Homework 1

Sidney Gerst

January 17, 2020

1. Three reasons why plants or animals might be patchily distributed:
   * Plants don’t have the ability to move away. They may drop seeds that fall directly by where they live forcing themselves to be clumped or patchy. Because they stay in one place, patches will not vary year to year.
   * Animals tend to live with their own species for protection such a herd of american bison. Herds of animals, while they will stay together, will not stay in one are, so patches will vary year to year.
   * Suitable habitat is essential for an ecosystem. Nutrient concentration may be a factor for plants where patched soil concentrations effect where the plants grow. Animals will follow that pattern as grazers will feast on the plant patches and carnivores will stay close to their prey. Patches may vary year to year.

**Eureka Dune Grass**

#Read in Swallenia Data  
swallenia <- read\_csv("Swallenia.csv")  
  
#Perform two sample t-test  
swallenia\_ttest <- t.test(swallenia$count\_2009, swallenia$count\_2010, var.equal = TRUE)  
  
swallenia\_ttest

##   
## Two Sample t-test  
##   
## data: swallenia$count\_2009 and swallenia$count\_2010  
## t = -0.81791, df = 20, p-value = 0.423  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -140.07803 61.16893  
## sample estimates:  
## mean of x mean of y   
## 61.90909 101.36364

1. Based on an = 0.05 the difference in mean abundance of swallenia does not have a significant change between 2009 and 2010 (p = 0.42). This alpha is appropriate because it is a 5% risk of concluding a difference exists when there is no actual difference.

#Paired t-test  
swallenia\_pairedtest <- t.test(swallenia$count\_2009, swallenia$count\_2010, paired = TRUE)  
  
swallenia\_pairedtest

##   
## Paired t-test  
##   
## data: swallenia$count\_2009 and swallenia$count\_2010  
## t = -2.4508, df = 10, p-value = 0.03421  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -75.324830 -3.584261  
## sample estimates:  
## mean of the differences   
## -39.45455

1. The true difference in means of abundance of swallenia between 2009 and 2010 does not equal to 0, so there is a difference in abundance (p = 0.03).
2. The second analysis with the paired t-test is more accurate than a two sample t-test because each plot is not independent of eachother between the two years.
3. Conservation and continuous monitoring efforts should be focused on swallenia dune grass species. Between the years 2009 and 2010 there was a change in dune grass population. Through a Paired t-test calculation, a p-value of 0.03 was found. This means that there is a 3% chance of not finding a difference in the means, low enough to assume there is a difference. More years with data would be effective to show how trends in growth will change.

**Yellowstone Grizzly Bears**

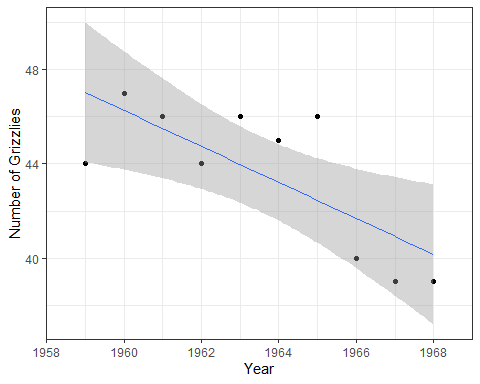
grizzly <- read\_csv("grizzlydata.csv")   
  
#Filter years 1959 to 1968  
grizzly5968 <- grizzly %>%  
 filter(Year >= "1959" & Year <= "1968")  
  
  
#perform linear regression  
grizzly5968\_model <- lm(N ~ Year, data = grizzly5968)  
  
  
grizzly5968\_model

##   
## Call:  
## lm(formula = N ~ Year, data = grizzly5968)  
##   
## Coefficients:  
## (Intercept) Year   
## 1543.0000 -0.7636

# N(grizzlies) = 1543 - 0.7636(Year)  
  
#summary function  
griz59\_sum <- summary(grizzly5968\_model)  
griz59\_sum

##   
## Call:  
## lm(formula = N ~ Year, data = grizzly5968)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.0364 -1.5591 -0.1273 1.5182 3.5455   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1543.0000 472.9413 3.263 0.0115 \*  
## Year -0.7636 0.2409 -3.170 0.0132 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.188 on 8 degrees of freedom  
## Multiple R-squared: 0.5568, Adjusted R-squared: 0.5014   
## F-statistic: 10.05 on 1 and 8 DF, p-value: 0.01319

#standard error: 2.188, p-value: 0.01319  
  
#Visualize the data  
girzzly5968\_plot <- ggplot(grizzly5968, aes(x = Year, y = N)) +  
 geom\_point() +  
 geom\_smooth(method = lm, se = T, size = 0.5)+  
 theme\_bw() +  
 scale\_x\_continuous(expand = c(0,0), limits = c(1958, 1969)) +  
 labs(x = "Year", y = "Number of Grizzlies")  
  
  
girzzly5968\_plot



1. Based on a linear regression, the population of grizzly bears decreases 0.7636 bears per year between the years 1959 and 1968. This model has significant evidence that there is predictability (S = 2.188, p = 0.013).

