

IEHC0046 BASIC STATISTICS FOR MEDICAL SCIENCES

Analysis of Continuous Data III: Practical

12 October, 2020

In this practical session, we will learn how to use Stata to compare several means and to examine the association between two or more continuous variables. Up to now, we have dealt with the difference between two means and in this session we will extend to the comparison of more than two means.

1. Assess the correlations between two continuous variables (correlation).
2. Examine whether there are differences by age groups in BMI We also need to check the assumption of the equality of variances between the groups (sdtest).
3. Examine whether there are differences in cholesterol and BMI according to physical health which is captured using three groups (anova).

We'll be using the `elsa` dataset again.

```
load("elsa.Rdata")
```

1. Assess the correlations between two continuous variables

Q. Calculate the 95% confidence intervals for the mean of the variables we will be using in this exercise cholesterol, blood CRP and bmi (`chol`, `crp`, `bmi`, `age`)

```
elsa %>%  
  select(chol, crp, bmi, age) %>%  
  map(t.test)
```

```
## $chol  
##  
## One Sample t-test  
##  
## data: .x[[i]]  
## t = 269.34, df = 2474, p-value < 2.2e-16  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## 5.936061 6.023131  
## sample estimates:  
## mean of x  
## 5.979596  
##  
##  
## $crp  
##  
## One Sample t-test  
##  
## data: .x[[i]]  
## t = 28.874, df = 2170, p-value < 2.2e-16  
## alternative hypothesis: true mean is not equal to 0
```

```
## 95 percent confidence interval:
## 3.534959 4.050117
## sample estimates:
## mean of x
## 3.792538
##
##
## $bmi
##
## One Sample t-test
##
## data: .x[[i]]
## t = 331.71, df = 2929, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 27.36890 27.69438
## sample estimates:
## mean of x
## 27.53164
##
##
## $age
##
## One Sample t-test
##
## data: .x[[i]]
## t = 350.33, df = 3128, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 59.78400 60.45697
## sample estimates:
## mean of x
## 60.12049
```

The mean age of this sample is 60.1, the mean BMI is 27.5 kg/m², the mean (total) cholesterol is 6.0 mmol/l and the mean blood CRP level is 3.8 mg/l.

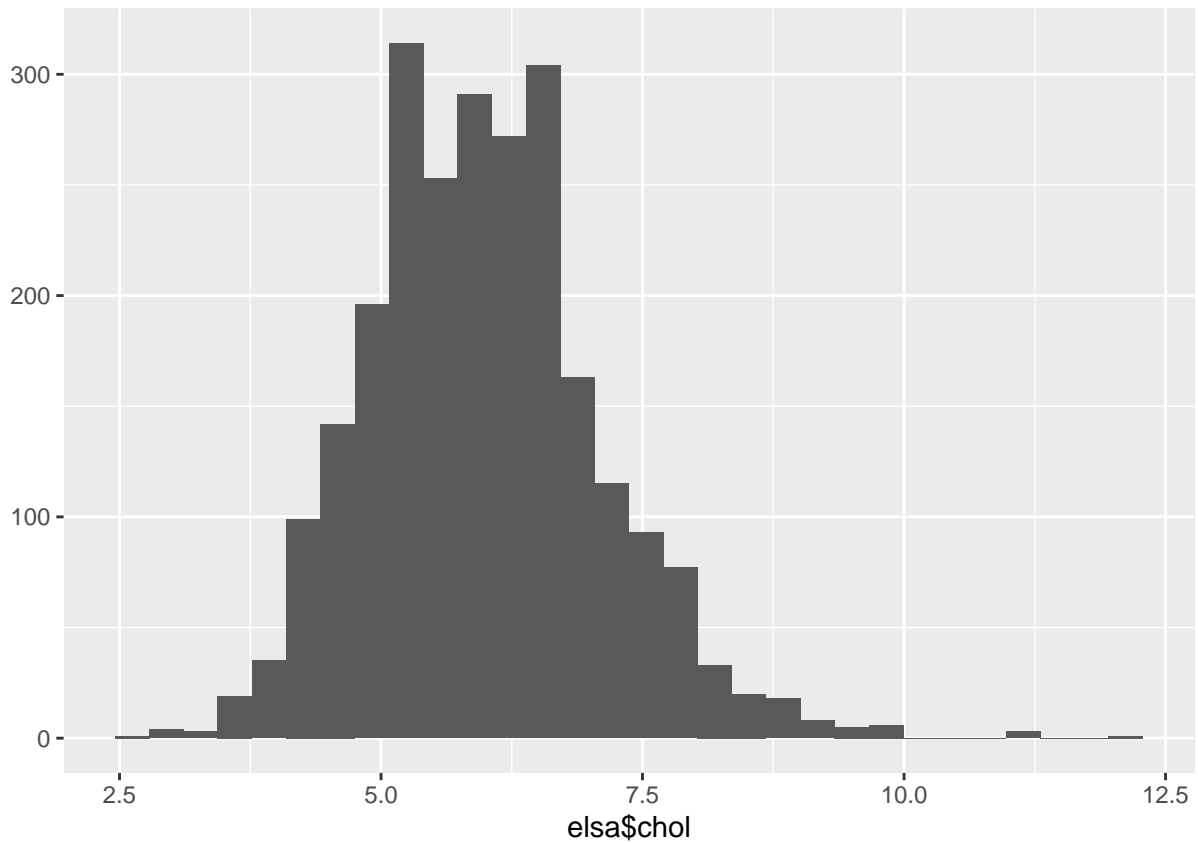
Q. Assess the correlation coefficient for the associations of each pair of variables specified

