Assignment 1 Non-Programming

Graded

Student

Nakul Kuttua

Total Points

87.5 / 141 pts

Question 1

Assignment Directions

0 / 0 pts

✔ **+ 0 pts** Correct

+ 0 pts Incorrect

Question 2

Problem 1.1: Determinant and Inverse of a Matrix **15** / 15 pts

2.1 **Problem 1.1a** **4** / 4 pts

✔ **– 0 pts** Correct

– 0.5 pts Incorrect Equation

– 1 pt Incorrect determinant

– 3 pts No work shown

– 4 pts No attempt or missing solution

2.2 **Problem 1.1b** **3** / 3 pts

✔ **– 0 pts** Correct

– 0.5 pts Did not discuss singularity or incorrect singularity

– 0.5 pts Did not discuss rank or incorrect rank

– 2 pts Value of r is wrong

– 3 pts No attempt or missing solution

2.3 **Problem 1.1c** **3** / 3 pts

✔ **– 0 pts** Correct- Only 1 linear combination required

– 2 pts Incorrect answer on r & linear combinations

– 1 pt Incorrect linear combination

– 1 pt Example of linear combination not provided, but argued through another method (e.g. RREF has zeros in last row, similar rows after row operations etc.)

– 2 pts Did not provide any linear combination

– 3 pts No attempt or missing solution

2.4 **Problem 1.1d** **2** / 2 pts

✔ **– 0 pts** Correct

– 1 pt Incorrect inverse, but work shown

– 2 pts Incorrect inverse, no work shown

– 2 pts No attempt or missing solution

2.5 **Problem 1.1e** **3** / 3 pts

✔ **– 0 pts** Correct

– 1.5 pts Incorrect or missing determinant for M^{-1}

– 1.5 pts Incorrect relationship or did not discuss the relationship between determinant of M and M^{-1}

– 3 pts No attempt or missing solution

Question 3

Problem 1.2: Characteristic Equation **8** / 8 pts

✔ **– 0 pts** Correct

– 2 pts Generally correct but missing one or two steps

– 3 pts Does not sufficiently explain why $|A - \lambda I| = 0$

– 4 pts Did not explain why $|A - \lambda I| = 0$

– 4 pts Vague or circular definition or proof

– 8 pts No attempt or missing solution

– 1 pt Minor error

Question 4

Problem 1.3: Eigenvalues and Eigenvectors **4** / 5 pts

4.1 **Problem 1.3.1a** **3** / 4 pts

– 0 pts Correct

✔ **– 1 pt** Minor error(s) in calculation/final answer

– 2 pts Major error(s) in calculation/final answer

– 2 pts Did not solve for λ

– 2.5 pts No work shown

– 4 pts No attempt or missing solution

did not solve for lamda

4.2 **Problem 1.3.1b** **1** / 1 pt

✔ **– 0 pts** Correct

– 0.25 pts one λ expression is incorrect

– 0.5 pts Incorrect or missing λ expression

– 1 pt No attempt or missing solution

Question 5

Problem 1.3.2 Eigenvectors **10** / 10 pts

5.1 **Problem 1.3.2a** **3** / 3 pts

✔ **– 0 pts** Correct

– 1 pt Incorrect eigenvalue(s)

– 3 pts No work or explanation shown

– 3 pts No attempt or missing solution

5.2 **Problem 1.3.2b** **7** / 7 pts

✔ **– 0 pts** Correct

– 0.5 pts The normalized eigenvectors for first lambda are actually representing the same thing.

– 1 pt Did not normalize eigenvectors at all or correctly

– 2 pts Small error in calculating eigenvectors or missing one or two steps

– 4 pts Incomplete or incorrect work for calculating eigenvectors

– 5 pts Incorrect eigenvectors and approach

– 7 pts No work shown

– 7 pts No attempt or missing solution

Question 6

Problem 2: Expectation, Co-variance, and Statistical Independence **3** / 9 pts

6.1 **Problem 2.1: Covariance** **3** / 5 pts

- **– 0 pts** Correct
- **– 1 pt** Minor error
- **– 1 pt** Incorrect answer about values of c
- **– 2 pts** Insufficient work shown
- ✔ **– 2 pts** Incorrect value for $Cov(Y, Z)$
- **– 3 pts** Missing calculation for $Cov(Y, Z)$
- **– 5 pts** No attempt or missing solution
- **– 0 pts** Click here to replace this description.

