PSYCGR01: Statistics

Statistical inference worksheet

Maarten Speekenbrink *UCL Experimental Psychology*
October, 2016

Remember: questions marked with an asterisk ($*$) are more advanced and optional.

1 General exercises

1. A doctor believes a patient may have contracted a rare but dangerous disease. There is a treatment for the disease, but it has serious side-effects, so the doctor would like to avoid treatment if possible. Based on the severity of the disease and side-effects, she assigns the following utilities:

		Patient	
		no disease	disease
Doctor	no treatment	0	-15
	treatment	-5	0

The doctor estimates the probability that the patient has the disease as $p(\text{disease}) = .2$.

(a) Should the doctor treat the patient?

Answer: We should work out the expected utility of treating and not treating. The expected utility of treating is

$$\begin{aligned} EU[\text{treatment}] &= -5 \times p(\text{no disease}) + 0 \times p(\text{disease}) \\ &= -5 \times (1 - .2) + 0 \times .2 \\ &= -4 \end{aligned}$$

The expected utility of not treating is

$$\begin{aligned} EU[\text{no treatment}] &= 0 \times p(\text{no disease}) + -15 \times p(\text{disease}) \\ &= 0 \times (1 - .2) - 15 \times .2 \\ &= -3 \end{aligned}$$

So, according to the principle of maximum expected utility, the doctor should decide not to treat, as $-3 > -4$.

*(b) Find out what the critical value of $p(\text{disease})$ is at which the doctor should switch her decision.

Answer: The critical value can be worked out from the point at which the doctor will be indifferent, i.e., the point at which the expected utility of the actions is identical. So we solve the following equality for $p(\text{disease})$:

$$
\begin{aligned}
EU[\text{no treatment}] &= EU[\text{treatment}] \\
0 \times p(\text{no disease}) + (-15) \times p(\text{disease}) &= (-5) \times (1 - p(\text{disease})) + 0 \times p(\text{disease}) \\
(-15) \times p(\text{disease}) &= -5 + 5 \times p(\text{disease}) \\
(-15) \times p(\text{disease}) - 5 \times p(\text{disease}) &= -5 \\
(-20) \times p(\text{disease}) &= -5 \\
p(\text{disease}) &= \frac{-5}{-20} = .25
\end{aligned}
$$

So if $p(\text{disease}) = .25$, then the doctor should be indifferent between treating and not treating. If $p(\text{disease}) > .25$, the doctor should decide to treat and if $p(\text{disease}) < .25$ the doctor should decide not to treat.

2. The confidence interval around the parameter β_0 in the simple model

$$Y_i = \beta_0 + \epsilon_i \qquad \epsilon_i \sim N(0, \sigma)$$

can be computed as

$$b_0 \pm \sqrt{F_{1,n-1;\alpha}} \frac{\sqrt{\text{MSE}}}{\sqrt{n}}$$

where b_0 is the estimated value of β_0, $F_{1,n-1;\alpha}$ is critical value of F when $\text{df}_1 = 1$, $\text{df}_2 = n - 1$, and significance is α (e.g., $\alpha = .05$ for a 95% confidence interval). Critical values for the F distribution can be found in statistical tables (in e.g. the appendix of Judd, McClelland & Ryan) or via statistical calculators such as the one at `http://watch.psychol.ucl.ac.uk:3838/shiny_apps/PSYCGR01/CriticalValues`.

A researcher measures IQ scores of 40 students and assumes the following model holds:

$$\text{MODEL C} : Y_i = \beta_0 + \epsilon_i \qquad \epsilon_i \sim N(0, \sigma)$$

The researcher estimates the value of β_0 as $b_0 = 115.14$. Predicting IQ scores with this estimate, the Sum of Squared Error (SSE) is

$$\text{SSE} = \sum_{i=1}^{n} e_i^2 = 5850$$

(a) Compute the 90% confidence interval for b_0.

