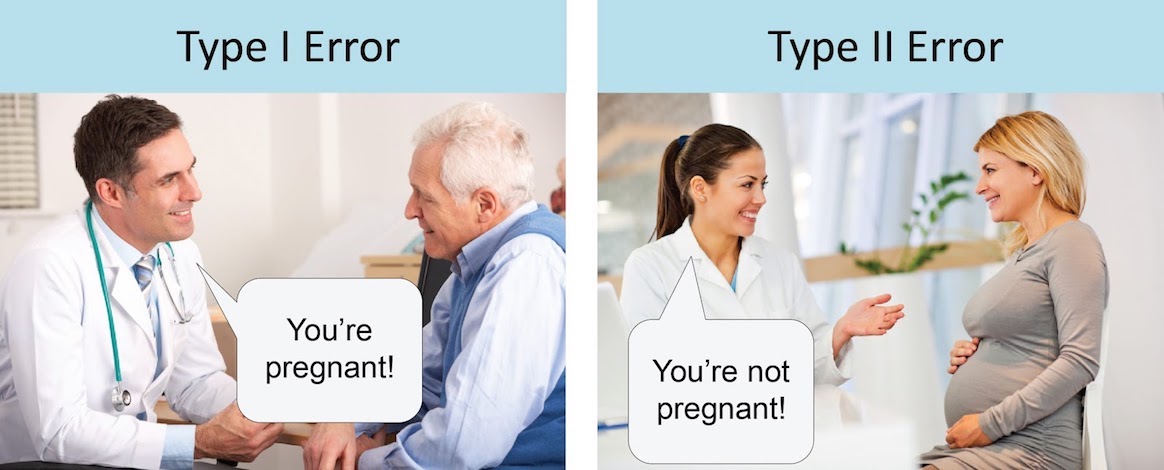
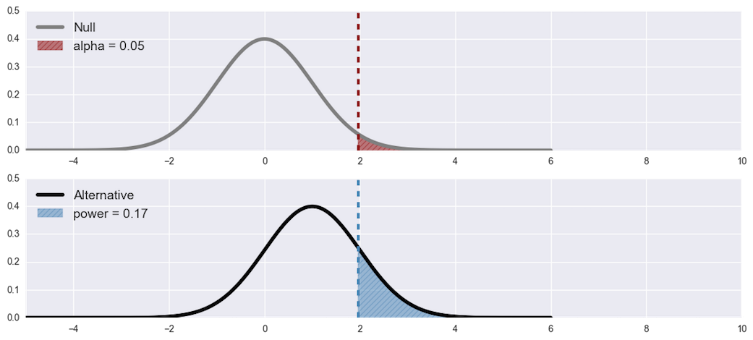
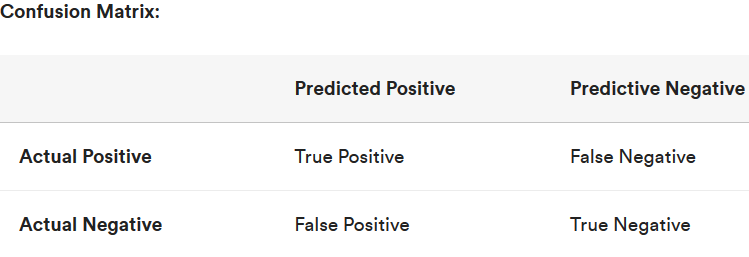
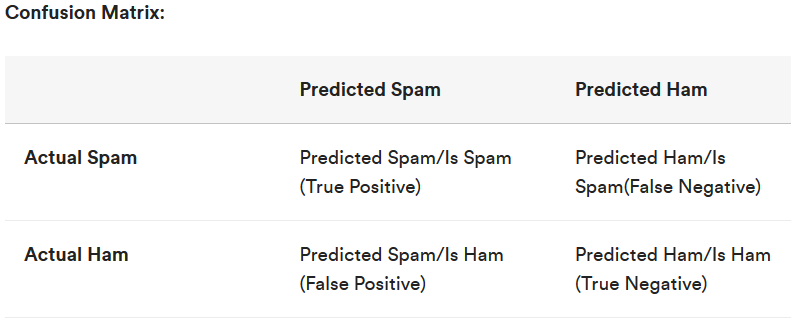
Unit 3-3 Type I and Type II Errors

* Motivating Example: Spam Detection
  + We're trying to build a spam detector, and we want to measure the accuracy of our model. We have a data set of emails that are labeled either "spam" or "ham." The label spam is encoded with a 1 and the label ham is encoded with a 0. Let's think about the categories emails fall into once they've been through our spam detector model:
    - 1) Is spam/marked spam (Is 1/classified as 1).
    - 2) Is ham/marked spam (Is 0/classified as 1).
    - 3) Is ham/marked ham (Is 0/classified as 0).
    - 4) Is spam/marked ham (Is 1/classified as 0).
  + So, 1 and 3 are correctly classified, but 2 and 4 are incorrectly classified. Let's talk about how we can describe these correct and incorrect classification
* Classification Categories
  + - **True positive**: A positive class observation (1) is correctly classified as positive by the model. An example is our "is spam/marked spam" category.