E[Y^2] is not zero (and not the same as E[Y]^2). You can use the formula for variance to calculate E[Y^2]. Here are the calculations.

Var(Y) = E[Y^2] - E[Y]^2
1 = E[Y^2] - 0
E[Y^2] = 1

6.2 **Problem 2.2: Correlation Coefficient** **0** / 4 pts

- **– 0 pts** Correct
- **– 1 pt** Incorrect or missing value of $\rho(X, Z)$
- **– 1 pt** Minor calculation error
- **– 2 pts** Major calculation error
- **– 3 pts** No work shown
- ✔ **– 4 pts** Missing solution
- **– 0.25 pts** Very minor calculation error

Question 7

Problem 3: Optimization **2.5** / 19 pts

7.1 **Problem 3a** **2** / 2 pts

✔ **– 0 pts** Correct

– 0.5 pts Incorrect maximization function

– 0.5 pts Incorrect minimization function

– 2 pts No attempt or missing solution

lagrange multipliers can be 0

7.2 **Problem 3b** **0.5** / 2 pts

– 0 pts Correct

– 0 pts Correct KKT conditions based on Lagrange function in part a

✔ **– 0.5 pts** Missing or incorrect stationarity condition

– 0.5 pts Missing or incorrect complementary slackness condition

✔ **– 0.5 pts** Missing or incorrect primal feasibility condition

✔ **– 0.5 pts** Missing or incorrect dual feasibility condition

– 2 pts No attempt or missing solution

7.3 **Problem 3c** **0** / 7 pts

– 0 pts Correct

– 1 pt Error in λ_1 active and λ_2 active

– 1 pt Error in λ_1 active and λ_2 inactive

– 1 pt Error in λ_1 inactive and λ_2 active

– 1 pt Error in λ_1 inactive and λ_2 inactive

– 2.5 pts Work shown but does not explore all 4 possibilities

– 2.5 pts Generally correct method but incorrect calculations due to initial error in Lagrange function

– 4 pts Error in method and incorrect calculations due to initial error in Lagrange function

– 4 pts No work shown solving 4 possibilities

✔ **– 7 pts** No attempt or missing solution

7.4 **Problem 3d** **0** / 2 pts

- **– 0 pts** Correct
- **– 1 pt** Missing correct candidate point
- **– 1 pt** Extra candidate points
- **– 1 pt** Incorrect candidate points due to incorrect initial Lagrange Function
- ✔ **– 2 pts** No attempt or missing solution

7.5 **Problem 3e** **0** / 2 pts

- **– 0 pts** Correct
- **– 0 pts** Correct maximizing point based on listed points in part d
- **– 1 pt** Incorrect maximizing point
- **– 0.5 pts** Missing or incorrect calculations for Hessian and Second Partial Derivative Test
- **– 0.5 pts** Missing or incorrect convexity for $L(x, y)$
- ✔ **– 2 pts** No attempt or missing solution

7.6 **Problem 3f - BONUS FOR ALL** **0** / 4 pts

- **– 0 pts** Correct
- **– 1 pt** No explanation
- **– 2 pts** Incorrect graph due to issues in previous sections
- **– 2 pts** No code/incorrect graph
- ✔ **– 4 pts** Did not attempt/no solution

Question 8

Problem 4.1: Discrete Example **10** / 10 pts

- ✔ **– 0 pts** Correct
- **– 1 pt** Minor error.
- **– 2 pts** Work shown, incorrect final answer.
- **– 2 pts** Did not set the derivative to zero
- **– 3 pts** Wrong or missed log-likelihood function.
- **– 3 pts** Wrong or missed derivative of the log-likelihood function.
- **– 9 pts** No work shown, correct answer.
- **– 10 pts** No attempt or missing solution

Question 9

Problem 4.2: Poisson Distribution - BONUS FOR UNDERGRADS **0** / 15 pts

9.1 **Problem 4.2a BONUS FOR UG** **0** / 2 pts

- **– 0 pts** Correct
- **– 1 pt** work shown, incorrect likelihood.
- **– 1 pt** Expression missing
- **– 2 pts** No work, incorrect likelihood.
- ✔ **– 2 pts** Not attempted.