Answer: There was no assumption made about the variance of ϵ, so we'll estimate it as the Mean Squared Error (this is an unbiased estimate of the population error variance). $\text{MSE} = \frac{\text{SSE}}{n-\text{PC}} = \frac{5850}{40-1} = 150$. The confidence interval can now be computed as $b_0 \pm \sqrt{\frac{F_{1,n-1;\alpha}\text{MSE}}{n}}$. The critical value $F_{1,n-1;\alpha}$ is $F_{1,39;.10} = 2.839$. So the confidence interval is $115.14 \pm \sqrt{\frac{2.839 \times 150}{40}} = 115.14 \pm 3.263$, i.e. the range from

111.877 to 118.403. Note that the critical F value I used is the precise value for the given degrees of freedom. This value may not be in a table of the F distribution. For instance, Exhibit A.3 (page 321) in Judd, McClelland & Ryan (2009) has entries for $\text{df}_2 = n - \text{PA} = 35$ and $\text{df}_2 = 40$, but not $\text{df}_2 = 39$. In that case, it is best to use the value for $\text{df}_2 = 35$. This will give a conservative confidence interval, such that the confidence is *at least* 90%. Using $F_{1,35;.10} = 2.85$ results in a confidence interval $115.14 \pm \sqrt{\frac{2.85 \times 150}{40}} = 115.14 \pm 3.269$. But a better option is, when possible, to obtain the correct value $\text{df}_2 = 39$. For instance, you can use an online calculator such as the one on `https://mspeekenbrink.shinyapps.io/CriticalValues`. However, I recommend you also practice with using the tables, as this is what you'll need to use during the exam.

(b) Compute the 95% confidence interval for b_0.

Answer: $F_{1,39;.05} = 4.091$. Confidence interval is $115.14 \pm \sqrt{\frac{4.091 \times 150}{40}} = 115.14 \pm 3.917$. Or, using $F_{1,35;.05} = 4.12$ from a table, the confidence interval is $115.14 \pm \sqrt{\frac{4.12 \times 150}{40}} = 115.14 \pm 3.931$.

(c) Suppose the researcher actually used a sample of $n = 79$ students. What is the value of the 95% confidence interval now? (use the same value for the SSE).

Answer: $\text{MSE} = \frac{5850}{78} = 75$. $F_{1,78;.05} = 3.963$. The confidence interval is $115.14 \pm \sqrt{\frac{3.963 \times 75}{79}} = 115.14 \pm 1.940$. Or, using a conservative $F_{1,60;.05} = 4.00$ from a table, the confidence interval is $115.14 \pm \sqrt{\frac{4.00 \times 75}{79}} = 115.14 \pm 1.949$.

(d) Suppose the estimated value of β_0 was actually $b_0 = 105.14$. What is the 95% confidence interval for this estimate (assuming $n = 40$ and the same SSE)?

Answer: This is easy, as we have worked out the ingredients already. All we need to do is "shift" the confidence interval, which now becomes 105.14 ± 3.917, or 105.14 ± 3.931 if you use the F value from a table. In other words, the *location* of the confidence depends on the sample mean, but the width depends on the sample variance and sample size.

3. Last week, we looked at the datafile `IQscores.sav`, which contains IQ scores of 100 fictitious participants in an experiment. Let's assume the participants are a random sample from the general population. So, we assume the IQ scores (Y_i) can be modelled as MODEL C:

$$Y_i = 100 + \epsilon_i \qquad \epsilon_i \sim N(0, \sigma) \tag{1}$$

(so ϵ_i is a normally distributed variable with mean 0 and standard deviation σ).

Alternatively, the sample may have come from a special subpopulation, where average IQ is either smaller or larger than 100. This can be implemented as the "augmented" MODEL A:

$$Y_i = \mu + \epsilon_i \qquad \epsilon_i \sim N(0, \sigma) \tag{2}$$

(so ϵ_i is a normally distributed variable with mean 0 and standard deviation σ).

Last week, you computed the Sum of Squared Error (SSE) of MODEL C and of MODEL A when using the sample mean as predicted value.

(a) Compute the Proportional Reduction in Error (PRE) for this MODEL C compared to MODEL A.

Answer: The Sum of Squared Error is $\text{SSE(C)} = 25,988$ for MODEL C and $\text{SSE(A)} = 23,072$ for MODEL A. The PRE is then

$$\begin{aligned}\text{PRE} &= \frac{\text{SSE(C)} - \text{SSE(A)}}{\text{SSE(C)}} \\ &= \frac{25,988 - 23,072}{25,988} \\ &= 0.112\end{aligned}$$

(b) Is the value of PRE computed above "surprising" (i.e., significant at $\alpha = .05$)?

