# Max Greene HW11

```r
library(tidyverse)
```

## 7.3.3

Could 65,30,55 come from $N(\mu = 50, \sigma = 10)$? Answer using $\chi^2$ distribution.
Hint: Let $z_i = \frac{Y_i - 50}{10}$.

$$U = \Sigma_{i-1}^{\infty} Z_i^2 = \Big(\frac{65-50}{10}\Big)^2 + \Big(\frac{30-50}{10}\Big)^2 + \Big(\frac{55-50}{10}\Big)^2$$

belongs to $\chi^2$ with 3 df.

## 7.3.4

Use the facts that

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2} Var(\chi^2_k) = 2k$$

is a chi square random variable with $(n-1)$ degrees of freedom to prove that:

$$Var(S^2) = \frac{2\sigma^4}{n-1}$$

$$\begin{aligned}
Var(\chi^2_{n-1}) &= Var\Big(\frac{(n-1)S^2}{\sigma^2}\Big) \\
2(n-1) &= \frac{(n-1)^2}{\sigma^4} Var(S^2) \\
\frac{2(n-1)\sigma^4}{(n-1)^2} &= Var(S^2) \\
Var(S^2) &= \boxed{\frac{2\sigma^4}{(n-1)}}
\end{aligned}$$

## 7.3.6

Use the assymptotic normality of $(Y-n)/\sqrt{2n}$ to approximate the 40th percentile of chi square r.v. with $df = 200$.

$$z = \frac{y - 200}{\sqrt{2 * 300}} \approx Norm(0, 1)$$

```r
qchisq(.40,df=200)
```

```
## [1] 194.3193
```

## 7.3.14

Evaluate the following using the Student t distribution:

$$\int_0^{\infty} \frac{1}{1+x^2} dx$$

$$f_{T_n}(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{(n\pi)}\Gamma(\frac{n}{2})(1+\frac{x^2}{n})^{\frac{n+1}{2}}} 1 = \int_{-\infty}^{\infty} f_{T_n}(x)dx = \int_{-\infty}^{\infty} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{(n\pi)}\Gamma(\frac{n}{2})(1+\frac{x^2}{n})^{\frac{n+1}{2}}}dx$$

Now, using $n = 1$,

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} \frac{\Gamma(1)}{\sqrt{\pi}\Gamma(\frac{1}{2})(1+x^2)}dx \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{(1+x^2)}dx = \frac{2}{\pi}\int_{0}^{\infty} \frac{1}{(1+x^2)}dx = \\ &\therefore \int_{0}^{\infty} \frac{1}{(1+x^2)}dx = \boxed{\frac{\pi}{2}} \end{aligned}$$

## 7.4.2

(a):$P(-x \leq T_{22} \leq x) = 0.98$

```r
alpha=1-0.98
qt(c(alpha/2,1-alpha/2),df=22)
```

```
## [1] -2.508325  2.508325
```

(b):$P(T_{13} \geq x) = 0.85$

```r
alpha=1-0.85
qt(alpha,df=13)
```

```
## [1] -1.079469
```

(c):$P(T_{26} < x) = 0.95$

```r
alpha=1-0.95
qt(alpha,df=26)
```

```
## [1] -1.705618
```

(d):$P(T_2 \geq x) = 0.025$

```r
alpha=1-0.025
qt(alpha,df=2)
```

```
## [1] 4.302653
```

## 7.4.4

```r
pt(c(.8,.9),df=8)
```

```
## [1] 0.7765933 0.8027979
```

## 7.4.10

```r
mu <- 8622/83
```

## 7.4.16

(a) Find 95% confidence interval for mean.

```
mu <- 1392.6/336
s <- sqrt((336*10518.84 -1392.6^2)/(336*(336-1)))
ci <- qt(0.975,df=335)
mu - ci*(s/sqrt(335))
```