9.2 **Problem 4.2b BONUS FOR UG** **0** / 3 pts

- **– 0 pts** Correct
- **– 0.5 pts** minor error
- **– 1 pt** Work shown, incorrect likelihood.
- **– 2 pts** No work, incorrect likelihood.
- ✔ **– 3 pts** Not attempted.

9.3 **Problem 4.2c BONUS FOR UG** **0** / 10 pts

- **– 0 pts** Correct
- **– 1 pt** minor error
- **– 2 pts** work shown, incorrect MLE.
- **– 3 pts** Inadequate work shown. Correct MLE.
- **– 5 pts** Inadequate work shown. Incorrect MLE.
- **– 8 pts** Missing steps and incorrect approach
- **– 9 pts** No work. Correct answer.
- ✔ **– 10 pts** Missing submission.

Question 10

Problem 5.1: Marginal Distribution **6** / 6 pts

10.1 **Problem 5.1a** **3** / 3 pts

✔ **– 0 pts** Correct

– 3 pts Incorrect

– 0.75 pts Incorrect $P(X = 0)$

– 0.75 pts Incorrect $P(X = 1)$

– 0.75 pts Incorrect $P(Y = 0)$

– 0.75 pts Incorrect $P(Y = 1)$

– 3 pts No partial credit

10.2 **Problem 5.1b** **3** / 3 pts

✔ **– 0 pts** Correct

– 0.5 pts Minor math error

– 1 pt Partial work shown, Correct Answer

– 1.5 pts Work shown, but incorrect/missing final answer

– 3 pts No attempt or no work shown

– 3 pts Incorrect

– 0 pts Incorrect due to Incorrect Marginal Distributions from previous section

Question 11

Problem 5.2: Mutual Information and Entropy **19** / 19 pts

11.1 **Problem 5.2a** **3** / 3 pts

✔ **– 0 pts** Correct

– 1 pt Used log10 instead of log2

– 1 pt Minor Error

– 2 pts 5.2.a Work shown but Incorrect Final Answer

– 3 pts 5.2.a Incorrect/Missing H(Y)

11.2 **Problem 5.2b** **8** / 8 pts

✔ **– 0 pts** Correct

– 3 pts Used log10 instead of log2

– 1 pt Mostly correct method with minor error

– 2 pts Missing step

– 4 pts Incorrect answer, work shown

– 4 pts 5.2.b One incorrect conditional entropy

– 8 pts 5.2.b Incorrect/Missing conditional entropies for both x's

11.3 **Problem 5.2c** **4** / 4 pts

✔ **– 0 pts** Correct

– 1 pt Used log10 instead of log2

– 1 pt Math error or minor error

– 1 pt 5.2.c Incorrect/missing claim of which x is more informative

– 1.5 pts 5.2.c One incorrect mutual information value

– 2 pts 5.2.c Incorrect mutual informations, but work shown

– 3 pts Incorrect/Missing mutual information

– 4 pts 5.2.c Both incorrect mutual informations and incorrect/missing decision on which x is more informative

11.4 **Problem 5.2d** **4** / 4 pts

✔ **– 0 pts** Correct

– 1 pt Used log10 instead of log2

– 1 pt Minor Error

– 2 pts Incorrect answer, work shown

– 2 pts 5.2.d Incorrect mutual entropy but correct formula setup

– 4 pts 5.2.d Incorrect/Missing Joint Entropy

Question 12

Problem 5.3: Entropy Proofs **2** / 10 pts

12.1 **Problem 5.3a** **2** / 3 pts

– 0 pts Correct

✔ **– 1 pt** (a) Error in definition or not rigorous/formal enough

– 2 pts (a) Error in definitions

– 3 pts (a) Incorrect definitions/ Not attempted.

p_+ is not clearly defined
H(X|Y) should have p(y) in numerator

12.2 **Problem 5.3b** **0** / 7 pts

– 0 pts Correct

– 1 pt (b) Minor error or missing step in proof.

