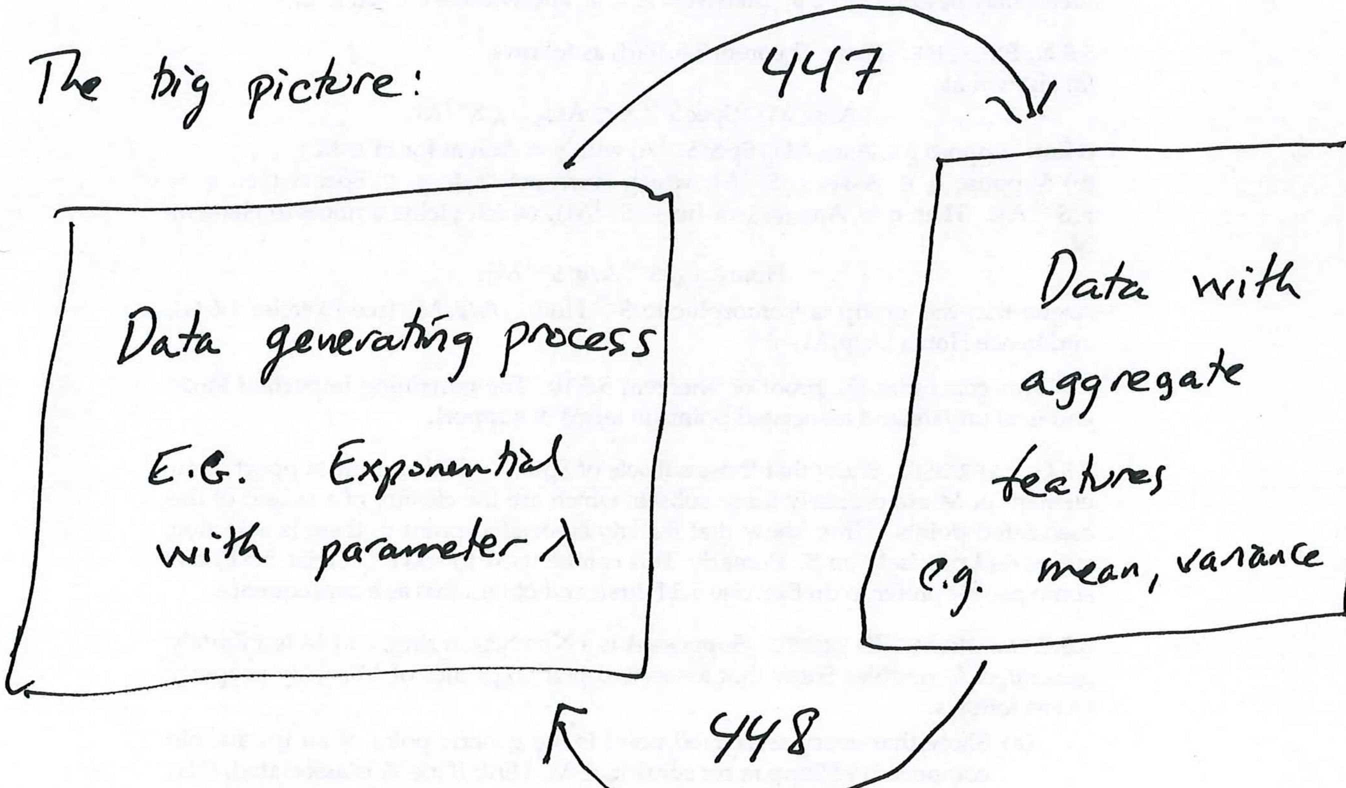
Math 455 Class 3 January 26 1

Homework 3: review of Hypothesis tests (Wackerly Ch. 10) [Not yet on Gradescope]

The big picture:



In 455: Our data generating process:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Where β_0, β_1 are fixed unknown parameters.

$\varepsilon_i \sim N(0, \sigma^2)$ are independent normal RVs with fixed but unknown σ^2.

The above statements about the data generation process in 455 are the standard axioms of linear regression.

Our data is a list of points: $(X_1, Y_1) \cdots (X_n, Y_n)$

We want to test hypotheses about β_0 and β_1

Thus our interest in Hypothesis Testing.

(Wackerly Ch. 10)

The output of a regression analyis in R:

```
lmod <- lm(Y ~ X)
summary(lmod)
```

Output of summary has p-values, which are defined in Ch. 10. This is a 3-layer defn.

