EXERCISE ON BOOTSTRAP, RANDOMIZATION INFERENCE

(1) Here's an i.i.d. sample from an unknown distribution:
3.66 1.00 -0.87 2.90 -0.80 3.20 1.69 -3.53 3.22 3.53

(a) Make 1000 bootstrap draws from this sample of size 10 and use them to construct an estimate of the pdf of the sample mean and the standard deviation of the sample mean.
(b) Plot the estimated pdf, which you can do in R with `plot(bkde())` using the KernelSmooth package or with the `hist()` function.
(c) On the assumption that when the mean of the population distribution changes, it changes the distribution by a pure location shift, use your bootstrapped sample to construct a 95% confidence interval for the mean (Should it be "flipped"?)
(d) The data were actually generated from an equally weighted mixture of two normals, one with mean 0, standard deviation 3, one with mean 3, standard deviation 0.2. Suppose we knew that the distribution had this form, except for the means of the two components. If the first component has mean $\mu - 1.5$ and the second mean $\mu + 1.5$, with μ unknown, we have a pure location shift model of the data. Plot the likelihood for this sample as a function of μ. Does it imply a 95% credible set (under a flat prior) similar to what you got from the bootstrap?
(e) Now suppose instead that we know the first component has mean 0, but don't know the mean of the second component. The mean of the second component is 2μ, so that μ is still the mean of the distribution, but μ no longer produces a pure location shift. Plot the likelihood and find a 95% credible set for μ for this case.

(2) **Lady tasting tea** A lady claims to be able to tell whether she has been served tea with the milk put in the cup before, or after, the tea. She is given a sequence of cups of tea, some of which (or perhaps all, or none of which) have had the milk put in first. She announces her guesses, we find out how many of them were correct. How do we assess the evidence for her abilities?

This is a classic example, probably the simplest and most appealing case for randomization inference. There is no need to appeal to hypothetical repetitions of the lady's test, as in the usual frequentist framework. Here there 2^N possible sequences of milk-first/milk-second, where N is the number of trials. So if she is given, say, 5 cups, there are 32 possibilities. If we choose the sequence of cups she is presented with "at random" from these 32 possibilities, the probability that she gets them all right if she is "just guessing" is 1/32, and that she gets at least 4 out of 5 right is 6/32. This lets us construct an exact, finite-sample test of "H0: she's guessing" at a .03125 or .1875 significance level.

Date: February 3, 2021.

But what is the alternative hypothesis? If it is that her guessing is i.i.d. across cups, with probability p of being right on each cup, we can construct a likelihood function for p and ask what is the flat-prior probability of $p > .5$ given that she has made 4, or 5, correct guesses. Do that. (The posterior is a Beta distribution.)

Suppose it happens that every one of the 5 test cups of tea has milk first, and she guesses milk first on every one of them. Does this seem like less strong evidence of her ability than if the cups she had been presented had at least some of each type? How might one justify a claim that the evidence is less strong in this case?