– 2 pts (b) Errors or missing steps in proof

– 4 pts (b) Error in proof. Refer to comment.

✔ **– 7 pts** (b) Incorrect/ Not attempted.

Question 13

Problem 6: Programming **0** / 0 pts

✔ **+ 0 pts** Correct

+ 0 pts Incorrect

Question 14

Problem 7: Bonus for All **8** / 15 pts

14.1 **Problem 7a BONUS FOR ALL** **0** / 5 pts

- **– 0 pts** Correct
- **– 1 pt** Minor error
- **– 2.5 pts** Partially correct
- **– 4 pts** Attempted problem, incorrect
- ✔ **– 5 pts** Not attempted or incorrect

14.2 **Problem 7b BONUS FOR ALL** **5** / 5 pts

- ✔ **– 0 pts** Correct
- **– 1 pt** Minor error
- **– 2.5 pts** Partially correct
- **– 4 pts** Attempted problem, incorrect
- **– 5 pts** Not attempted or incorrect

14.3 **Problem 7c BONUS FOR ALL** **3** / 5 pts

- **– 0 pts** Correct
- **– 1 pt** Minor error
- ✔ **– 2 pts** Incorrect variance
- **– 2 pts** Incorrect or missing probability distribution
- **– 2.5 pts** Partially correct
- **– 4 pts** Attempted problem, incorrect
- **– 5 pts** Not attempted or incorrect

Q1 Assignment Directions
0 Points

Class Policies

- No unapproved extension of the deadline is allowed. For late submissions, please refer to the course website.
- Discussion is encouraged on Ed as part of the Q/A. However, all assignments should be done individually.
- **Plagiarism is a serious offense**. You are responsible for completing your own work. You are not allowed to copy and paste, or paraphrase, or submit materials created or published by others, as if you created the materials. All materials submitted must be your own.
- All incidents of suspected dishonesty, plagiarism, or violations of the Georgia Tech Honor Code will be subject to the institute's Academic Integrity procedures. If we observe any (even small) similarities/plagiarisms detected by Gradescope or our TAs, **WE WILL DIRECTLY REPORT ALL CASES TO OSI**, which may, unfortunately, lead to a very harsh outcome. **Consequences can be severe, e.g., academic probation or dismissal, grade penalties, a 0 grade for assignments concerned, and prohibition from withdrawing from the class.**

Instructions

- We will be using Gradescope for submission and grading of assignments.
- **Unless a question explicitly states that no work is required to be shown, you must provide an explanation, justification, or calculation for your answer.**
- Throughout this assignment, you will have the opportunity to show your work by attaching screenshots. You may use either Latex, markdown, or any word processing software to show your work. We will **NOT** accept handwritten work. Make sure that your work is formatted correctly, for example submit $\sum_{i=0} x_i$ instead of \text{sum_{i=0} x_i}.
- All assignments should be done individually, each student must write up and submit their own answers.
- **Graduate Students**: You are required to complete any sections marked as Bonus for Undergrads

Tips

- **A useful video tutorial on LaTeX has been created by our TA team** and can be found here and an Overleaf document with the commands can be found here.
- Utilize Gradescope's guides for LaTeX and Markdown to properly format your answers.

◉ Click here to confirm your understanding of the directions and/or notes above

Q2 Problem 1.1: Determinant and Inverse of a Matrix

15 Points

Given a matrix $\boldsymbol{M}$:

$$\boldsymbol{M} = \begin{bmatrix} 2 & 1 & -3 \\ -6 & r & 2 \\ -2 & 4 & 2 \end{bmatrix}$$

Q2.1 Problem 1.1a
4 Points

Calculate the determinant of $\boldsymbol{M}$ in terms of r. (Calculation process required.)

