

Describing Distributions with Numbers*

Worksheet 2

1. The University of Arizona Wildcat varsity football team plays 12 regular season games each year. The data set `uafootball` contains the difference in scores for the five seasons from 2104 to 2019. These data can be downloaded using the R command

```
uafootball<-read.csv("http://math.arizona.edu/~jwatkins/uafootball.csv")
```

- (a) How many games did the Wildcats win each year?
 - (b) Give 5 number summaries of the run differences.
 - (c) Draw a side-by-side boxplots of the difference in scores for the five seasons ?
 - (d) What is the mean difference in scores for the five seasons ?
 - (e) What is the standard deviation for the difference in scores for the five seasons?
 - (f) Write a short paragraph to describe how the Wildcat football team has performed over the past five seasons. Use some of the summaries in yourr description.
2. The life span in days of 88 wildtype and 99 transgenic mosquitoes is given in `mosquitoes.txt`. Download these data using

```
mosquitoes<-read.csv("http://math.arizona.edu/~jwatkins/mosquitoes.csv")
```

- (a) Give a summary of the life span of both types of mosquitoes.
- (b) Give side by side box plots of the life span of both types of mosquitoes.
- (c) Show the data for each column (`wildtype` and `transgenic`) using `attach(mosquitoes)` and `wildtype<-na.omit(wildtype)`. (This removes the “not available” data in R) Place on one graph the empirical survival functions by using the command `par(new=TRUE)`. Be sure to give them the same limits for the values on each of the axes.
- (d) Give the Q-Q plot of the two types of mosquitoes. Indicate the median and the first and third quartiles on the graph. (See the `points` command.)
- (e) One genotype of mosquito lives longer, on average, than the other. Explain how this can be seen in the boxplots, in the survival function and on the Q-Q plot.

*©2019 Joseph C Watkins