We are now interested in examining the association between two continuous variables. This is done through the calculation of a correlation coefficient. The correlation is a measure (single number) to assess the association between two variables, i.e. if the values of one variable tend to be higher (or, alternatively, lower) for higher values of the other variable. The Pearson's correlation coefficient (r) is used to assess the association between two continuous variables that either one or both of them are normally distributed.

First let's check the distribution of the three variables, `bmi`, `chol`, and `age`. We can create histograms of the variables using the `qplot()` function from the `ggplot2` packages loaded with the `tidyverse`. If you pass a numeric vector to the `qplot()` function, it produces a histogram by default, by other plots are possible (e.g. boxplots).

```
qplot(elsa$chol)
```

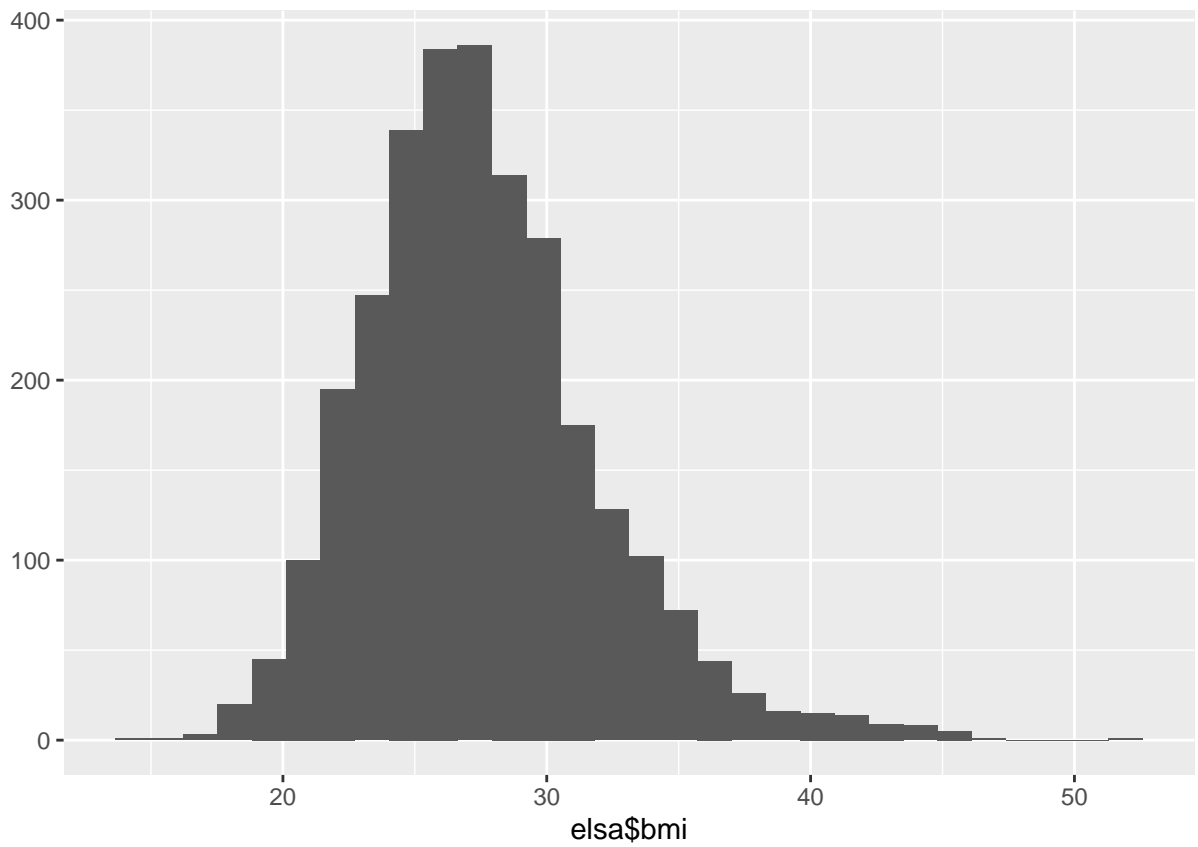
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 654 rows containing non-finite values (stat_bin).
```