    - **False positive**: A negative class observation (0) is incorrectly classified as positive. An example is our "is ham/marked spam" category.
    - **True negative**: A negative class observation (0) is correctly classified as negative. An example is our "is ham/marked ham" category.
    - **False negative**: A positive class observation (1) is incorrectly classified as negative. An example is our "is spam/marked ham" category.
  + Note: The first word (true or false) indicates whether or not the model was correct. The second word (positive or negative) indicates the model's guess — not the actual label.
  + In the context of hypothesis testing, false positives and false negatives are referred to as type I and type II errors, respectively.
* Statistical Significance and P Values
  + Recall that a p value is a metric indicating the probability that our measured difference was the result of random chance in the sampling of our subjects. It's quite common to see frequentist tests reported as "statistically significant with p < 0.05" or "with p < 0.01."
  + So, what does it mean for a test to be **statistically significant**? On the surface, these statements are saying that a calculated p value is less than a specific value. In this case, the values of 0.05 and 0.01 are arbitrary numbers commonly used in academic research.
  + And what is p < 0.05 actually telling us? It's saying that, in hypothetical repetitions of this experiment with the same sample size, fewer than 5 percent of the experiments would have a measured difference to this degree that was the result of chance.
  + The same explanation applies to p < 0.01. However, here we can frame it in the context of null and alternative hypotheses.
  + p < 0.01: There is less than a 1 percent chance of accepting the alternative hypothesis when the null hypothesis is in fact true.
* Type I and Type II Errors
  + A type I error is the incorrect rejection of the null hypothesis when it is in fact true. This is equivalent to the false positive rate in classification — the rate of a model labeling an observation as "true" when in fact it is "false."
  + A type II error, on the other hand, directly corresponds to false negatives. A type II error in the context of hypothesis testing would be to accept the null hypothesis when the alternative hypothesis is true.
  + 
* Type I Errors and the Alpha Threshold
  + In hypothesis testing, we set a threshold for how likely we are to falsely accept the alternative hypothesis prior to running our experiment. This is known as the type I error.
  + It's important that this threshold for type I error is set before experiments begin so we aren’t tempted to set one after seeing the p value.
  + We denote this threshold as α\alphaα.
  + For example, α=0.05 corresponds to a p < 0.05 significance level. α=0.01 corresponds to a p < 0.01 significance level.
* Type II Errors and Statistical Power
  + A type II error, on the other hand, directly corresponds to false negatives. A type II error in the context of hypothesis testing would be to accept the null hypothesis when in fact the alternative hypothesis is true.
  + β (the Type II counterpart to α\alphaα is the probability that we accepted the null hypothesis by chance when the alternative hypothesis is true.
  + Whereas type I errors correspond to the concept of statistical significance, type II errors correspond to the concept of statistical power. The power of a test is:
  + Power=(1 − type II error)
    - , or
    - Power=(1−β)
* Statistical Power
  + 1−β is known as the statistical power of a test (and is used more commonly than β on its own).
  + The power of a test is the probability of correctly accepting the alternative hypothesis when it is true.
  + More intuitively, power measures our ability to detect an effect that's present.
* Alpha and Statistical Power
  + 
  + Here's a plot that helps visualize the relationship between the α\alphaα threshold and the power 1−β of the experiment.
  + **Decreasing α\alphaα (making the threshold more strict) necessarily reduces power. Alternatively, the more lenient we are about type I error rates, the more we reduce type II error rates**.
* Type I and Type II Errors in the Real World
  + Let's review the idea of type I and type II errors with a real-world application: A new drug for cancer that's being tested.
  + With a type I error, the results show that the drug is effective when, in reality, it's not. The drug has many side effects, but because it's believed to be effective in treating cancer, it's given to patients on a regular basis. Consequently, people's cancer is not being treated and they're experiencing more health issues that can lead to an early death. This is an example of how serious the type 1 error can be.
  + With a type II error, the results might show that the drug is not effective in treating cancer when in reality it is. Therefore the drug is not administered to people and their cancer is not treated.
  + It could be argued that the type I error is worse than the type II error as it not only wastes time, resources, and money on a fals4e positive, but also has a greater negative impact on patients than if they were never treated at all
* Confusion Matrix
  + The confusion matrix is a table that allows you to visualize the performance of your test or model.
  + 
  + Let's remember the definitions of each of these.
    - True positive: A positive class observation (1) is correctly classified as positive by the model.
    - False positive: A negative class observation (0) is incorrectly classified as positive.
    - True negative: A negative class observation (0) is correctly classified as negative.
    - False negative: A positive class observation (1) is incorrectly classified as negative.
* Spam Confusion Matrix
  + The confusion matrix can also represent your model's ability to classify labels correctly. Here we've applied it to our spam detection example:
  + 
* Power and The Confusion Matrix
  + We can visualize the ideas of significance, power, and error types in a matrix similar to the confusion matrix from above:
  + Confusion Matrix:
  + 