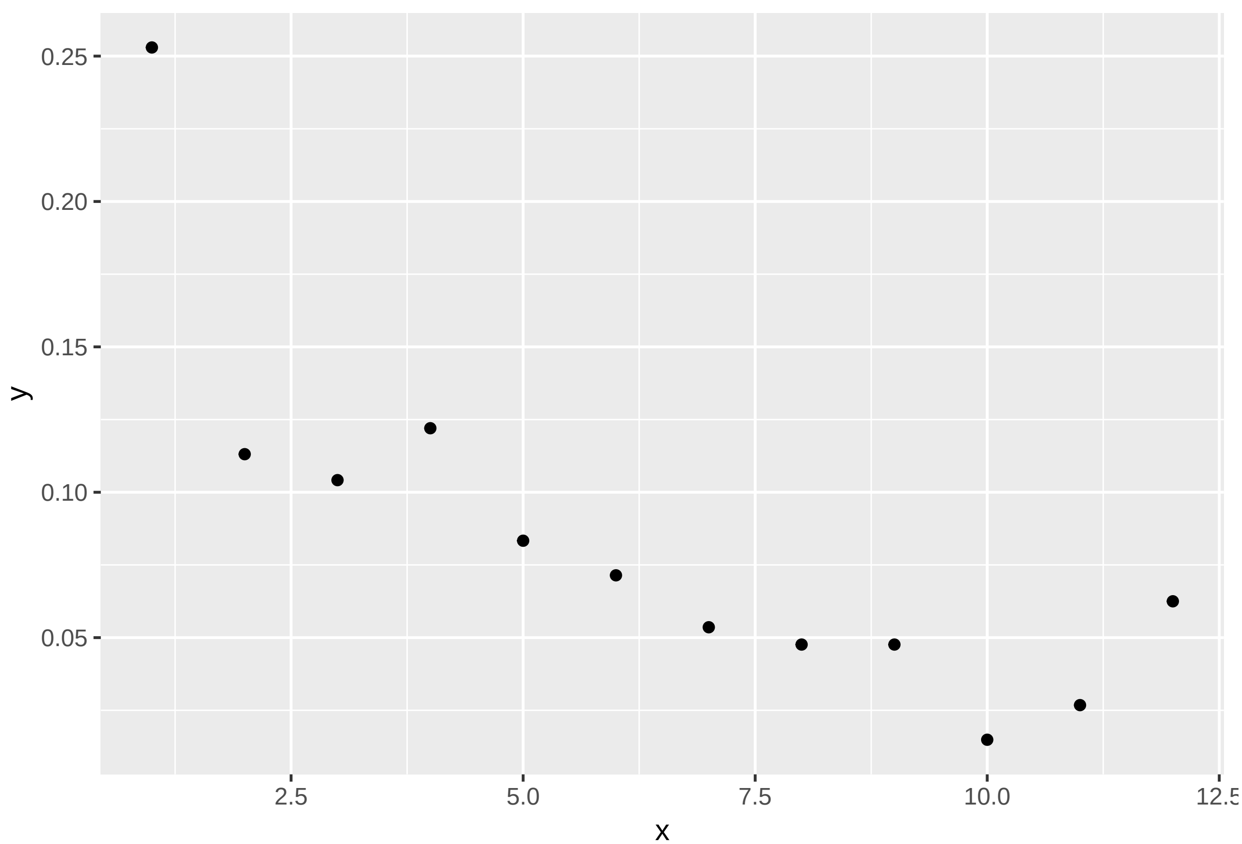
```
## [1] 3.740081
```

```
mu + ci*(s/sqrt(335))
```

```
## [1] 4.549205
```

(b) Is the data representative of a normal distribution sample?

```
n <- 336
probs <- c(85,38,35,41,28,24,18,16,16,5,9,21)/n
ggplot(data.frame(y=probs,x=1:length(probs))) + geom_point(mapping=aes(x=x,y=y))
```




No, this data is not representative of a normal distribution sample.

## 7.4.18

Use theorem 7.4.2. Read case study 5.3.1, does data come from normal distribution?

```
data <- c(141,146,144,141,141,136,137,149,141,142,142,147,148,155,150,144,140,140,139,148,143,143,149,1
y_bar <- mean(data)
mu <- 132.4
n <- length(data)
s <- sqrt((n*sum(data^2)-sum(data)^2)/(n*(n-1)))
```

```r
(t <- (y_bar-mu)/(s/sqrt(n)))
```

