Day 26

Hypothesis:
an idea or explanation for something that is based on known facts but has not yet been proved

Statistical hypothesis:
Statistical hypotheses must be mathematically precise, and they must correspond to specific claims about the characteristics of the data generating mechanism

Null Hypothesis (H0):
The null hypothesis is the statement or claim being made (which we are trying to disprove)
Alternate Hypothesis (H1): the alternative hypothesis is the hypothesis that we are trying to prove and which is accepted if we have sufficient evidence to reject the null hypothesis.

Errors: Type I and Type II

	retain H_0	retain H_0
H_0 is true	correct decision	error (type I)
H_0 is false	error (type II)	correct decision

If we reject a null hypothesis that is actually true, then we have made a Type I error. On the other hand, if we retain the null hypothesis when it is in fact false, then we have made a Type II error

Statistical Significance:
If the data allow us to reject the null hypothesis, we say that "the result is *statistically significant*", which is often shortened to "the result is significant".

Note:
a lot of modern readers get very confused when they start learning statistics, because they think that a "significant result" must be an important one. It doesn't mean that at all. All that "**statistically significant"** means is that the data allowed us to **reject a null hypothesis.**

P value:

p is defined to be the smallest Type I error rate ($\alpha\alpha$) that you have to be willing to tolerate if you want to reject the null hypothesis.
If it turns out that pp describes an error rate that you find intolerable, then you must retain the null. If you're comfortable with an error rate equal to pp, then it's okay to reject the null hypothesis in favour of your preferred alternative.

A p-value less than **0.05 is typically considered to be statistically significant, in which case the null hypothesis should be rejected**. A p-value greater than 0.05 means that deviation from the null hypothesis is not statistically significant, and the null hypothesis is not rejected.

if you get $p=.062p=.062$, then it means that you'd have to be willing to tolerate a Type I error rate of 6.2% to justify rejecting the null. If you personally find 6.2% intolerable, then you retain the null.

Approaches to hypothesis testing:
one proposed by **Sir Ronald Fisher** and the other proposed by **Jerzy Neyman**

First, let's talk about Fisher's approach. As far as I can tell, Fisher assumed that you only had the one hypothesis (the null), and what you want to do is find out if the null hypothesis is inconsistent with the data. From his perspective, what you should do is check to see if the data are "sufficiently unlikely" according to the null. In fact, if you remember back to our earlier discussion, that's how Fisher defines the p-value. According to Fisher, if the null hypothesis provided a very poor account of the data, you could safely reject it. But, since you don't have any other hypotheses to compare it to, there's no way of "accepting the alternative" because you don't necessarily have an explicitly stated alternative. That's more or less all that there was to it.

In contrast, Neyman thought that the point of hypothesis testing was as a guide to action, and his approach was somewhat more formal than Fisher's. His view was that there are multiple things that you could *do* (accept the null or accept the alternative) and the point of the test was to tell you which one the data support. From this perspective, it is critical to specify your alternative hypothesis properly. If you don't know what the alternative hypothesis is, then you don't know how powerful the test is, or even which action makes sense. His framework genuinely requires a competition between different hypotheses. For Neyman, the p value didn't directly measure the probability of the data (or data more extreme) under the null, it was more of an abstract description about which "possible tests" were telling you to accept the null, and which "possible tests" were telling you to accept the alternative.

Having both a null hypothesis and an alternative (Neyman), but usually define the p value in terms of exreme data (Fisher), but we still have α values (Neyman). Some of the statistical tests have explicitly

specified alternatives (Neyman) but others are quite vague about it (Fisher). And, according to some people at least, we're not allowed to talk about accepting the alternative (Fisher)