Answer: Exhibit A.1 (page 319) in Judd, McClelland & Ryan (2009) contains a table with critical values for the PRE. We have $\text{PA} - \text{PC} = 1 - 0 = 1$ and $n - \text{PA} = 100 - 1 = 99$. The table does not contain an entry for this last value, so we'll use a conservative criterion (one that is higher). The critical value for $\text{PA} - \text{PC} = 1$, $n - \text{PA} = 80$, and $\alpha = .05$, is .047. As the value we just computed is larger than the critical value, the PRE is significant and we can reject MODEL C in favour of MODEL A.

The distribution of the PRE does not have a "nice" form and the table for the PRE is somewhat limited. Therefore, I would rather test using the F statistic:

$$\begin{aligned}F &= \frac{(\text{SSE(C)} - \text{SSE(A)})/(\text{PA} - \text{PC})}{\text{SSE(A)}/(n - \text{PA})} \\ & \frac{(25988 - 23072)/1}{23072/99} \\ &= 12.51\end{aligned}$$

Using an online calculator (for instance `http://watch.psychol.ucl.ac.uk:3838/shiny_apps/PSYCGR01/Distributions`) we can work out that $P(F_{1,99} > 12.51) = .000617081$, which seems "surprising" under the assumption that the true population mean is 100. Alternatively, we can look up the critical F value (for instance `http://watch.psychol.ucl.ac.uk:3838/shiny_apps/PSYCGR01/CriticalValues`) as $F_{1,99;.05} = 3.9371$, or a conservative value in Exhibit A.3 as $F_{1,80;.05} = 3.96$ and as the value we found is larger, the test is significant and we can reject MODEL C in favour of MODEL A.

*4. In their book, Judd, McClelland and Ryan (2009) define the F^* statistic as

$$F^* = \frac{\text{PRE}/(\text{PA} - \text{PC})}{(1 - \text{PRE})/(n - \text{PA})} \tag{3}$$

where the Proportional Reduction in Error (PRE) is

$$\text{PRE} = \frac{\text{SSE(C)} - \text{SSE(A)}}{\text{SSE(C)}}$$

This statistic is usually just called F (without the *) and defined as

$$\begin{aligned} F &= \frac{(\text{SSE(C)} - \text{SSE(A)})/(\text{PA} - \text{PC})}{\text{SSE(A)}/(n - \text{PA})} \\ &= \frac{\text{MSR}}{\text{MSE}} \end{aligned}$$

Show that these formulations are equivalent.

Answer: We need to show that

$$\begin{aligned} \frac{\text{PRE}/(\text{PA} - \text{PC})}{(1 - \text{PRE})/(n - \text{PA})} &= \frac{(\text{SSE(C)} - \text{SSE(A)})/(\text{PA} - \text{PC})}{\text{SSE(A)}/(n - \text{PA})} \\ \frac{\text{PRE}}{(1 - \text{PRE})}/\frac{(\text{PA} - \text{PC})}{(n - \text{PA})} &= \frac{\text{SSE(C)} - \text{SSE(A)}}{\text{SSE(A)}}/\frac{(\text{PA} - \text{PC})}{(n - \text{PA})} \\ \frac{\text{PRE}}{1 - \text{PRE}} &= \frac{\text{SSE(C)} - \text{SSE(A)}}{\text{SSE(A)}} \end{aligned}$$

We'll do this by rewriting the left hand side, using the formulation of the PRE in terms of the SSE:

$$\begin{aligned} \frac{\text{PRE}}{1 - \text{PRE}} &= \frac{\frac{\text{SSE(C)}-\text{SSE(A)}}{\text{SSE(C)}}}{1 - \frac{\text{SSE(C)}-\text{SSE(A)}}{\text{SSE(C)}}} \\ &= \frac{\frac{\text{SSE(C)}-\text{SSE(A)}}{\text{SSE(C)}}}{\frac{\text{SSE(C)}}{\text{SSE(C)}} - \frac{\text{SSE(C)}-\text{SSE(A)}}{\text{SSE(C)}}} \\ &= \frac{\frac{\text{SSE(C)}-\text{SSE(A)}}{\text{SSE(C)}}}{\frac{\text{SSE(A)}}{\text{SSE(C)}}} \\ &= \frac{\text{SSE(C)} - \text{SSE(A)}}{\text{SSE(A)}} \end{aligned}$$

So yes, the formulations are indeed equivalent. However, in practice, you should **not** compute the F value from the PRE. Instead, you should compute the F directly from the SSE. The problem is that first computing the PRE and then the F from the PRE, you are likely to introduce additional rounding errors. As a result, the F value you calculate can be very far off what it should be.

