

Midterm practice (2024)

Name: _____

Question 1: R code interpretation

Consider this R code.

```
gender <- c("male", "female", "female", "male", "female")
educ <- c(rep("HS", 3), "college", "college")
earnings <- c(1, 0, 2, 8, 4)
```

(1a) What is `mean(earnings)`?

(1b) What is `mean(earnings[gender == "female"])`?

(1c) What is `mean(earnings[gender == "female" & educ == "HS"])`?

(1d) What is `mean(earnings[earnings <= 4])`?

Question 2: Proof interpretation

Theorem: If events A and B are independent and $P(B) > 0$, then $P(A \mid B) = P(A)$.

Proof:

$$P(A \cap B) = P(A)P(B) \quad (\text{Step 1})$$

$$P(A \mid B)P(B) = P(A)P(B) \quad (\text{Step 2})$$

$$P(A \mid B) = P(A) \quad (\text{Step 3})$$

Explain what definition/property/mathematical operation is being used in each step of the proof.

(2a) Step 1:

(2b) Step 2:

(2c) Step 3:

(2d) Explain in words what $P(A \mid B) = P(A)$ means.

Question 3: Joint distribution of two random variables

Consider the joint PMF of two random variables, X and Y :

x	y	$f(x, y)$
0	0	1/4
0	1	1/5
1	0	1/5
1	1	1/10
1	2	1/4

(3a) What is the marginal distribution of X , i.e. $f_X(x)$?

x	$f_X(x)$
0	
1	

(3b) What is the expectation of X , i.e. $E[X]$?

(3c) What is the standard deviation of X , i.e. $\sigma[X]$?

(3d) What is the conditional distribution of Y given X , i.e. $f_{Y|X}(y | x)$?

x	y	$f_{Y X}(y x)$
0	0	
0	1	
1	0	
1	1	
1	2	