{L}(\{\hat{\alpha},\hat{\beta}\}|\mathbf{x})\right]}$$

   (d) Compute the likelihood ratio to compare the Weibull and the Log-Normal distribution.

   $$\text{Likelihood Ratio} = 2.370668e + 16$$

   Which has a better fit according to the likelihood ratio?
   Since the overall result is greater than one, the numerator is greater than the denominator Therefore, the Weibull (numerator) has a greater likelihood ratio and is a better fit.

   $$Q = \frac{L(\{\hat{a}, \hat{\sigma}\}|\mathbf{x})}{L(\{\hat{\mu}, \hat{\sigma}\}|\mathbf{x})} = e^{[\mathcal{L}(\{\hat{a},\hat{\sigma}\}|\mathbf{x}) - \mathcal{L}(\{\hat{\mu},\hat{\sigma}\}|\mathbf{x})]}$$

(e) Compute the likelihood ratio to compare the Gamma and the Log-Normal distribution.

$$\text{Likelihood Ratio} = 1.096862e + 23$$

Which has a better fit according to the likelihood ratio? **Since the overall result is greater than one, the numerator is greater than the denominator. Therefore, the Gamma (numerator) has a greater likelihood ratio and is a better fit.**

$$Q = \frac{L(\{\hat{\alpha}, \hat{\beta}\}|\mathbf{x})}{L(\{\hat{\mu}, \hat{\sigma}\}|\mathbf{x})} = e^{\left[\mathcal{L}(\{\hat{\alpha},\hat{\beta}\}|\mathbf{x}) - \mathcal{L}(\{\hat{\mu},\hat{\sigma}\}|\mathbf{x})\right]}$$

2. Optional Coding Challenge. Choose the "best" distribution and refit the model by season.

   (a) Fit the Distribution for Winter (December-February).

   (b) Fit the Distribution for Spring (March-May).

   (c) Fit the Distribution for Summer (June-August).

   (d) Fit the Distribution for Fall (September-November).

   (e) Plot the four distributions in one plot using `cyan3` for Winter, `chartreuse3` for Spring, `red3` for Summer, and `chocolate3` for Fall. Note any similarities/differences you observe across the seasons.