```
qplot(elsa$chol)
```

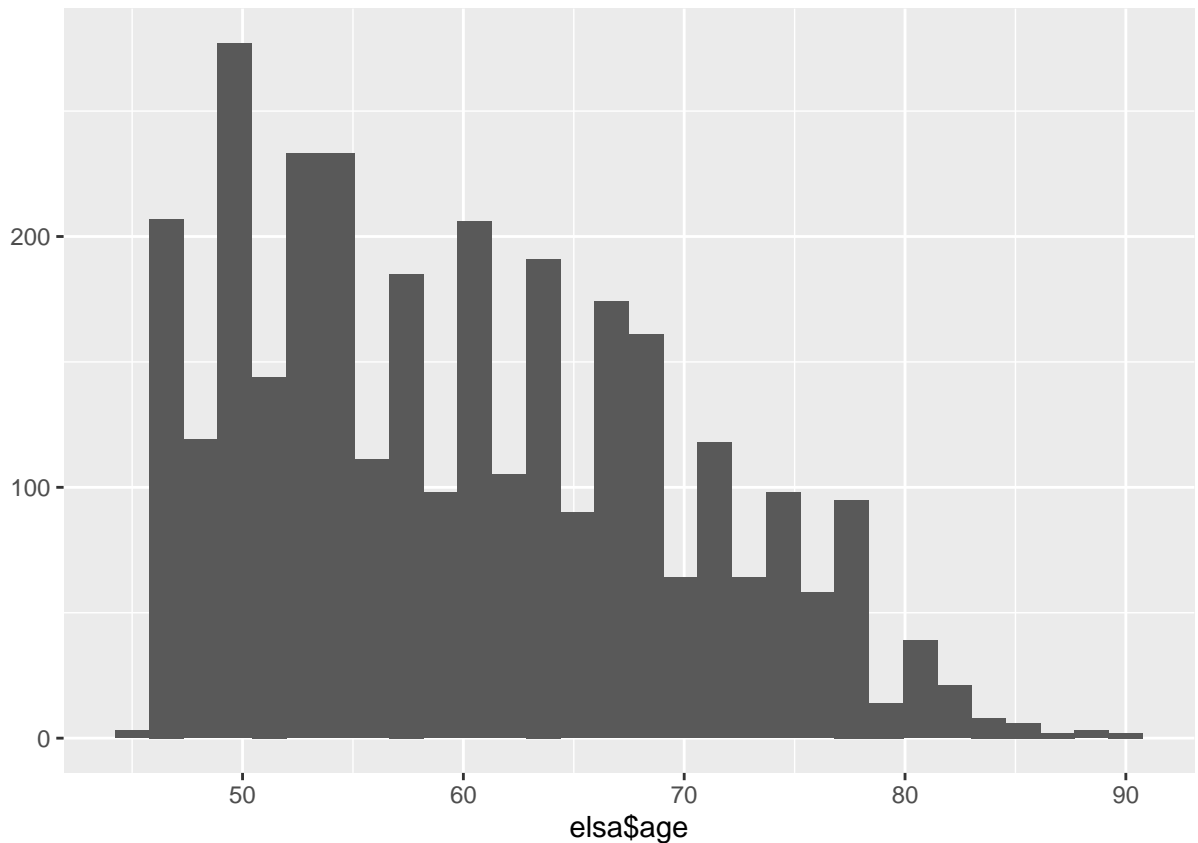
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 199 rows containing non-finite values (stat_bin).
```



```
qplot(elsa$age)
```

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



We can see that although age is not normally distributed, the others are sufficiently enough to enable us to proceed.

