C.1 One Sample

How Long does the Gum Last?

Problem:

We want to determine the flavor duration, in minutes, for a certain brand of chewing gum. However, some of our data values have been censored at 40 minutes. How can we analyze this data?

Background: (Collett (1996); Lang and Secic (1997))

In analyzing the gum flavor duration data, we are analyzing the times to an event (the event that the gum flavor has expired). *Survival analysis* is a common application of time-to-event analysis. Estimates can be obtained of the probability of survival (the event does not occur) as a function of time from a starting point. Any event occurring at the end of some time interval, such as the death of a medical patient, or the failure of a part in a piece of equipment, can be viewed as the event in a survival analysis.

Survival analysis requires special statistical methods due to the fact that the event of interest may not yet have occurred when the data analysis is performed. When data are collected on occurrence times for an event of interest and the data includes events that have not yet occurred, these data are said to be *right-censored*. Survival analysis methodology can incorporate right-censored data. Note that the gum flavor duration values are right-censored (at 40 minutes).

One way to summarize survival data is through estimates of the *survival function*. The survival function, $S(t)$, is the probability that the survival time is greater than or equal to time, t. That is $S(t) = P(\text{survival time} \geq t)$.

The Kaplan-Meier procedure is a *nonparametric* (*distribution-free*) method of estimating survival rates at each point in time. Kaplan-Meier is said to be nonparametric since it does not require specific assumptions to be made about the underlying distribution of the survival times.

Instructions:

1. In the space provided below, record the gum flavor longevity data values that were collected in class. Recall, that '40c' denotes right-censoring at 40 minutes.

2. Enter the data into a statistical software package (be careful and make sure that you have correctly entered the data -- i.e., make sure that you have handled the censored data values properly) and generate the Kaplan-Meier life table. Use complete sentences to interpret the life table. Be sure to state the approximate median flavor duration value.

3. Use a statistical software package to generate a plot of the Kaplan-Meier survival function. Use complete sentences to give an interpretation of the survival function.