Provide your equation using LaTeX in the box below or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 4.34.51 PM.png Download

Given a matrix M:

$$M = \begin{bmatrix} 2 & 1 & -3 \\ -6 & r & 2 \\ -2 & 4 & 2 \end{bmatrix}$$

(a) Calculate the determinant of $\boldsymbol{M}$ in terms of r (calculation process is required). [4pts]

$$\begin{aligned} \det(M) &= 2(r \cdot 2 - 2 \cdot 4) - (-6 \cdot 2 + 2 \cdot 2) - 3(-6 \cdot 4 - r \cdot 2) \\ &= 2(2r - 8) - (-8) - 3(-24 - 2r) \\ &= 4r - 16 + 8 + 12 + 6r \\ &= 64 - 2r \end{aligned}$$

Show your work using LaTeX in the provided box or upload a screenshot of your work. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 4.34.51 PM.png Download

Given a matrix M:

$$M = \begin{bmatrix} 2 & 1 & -3 \\ -6 & r & 2 \\ -2 & 4 & 2 \end{bmatrix}$$

(a) Calculate the determinant of $\boldsymbol{M}$ in terms of r (calculation process is required). [4pts]

$$\begin{aligned} \det(M) &= 2(r \cdot 2 - 2 \cdot 4) - (-6 \cdot 2 + 2 \cdot 2) - 3(-6 \cdot 4 - r \cdot 2) \\ &= 2(2r - 8) - (-8) - 3(-24 - 2r) \\ &= 4r - 16 + 8 + 12 + 6r \\ &= 64 - 2r \end{aligned}$$

Q2.2 Problem 1.1b
3 Points

For what value(s) of r does $\boldsymbol{M}^{-1}$ not exist?

If there are multiple values, enter your answer in ascending order as a comma separated list with no spaces up to 3 decimal places (e.g. 1,2,3).

32

Show your work using LaTeX in the provided box or upload a screenshot of your work. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 4.42.25 PM.png Download

(b) For what value(s) of r does $\boldsymbol{M}^{-1}$ not exist?
64 - 2r = 0
64 = 2r
r = 32
The matrix inverse does not exist for r = 32. Why doesn't $\boldsymbol{M}^{-1}$ exist in this case? What does it mean in terms of rank and singularity for these values of r? [3pts]
For this value, the matrix is singular and rank is less than 3

Why doesn't $\boldsymbol{M}^{-1}$ exist in this case? What does it mean in terms of rank and singularity for these values of r?

The matrix is singular for that value, and this means the matrix is singular and the rank is less than 3.

Q2.3 Problem 1.1c
3 Points

Will all values of r found in part (b) allow for a column (or row) to be expressed as a linear combination of the other columns (or rows)?

(●) Yes

( ) No

If yes, provide:

- **either** the linear equation of the third column $\boldsymbol{C_3}$ as a linear combination of the first column $\boldsymbol{C_1}$ and second column $\boldsymbol{C_2}$
- **or** the linear equation of the second row $\boldsymbol{R_2}$ as a linear combination of the first row $\boldsymbol{R_1}$ and third row $\boldsymbol{R_3}$.

If no, explain why.

If yes, provide your equation using LaTeX in the box below or upload a screenshot of your work. Blurry screenshots or screenshots of handwritten work will result in 0 credit. If no, type your answer in the box below.

Screenshot 2023-02-09 at 5.43.15 PM.png Download

(c) Will all values of r found in part (b) allow for a column (or row) to be expressed as a linear combination of the other columns (or rows)? **If yes**, provide

- **either** the linear equation of the third column $\boldsymbol{C_3}$ as a linear combination of the first column $\boldsymbol{C_1}$ and second column $\boldsymbol{C_2}$
- **or** the linear equation of the second row $\boldsymbol{R_2}$ as a linear combination of the first row $\boldsymbol{R_1}$ and third row $\boldsymbol{R_3}$.

If no, explain why. [3pts]

$$\begin{bmatrix} 2 & 1 & -3 \\ -6 & 32 & 2 \\ -2 & 4 & 2 \end{bmatrix}$$

$4R_1 + 7R_3 = R_2$

Q2.4 Problem 1.1d
2 Points

Write down $\boldsymbol{M}^{-1}$ for $r = 0$. (calculation process is **NOT** required)

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 5.49.33 PM.png Download

(d) Write down $\boldsymbol{M}^{-1}$ for $r = 0$ (calculation process is **NOT** required). [2pts]

$$M^{-1} = \begin{bmatrix} 2 & 1 & -3 \\ -6 & 0 & 2 \\ -2 & 4 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -1/8 & -7/32 & 1/32 \\ 1/8 & -1/32 & -1/32 \\ -3/8 & -5/32 & 3/32 \end{bmatrix}$$

Q2.5 Problem 1.1e
3 Points

Find the mathematical equation that describes the relationship between the determinant of $\boldsymbol{M}$ and the determinant of $\boldsymbol{M}^{-1}$.

