Question #3

3. A random variable with *geometric distribution* is a discrete variable that can attain values of 1, 2, ... It is used to model the number of trials of a Bernoulli (binary) test is repeated until the first success occurs. So, if a Bernoulli trial can have "success" with probability p and "failure" with probability $1-p$, then the number of iid[1] trials before a success happens

[1] The term 'iid' means *identically and independently distributed.*

follows the geometric distribution. The pmf for this distribution is given by

$$f(k) = (1-p)^{k-1}p$$

This is because if the first success is seen at the k^{th} trial, it means that we have seen $k-1$ failures in a row and then saw success in the k^{th} trial.

$$f(k) = (1-p)^{k-1} p$$

3a) Given = fair coin = 50% of heads
= fake coin = 20% of heads
Toss coin 3 times all tail with heads on 4th toss

$$\text{fair coin} = (1-0.50)^{4-1}\,0.50 = \underline{0.0625}$$

<u>6.25%</u> probability of 3 tail then 1 head with fair coin

$$\text{Fake coin} = (1-0.20)^{4-1}\,0.20 = 0.1024$$

10.24% probability of 3 tail then 1 head with fake coin

3a) using the geometric distribution we are able to see probability of 3 tails, then 1 head with fair coin is 6.25%. While 3 tails, then 1 head with fake coin is 10.24% Since it is more likely that it is a fake coin then fair coin I would say that based on the current number of trials we cannot definitively say if it is fair or fake but currently there is a 10.24% chance that we got 3 tails and then 1 head with the fake coin. which is ~4% increase in probability then that occuring with real coin which was 6.25%. Therefore, it is around 4% more likely this was fake coin compared to fair but we cannot guarantee it is fake. We need more trials to confirm if it is fake

3a Answer → (More likely it is fake but no guarantee)

3b

3b) Suppose we have two scenarios: the probability of success is p_1, or probability of success is p_2. Given values of $0 \leq p_1 < p_2 \leq 1$, what is the minimum number of times failures should occur before the first success in order for you to select scenario 1 over scenario 2 using Bayes decision rule?

Based on Bayes we will choose the hypothesis that maximizes the posterior prob. H_1 = coin has probability of success P_1

H_2 = coin has probability of success P_2

Assuming equal priors $P(H_1) = P(H_2)$ we will only need to evaluate the likelihood ratio

$$\frac{P(\text{data}|H_1)}{P(\text{data}|H_2)} = \frac{(1-P_1)^{n-1} P_1}{(1-P_2)^{n-1} P_2} = \frac{P_1}{P_2} * \left(\frac{1-P_1}{1-P_2}\right)^{n-1}$$

We should choose H_1 if

$$\frac{P(\text{data}|H_1)}{P(\text{data}|H_2)} > 1$$

$$\frac{p_1}{p_2}\cdot\left(\frac{1-p_1}{1-p_2}\right)^{n-1} > 1$$

$$\downarrow$$

$$\left(\frac{1-p_1}{1-p_2}\right)^{n-1} > \frac{p_2}{p_1}$$

$$\downarrow$$

$$(n-1)\log\left(\frac{1-p_1}{1-p_2}\right) > \log\left(\frac{p_2}{p_1}\right)$$

$$\downarrow$$

$$n-1 > \frac{\log\left(\frac{p_2}{p_1}\right)}{\log\left(\frac{1-p_1}{1-p_2}\right)}$$

$$\downarrow$$

3b) Answer

$$n > \frac{\log\left(\frac{p_1}{p_2}\right)}{\log\left(\frac{1-p_1}{1-p_2}\right)} + 1$$

This gives the smallest number of trials n, where the difference between the two hypothesis becomes distinguishable using Bayes Decision Rule, in order to select scenario 1 over scenario 2.