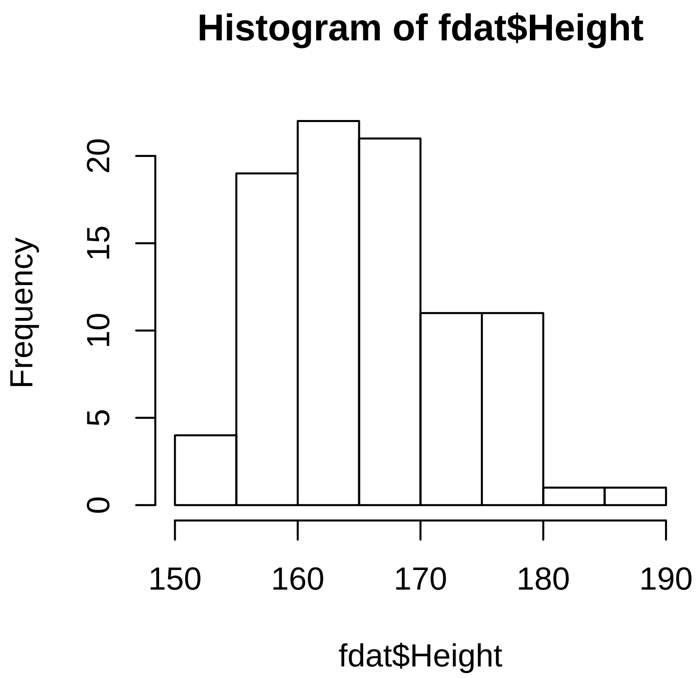
2 Computer exercises

1. The datafile `2016_Questionnaire_1_cleaned.sav` contains the answers of those students who filled in the first questionnaire of this term (and didn't make any obvious mistakes).

(a) Create a histogram for the height of all females in the sample. In 2013, the average height of females over 16 in England was determined to be 164.1 cm.[1] Assess whether the hypothesis that the female students in this course are a random sample from the population of females over 16 in England is tenable. For this exercise, compare the SSE for MODEL C in which you predict the height with the theoretical value to that of MODEL A in which you estimate the height from the sample.

```
> # read the data
> library(foreign)
> dat <- as.data.frame(read.spss("2016_Questionnaire_1_cleaned.sav"))

Warning in read.spss("2016_Questionnaire_1_cleaned.sav"):  2016_Questionnaire
File-indicated value is different from internal value for at least one
of the three system values.  SYSMIS: indicated -1.79769e+308, expected
-1.79769e+308; HIGHEST: 1.79769e+308, 1.79769e+308; LOWEST: -1.79769e+308,
-1.79769e+308

Warning in read.spss("2016_Questionnaire_1_cleaned.sav"):  2016_Questionnaire
Unrecognized record type 7, subtype 18 encountered in system file

> # select the females
> fdat <- subset(dat,Gender == "Female")
> # create a histogram
> hist(fdat$Height)
```

[1] http://www.hscic.gov.uk/catalogue/PUB16076/HSE2013-Ch10-Adult-anth-meas.pdf




```
> # compute SSC
> SSC <- sum((fdat$Height-164.1)^2)
> SSC

[1] 5352.5

> # sample mean of the height
> mean(fdat$Height)

[1] 166.5222

> # compute SSA
> SSA <- sum((fdat$Height-mean(fdat$Height))^2)
> SSA

[1] 4824.456

> # compute F
> sampleF <- (((SSC-SSA)/(1-0))/(SSA/(nrow(fdat)-1)))
> sampleF

[1] 9.741194

> # compute the exceedance probability
> 1-pf(sampleF,1,89)

[1] 0.00242979
```