We can obtain correlation coefficients for the associations of each pair of variables specified using the function `cor.test()`. The function takes two vectors as inputs and returns the correlation coefficient (r), the 95% CI for r , and p-values for null hypothesis that the true (population) correlation is zero.

```
cor.test(elsa$chol, elsa$age)
```

```
##
## Pearson's product-moment correlation
##
## data:  elsa$chol and elsa$age
## t = 8.6716, df = 2473, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1332852 0.2097638
## sample estimates:
##      cor
## 0.1717833
```

```
cor.test(elsa$bmi, elsa$age)
```

```
##
## Pearson's product-moment correlation
##
## data:  elsa$bmi and elsa$age
## t = 1.3428, df = 2928, p-value = 0.1795
```

```
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.01141464 0.06096390
## sample estimates:
##      cor
## 0.02480714
```

```
cor.test(elsa$chol, elsa$bmi)
```

```
##
## Pearson's product-moment correlation
##
## data:  elsa$chol and elsa$bmi
## t = 4.455, df = 2341, p-value = 8.784e-06
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.05138408 0.13169455
## sample estimates:
##      cor
## 0.09168841
```

Q: Are there any coefficients that are strong, i.e. above 0.7?

We can see that although there were some correlations where the p value was <0.05 the size of these associations were usually very low.

Like in all correlations, the values range between -1 and 1, where -1 is indicative of the perfect negative association and +1 indicates the perfect positive association, while a value of 0 indicates no association whatsoever. By “negative correlation”, we mean a correlation whereby lower values of one variable are associated with higher values of the other variable. On the other hand, a “positive correlation” is when higher values of one variable are associated with higher values of the other variable.

In terms of further interpretation of the value of the correlation coefficient (r), this is in essence a measure of the scatter of the points around an underlying linear trend; the greater the spread of the points the lower the correlation. A useful practical guide can be provided by multiplying the squared value of r with 100 ($100 \cdot r^2$); this gives the percentage of variability of the data that is “explained” by the linear association between the two variables. while a higher absolute value of r , i.e. one that is closer to either +1 or -1, indicates a stronger association, a small r does not necessarily mean lack of association overall. There can be other forms of association, not just linear, between the two variables.

The largest association was between age and cholesterol ($r=0.17$) indicating that 2.9% ($100 \cdot 0.172^2$) of the variability between the two variables is explained by their linear association. This could be because the association between these two variables is actually non-linear, that there is not really a strong association between age and bmi, or that we don’t see a difference when we use age as a continuous variable in this way. It is also important to make a clear conceptual distinction between association and causation. While a high absolute value for a correlation coefficient shows a strong linear association between two variables, this association is not necessarily causal. Establishing causation between two variables requires a judgement that takes into account many other important conditions, not just a strong association.

2. Examining whether there are differences in BMI by age groups

We will look in more detail at the association between age and BMI, creating a new binary variable to do this and then use an unpaired t-test. We will also check the assumption of the equality of variances between the groups using the standard deviation test.

Q. Create a new variable which recodes the variable age into two age groups (50-64 and 65+)

We can use the `case_when()` and `factor()` functions within `mutate()` to create a new categorical age variable.

```

elsa <- elsa %>%
  mutate(age_grps = case_when(age >= 65 ~ "65+", # 65+ if age >= 65
                              age >= 50 & age < 65 ~ "50-64"), # 50-64 is age >= 50 & age < 65
  age_grps = factor(age_grps, levels = c("50-64", "65+"))) # Convert to factor!
by(elsa$age, elsa$age_grps, summary) # Check worked correctly

```

```

## elsa$age_grps: 50-64
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   50.00  53.00   56.00   56.49  60.00   64.00
## -----
## elsa$age_grps: 65+
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   65.00  67.00   71.00   71.67  75.00   90.00

```

Q. Assess the equality of variances assumption for a t-test testing the association between BMI and age.

The default behaviour of the `t.test()` function is to assume unequal variances when conducted an upaired two-sample t-test. We can test this assumption using the `var.test()` function. It has a similar syntax to `t.test()`.