NOTE: It may be helpful to find the determinant of $\boldsymbol{M}$ and $\boldsymbol{M}^{-1}$ for $r = 0$.

Provide your equation using LaTeX in the box below or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

▾ **Screenshot 2023-02-09 at 6.02.21 PM.png** Download

(e) Find the mathematical equation that describes the relationship between the determinant of $\boldsymbol{M}$ and the determinant of $\boldsymbol{M}^{-1}$. [3pts]

NOTE: It may be helpful to find the determinant of $\boldsymbol{M}$ and $\boldsymbol{M}^{-1}$ for $r = 0$.

$$\det^{-1}(M) = \frac{-1}{8}\left(\frac{1}{32}\right) + \frac{7}{32}\left(\frac{27}{256}\right) + \frac{1}{32}\left(\frac{-1}{32}\right)$$
$$= \frac{1}{64}$$

For $r = 0, \det\left(M^{-1}\right) = \frac{1}{64}$ The relation between $\det(M)$ and $\det\left(M^{-1}\right)$ is $\det(M) = \frac{1}{\det(M)}$

Show your work using LaTeX in the provided box or upload a screenshot of your work. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

▾ **Screenshot 2023-02-09 at 6.02.21 PM.png** Download

(e) Find the mathematical equation that describes the relationship between the determinant of $\boldsymbol{M}$ and the determinant of $\boldsymbol{M}^{-1}$. [3pts]

NOTE: It may be helpful to find the determinant of $\boldsymbol{M}$ and $\boldsymbol{M}^{-1}$ for $r = 0$.

$$\det^{-1}(M) = \frac{-1}{8}\left(\frac{1}{32}\right) + \frac{7}{32}\left(\frac{27}{256}\right) + \frac{1}{32}\left(\frac{-1}{32}\right)$$
$$= \frac{1}{64}$$

For $r = 0, \det\left(M^{-1}\right) = \frac{1}{64}$ The relation between $\det(M)$ and $\det\left(M^{-1}\right)$ is $\det(M) = \frac{1}{\det(M)}$

Q3 Problem 1.2: Characteristic Equation

8 Points

Consider the eigenvalue problem:

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}, \boldsymbol{x} \neq 0$$

where $\boldsymbol{x}$ is a non-zero eigenvector and λ is an eigenvalue of $\boldsymbol{A}$. Prove the determinant $|\boldsymbol{A} - \lambda\boldsymbol{I}| = 0$.

NOTE: There are many ways to solve this problem. You are allowed to use linear algebra properties as part of your solution.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 6.11.39 PM.png Download

1.2 Characteristic Equation [8pts]

Consider the eigenvalue problem:

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}, \boldsymbol{x} \neq 0$$

where $\boldsymbol{x}$ is a non-zero eigenvector and λ is an eigenvalue of $\boldsymbol{A}$. Prove the determinant $|\boldsymbol{A} - \lambda\boldsymbol{I}| = 0$.

NOTE: There are many ways to solve this problem. You are allowed to use linear algebra properties as part of your solution.

$$Ax = \lambda x$$
$$Ax - \lambda x = 0$$
$$x(A - I\lambda) = 0$$

$Ax = 0$ means that he have a homogenous system but since $x \neq 0$, it must be a mon-trivial solution Thus $|A - \lambda I| = 0$.

Q4 Problem 1.3: Eigenvalues and Eigenvectors

5 Points

Given a matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

Q4.1 Problem 1.3.1a

4 Points

Find an expression for the eigenvalues λ of $\boldsymbol{A}$ in terms of a, b, and c.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 6.24.32 PM.png Download

Given a matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

(a) Find an expression for the eigenvalues λ of $\boldsymbol{A}$ in terms of a, b, and c. [4pts]

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} a-\lambda & b \\ b & c-\lambda \end{bmatrix}$$

$$(a-\lambda)(c-\lambda) - b^2 = 0$$

$$\lambda^2 - a\lambda - c\lambda = b^2 - ac$$

Q4.2 Problem 1.3.1b
1 Point

Find a simple expression for the eigenvalues if $c = a$.