Answer: Using as MODEL C:

$$Y_i = 164.19 + \epsilon_i \qquad \epsilon_i \sim N(0, \sigma)$$

we get, for the 90 females in the sample, an SSE(C) = 5352.5. The average height of the female students in the sample is $\overline{Y} = 166.52$. Using this value as the predictions of MODEL A, we get SSE(A) = 4824.46. The F statistic is then computed as

$$\begin{aligned} F &= \frac{(\text{SSE(C)} - \text{SSE(A)})/(\text{PA} - \text{PC})}{\text{SSE(A)}/(n - \text{PA})} \\ &\frac{(5352.5 - 4824.46)/1}{4824.46/(90 - 1)} \\ &= 9.741 \end{aligned}$$

with $P(F_{1,89} \geq 9.741) = 0.002$. Thus, the probability of obtaining this F-value (or a more extreme one) is very small under the null hypothesis (MODEL C). As such, it seems unlikely that the null hypothesis is correct, and we reject it. Alternatively, we can use a statistical table to find a (conservative) critical F-value of 3.960. As the F-value we computed is much larger than the critical value, we reject MODEL C. So, the hypothesis that the female students are a random sample from the same population as the English females in 2013 is untenable.

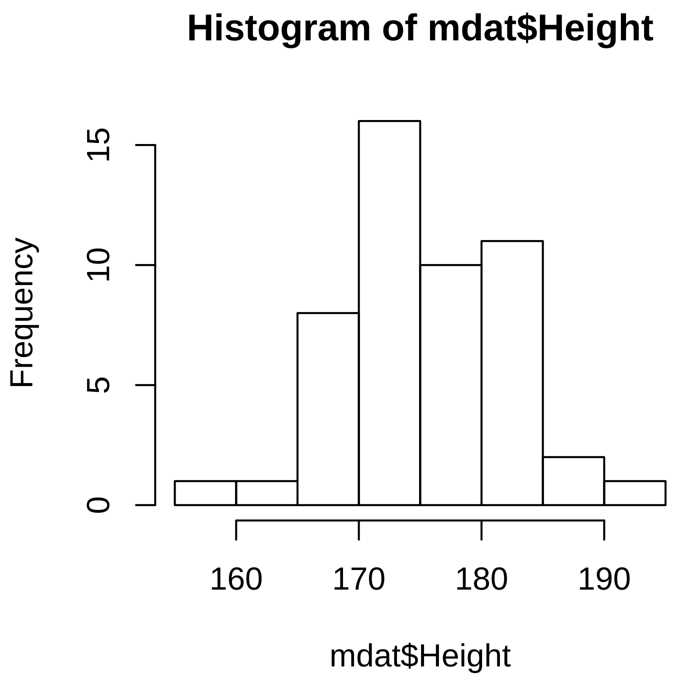
(b) Create a histogram for the height of all males in the sample. In 2013, the average height of males over 16 in England is was determined to be 175.6 cm. Assess whether the hypothesis that the male students are a random sample from the population of males over 16 in England is tenable. Again, compare the SSE for MODEL C in which you predict the height with the theoretical value to that of MODEL A in which you estimate the height from the sample. To compute the SSE for model A, you can use the sample variance, computed as

$$S^2 = \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{n - 1}.$$

Because the top part (numerator) is equal to the SSE of this MODEL A

$$\text{SSE}(A) = (n - 1) \times S^2.$$

```
> # select the males
> # note that there are, for some reason, some spaces in the Male level
> mdat <- subset(dat,Gender == "Male  ")
> # create a histogram
> hist(mdat$Height)
```




```
> # compute SSC
> SSC <- sum((mdat$Height-175.6)^2)
> SSC

[1] 2361

> # sample mean of the height
> mean(mdat$Height)

[1] 176.5

> # compute SSA
> SSA <- sum((mdat$Height-mean(mdat$Height))^2)
> SSA

[1] 2320.5

> # can also compute this from the variance as
> (nrow(mdat)-1)*var(mdat$Height)

[1] 2320.5

> # compute F
> sampleF <- (((SSC-SSA)/(1-0))/(SSA/(nrow(mdat)-1)))
> sampleF

[1] 0.8552036
```

```
> # compute the exceedance probability
> 1-pf(sampleF,1,49)

[1] 0.3596165
```

Answer: Using as MODEL C:

$$Y_i = 175.6 + \epsilon_i \qquad \epsilon_i \sim N(0, \sigma)$$

we get, for the 50 males in the sample, a value of SSE(C) = 2361. The average height in the sample is $\overline{Y} = 176.5$. Using this value as the predictions of MODEL A, we get SSE(A) = 2320.5. Alternatively, we can compute this from the variance ($S^2 = 47.3571429$) as $(50 - 1) \times 47.3571429$. The F statistic is then computed as

$$F = \frac{(\text{SSE(C)} - \text{SSE(A)})/(\text{PA} - \text{PC})}{\text{SSE(A)}/(n - \text{PA})}$$
$$\frac{(2361 - 2320.5)/1}{2320.5/(50 - 1)}$$
$$= 0.855$$

with $P(F_{1,49} \geq 0.855) = 0.36$. Alternatively, we can use a statistical table to find a (conservative) critical F-value of 4.085. As the F-value we computed is smaller than the critical value, we do not reject MODEL C in favour of MODEL A. So, the hypothesis that the male students are a random sample from the same population as the English males in 2013 is tenable.