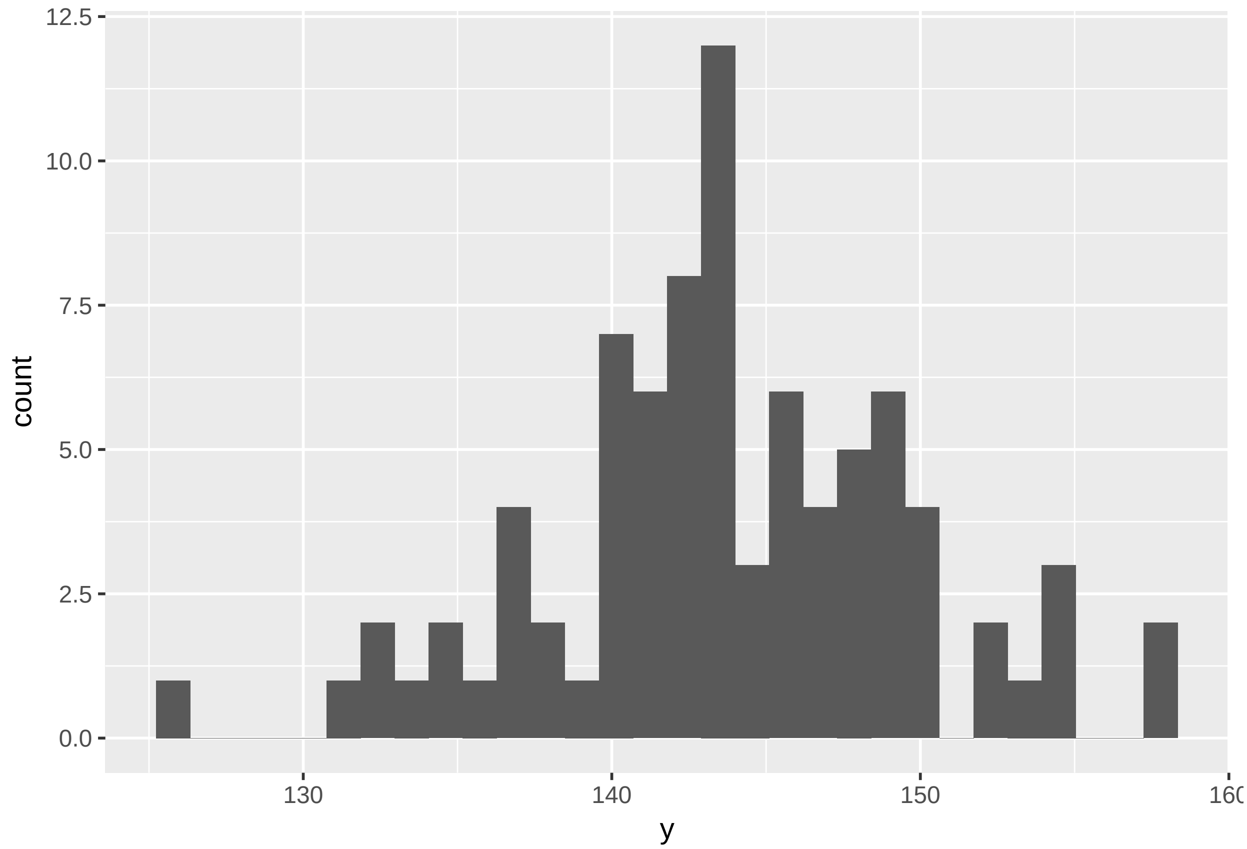
```
## [1] 17.45959
```

```r
qt(c(0.025,0.975),df=n-1)
```

```
## [1] -1.98896  1.98896
```

```r
ggplot(data.frame(y=data)) + geom_histogram(aes(y))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




### 7.4.22

```r
y_bar <- 4.57
mu <- 0
n <- 317
s <- 18.29
(t <- (y_bar-mu)/(s/sqrt(n)))
```

```
## [1] 4.44869
```

```r
qt(0.95,df=n-1)
```

```
## [1] 1.64969
```

## 7.5.2

(a)

```
1-pchisq(8.672,df=17)
```

```
## [1] 0.9499933
```

(b)

```
pchisq(10.645,df=6)
```

```
## [1] 0.9000124
```

(c)

```
pchisq(34.170,df=20)-pchisq(9.591,df=20)
```

```
## [1] 0.9499992
```

(d)

