There is not an even spread of male and females.

There is more pain recorded for males compared to females for both control and meditation.

But there were more males in the control cases than meditation so this will drive down the average more than they will the meditiation only cases.

It is assumed that age has no factor on pain tolerance as it is not a recorded variable.

It is assumed that the treatments for these patients are similiar as minor surgery v major surgery would have different responses.

Chapt7(p251)

Before getting confidential interval.

To use confidential intervals and such we will use the central Limit Theorem but also check the data can be used.

To apply the Central Limit Theorem for a sample mean x, the sample must be deemed independent. If the sample is selected at random.

We also need the data to come from a normally distributed population. If n is large enough we can take it that the data fits a normal distributed population and doesn’t have significant outliers

The rule of thumb is to take N as >=30.

We also have to use the sample SD (s) instead of the population standard deviation because the pop standard deviation is unknown. This is an issue with a small sample but if the sample size is large enough it is ok.

The T Distribution is used instead of the Normal Distribution when we use s instead of sd.

This T distribution has longer tails than the Normal distribution to compensate for not having the true SD. So more samples are likely to fall between 2 deviations of the mean value.

The T-distribution is always centred at 0 and has a single parameter – degrees of freedom. This degree of freedom (dof) is got by subtracting 1 from the sample size.

E.G n = 100 so df = 100-1 = 99.

The larger the dof the more it resembles the normal distribution.

Chart, histogram

Description automatically generated

1. Determine whether the data provided is appropriate for the test(s) available and that any analysis is achievable.

We can assume a normal distribution as n>=30 and there are no significant outliers. The data is also randomly selected.

We have to assume that there is no difference in the average age group and that the samples are relatively spread across the ages.

Though there were 50 males and 50 females, there was more females (27) in the sample who meditated against not (23). Alternatively, there was less males (23) in the sample who meditated than not (27).

We can use the t-distribution tests if normal distribution can be assumed.

We have a confounding variable. A variable which could cause the variability.

Response variable/ blocking variable/ explanatory variable

Is there a difference between scores from tests at galvanic responses v perceived Pain.

Is there a difference between scores from males v females.

One thing I considered was whether to treat

The data here is Paired data see p262 of notes.

The pain scores are recorded using different methods – verbal v galvanic responses but have a common relationship – both regarding the pain levels of the same patient.

We will record the differences between the data sets to get a new table of data scores.

Test is there a difference between scores.

Is the difference between scores normal? If yes can we test control v normal?

Can we test control v normal twice using the different scores. Are these scores independent of each other?

See page 20 lecture slides for paired data example

1. Formulate a hypothesis test to be used to compare the effectiveness of the two approaches (control, meditation) used during dental surgery.

H0: µdiff = 0. There is no difference in the average pain recorded between the Meditation group and the control group.

HA: µdiff 6= 0. There is a difference between the pain recorded from those who did meditation and those who were in control group.

1. Analyse the data to provide the hypothesis testing conclusion.
2. Provide descriptive statistics (graphs and tables) of the data.

Whisker plots – median – IQ range – outliers. Would be good to show outliers in seeing does it fit normal distribution.

Histograms of data to show if normal.

1. Determine the 95% confidence interval for the population mean of each group, and the 95% confidence interval for the difference between the means of the two groups.

Confidence Interval of mean:

Use t test as we don’t know SD for population. Have SD for sample.

X(mean) +\_ t(s/n^.5)

Then calculate p:

e.g. using R:

>2\*(pt(t value, df (n-1), lower.tail=FALSE)

If the p value is low, we conclude that the data provides strong evidence means.

Using T test:

Confidence Interval of difference of means

X(mean1)-x(mean2/SE

Where SE=(s1^2/n1+s2^2/n2)^.5

Again get p value –

Pt(q=t value,df = n-1)

If p value is small reject h0 that there is no difference.