Midterm 2: Practice Test

This exam will be held on Tuesday, April 8th. It is 75 minutes long and is made up of 20 short-answer questions, each worth 5 points for 100 points in total. Grab your pencil (no calculators needed!) and get ready to show what you know. Good luck—you're going to do great!

1) Quiz 1B Question 4

If you wanted to find the median GDP per capita for each continent, which two dplyr functions would you need to use together?

2) Quiz 1D Question 1

Write an anonymous function that multiplies its input by 3.

3) Quiz 2B Question 1

In the equation $y_i = \beta_0 + \beta_1 x_i + u_i$, explain what β_0 and β_1 represent. How are they different from $\hat{\beta}_0$ and $\hat{\beta}_1$?

4) Assignment 2.1 Question 6)

Let X be the random variable "the number of likes you get on a social media post" where you get 0-4 likes per post each with equal probability. What is the variance of X?

5) Assignment 2.3 Question 10)

Consider the linear model without an intercept $y = \beta_0 x + u$. Recall that OLS minimizes the sum of squared residuals. Write down what that means in this context mathematically, and then take first order conditions. Show that the OLS estimator is $\hat{\beta}_0 = \frac{\sum_i x_i y_i}{\sum_i x_i^2}$.

6) Assignment 3.1

Explain what a linear probability model is. What are the weaknesses of this approach, and what alternatives exist?

Answer:

A linear probability model has a binary dependent variable (like yes/no or win/lose), and you use a linear regression to see the effects of explanatory variables on the probability the dependent variable takes on one value versus the other. The problems are that the predicted values can be less than 0 or more than 1 (which are not valid probabilities), marginal effects are assumed to be linear, and the disturbance term is only ever 0 or 1 (therefore not distributed normally), which makes standard errors invalid.

Alternatives are any other classification method, including logits, probits, K Nearest Neighbors, and decision trees.

7) Assignment 3.1

The logistic regression (logit) models the log odds of the binary dependent variable as a linear function of explanatory variables:

$$\log_e \left(\frac{Prob(y = Y)_i}{1 - Prob(y = Y)_i} \right) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$$

Show that if you solve for $Prob(y = Y)_i$, you get $Prob(y = Y)_i = \frac{\exp(X'\beta+u)}{1+\exp(X'\beta+u)}$. Then explain why logits never predict a probability outside of the [0, 1] range.

8) Assignment 3.2

If $X \sim N(\mu, 1)$ and X is observed to take on values 0 and 2, show how to use the method of maximum likelihood to find that the most likely value for μ is 1.

9) Assignment 3.3

Explain what the Bayes classifier is and why it is useful.

10) Assignment 3.4

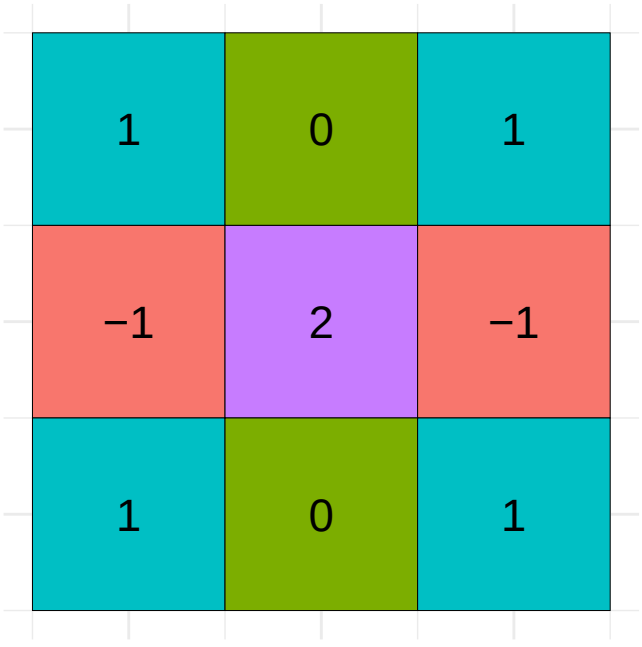
Explain the Bias-Variance trade-off.

11) Assignment 4.1

What is a probability transition matrix? Give an example.

12) Assignment 4.2

Consider this GridWorld. Assuming an agent can move in any direction or stay in the same place, and find the policy function and the value function given a discount rate of $\beta = 0.9$.



position	policy	value
1 (upper left)		
2 (upper middle)		
3 (upper right)		
4 (center left)		
5 (center)		
6 (center right)		
7 (lower left)		
8 (lower middle)		
9 (lower right)		

13) Assignment 4.2

What is a discount rate? Explain the two different interpretations.

14) Unit 4

Summarize the Markov Decision Process in John Rust's 19987 paper *Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher*.

15) Unit 4

Summarize the Nested Fixed Point algorithm.

Answer:

For each guess for the payoff function, you:

- Use value function iteration to find the value function
- Given the value function and a choice for the functional form of the unobservables (extreme value –> logit; normal –> probit), calculate the choice probabilities (probability of choosing action A in state S)
- Calculate the likelihood function (number of times the agent was in state S and chose action A times the choice probability they chose action A in state S)
- Update your guess for the payoff function based on maximizing the likelihood function.

Continue until the likelihood function is maximized.

16) Assignment 5.1

Take this decision tree:

What is the predicted log salary for a baseball player who has been playing in the major leagues for 3 years and has 120 hits the previous season?

Answer: 5.11

17) Assignment 5.1

Explain: How do you construct a decision tree?

18) Assignment 5.2

What is the major strength and weakness of decision trees?

Answer:

- Strength: easy to interpret
- Weakness: high variance. Different training data yields very different decision trees.

19) Assignment 5.3

How does bagging address the major issue with decision trees?

20) Assignment 5.4

What is the difference between bagging and random forests?