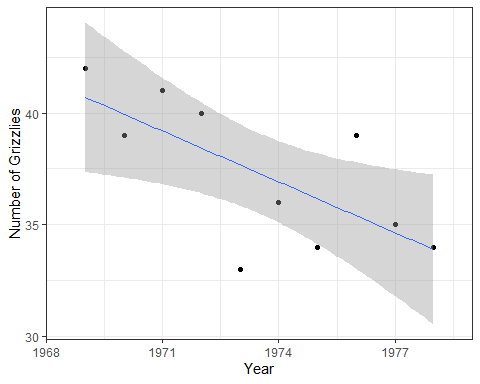
#filter years 1969 to 1978  
grizzly6978 <- grizzly %>%  
 filter(Year >= "1969" & Year <= "1978" )  
  
#perform linear regression   
grizzly6978\_model <- lm(N ~ Year, data = grizzly6978)  
  
grizzly6978\_model

##   
## Call:  
## lm(formula = N ~ Year, data = grizzly6978)  
##   
## Coefficients:  
## (Intercept) Year   
## 1532.3758 -0.7576

#N(grizzlies) = 1532.3758 - 0.7576(Year)  
  
#summary statistics  
griz69\_sum <- summary(grizzly6978\_model)  
griz69\_sum

##   
## Call:  
## lm(formula = N ~ Year, data = grizzly6978)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.6788 -0.9439 0.2303 1.4955 3.5939   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1532.3758 536.3641 2.857 0.0213 \*  
## Year -0.7576 0.2718 -2.787 0.0237 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.469 on 8 degrees of freedom  
## Multiple R-squared: 0.4927, Adjusted R-squared: 0.4293   
## F-statistic: 7.77 on 1 and 8 DF, p-value: 0.02365

#S = 2.469, p = 0.023  
  
#visuallize data  
grizzly6978\_plot <- ggplot(grizzly6978, aes(x = Year, y = N)) +  
 geom\_point() +  
 geom\_smooth(method = lm, se = T, size = 0.5)+  
 theme\_bw() +  
 scale\_x\_continuous(expand = c(0,0), limits = c(1968, 1979)) +  
 labs(x = "Year", y = "Number of Grizzlies")  
  
  
grizzly6978\_plot



#comparing the two models: ttest on the slopes  
grizzly\_ttest <- t.test(grizzly5968$N, grizzly6978$N, var.equal = T)  
grizzly\_ttest

##   
## Two Sample t-test  
##   
## data: grizzly5968$N and grizzly6978$N  
## t = 4.4242, df = 18, p-value = 0.0003275  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 3.308286 9.291714  
## sample estimates:  
## mean of x mean of y   
## 43.6 37.3

1. Based on a linear regression, the population of grizzly bears decreases 0.7576 bears per year between the years 1969 and 1978. This model has significant evidence that there is predictability (S = 2.469, p = 0.023). This is a significantly slower decline than the previous ten years (p = 0.0003).

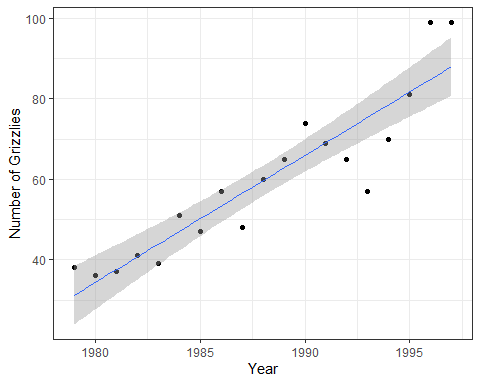
#filter the rest of the data  
grizzly79 <- grizzly %>%  
 filter(Year >= "1979")  
  
#perform linear regression  
grizzly79\_model <- lm(N ~ Year, data = grizzly79)  
  
grizzly79\_model

##   
## Call:  
## lm(formula = N ~ Year, data = grizzly79)  
##   
## Coefficients:  
## (Intercept) Year   
## -6211.288 3.154

# N(grizzly) = -6211.288 + 3.14(Years)  
  
#summary statistics  
griz79\_sum <- summary(grizzly79\_model)  
griz79\_sum

##   
## Call:  
## lm(formula = N ~ Year, data = grizzly79)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -18.4035 -4.0140 0.2947 3.8316 14.1333   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -6211.2877 650.9227 -9.542 3.06e-08 \*\*\*  
## Year 3.1544 0.3274 9.634 2.67e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7.817 on 17 degrees of freedom  
## Multiple R-squared: 0.8452, Adjusted R-squared: 0.8361   
## F-statistic: 92.81 on 1 and 17 DF, p-value: 2.667e-08

#Visualize   
grizzly79\_plot <- ggplot(grizzly79, aes(x = Year, y = N)) +  
 geom\_point() +  
 geom\_smooth(method = lm, se = T, size = 0.5)+  
 theme\_bw() +  
 scale\_x\_continuous(expand = c(0,0), limits = c(1978, 1998)) +  
 labs(x = "Year", y = "Number of Grizzlies")  
  
  
grizzly79\_plot



1. The trend in the population continued to change, but in the other direction from the previous two decades. The population size of grizzlies from 1979 to 1997 increases 3.14 bears per year. The trend in population growth is positive. This model has significant evidence that there is predictability (S = 7.817, p < 0.001).
2. During the 1960’s there was a decline in the population of female grizzly bears. Starting in 1969, the decline started to slow down, where the population of grizzlies did not decline as quickly as they did the decade before. The dump was taken out of the park in 1968. Starting in the year 1979, the trend in grizzly bear population switches from decreasing to increasing. The only the females with cubs are counted in this analysis, so it is safe to say that keeping the dump within the park closed has a positive effect on the female bear population.