```
pchisq(9.21,df=2)
```

```
## [1] 0.9899983
```

## 7.5.4

(a) $df = 1$

```
pchisq(5.009,df=1)
```

```
## [1] 0.9747841
```

(b) $df = 19$

```
df=19
pchisq(30.144,df=df)-pchisq(27.204,df=df)
```

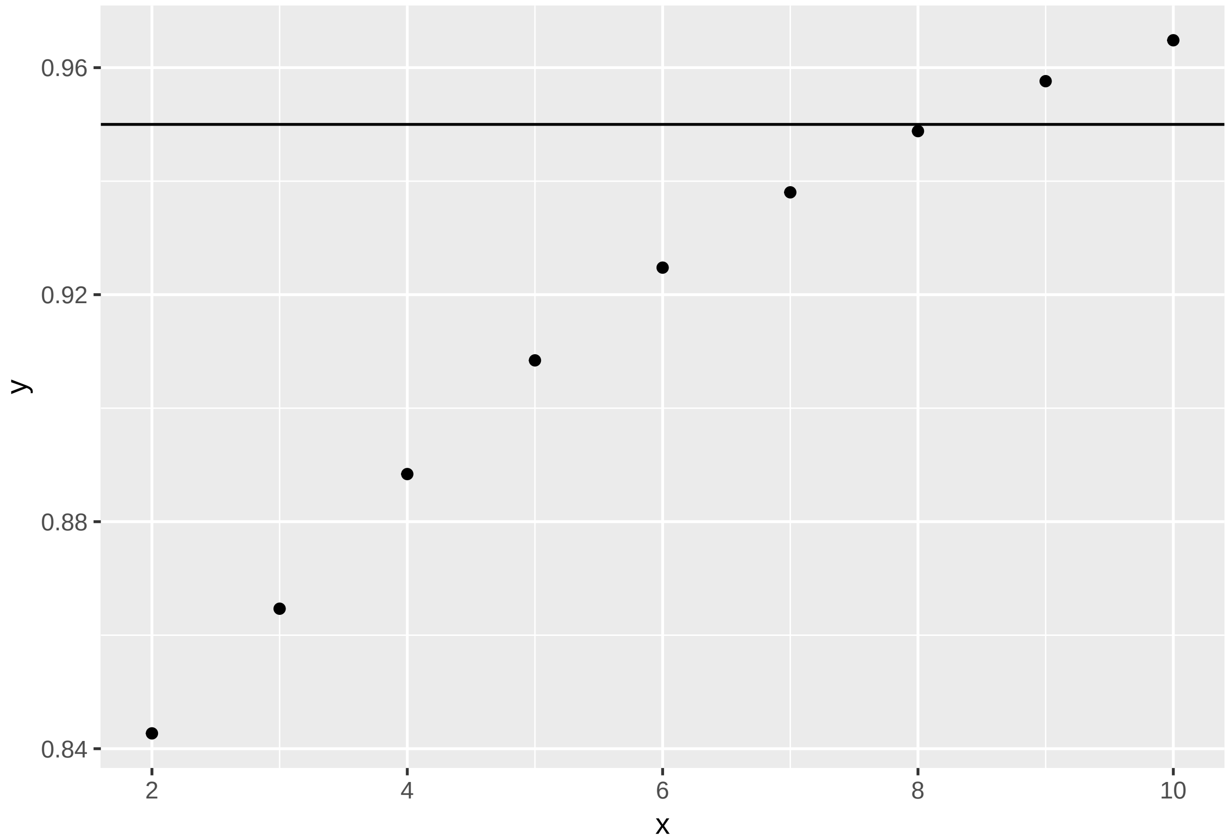
```
## [1] 0.04999621
```

## 7.5.6

Random variable of size $n$ from a normal distribution having mean $\mu$ and variance $\sigma^2$. What is the smallest value of $n$ for which the following is true?

$$P\left(\frac{S^2}{\sigma^2} < 2\right) \geq 0.95 S^2 = \frac{1}{1-n}\Sigma_{i=1}^{n}(Y_i - \bar{Y})^2 \chi^2_{n-1} \approx \frac{(n-1)S^2}{\sigma^2} \Rightarrow \boxed{\frac{1}{n-1}\chi^2_{n-1} \approx \frac{S^2}{\sigma^2}}$$

```
n <- c(2:10)
probs <- pchisq(2*(n-1),df=n-1)
ggplot(data.frame(x=n,y=probs)) + geom_point(mapping=aes(x=x,y=y)) + geom_hline(mapping=aes(yintercept=
```




### 7.5.8

```
n <- 19
sigma2 <- 12
alpha <- 0.05
(a <- qchisq(alpha/2,df=n-1)*(12/(n-1)))
```

```
## [1] 5.487164
```

```
(b <- qchisq(1-(alpha/2),df=n-1)*(12/(n-1)))
```

```
## [1] 21.01759
```