

KJPipho_PS_4

Krista Pipho

September 22, 2018

Question 1.

Has the storm driven stabilizing selection on the humerus length in this population?

```
# First we open the bumpus data set and save it as 'bumpus'
bumpus <- read.csv("C:\\Users\\xenon\\Desktop\\R Studio 2018\\bumpus.csv")

# Next we will view summary statistics of all humerus length values.
summary(bumpus$length.humerus.in.)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.6590  0.7177  0.7330  0.7319  0.7482  0.7800
```

```
sd_total <-sd(bumpus$length.humerus.in.)
print(sd_total)
```

```
## [1] 0.02307821
```

Considering all of the humerus length values shows us the variation in the population as it existed before the storm. To answer our question we must divide the data by survival outcome.

```
# Here we subset the data into observations from surviving individuals, 'survived', and observations from deceased individuals, 'died'.
survived <-subset(bumpus, bumpus$survival == TRUE)
died <-subset(bumpus, bumpus$survival == FALSE)

# This will print statistics for the surviving subset
print("Survived")
```

```
## [1] "Survived"
```

```
summary(survived$length.humerus.in.)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.6850  0.7260  0.7355  0.7359  0.7512  0.7800
```

```
sd_survived <- sd(survived$length.humerus.in.)
print(sd_survived)
```

```
## [1] 0.02033445
```

```
# This will print statistics for the deceased subset
print("Died")
```

```
## [1] "Died"
```

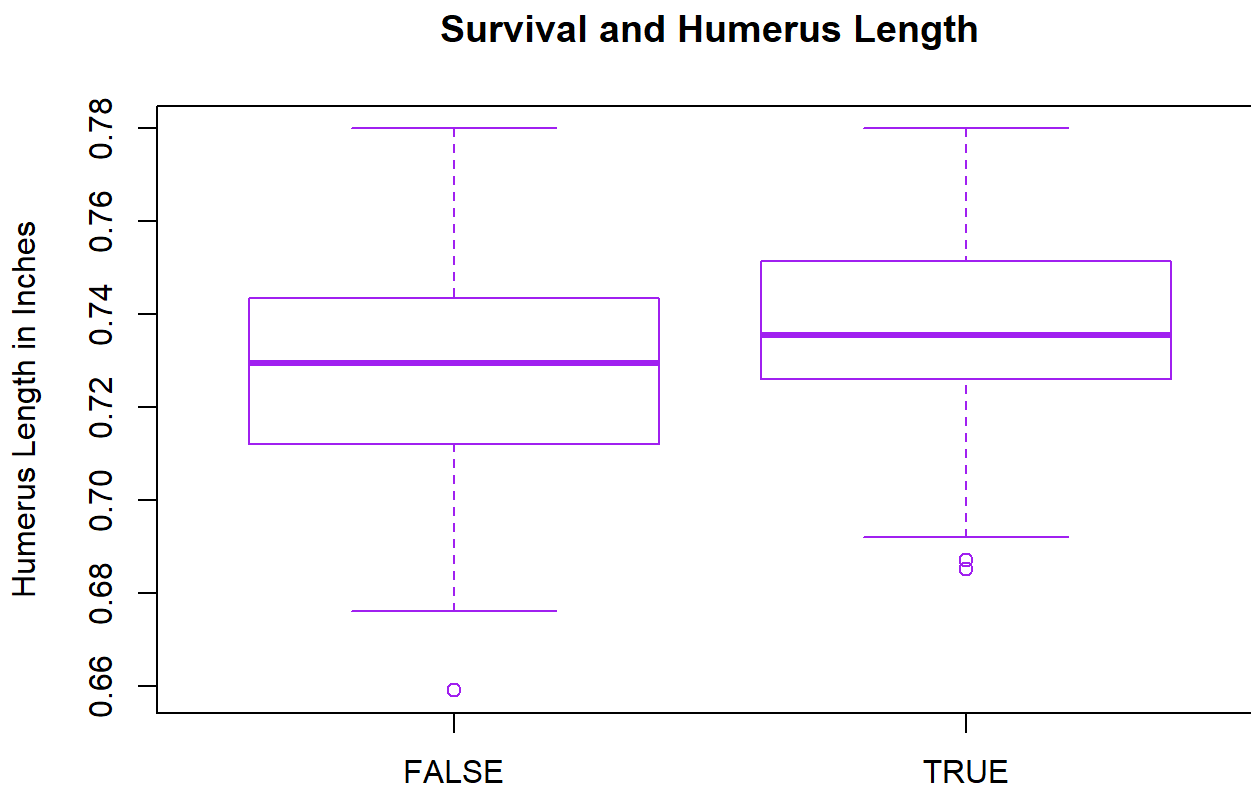
```
summary(died$length.humerus.in.)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.6590  0.7125  0.7295  0.7274  0.7432  0.7800
```

```
sd_died <- sd(died$length.humerus.in.)
print(sd_died)
```

```
## [1] 0.02522651
```

```
# This will create a boxplot comparing surviving and deceased individuals.
boxplot(bumpus$length.humerus.in.~bumpus$survival, main="Survival and Humerus Length", border="purple", ylab = "Humerus Length in Inches")
```



