Philosophy of Statistics: Homework 2

due on Gradescope by 11am on Wednesday January 20

Guidelines. Some questions ask you to justify your answers. For these questions, credit will be based on how well you justify your answers, not on whether your answers are correct. (There's often no consensus on the correct answers, even among statisticians.) However, that doesn't mean that anything goes: some answers will be hard to justify well. I give suggested word counts but these are just ballpark numbers. Don't sweat them too much. Collaboration is encouraged, but make sure to write up your answers by yourself and list your collaborators.

Problem 1 (15 points). Gemma the geologist has a come up with a new model of tectonic plate movements on the west coast. To test the hypothesis, she uses it to predict how many earthquakes of magnitude at least 1.8 there will be in LA tomorrow. Her prediction:

0	1	2	3	4	5	6	7
.002	.412	.287	.140	.040	.056	.051	.012

So, for example, according to her model the probability there will be exactly 1 earthquake tomorrow is .412.

For each X from 0 thru 7, calculate the p-value if there are X earthquakes tomorrow. If we chose a significance level of .05, which outcomes would lead us to reject Gemma's model?

Problem 2 (10 points). I challenged you to modify a coin to make it 75%-heads biased. Let's test your coin using a significance test. Flip it 10 times and count the number of heads. The probabilities of the possible outcomes, assuming your coin is indeed 75%-heads biased, are in the slides for Class 3. How many heads did you get what is the p-value?

Problem 3 (20 points). In Problem 3 of Homework 1, we imagined testing the hypothesis that James's coin is 75%-heads biased by flipping it 100 times and counting the number of heads. We focused on four cases:

1. We get 13 heads.
2. We get 66 heads.
3. We get 75 heads.
4. We get 84 heads.

And I asked how you would answer Royall's three questions in each case.

You might answer Royall's questions using significance tests. To do that, you need to choose a significance level (for v.2 and v.3) and calculate the p-values. The choice of significance level is up to you. To save you doing the calculations, the p-values to three decimal places are: .000, .049, 1, .038.

First, given the p-values and your choice of significance level how do the three versions of significance test answer Royall's questions? Second, how do these answers compare with your own from Homework 1? For any answers which disagree, who got it wrong: you or the significance test? Justify your answer. (200 words)

Problem 4 (15 points). Remember a key concept from class: given an experiment, a *probability distribution* is an assignment of probabilities to each possible outcome. We can represent a probability distribution in a table. For example:

0	1	2	3	4	5
.001	.015	.088	.264	.396	.237

Here's another key concept: the *mean*, or *expectation*, of a probability distribution is the sum of the outcomes multiplied by their probabilities. For example, the mean, or expectation, of the probability distribution above is:

$$0 \cdot .001 + 1 \cdot .015 + 2 \cdot .088 + 3 \cdot .264 + 4 \cdot .396 + 5 \cdot .237$$

which equals 3.75 (to two decimal places).

The aim of this question is to give you practice thinking about means, which will be a key concept in later weeks and also come up in the next problem. First, what is the mean of the probability distribution below:

1	5	6	9
.05	.6	.123	.227

Second, write down a probability distribution such that: the mean is 4, there are three possible outcomes, and no two outcomes have the same probability.

Third, write down three different probability distributions which all have the same mean.

Problem 5 (20 points). The p-value of an outcome is the probability, assuming the null hypothesis is true, of getting an outcome as extreme as the actual outcome. But what does *as extreme* mean? In class, we defined it like this:

1. o_1 is as extreme as o_2 just if $\mathrm{P}_h(X = o_1) \leq \mathrm{P}_h(X = o_2)$

But we could have defined it differently. Two alternatives: o_1 is as extreme as o_2 just if...

2. o_1 is at least as far away from the mean μ as o_2
3. $o_1 \leq o_2 \leq \mu$ or $o_1 \geq o_2 \geq \mu$

Show that the three definitions can all yield different p-values. Your answer should specify: a probability distribution, the actual outcome, and a calculation of the three p-values. You don't need to worry about where the probability distribution comes from.

Problem 6 (20 points). In the previous question, we saw that p-values depend on how we define *as extreme*: define it one way, you'll get one p-value; define it another way, you'll get a different p-value.

So how should we define *as extreme*? Maybe you endorse one of the definitions in the previous question, or maybe you endorse a definition of your own. If you think which definition we should use depends on the context, then what aspects of the context does it depend on? If you think there's no good answer to the question of how we should define it, then is that a problem for significance tests? Justify your answer. (300 words)