4211 Homework 5

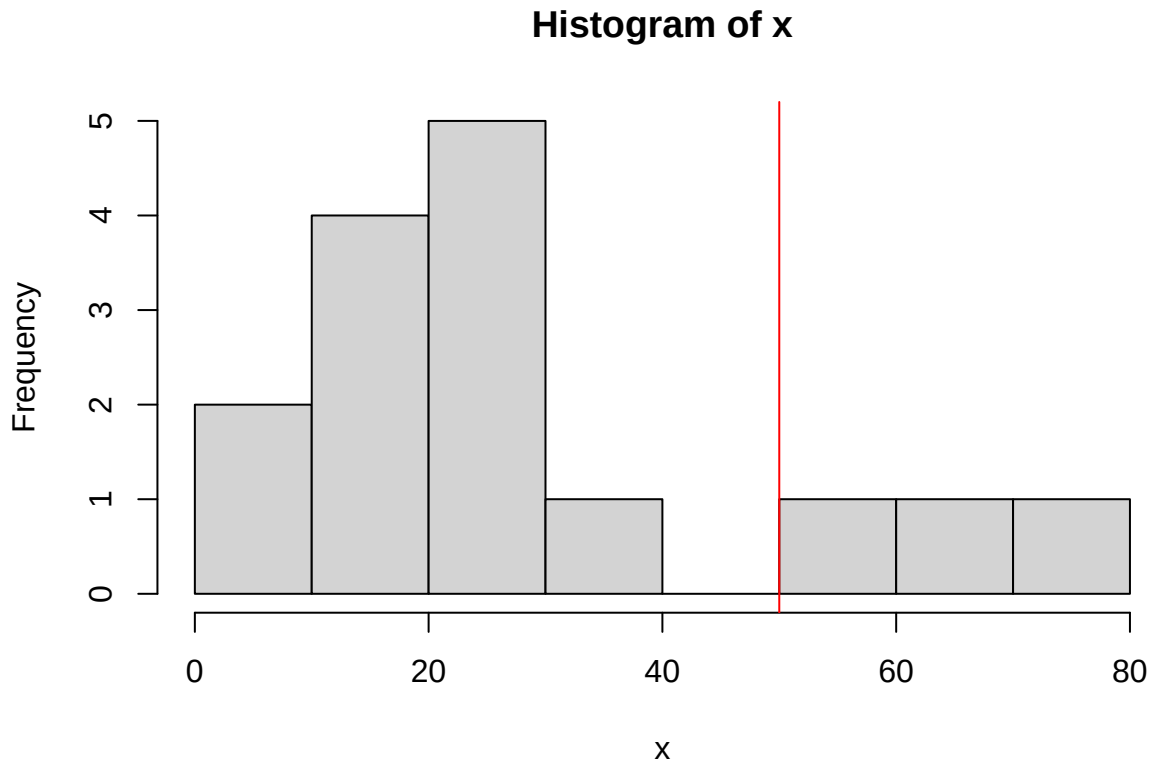
Matthew DeSantis

2023-03-09

1

(a)

```
x = c(19,15,76,23,24,66,27,12,25,7,6,16,51,26,39)
hist(x)
abline(v = 50, col = 'red')
```



(b)

```
teststat = (2/50)*sum(x)
qchisq(0.95, df = 30) #upper
```

```
## [1] 43.77297
```

```
qchisq(0.05, df = 30) #lower
```

```
## [1] 18.49266
```

At the 0.10 level of confidence, there is sufficient evidence to reject the claim that theta = 50.