We can apprehend visually that survivors of the storm have a narrower range of humerus lengths than the group of individuals that died, and have a distribution tending more towards longer lengths.

Numerically we see that every summary value is higher for the survivors than for the deceased group, except for standard deviation, which is appreciably lower. The surviving population exhibits a ~ 0.003 inch decrease in standard deviation, which suggests that the storm may have exerted stabilizing selection on humerus length in this population.

Question 2.

Is there a difference in humerus length variance between males and females?

```
# This will compute variance by sex
print("Standard deviation and summary statistics of humerus length for female(f) and male(m) sparrows")
```

```
## [1] "Standard deviation and summary statistics of humerus length for female(f) and male(m) sparrows"
```

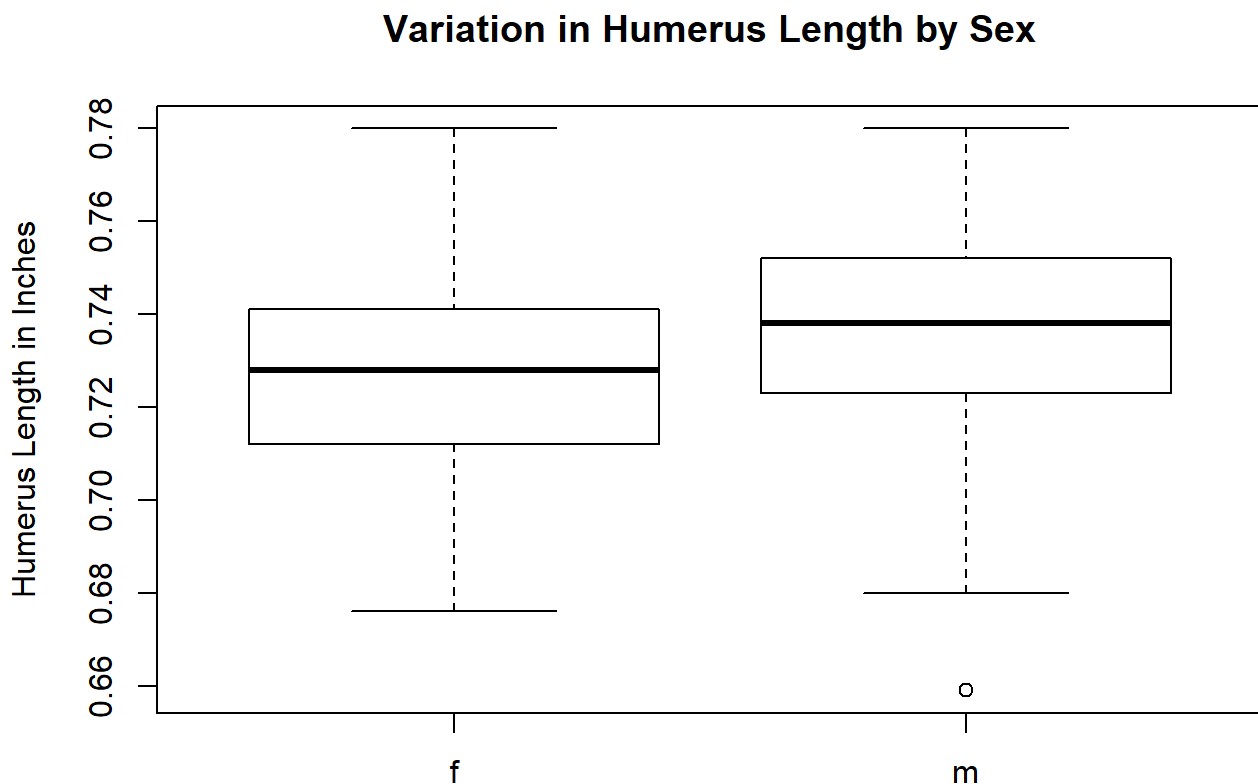
```
tapply(bumpus$length.humerus.in.,bumpus$sex,sd)
```

```
##           f           m
## 0.02215898 0.02324736
```

```
# This will give summary statistics by sex
tapply(bumpus$length.humerus.in.,bumpus$sex,summary)
```

```
## $f
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.676  0.712   0.728   0.727   0.741   0.780
##
## $m
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.6590 0.7230   0.7380   0.7347 0.7520   0.7800
```

```
# This will create a boxplot illustrating humerus length by sex
boxplot(bumpus$length.humerus.in.~bumpus$sex, main="Variation in Humerus Length by Sex", ylab="Humerus Length in Inches")
```