Note, we do not really use this test that often-in practice, because the t-test is very robust and can give trustworthy results even when there is a slight violation of assumptions (i.e. we could have assumed equal variances and got similar t-test results).

```
var.test(bmi ~ age_grps, elsa)
```

```

##
## F test to compare two variances
##
## data:  bmi by age_grps
## F = 1.1585, num df = 1552, denom df = 934, p-value = 0.01277
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  1.031927 1.298383
## sample estimates:
## ratio of variances
##          1.158475

```

Q: Is the assumption of unequal variances correct?

This is a two-sided test of the (null) hypothesis that the variances are equal, the alternative hypothesis is that they are not equal. If we decide to use a 5% significance level then we can see that we must reject null hypothesis and so the assumption of unequal variances is in the `t.test()` function is correct ($p < 0.05$).

Q. Use the t-test to assess whether there are differences in bmi between the two age groups.

We do not need to change the t-test to assume unequal variances. However, we can see that since the t-test is very robust and you would have got practically the same result even if you had ignored the violation of its assumptions and did not employ the “unequal” feature.

```
t.test(bmi ~ age_grps, elsa)
```

```

##
## Welch Two Sample t-test
##
## data:  bmi by age_grps
## t = -1.596, df = 2081.8, p-value = 0.1106
## alternative hypothesis: true difference in means is not equal to 0

```

```
## 95 percent confidence interval:
## -0.63477016 0.06515785
## sample estimates:
## mean in group 50-64    mean in group 65+
##      27.43013          27.71494
t.test(bmi ~ age_grps, elsa, var.equal = TRUE)

##
## Two Sample t-test
##
## data:  bmi by age_grps
## t = -1.5671, df = 2486, p-value = 0.1172
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.64118056 0.07156826
## sample estimates:
## mean in group 50-64    mean in group 65+
##      27.43013          27.71494
```

Q: Are there differences in bmi between the two age groups?

The t-test (using the “unequal” feature to account for the non-equal variances between the two groups) shows that there is little difference in the mean BMI between adults aged 50-64 and those aged 65 or older (p=0.11). Results are differ little when changing the assumptions.

Formative Exercise: Test the equality of variances assumption and then running the appropriate t-test for cholesterol (chol) and age

3. Assessing differences between three or more groups: the ANOVA

Now, we are ready to assess whether there are differences in the three outcome variables by physical activity groups. First we will look at this new exposure variable (physact) and then because it is a categorical variable with more than two groups, we will use one way Anova to examine whether there are differences in mean bmi, cholesterol and crp by these three groups of physical activity.

3.2 Examining physical activity

```
str(elsa$physact)

## Factor w/ 3 levels "Group 1 -low",...: 1 1 1 1 1 2 2 2 2 2 ...
## - attr(*, "label")= chr "3 categories of physical activity (1=low, 3=high)"
table(elsa$physact) # To get frequencies by group

##
##      Group 1 -low Group 2 - medium   Group 3 - high
##      1321          1051             756
prop.table(table(elsa$physact)) # To get proportions in each group

##
##      Group 1 -low Group 2 - medium   Group 3 - high
##      0.4223146    0.3359974         0.2416880
```

We can see that there are three categories of physical activity assessed in our dataset. We can see that most people (42%) report low levels of physical activity in our sample, which is not surprising perhaps, since this is an ageing study.


```
by(elsa$bmi, elsa$physact, mean, na.rm = TRUE)
```

```
## elsa$physact: Group 1 -low
## [1] 28.33287
## -----
## elsa$physact: Group 2 - medium
## [1] 27.25954
## -----
## elsa$physact: Group 3 - high
## [1] 26.56143
```

```
aov_res <- aov(bmi ~ physact, elsa)
summary(aov_res)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## physact        2   1526    763.1    38.82 <2e-16 ***
## Residuals    2926   57517     19.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 200 observations deleted due to missingness
```

Examine whether there are differences in cholesterol, blood CRP and bmi according to physical health, which is assessed using three groups (anova).