(c) The tests you performed in part (a) and (b) of this exercise are equivalent to a one-sample t test. This test is based on the t statistic, but for the model comparisons you just performed, $t = \sqrt{F}$.[2] Use this test to check your answers to part (a) and (b). Do you get the same answers?

```
> # t-test for females
> t.test(fdat$Height,mu=164.1)


One Sample t-test

data:  fdat$Height
t = 3.1211, df = 89, p-value = 0.00243
alternative hypothesis: true mean is not equal to 164.1
95 percent confidence interval:
 164.9802 168.0643
sample estimates:
mean of x
 166.5222
```

[2]This relation between the t statistic and F statistic only holds when $\text{df}_1 = 1$ for the F statistic, in which case df_2 is equal to the df of the t statistic.

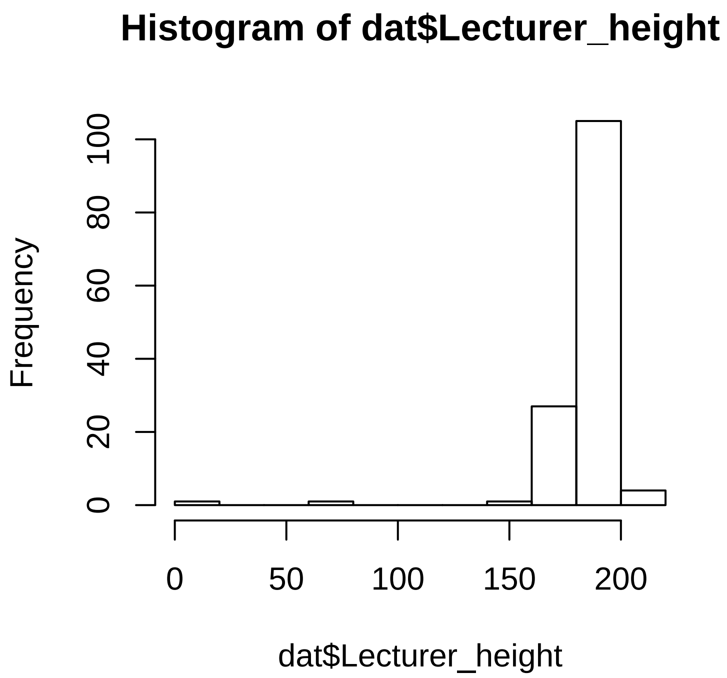
```
> # t-test for males
> t.test(mdat$Height,mu=175.6)


One Sample t-test

data:  mdat$Height
t = 0.92477, df = 49, p-value = 0.3596
alternative hypothesis: true mean is not equal to 175.6
95 percent confidence interval:
 174.5443 178.4557
sample estimates:
mean of x
    176.5
```

Answer: Computing a one-sample t-tests for females, testing against the value of 164.1, gives $t_{89} = 3.121$. This is the same value as $\sqrt{0.855}$; the F test performed here (comparing a model with 0 parameters to one with 1 parameter) is equivalent to a one-sample t-test! We also get the same results in terms of significance, as $p(|t_{89}| \geq 3.121) = p(t_{89} \geq 3.121) + p(t_{89} \leq -3.121) = 0.001 + 0.001 = 0.002$ (note that this is a two-sided test). For the males, we get $t_{49} = 0.925$ which, if we raise this to the power of 2, gives $(0.925)^2 = 0.856$ which is equal to the F value found earlier (up to rounding error).