Males appear to have slightly more variance in humerus length and slightly longer humerus lengths on average.

Question 3.

Were juvenile and adult males equally likely to survive the storm?

1. Hypothesis $H_0: p(\text{juvenile males surviving}) = p(\text{adult males surviving})$ $H_a: p(\text{juvenile males surviving}) \neq p(\text{adult males surviving})$
2. Testing Binomial test (assumes two mutually exclusive outcomes, random sampling) Equal Proportions test
3. Alpha
We want an $\alpha = .05$, so we set the confidence interval to .95

```
# Here we find the total numbers of male observations
Num_Male <- sum(bumpus$sex=='m')
Num_AM <- sum(bumpus$age=='a')
Num_JM <- sum(bumpus$age=='y')
print("Total Males")
```

```
## [1] "Total Males"
```

```
print(Num_Male)
```

```
## [1] 87
```

```
print("Total Adult Males")
```

```
## [1] "Total Adult Males"
```

```
print(Num_AM)
```

```
## [1] 59
```

```
print("Total Juvenile Males")
```

```
## [1] "Total Juvenile Males"
```

```
print(Num_JM)
```

```
## [1] 28
```

```
# Here we find the raw numbers of young and adult males that died / survived
AM_Survival <- sum(bumpus$survival==TRUE&bumpus$age=="a")
print("Adult Males Survived ")
```

```
## [1] "Adult Males Survived "
```

```
print(AM_Survival)
```

```
## [1] 35
```

```
JM_Survival <- sum(bumpus$survival==TRUE&bumpus$age=="y")
print("Juvanile Males Survived ")
```

```
## [1] "Juvanile Males Survived "
```

```
print(JM_Survival)
```

```
## [1] 16
```

```
AM_Death <- sum(bumpus$survival==FALSE&bumpus$age=="a")
print("Adult Males Died ")
```

```
## [1] "Adult Males Died "
```

```
print(AM_Death)
```

```
## [1] 24
```

```
JM_Death <- sum(bumpus$survival==FALSE&bumpus$age=="y")
print("Juvanile Males Died ")
```

```
## [1] "Juvanile Males Died "
```

```
print(JM_Death)
```

```
## [1] 12
```

```
# Making sure the categories add up to the total
Check_AJM <- sum(AM_Survival,AM_Death,JM_Survival,JM_Death)
print("Check to see it all adds up")
```

```
## [1] "Check to see it all adds up"
```

```
print(Check_AJM)
```

```
## [1] 87
```

```
# Binomial test on both groups
AM_Prop_bin <- binom.test(AM_Survival,Num_AM,p=0.5,alternative="two.sided",conf.level = 0.95)
print("Proportion Adult males Surviving")
```

```
## [1] "Proportion Adult males Surviving"
```

```
print(AM_Prop_bin)
```

```
##
## Exact binomial test
##
## data: AM_Survival and Num_AM
## number of successes = 35, number of trials = 59, p-value = 0.1925
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.4574519 0.7193038
## sample estimates:
## probability of success
##      0.5932203
```

```
JM_Prop_bin <-binom.test(JM_Survival,Num_JM,p=0.5,alternative="two.sided",conf.level = 0.95)
print("Proportion Juvenile males Surviving")
```

```
## [1] "Proportion Juvenile males Surviving"
```

```
print(JM_Prop_bin)
```

```
##
## Exact binomial test
##
## data: JM_Survival and Num_JM
## number of successes = 16, number of trials = 28, p-value = 0.5716
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.3717936 0.7553761
