18.05 Problem Set 8, Spring 2025

(due on Gradescope Monday, Apr. 14 at 9:00 PM ET)

Problem 1. (35: 5, 5, 5, 5, 5, 5, 5 pts.) **Spinning gold**
When spun on edge 250 times a certain coin came up heads 140 times and tails 110. We can make the statement: 'if the coin is unbiased then the probability of getting a result at least this extreme is 7%.'

(a) Let θ be the probability of coming up heads. Consider the null hypothesis that the coin is fair, $H_0 : \theta = 0.5$. Carefully explain how the 7% value arises. What term describes this value in NHST? Does this value correspond to a one-sided or two-sided test? (Assume this sided-ness in all remaining parts.)

(b) Should we reject H_0 at a significance level of $\alpha = 0.1$? What about $\alpha = 0.05$?

(c) How many heads would we need to have observed out of 250 spins to reject at a significance of $\alpha = 0.01$?

(d) (i) Fix the significance level at $\alpha = 0.05$. Compute the power of the test above for these particular values of the composite H_A: $\theta = 0.55$ and $\theta = 0.6$.

(ii) On one plot, sketch the pmf of the test statistic given the null hypothesis and given each of these two values for the alternative hypothesis. Use the sketch to explain the change in power observed in part (i).

(e) (i) Again fix $\alpha = 0.05$. What is the *fewest* number of spins necessary for the test to have a power of 0.9 for the value $\theta = 0.55$ of the composite H_A?

(ii) As in part (d), draw sketches and explain how they illustrate the change in power.

(f) Suppose we limit ourselves to two hypotheses, $H_0 : \theta = 0.50$ and $H_1 : \theta = 0.55$, with a flat (Bayesian) prior $P(H_0) = P(H_1) = 0.5$. What is the (Bayesian) posterior probability of H_1 given the data?

(g) Suppose instead we have a flat prior on θ over the interval [0,1]. What is the posterior probability that the coin is biased towards heads? Hint: use `pbeta`.

Problem 2. (10: 5, 5 pts.) **Polygraph analogy.**
In an experiment on the accuracy of polygraph tests, 140 people were instructed to tell the truth and 140 people were instructed to lie. Testers used a polygraph to guess whether or not each person was lying. By analogy, let's say H_0 corresponds to the testee telling the truth and H_A corresponds to the testee lying.

(a) Describe the meaning of type I and type II errors in this context, and estimate their probabilities based on the table.

	Testee is truthful	Testee is lying
Tester thinks testee is truthful	131	15
Tester thinks testee is lying	9	125

(b) In NHST, what relationships exist between the terms significance level, power, type I error, and type II error?

Problem 3. (25: 5, 10, 10 pts.) **z-test**
Suppose three radar guns are set up along a stretch of road to catch people driving over the speed limit of 40 mph. Each radar gun is known to have a normal measurement error modeled on $N(0, 5^2)$. For a passing car, let $\bar{x}$ be the average of the three readings, which we assume have independent errors. Our default assumption for a car is that it is not speeding.

(a) Describe the above story as a NHST. Are the most natural null and alternative hypotheses simple or composite? Is the test one- or two-sided?

(b) The police would like to set a threshold on $\bar{x}$ for issuing tickets so that no more than 4% of non-speeders are given tickets. Assume the threshold is set conservatively, so that no more than 4% of drivers going exactly 40 mph get a ticket.

(i) Use the NHST description in part (a) to determine the correct threshold.

(ii) Sketch a graph illustrating your reasoning in part (i).

(iii) What is the probability that a person given a ticket was in fact not speeding?

(iv) Suppose word gets out about the speed trap and no one speeds along that road anymore. Now, what percentage of tickets are given in error?

(c) Consider the specific value of the alternative hypothesis: $\mu = 45$.

(i) Use R to determine the power of this test.

Next, determine how many cameras are needed to achieve a power of 0.9 with $\alpha = 0.04$ **in two ways**.

(ii) First, compute the power for increasing integer values of n using `qnorm` and `pnorm` until you first exceed 0.9.

(iii) Second, work out the minimum real number n algebraically using standardization and `qnorm`, and then round up.

Problem 4. (25: 5, 5, 5, 5, 5 pts.) **Climate change in Massachusetts**
Download and view the file `climate-data-MA.csv` from the Files/Other folder on Canvas. The second column gives the average annual temperature in Fahrenheit in Massachusetts for each year from 1895 to 2023. Source:
`https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/statewide/time-series`

Load this data into R using `data=read.csv('path/to/file')` and plot it with `plot(data)`.

(a) Compute the sample mean $\overline{x}$ and sample variance s^2 (with denominator $n-1$, not n) for the 20 years from 1895 to 1914. Do the same for the 20 years from 2004 to 2023.

(b) Using the pooled variance formula (assuming equal variance), perform a two-sample t-test of whether temperatures from 2004 to 2023 were *significantly higher* on average than temperatures from 1895 to 1914. What is the value of the t statistic? What is the p-value? Should you reject the null at a significance level of $\alpha = 0.001$?

(c) Use the R command `t.test` to obtain the same values as in (b). Be sure to read the documentation to order the arguments and set `alternative` and `var.equal` correctly. Include your code and output here.

Note: it'd be safer to not assume equal variances, but you can check with R that the conclusion is robust to this choice. See here for more background:
`https://en.wikipedia.org/wiki/Welch%27s_t-test`

(d) As in Studio 6, run `lm = lm(AverageTempMA ~ Year, data)` to model the average annual temperature as a linear function of the year. Use `abline(lm)` to add the regression line to the scatter plot and include your plot here.

Using `summary(lm)`, what t statistic and p-value are associated with the null hypothesis that the slope (coefficient of Year) is zero? Based on this p-value, should you reject the null in favor of the alternative of a non-zero slope? (We'll dig into this info more later.)

(e) Based on (d), predict the average annual temperature in 2100.

Problem 5. (5 (extra credit): 2, 3 pts.) **Interpreting XKCD**
Give an **accurate and precise** explanation of the following XKCD comics.

(a) http://xkcd.com/1132/

(b) http://xkcd.com/882/ Hint: look up the "multiple comparisons problem."