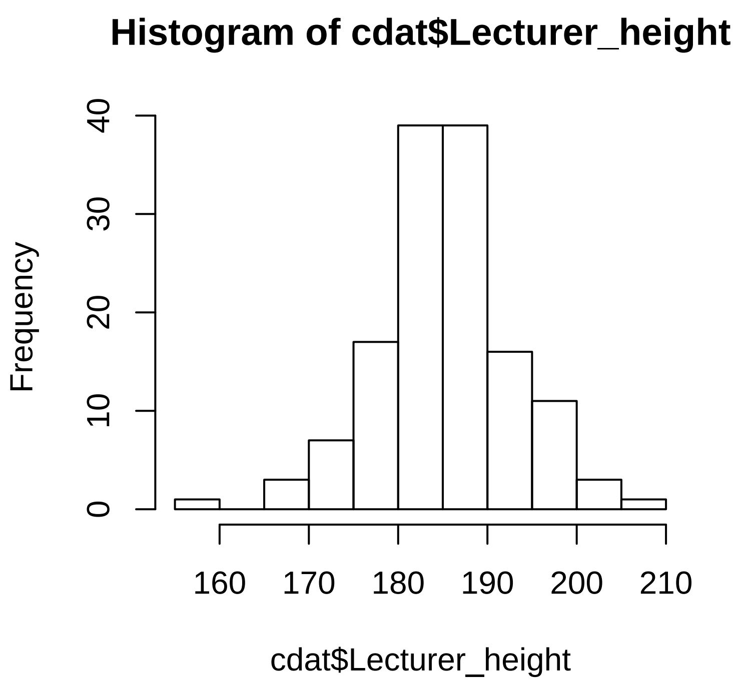
(d) In 1907, Francis Galton famously observed that the crowd at a county fair accurately guessed the weight of an ox when their individual guesses were averaged (the average was closer to the ox's true butchered weight than the estimates of most crowd members, and also closer than any of the separate estimates made by cattle experts).[3] This "wisdom of the crowds" effect has been observed repeatedly for different phenomena and led James Surowiecki to devote a whole book about it in 2004. In the questionnaire, you were asked to guess the height of your lecturer. Create a histogram of the guesses of all respondents in the sample. Maarten's real height is 196 cm. Test the hypothesis that the wisdom of crowds effect is real, i.e., that your guesses were drawn from a Normal distribution in which the mean is equal to the true height of your lecturer (i.e., $\mu = 196$). In this case, would the sample mean, median, or mode have given the best prediction? If you feel like it, you can repeat the analysis for the guesses of his age (Maarten's age is 41 years).

```
> # create a histogram
> hist(dat$Lecturer_height)
```

[3] Galton originally used the median of the crowd's guesses, but later observed that the mean was even closer to the true value.