## sample estimates:
## probability of success
##      0.5714286
```

```
# Equal proportions test on both groups
AM_Prop_prop <-prop.test(AM_Survival,Num_AM,p=0.5,alternative="two.sided",conf.level = 0.95)
print("Proportion Adult males Surviving")
```

```
## [1] "Proportion Adult males Surviving"
```

```
print(AM_Prop_prop)
```

```
##
## 1-sample proportions test with continuity correction
##
## data: AM_Survival out of Num_AM, null probability 0.5
## X-squared = 1.6949, df = 1, p-value = 0.193
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
##  0.4576220 0.7166757
## sample estimates:
##      p
## 0.5932203
```

```
JM_Prop_prop <-prop.test(JM_Survival,Num_JM,p=0.5,alternative="two.sided",conf.level = 0.95)
print("Proportion Juvenile males Surviving")
```

```
## [1] "Proportion Juvenile males Surviving"
```

```
print(JM_Prop_prop)
```

```
##
## 1-sample proportions test with continuity correction
##
## data: JM_Survival out of Num_JM, null probability 0.5
## X-squared = 0.32143, df = 1, p-value = 0.5708
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
## 0.3743185 0.7497305
## sample estimates:
## p
## 0.5714286
```

4. Conclude For the binomial test, the probability of adult male survival, 0.5932203, falls within the 95% confidence interval for juvenile male survival (0.3717936 0.7553761). Likewise the probability of juvenile male survival, 0.5714286, falls within the 95% confidence interval for adult male survival (0.4574519 0.7193038). For the equal probability test, the probability of adult male survival, 0.5932203, falls within the 95% confidence interval for juvenile male survival (0.3743185 0.7497305). Likewise the probability of juvenile male survival, 0.5714286, falls within the 95% confidence interval for adult male survival (0.4576220 0.7166757) Neither of these tests provide us with grounds to reject the null hypothesis, so we accept H_0 .

The p-value obtained from binomial test was 0.1925 for adult males and 0.5716 for juvenile males The p-value obtained from proportions test was 0.193 for adult males and 0.5708 for juvenile males This means that there is no clear trend for which p-value is smaller and better. These tests have disparate underlying mechanics but as expected perform somewhat comparably and are both valid explorations of the data.

Question 4. What is the probability of having 84 or fewer heads in 1000 coin flips?

We will use

```
# This will tell the simulation to repeat 100,000 times
runs <- 100000

# In it the sample function performs 1000 flips. If there are >84 heads 0 is added to the sum. If 85> then
1 is added to the sum. This loop will be repeated 100000 times.
one.trial <- function()
{
  sum((sample(c(0,1),1000,replace=TRUE)) < 85)
}
# We call the function and divide the sum of runs with <85 by the number of trials to get the probability
Toss_Results<- sum(replicate(runs,one.trial()))/runs
Toss_Results
```

```
## [1] 1000
```