Welcome to Week 3!

Welcome to Week 3!

And so it begins! Our journey into the world of hypothesis testing. I thought I would start us off with two of the most common hypothesis tests, **t-tests and chi-square tests**. The t-test is a statistical test used to **compare the means of two different groups**. These groups may be **independent** or they may be **paired** (related). For example, two independent groups are when changing the position of a single value in one group has no influence on the other group. For example, say, you have two groups, one is diabetic and the other is non-diabetic (also known as the control) and you have recorded each patient's blood glucose level in mcg/dL after fasting for 8 hours as shown below - 9 patients who are diabetic and 8 who aren't:

Blood Glucose Level Diabetic (mcg/dL)	Blood Glucose Level Non-diabetic (mcg/dL)
136	86
177	85
146	83
128	91
179	84
144	90
134	99
167	86
129	

If I moved the 91 mcg/dL data point and swapped it with the 83 mcg/dL data point (as indicated by the blue arrow) then we can see that the diabetic group is unaffected. This tells us that we have independent data. Paired data on the other hand is when data are related or in pairs. I happen to be a massive fan of a malaysian noodle dish called "char kuey teow" and often like to go out with my friends to eat large dishes of it. Admittedly, I tend to put on more weight after having a few courses...

Person	Weight before eating char kuey teow (kg)	Weight after eating char kuey teow (kg)
Brandon	77.35	78.96
Oliver	80.12	82.03
Barry	71.09	71.99
Kara	63.53	65.16

Here we can see that if I was to move one data point in the "after eating char kuey teow" group this would change the data in the "before eating char kuey teow" group. Hence, we realise that in this case the data is in related pairs.

There are three types of t-test that you will learn in this unit, these are: **The one-sample t-test, the independent t-test and the paired t-test**. Each one serves their own purpose depending on the type of data we have and the research question we are trying to answer.

The chi-square test (pronounced "ky" - as in the "ky" in "Kylo Ren") is a statistical test used to **analyse frequency data**, in fact, you should take note that a t-test works with continuous data but a chi-square test only works with data that is categorical (hence the importance of understanding content in weeks 1 and 2!). Since I mentioned a Star Wars character, an example of this might be taking a survey on "favourite coloured light sabers" and seeing if there is an association with sex (male or female). Let's say we have green, blue, purple, red, yellow coloured light sabers, realising that both light saber colour and sex are categorical variables we can create what's called a 2 (rows) x 5 (columns) contingency table as shown below (results from a fictional survey in a galaxy far far away...)

Sex	Green	Blue	Purple	Red	Yellow
Male	23	15	7	22	9
Female	10	17	20	16	21

For a bit more visual appeal, I even colour coded it! When we have a contingency table like this, we can see whether or not there is an association (a relationship) between our two categorical variables (favourite light saber colour and sex). We will cover this in more detail in this week's lecture.

Obviously Star Wars is a fictional franchise...so let's translate this idea of the chi-square test to a public health context. Suppose instead of light saber colours we had socioeconomic status instead (low, middle, high)? Or we had different categories of the Gross Motor Function Classification System (GMFCS) or stages of cancer (Stage 0 to IV)? What about types of cardiac surgery (CABG, LVAD, TMR etc.) or cardiovascular disease (coronary, congenital, rheumatic, DVT etc)? I hope that you are realising that the power of statistics lies in its versatility, it is not just something that is used in health research but can be used widely in other disciplines. There are also three types of chi-square tests that we will learn in this unit, these are: **the chi-square test for goodness of fit, chi-square test of association and chi-square test of association with Yate's correction method**.