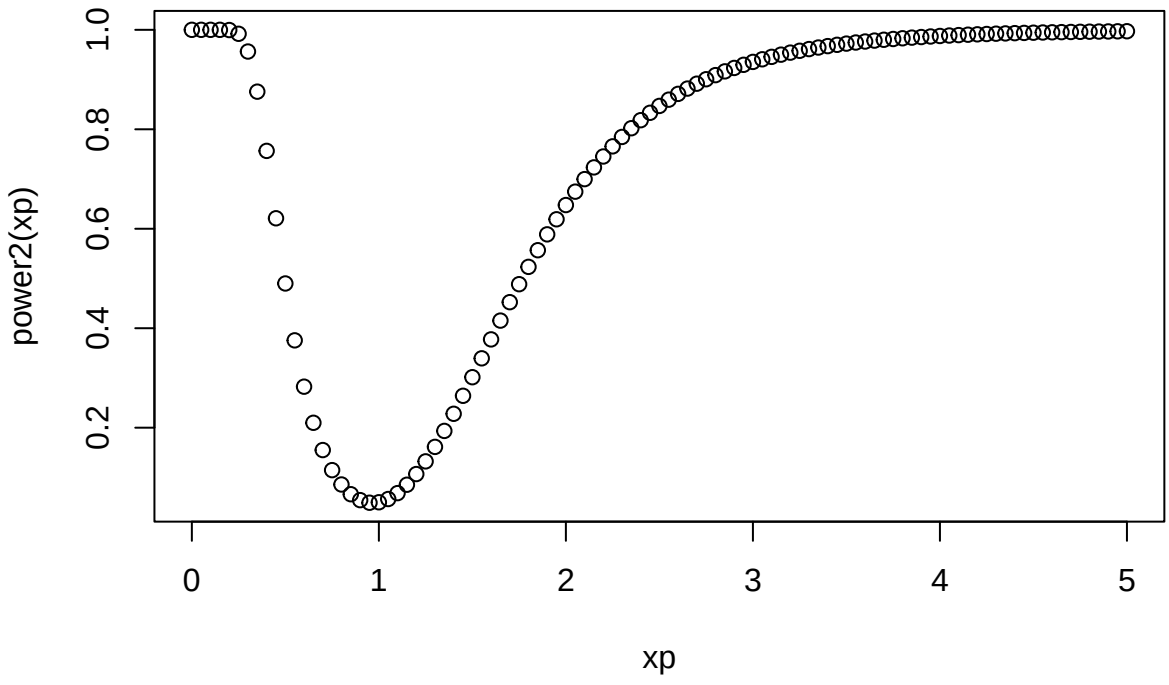
2

```
power2 = function(theta1){
  a = 0.05
  n = 10
  theta0 = 1
  thing = theta0/theta1
  upper = thing*qchisq(1-a/2, df = 2*n)
  lower = thing*qchisq(a/2, df = 2*n)

  power = 1 - (pchisq(upper, df = 20)-pchisq(lower, df = 20))

  return(power)
}

xp = seq(from = 0, to = 5, by = 0.05)
plot(xp, power2(xp))
```



3

(a)

(sqrt(n/xbar)*(xbar-theta0))^2

(b)

```
poiswald = function(x, theta0){
  n = length(x)
  u = mean(x)
  return((sqrt(n/u)*(u-theta0))^2)
}
```

(c)

```
x = c(27,13,21,24,22,14,17,26,14,22,21,24,19,25,15,25,23,16,20,19)
(stat = poiswald(x, 23))
```

```
## [1] 6.90172
```

```
(1 - pchisq(stat, df = 1))*2
```

```
## [1] 0.01722257
```

4

```
salmon <- read.csv("C:/Users/Hyrul/Desktop/School/6 Semester/Stats for DS 2/Homework/HW5/salmon.dat", s
salmon$R = 1/salmon$recruits
salmon$S = 1/salmon$spawners
lm4 = lm(R~S, data = salmon)
summary(lm4)
```

```
##
## Call:
## lm(formula = R ~ S, data = salmon)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -5.776e-04 -2.403e-04 -1.903e-05  1.755e-04  7.166e-04
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.013e-03  8.216e-05   24.50   <2e-16 ***
## S           6.978e-01  1.149e-02   60.72   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0003362 on 38 degrees of freedom
## Multiple R-squared:  0.9898, Adjusted R-squared:  0.9895
## F-statistic:  3687 on 1 and 38 DF,  p-value: < 2.2e-16
```

At the 5% confidence level (but also practically any confidence level, since the pvalues are so small), both beta1 and beta0 are positive.

5

(a)

```
data(sat, package="faraway")
afit = lm(total~ratio+salary+expend, data = sat)
summary(afit)
```

```
##
## Call:
## lm(formula = total ~ ratio + salary + expend, data = sat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -140.911  -46.740   -7.535   47.966  123.329
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1069.234    110.925   9.639 1.29e-12 ***
## ratio          6.330      6.542   0.968   0.3383
## salary        -8.823      4.697  -1.878   0.0667 .
## expend        16.469     22.050   0.747   0.4589
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 68.65 on 46 degrees of freedom
## Multiple R-squared:  0.2096, Adjusted R-squared:  0.1581
## F-statistic: 4.066 on 3 and 46 DF,  p-value: 0.01209
```

From the table, only the intercept is a significant term. Ratio has a positive correlation, salary has a negative, and expend has a positive, though none are significant.

(b)

```
sat$catratio=(cut(sat$ratio,breaks=c(13,16,18,25)))
bfit = lm(total~catratio+salary+expend, data = sat)
summary(bfit)
```

```
##
## Call:
## lm(formula = total ~ catratio + salary + expend, data = sat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -137.103  -47.497    6.392   46.365  136.296
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     1167.083     61.417  19.003   <2e-16 ***
```

```
## catratio(16,18]  -24.838     25.101  -0.990    0.328
## catratio(18,25]  -16.068     34.295  -0.469    0.642
## salary            -4.163      4.216  -0.988    0.329
## expend            -7.314     19.594  -0.373    0.711
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 69.35 on 45 degrees of freedom
## Multiple R-squared:  0.211,  Adjusted R-squared:  0.1409
## F-statistic: 3.008 on 4 and 45 DF,  p-value: 0.02779
```

Here, we treat small ratios (13-16) as the base group. Performing the regression, none of the predictors have significant pvalues, and they are all negative.

(c)

It changes slightly, in that expend is now negative rather than positive, but it doesn't matter much in practice since they are both insignificant models.