Compute the analysis of variance

```
res.aov <- aov(weight ~ group, data = my_data) # Summary of the analysis summary(res.aov)
```

1.1. Calculating means and 95% confidence intervals (CIs)

First, to recap from last week, please calculate the 95% confidence intervals for the mean for systolic blood pressure (**sbp**), diastolic blood pressure (**dbp**), BMI (**bmi**) and baseline cholesterol (**chol**).

```
t.test(elsa$sbp)
```

```
##
## One Sample t-test
##
## data:  elsa$sbp
## t = 370.7, df = 2691, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  139.6404 141.1256
## sample estimates:
## mean of x
##  140.383
```

```
t.test(elsa$dbp)
```

```
##
## One Sample t-test
##
## data:  elsa$dbp
## t = 337.22, df = 2691, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
```

```
## 76.87486 77.77409
## sample estimates:
## mean of x
## 77.32448

t.test(elsa$bmi)

##
## One Sample t-test
##
## data: elsa$bmi
## t = 331.71, df = 2929, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 27.36890 27.69438
## sample estimates:
## mean of x
## 27.53164
```

```
t.test(elsa$chol)

##
## One Sample t-test
##
## data: elsa$chol
## t = 269.34, df = 2474, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 5.936061 6.023131
## sample estimates:
## mean of x
## 5.979596
```

Note, what we have done here is repetitive. We could have made this simpler using the `map()` function from the package `purrr` which is loaded with `tidyverse`. `map()` takes a collection of objects and repeats a function on each object within the collection. If we pass `map()` a `data.frame`, it will carry out a function for each column (variable). The first input to `map()` is the collection of objects, the second input is the function to be repeated.

```
elsa %>%
  select(sbp, dbp, bmi, chol) %>% # Keep the relevant variables
  map(t.test)

## $sbp
##
## One Sample t-test
##
## data: .x[[i]]
## t = 370.7, df = 2691, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 139.6404 141.1256
## sample estimates:
## mean of x
## 140.383
##
##
```

```
## $dbp
##
## One Sample t-test
##
## data: .x[[i]]
## t = 337.22, df = 2691, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 76.87486 77.77409
## sample estimates:
## mean of x
## 77.32448
##
##
## $bmi
##
## One Sample t-test
##
## data: .x[[i]]
## t = 331.71, df = 2929, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 27.36890 27.69438
## sample estimates:
## mean of x
## 27.53164
##
##
## $chol
##
## One Sample t-test
##
## data: .x[[i]]
## t = 269.34, df = 2474, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 5.936061 6.023131
## sample estimates:
## mean of x
## 5.979596
```

Remember, the pipe operation, `%>%`, takes the object on the left hand side and puts it into the first argument of the function on the right hand side.

1.1. One sample t-test

Q: Compare whether the mean systolic blood pressure in this sample is significantly different from the population mean (estimated to be 130 mmHg).

The variable of interest is systolic blood pressure (`sbp`). Recall you can use the function `look_for()` from the package `janitor` to explore the `elsa` dataset.

```
t.test(elsa$sbp, mu = 130)
```

```
##
## One Sample t-test
```

```
##
## data:  elsa$sbp
## t = 27.417, df = 2691, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 130
## 95 percent confidence interval:
##  139.6404 141.1256
## sample estimates:
## mean of x
##  140.383
```

Mean systolic blood pressure in the sample is 140.4 mmHg (95% CI: 139.6, 141.1). This is higher than the reference value of 130 mmHg. The output shows a small p-value (< 0.01) and a large t-statistic. The results suggest that the mean systolic blood pressure in the sample from which ELSA participants are drawn is not 130 mmHg.

2. Comparison of two independent samples: the unpaired t-test

If we are asked to assess the potential differences in a continuous normally distributed outcome (systolic blood pressure at baseline) between two groups of a dichotomous variable (**sex**) then we can use the (unpaired) t-test. This test is only for examining differences between two groups, if you need to do some work in creating the appropriate variable for that, such as recoding an existing variable has more than two categories. The data that we collected in order to make the comparison are from independent samples; the measurements in men are not connected in any way to the measurements in women (they are different people).

The same principle previously discussed for hypothesis testing for comparing a mean with a hypothesised value applies to the comparison of two sample means. In this case, the null hypothesis states that there is no difference between the two means in the populations from which the sample means are drawn. If the null hypothesis is correct, then the difference between the two means that we have observed in the sample would be due to sampling variation. We then estimate the probability of observing this difference in the sample if the null hypothesis (of no difference in the population) were true. We do this by taking the difference between the two means in our sample and assessing its relative position in the respective sampling distribution of mean differences. In other words, we assess whether the difference between the sample means lies as far away from the hypothesised difference between the population means so that to be included in the margins of the sampling distribution that correspond to the extreme 5% of means (2.5% lower and 2.5% upper) or not.