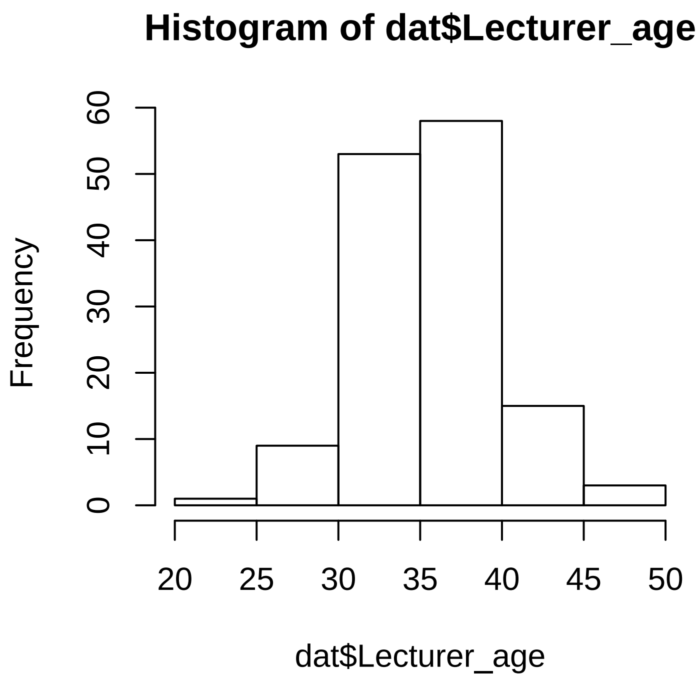



```
> # there are some very low values here, as well as a missing value
> # we'll get rid of these before analysing the data
> cdat <- subset(dat,!is.na(Lecturer_height) & Lecturer_height > 100)
> hist(cdat$Lecturer_height)
```




```
> # As these computations become a bit tedious
> # let's write a function for performing our own F test
> simpleFtest <- function(Y,value) {
+   # 'Y' is the dependent variable
+   # 'value' is the predicted value under model C
+   SSC <- sum((Y-value)^2)
+   SSA <- sum((Y-mean(Y))^2)
+   # compute the PRE
+   PRE <- (SSC-SSA)/SSC
+   # number of observations
+   n <- length(Y)
+   PA <- 1
+   PC <- 0
+   sampleF <- ((SSC-SSA)/(PA-PC))/(SSA/(n-PA))
+   pvalue <- 1-pf(sampleF,df1=PA-PC,df2=n-PA)
+   # use the 'cat' function to write the computed values on the console
+   cat("Simple F test, H0: mu =",value,"\n")
+   cat("sample mean =",mean(Y),"\n")
+   cat("SSC =",SSC," SSA =",SSA,"\n")
+   cat(paste("F(",PA-PC,",",n-PA,") =",sep=""),sampleF,
+       "p-value =",pvalue,"\n")
+   cat("PRE =",PRE,"\n")
+ }
> simpleFtest(cdat$Lecturer_height,value=196)
```

```
Simple F test, H0: mu = 196
sample mean = 186.7308
SSC = 20036.49  SSA = 8265.726
F(1,136) = 193.6701 p-value = 0
PRE = 0.5874664

> # alternatively, we can use a t-test
> t.test(cdat$Lecturer_height,mu=196)


One Sample t-test

data:  cdat$Lecturer_height
t = -13.917, df = 136, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 196
95 percent confidence interval:
 185.4136 188.0480
sample estimates:
mean of x
 186.7308

> # Now for age
> hist(dat$Lecturer_age)
```



```
> # histogram looks ok, but we'll need to get rid of missing values
> cdat <- subset(dat,!is.na(Lecturer_age))
> simpleFtest(cdat$Lecturer_age,value=41)

Simple F test, H0: mu = 41
sample mean = 36.59712
SSC = 5110  SSA = 2415.439
F(1,138) = 153.9469 p-value = 0
PRE = 0.5273114

> # the t-test allows missing values so we can use the original dataset
> t.test(dat$Lecturer_age,mu=41)


One Sample t-test

data:  dat$Lecturer_age
t = -12.408, df = 138, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 41
95 percent confidence interval:
 35.89547 37.29878
sample estimates:
mean of x
 36.59712
```

Answer: After removing the very small value of 1.86 and a missing value, there are $n = 138$ observations left in the data. Predicting the guesses of my height with my actual height of 196 cm gives $\text{SSE(C)} = 20,036.49$, while predicting with the sample mean of 183.0866 gives $\text{SSE(A)} = 8265.73$. The test statistic is thus

$$
\begin{aligned}
F &= \frac{(\text{SSE(C)} - \text{SSE(A)})/(\text{PA} - \text{PC})}{\text{SSE(A)}/(n - \text{PA})} \\
&\frac{(20036.49 - 8265.73)/1}{8265.73/(138 - 1)} \\
&= 193.67
\end{aligned}
$$

with $P(F_{1,137} \geq 193.67) < .001$. So the p-value is smaller than .05, and hence the hypothesis that your guesses are drawn from a distribution with a mean of 196 is untenable. The sample median and mode were 186 and 190 respectively, so, here, the sample mode would have given you the best prediction, although you should remember that sample means, medians, and modes, will vary from sample to sample, so for other samples, e.g. the sample median might have been closer to the true value.

Repeating this for your guesses of my age with a total of $n = 139$ observations (there was again one missing observation) gives $\text{SSE(C)} = 5110$, while predicting with the sample mean of 36.597 gives SSE(A) = 2415.44. The test statistic is thus

$$
\begin{aligned}
F &= \frac{(\mathrm{SSE(C)} - \mathrm{SSE(A)})/(\mathrm{PA} - \mathrm{PC})}{\mathrm{SSE(A)}/(n - \mathrm{PA})} \\
&\frac{(5110 - 2415.44)/1}{2415.44/(139 - 1)} \\
&= 153.947
\end{aligned}
$$

with $P(F_{1,138} \geq 153.947) < .001$. So the p-value is smaller than .05, and hence the hypothesis that your guesses are drawn from a distribution with a mean of 41 is untenable. How flattering :-)