What is a statistical test? (4 elements)

1. Null Hypothesis H_0.
2. Alternative Hyp. H_a.
3. Test statistic
4. Rejection region: Reject H_0 if stat in RR.

Example: A poll of voters do you favor candidate C? (Yes / No)

We sample 15 voters, and our test stat is # Yeses. (Call this Y)

$$\left.\begin{array}{l} H_0 = p = P(\text{Yes}) = 0.5. \\ H_a = p < 0.5. \\ \text{Rejection region: } Y \leq 2 \end{array}\right\} \text{Defn of our stat. test.}$$

This is a statistical test. We do not yet have a "level" or a "p-value".

The level of a test is the probability of a type I error. (Usually denoted α)

~~The~~ A type I error is the probability of rejecting H_0 when H_0 is true.

From the above example we may now compute the level of the test.

If H_0 is true, then Y in the ~~sd~~ test above is $Y \sim Bin(15, \frac{1}{2})$

(Binomial with $n = 15$ $p = \frac{1}{2} = 0.5$)

$P($Reject H_0 when H_0 is true$)$

$$= P(Y \le 2) = \binom{15}{0}\left(\tfrac{1}{2}\right)^0\left(\tfrac{1}{2}\right)^{15} + \binom{15}{1}\left(\tfrac{1}{2}\right)^1\left(\tfrac{1}{2}\right)^{14} + \binom{15}{2}\left(\tfrac{1}{2}\right)^2\left(\tfrac{1}{2}\right)^{13}$$

(By Math 447).

$\simeq 0.004$.

Thus the <u>level</u> of this test is $\simeq 0.004$.

Notice that there is no p-value yet.

We can't compute a p-value, because the p-value requires <u>actual data</u>.

Def'n of p-value: ~~Th~~ The <u>p-value</u> or <u>attained significance level</u>, is the smallest level of significance α for which the observed data indicate that the null hypothesis should be rejected.

In order to compute a p-value, we need data.

We could compute the level, because the level depends only on the test, not the observed data.

Let's suppose that we actually do the survey of 15 voters. We find $Y=3$.

What is the p-value?

Now we can compute the p-value:

We know that $P(Y \leq 3 \mid H_0)$

$$= \binom{15}{0}\left(\frac{1}{2}\right)^{15} + \binom{15}{1}\left(\frac{1}{2}\right)^{15} + \binom{15}{2}\left(\frac{1}{2}\right)^{15} + \binom{15}{3}\left(\frac{1}{2}\right)^{15}$$

$\simeq 0.018$ (Math 447)

The p-value is the smallest level of significance for which the observed data indicate "reject H_0"

If the test were instead to have rejection region $Y \leq 3$, in place of $Y \leq 2$.

Then: the observed data $Y = 3$ would indicate "reject H_0"

And the level of the test would be 0.018.

Thus, the p-value associated to the data $Y = 3$ is 0.018.

Very short version: low p-values good.

The lower the p-value, the more stars R gives in the regression summary.

lower p-values constitute stronger evidence for rejecting H_0.

In the regression summary there is one p-value for each regression coeff β_0, β_1, etc.

These p-values are associated to a statistical test of $H_0: \beta_i = 0$ vs. H_a $\beta_i \neq 0$.

The test statistic in this case has the t-distribution. (There is another test stat in regression output with the F-distribution.)

This is why HW 2 is a review of the t- and F- distributions.

Q: Why are low p-values good?

A: let's go back to our hypothetical drug that we want the FDA to approve

Our H_0 is "drug doesn't work".

Note that we don't <u>believe</u> H_0.

We <u>assume</u> H_0 because in trying to demonstrate that the drug works, we put on our skeptical statistician hats and say let's assume H_0 and gather data, to see if we can reject it.

We get our data: how strongly do our data reject H_0?

The level of a test is the probability of rejecting H_0, given that H_0 is true.

For our data to convince the FDA that the drug works we want the probability of "accidental" rejection of H_0 to be low.

Thus a low level is more convincing to the FDA.

So if we have a p-value of 0.04 this says, if the drug ~~was~~ doesn't work (H_0), there is a 4% chance that we could get data this good by sheer luck.

If the p-value is 0.001, this says that if H_0 (drug doesn't work), there is a 0.1% chance we could get data this good by sheer luck.

The second case is more convincing, so lower p-values are better.