Provide your equation using LaTeX in the box below or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 6.25.02 PM.png Download

(b) Find a simple expression for the eigenvalues if $c = a$. [1pt]

$$a = c$$
$$a^2 - 2\lambda a + \lambda^2 - b^2 = 0$$
$$\lambda_1 = a - b$$
$$\lambda_2 = a + b$$

Q5 Problem 1.3.2 Eigenvectors

10 Points

Given a matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{bmatrix} x & 12 \\ 3 & x \end{bmatrix}$$

Q5.1 Problem 1.3.2a

3 Points

Calculate the eigenvalues of $\boldsymbol{A}$ as a function of x (calculation process required).

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 6.28.36 PM.png Download

1.3.2 Eigenvectors [10pts]

Given a matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{bmatrix} x & 12 \\ 3 & x \end{bmatrix}$$

(a) Calculate the eigenvalues of $\boldsymbol{A}$ as a function of x (calculation process required). [3pts]

$$\begin{bmatrix} x & 12 \\ 3 & x \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} x-\lambda & 12 \\ 3 & x-\lambda \end{bmatrix}$$

$$(x-\lambda)^2 - 36 = 0$$

$$x^2 - 2\lambda x + \lambda^2 - 36 = 0$$

$$\lambda_1 = x - 6$$

$$\lambda_2 = x + 6$$

Q5.2 Problem 1.3.2b
7 Points

Find the normalized eigenvectors of matrix $\boldsymbol{A}$ (calculation process required).

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 6.38.38 PM.png Download

(b) Find the normalized eigenvectors of matrix $\boldsymbol{A}$ (calculation process required). [7pts]

$$\lambda = x - 6$$

$$\begin{bmatrix} x & 12 \\ 3 & x \end{bmatrix} - \begin{bmatrix} x-6 & 0 \\ 0 & x-6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 12 \\ 3 & 6 \end{bmatrix} \quad \text{Eigenvector} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\text{Normalized Eigenvector} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{-1}{\sqrt{5}} \end{bmatrix}$$

$$\lambda = x - 6$$

$$\begin{bmatrix} x & 12 \\ 3 & x \end{bmatrix} - \begin{bmatrix} x+6 & 0 \\ 0 & x+6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -6 & 12 \\ 3 & -6 \end{bmatrix} \quad \text{Eigenvector} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\text{Normalized Eigenvector} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

Q6 Problem 2: Expectation, Co-variance, and Statistical Independence

9 Points

Q6.1 Problem 2.1: Covariance
5 Points

Suppose X, Y, and Z are three different random variables. Let X obey a Bernoulli Distribution. The probability mass function of X is:

$$p(x) = \begin{cases} 0.7 & x = c \\ 0.3 & x = -c \end{cases}$$

where c is a nonzero constant. Let Y obey the Standard Normal (Gaussian) Distribution, which can be written as $Y \sim N(0,1)$. X and Y are statistically independent (i.e. $P(X|Y) = P(X)$). Meanwhile, let $Z = XY$.

Calculate the covariance of Y and Z (i.e. $Cov(Y, Z)$).

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

▾ **Screenshot 2023-02-09 at 7.07.59 PM.png** Download

2 Expectation, Co-variance and Statistical Independence [5pts + 4pts]

2.1 Covariance [5pts]

Suppose X, Y, and Z are three different random variables. Let X obey a Bernoulli Distribution. The probability mass function for X is:

$$p(x) = \begin{cases} 0.7 & x = c \\ 0.3 & x = -c \end{cases}$$

where c is a nonzero constant. Let Y obey the Standard Normal (Gaussian) Distribution, which can be written as $Y \sim N(0,1)$. X and Y are statistically independent (i.e. $P(X|Y) = P(X)$). Meanwhile, let $Z = XY$.