We can compare independent means using the syntax `t.test(x ~ group, data)`. For the present question:

```
t.test(sbp ~ sex, elsa)
```

```
##
## Welch Two Sample t-test
##
## data:  sbp by sex
## t = 1.9418, df = 2667.9, p-value = 0.05227
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.01426405  2.92279692
## sample estimates:
## mean in group male mean in group female
##      141.1976      139.7434
```

Q: Are there differences in systolic blood pressure at baseline between men and women?

The output provides the mean systolic blood pressure in each group (males and females). It also provides a 95% CI for the difference in the means (-0.014, 2.922) and a p-value for whether the difference in means is statistically significantly different from zero ($p = 0.05227$). We can infer that the p value is greater than 0.05 if the 95% CIs overlap zero, which they do in this case - but only just. The results state that if we assume the

population mean in males and females is zero, we would get a difference in means at least as large as 5.227% of samples.

Q: Are there differences by BMI groups?

For this, we need to recode the grouped BMI (`bmi4`) into two groups: (i) people with BMI under 25, (ii) people with BMI over 25. Let's use the function `fct_collapse()` from the package `forcats` (part of the `tidyverse`) to do this.

```
levels(elsa$bmi4)

## [1] "Under 20" "Over 20-25" "Over 25-30" "Over 30"

elsa <- elsa %>%
  mutate(bmi_bin = fct_collapse(bmi4,
                                "Less than 25" = levels(bmi4)[1:2],
                                "Over 25" = levels(bmi4)[3:4]))
table(elsa$bmi_bin, elsa$bmi4, useNA = "ifany")

##
##           Under 20 Over 20-25 Over 25-30 Over 30 <NA>
## Less than 25      63        814          0         0    0
## Over 25           0          0       1336       717    0
## <NA>              0          0          0         0   199
```

I used the `table()` function to confirm that we recoded the BMI variable correctly. This is good practice.

Now, we can compare mean systolic blood pressure by BMI groups.

```
t.test(sbp ~ bmi_bin, elsa)

##
## Welch Two Sample t-test
##
## data: sbp by bmi_bin
## t = -8.9472, df = 1412, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -9.122468 -5.841646
## sample estimates:
## mean in group Less than 25      mean in group Over 25
##           134.9582                142.4403
```

The output provides the mean systolic blood pressure in each group (low and high BMI). It also provides a 95% CI for the difference in the means (-9.12, -5.84) and a p-value for whether the difference in means is statistically significantly different from zero ($p < 0.01$). We can infer that the p value is less than 0.05 because the 95% CIs do not overlap zero. The null hypothesis of no difference can be rejected. The results suggest that systolic blood pressure is higher among high BMI individuals.

3. Comparison of two dependent samples: the paired t-test

For the next few exercises, we are going to use a different dataset: `contin2.Rdata`. This is data extracted from a trial assessing the effectiveness of surgical periodontal treatment (treatment of the severe inflammation of the gums) using different oral and general health outcomes.

```
load("contin2.Rdata")
```

Up to now, the difference between the two means referred to independent observations (separate groups of subjects). However, there are cases where the data are paired. Data are considered to be **paired** in the

following circumstances:

1. When the **same individuals are studied more than once**, usually in different circumstances (**pre- and post-treatment** measurements). This happens, for example, in a clinical trial evaluating the effect of a treatment (e.g. periodontal treatment) on an outcome (e.g. quality of life); the quality of life of each patient is measured first before and then after treatment. Then, for each patient we have a pair of measurements of the outcome and we should consider all these pairs when we estimate the effect of treatment on the outcome.
2. When we have two **different groups of subjects who have been individually matched** (e.g. in a matched pair case-control study or in a clinical trial with matched controls). In this case, individuals are matched during the sample selection, so that they have key characteristics in common (e.g. age, sex, socioeconomic status etc.). Then, the data does not consist of two independent groups, but rather of case-control pairs and this pairing should be considered in the analysis.

When the data are paired, this should guide the analysis appropriately. In such cases, we need to focus on the **difference in the outcome measurement between each pair**. The analysis is effectively reduced to a one-sample problem: we calculate the difference between each pair ($x = x_1 - x_2$) and treat these differences as a single sample of differences.

