

ST511 HW #1

Reading: Browse Chapters 1- 3 of Ott & Longnecker.
See Canvas Calendar for due date.

Local Students: Submit your HW (1) at class, (2) to the file in Statistics 102 or (3) through Canvas in pdf format. Do not slide it under my office door!

Distance Students: Submit your HW through Canvas in pdf format.

12 points total, 2 points per problem part unless otherwise noted.

Note: The data sets from the Ott and Longnecker book can be downloaded from the book's companion website. From the link available from Canvas, download the **ASCII-comma** (CSV, comma delimited) files, then unzip the data. Confusingly, the file extension is .TXT even though the files are actually CSV! The quote option used below is because the column names in the original data are quoted; for example 'Number'.

1. Use the data described in Problem 3.30 regarding lumber. From the files you downloaded above, you will find the data under CH03, named ex3-30.txt. Use the following commands to import and summarize the data.

```
lumber <- read.csv(file.choose(), quote = "'")
View(lumber)
str(lumber)
hist(lumber$Number, col = "grey", main = "Lumber Histogram")
mean(lumber$Number)
median(lumber$Number)
```

- A. Include the histogram in your assignment.
- B. Give the mean and median of the sample.
- C. Does the data appear to be “normal” (bell-shaped)? Justify your response.

2. Use the data described in Problem 3.7 regarding survival times. You will find the data under CH03, named ex3-7.txt. Notice that the two therapies are in two different columns. Use the following code to get you started:

```
survival <- read.csv(file.choose(), quote = "'")
View(survival)
str(survival)
mean(survival$StandardTherapy)
sd(survival$StandardTherapy)
boxplot(survival)
```

- A. What is the sample mean and sample standard deviation for each of the therapies? (4 pts)
- B. Construct side-by-side boxplots showing the survival times for each therapy.