

Midterm Exercises

C. Durso

Winter, 2019

Instructions

Please work these problems on your own. You may use web searches, but not interactive methods such as asking others online or in person. In questions with code blocks, full credit will reserved for effective use of R to reach a correct solution.

Questions

1. Construct a reasonable, standard model for rolling a fair (each face is equally likely) ten-sided die twice and recording the values in order. You don't have to explain the model, just provide the values requested below.

What is the probability of single outcome (a, b) with $a, b \in \{1, 2, 3, \dots, 10\}$? (5 points)

What is the probability of the event $(a, b) | a \neq b$? (5 points)

2. The function defined by $f(x) = cx$ for $x \in [0, 1]$ and $f(x) = 0$ otherwise is the probability density function of a continuous random variable. What is the value of c ? (10 points)
3. Construct a reasonable model for flipping a fair coin 4 times and recording the results in order. You don't have to explain the model, just provide the information requested below.

Are the event that the number of heads is even and the event that the first flip is a head independent? (5 points)

Are the event that the number of heads is at least 3 and the event that the first flip is a head independent? (5 points)

4. Consider a continuous random variable X with the probability density function defined by $f(x) = \frac{3}{4}(1 - x^2)$ for $x \in [-1, 1]$ and $f(x) = 0$ otherwise. What is the value of the cumulative distribution of X , $F(t)$, at $t = \frac{1}{2}$? (5 points)
5. If you model data in which there are 100 trials, of which 36 are successes, as a sample from the Binomial distribution with the size parameter equal to 100, what is the maximum likelihood estimate for the probability parameter? (5 points)
6. If you model the "total" column in the "animal_stats_compact.csv" as a sample from a Normal distribution, what are the maximum likelihood values of μ and σ^2 ? (Note that this may not be a very accurate model.) (10 points)
7. In the data from #6, if you model the "dog_surg" variable as a linear function of the "dog" variable, what slope and intercept give the line with the least squares best fit? (You may use a built-in function for this or the formulas derived in class) (10 points)
8. Make a scatterplot with the "dog" variable on the horizontal axis and the "dog_surg" variable on the vertical axis. Add the line computed in #7. (If you used the built-in function for #7, please extract the values of the slope and intercept from the fitted model object, rather than using copy-paste.) (5 points)
9. What is the total number of dogs in each city? (To do this without looping through the city values, you can use "summarize" from dplyr, with group_by.) (5 points)

10. What is the probability that a Normal random variable with $\mu = 10$ and $\sigma^2 = 4$ lies between 11 and 12? (5 points)
11. At what value is the cumulative distribution of a Normal random variable with $\mu = 10$ and $\sigma^2 = 4$ equal to 0.80? (5 points)
12. Set the random seed equal to 12345. Sample the Binomial distribution with size equal to 12 and probability of success equal to 0.5, $\text{Binom}(12, 0.5)$, ten times. Display the resulting numbers. (5 points)
13. Again, set the random seed to 12345. Sample the integers between $\{1, 2, \dots, 50\}$ ten times without replacement. What is the last value in the sample? (5 points)
14. Plot the standard Normal density curve, the Student's t density curve (the function `dt`) for $\text{df}=5$ in green, and the Student's t density curve for $\text{df}=20$ in orange. (5 points)
15. Set the random seed to 12345 again. Generate a sample of size 100 from the uniform distribution on $[-10, 10]$ and plot the histogram using the `binwidth=1` and the `"..density.."` aes. On this, plot the horizontal line at height $1/20$. We haven't used the `geom` for a horizontal line. You will need to identify it and figure out the syntax yourself.) (5 points)