3.1. Differences before and after treatment for the whole sample

If data are paired (both outcomes are measured on the same person), then you use a paired t-test to examine differences. The code for this is below.

```
t.test(contin2$oidpsco2, contin2$oidpsco1, paired = TRUE)

##
## Paired t-test
##
## data: contin2$oidpsco2 and contin2$oidpsco1
## t = -3.1052, df = 44, p-value = 0.003322
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -5.851590 -1.245379
## sample estimates:
## mean of the differences
## -3.548485
```

The first two arguments are the two (paired) vectors and the `paired` argument is set to `TRUE`. The difference is the first vector minus the second (in this case, the change between measurement 1 and 2).

Q: Are there differences before and after treatment for the whole sample in terms of quality of life (oidpsco1 and oidpsco2)?

The sample had lower quality of life score, comparing measurement 2 (oidpsco2) to measurement 1 (oidpsco1). (A lower score on the measure indicates better quality of life.) The confidence interval does not overlap zero and the p-value is very small (< 0.01). This suggests that quality of life improved through time.

3.2. Differences before and after treatment between treatment (tx) groups.

Again, we use the paired t-test here but we need to carry out the test separately for different subgroups in the sample. We can use sub-setting as so:

```
levels(contin2$tx)

## [1] "control (scale+polish)" "test (intensive care + extractions)"
```

```

t.test(contin2$oidpsco2[contin2$tx == levels(contin2$tx)[1]],
       contin2$oidpsco1[contin2$tx == levels(contin2$tx)[1]],
       paired = TRUE)

##
## Paired t-test
##
## data: contin2$oidpsco2[contin2$tx == levels(contin2$tx)[1]] and contin2$oidpsco1[contin2$tx == levels(contin2$tx)[1]]
## t = -2.3922, df = 17, p-value = 0.02857
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -8.3309442 -0.5225912
## sample estimates:
## mean of the differences
## -4.426768

t.test(contin2$oidpsco2[contin2$tx == levels(contin2$tx)[2]],
       contin2$oidpsco1[contin2$tx == levels(contin2$tx)[2]],
       paired = TRUE)

##
## Paired t-test
##
## data: contin2$oidpsco2[contin2$tx == levels(contin2$tx)[2]] and contin2$oidpsco1[contin2$tx == levels(contin2$tx)[2]]
## t = -2.0171, df = 26, p-value = 0.05412
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -5.98241201 0.05648616
## sample estimates:
## mean of the differences
## -2.962963

```

Q: Are the results similar for the treatment groups? Are they similar with the ones shown for the whole sample?

The results for the quality of life indicate improvement in quality of life in the control group (difference of 4.4; 95% CI: 0.5, 8.3; $p=0.029$), while the test group reported some improvement in quality of life, though the respective result was marginally not significant (difference of 3.0; 95% CI: -0.1, 6.0; $p=0.054$).

3.3. Testing the difference in quality of life changes between treatment groups.

This is about assessing the difference between two groups in relation to difference before and after treatment. The data are NOT paired, since you only have one observation per participant, i.e. the difference between before- and after-treatment. Firstly we need to calculate this variable for every participant (difference in quality of life before and after treatment), then we will use that to assess differences in that variable between the subgroups (treatment and control).

Let's create a new variable using the `mutate()` function from the `tidyverse` and use the `t.test()` formula syntax for Section 2 to test for differences between groups.

```

contin2 <- contin2 %>%
  mutate(oidpsco_diff = oidpsco2 - oidpsco1)
t.test(oidpsco_diff ~ tx, contin2)

##
## Welch Two Sample t-test
##

```

```
## data: oidpsco_diff by tx
## t = -0.61956, df = 35.864, p-value = 0.5395
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -6.256093  3.328483
## sample estimates:
##          mean in group control (scale+polish)
##                                -4.426768
## mean in group test (intensive care + extractions)
##                                -2.962963
```

Q: Is there a difference between the treatment groups in relation to the before- and after-treatment differences in probing pocket depth?

The improvement in the quality of life score in the control group was greater (since lower scores indicates better quality of life, the higher the score the worse the quality of life) than the respective reduction in the treatment group: -2.96 (-5.98, 0.06) vs. -4.42 (-8.33, -0.52). The difference in these differences was -1.46 (-6.2, 3.3). The confidence intervals for this difference are wide. We can't reject the null hypothesis that there is actually no difference in quality of life between the treatment and the control group.

Formative Exercise

Practice more with the paired t-test. Please assess whether there are differences before and after treatment in probing pocket depth ($\text{ppd}_0 - \text{ppd}_{2m}$).