

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.

Calculating the MLEs results in $\hat{\alpha} = 4.174581$ and $\hat{\beta} = 1.189099$.

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

Calculating the MLEs results in $\hat{\mu} = 1.131361$ and $\hat{\sigma} = 0.5333432$.

- (c) Compute the likelihood ratio to compare the Weibull and the Gamma distribution. Which has a better fit according to the likelihood ratio?

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

Calculating the likelihood ratio yields the result $2.161318e - 07$, which indicates that gamma distribution is a better fit.

- (d) Compute the likelihood ratio to compare the Weibull and the Log-Normal distribution. Which has a better fit according to the likelihood ratio?

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

Calculating the likelihood ratio yields the result $2.37062e + 16$, which indicates that weibull distribution is a better fit.

- (e) Compute the likelihood ratio to compare the Gamma and the Log-Normal distribution. Which has a better fit according to the likelihood ratio?

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

Calculating the likelihood ratio yields the result $1.09684e + 23$, which indicates that gamma distribution is a better fit.

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.