Calculate the covariance of Y and Z (i.e. $Cov(Y, Z)$). Do values of c affect the covariance between Y and Z? [5pts]

$$Z = XY, Y \sim N(0,1)$$

X and Y are independent

$$\text{cov}(Y,Z) = \frac{\text{cov}(Y,Z)}{\sqrt{V(Y)}\sqrt{V(Z)}}$$
$$E(YZ) = P(X=1)E[YZ \mid X=1] + P(X=-1)E[YZ \mid X=-1]$$
$$E(YZ) = 0.7E\left[Y^2\right] + 0.3E\left[-Y^2\right] = 0$$

Y and Z are not co-related Yes values of c affect the covariance.

Do values of c affect the covariance between Y and Z?

- (●) Yes
- ( ) No

Q6.2 Problem 2.2: Correlation Coefficient
4 Points

Let X and Y be statistically independent random variables with $Var(X) = 7$ and $Var(Y) = 13$. We do not know $E[X]$ or $E[Y]$. Let $Z = 5X + 3Y$. Calculate the correlation coefficient defined as

$$\rho(X, Z) = \frac{Cov(X, Z)}{\sqrt{Var(X)Var(Z)}}$$

If applicable, please round your answer to 3 decimal places.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

No files uploaded

Q7 Problem 3: Optimization

19 Points

Optimization problems are related to minimizing a function (usually termed loss, cost or error function) or maximizing a function (such as the likelihood) with respect to some variable x. The Karush-Kuhn-Tucker (KKT) conditions are first-order conditions for a solution in nonlinear programming to be optimal, provided that some regularity conditions are satisfied. In this question, you will be solving the following optimization problem:

$$\begin{aligned} \underset{x,y}{max} \quad & f(x,y) = xy - 8y \\ \text{s.t.} \quad & g_1(x,y) = 2x^2 + y^2 \leq 20 \\ & g_2(x,y) = x \leq 1 \end{aligned}$$

HINT: Click here for an example maximization problem.
HINT: Click here to determine how to set up the problem for minimization in part (a) and for KKT conditions in part (b).

Q7.1 Problem 3a
2 Points

Write the Lagrange function for the maximization problem. Now change the maximum function to a minimum function (i.e. $\underset{x,y}{min}\ \ f(x,y) = xy - 8y$) and provide the Lagrange function for the minimization problem with the same constraints g_1 and g_2.

NOTE: The minimization problem is only for part (a).

Provide your equation using LaTeX in the box below or upload a screenshot of your answer. Screenshots of handwritten work will result in 0 credit.

Maximization Lagrange Function:

Screenshot 2023-02-10 at 3.17.48 PM.png Download

(a) Write the Lagrange function for the maximization problem. Now change the maximum function to a minimum function (i.e. $\underset{x,y}{min}\ \ f(x,y) = xy-8y$) and provide the Lagrange function for the minimization problem with the same constraints g_1 and g_2. [2pts]

NOTE: The minimization problem is only for part (a).

Maximization

$$L(x,y) = xy - 8y - \lambda_1\left(2x^2 + y^2 - 20\right) - \lambda_2(x-1), \lambda_1, \lambda_2 > 0$$

Minimization

$$L(x,y) = xy - 8y + \lambda_1\left(2x^2 + y^2 - 20\right) + \lambda_2(x-1), \lambda_1, \lambda_2 > 0$$

Minimization Lagrange Function:

Screenshot 2023-02-10 at 3.17.48 PM.png Download

(a) Write the Lagrange function for the maximization problem. Now change the maximum function to a minimum function (i.e. $\underset{x,y}{min}\ \ f(x,y) = xy-8y$) and provide the Lagrange function for the minimization problem with the same constraints g_1 and g_2. [2pts]

NOTE: The minimization problem is only for part (a).

Maximization

$$L(x,y) = xy - 8y - \lambda_1\left(2x^2 + y^2 - 20\right) - \lambda_2(x-1), \lambda_1, \lambda_2 > 0$$

Minimization

$$L(x,y) = xy - 8y + \lambda_1\left(2x^2 + y^2 - 20\right) + \lambda_2(x-1), \lambda_1, \lambda_2 > 0$$

Q7.2 Problem 3b
2 Points

List the names of all 4 groups of KKT conditions and their corresponding mathematical equations or inequalities for this specific maximization problem.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-10 at 3.20.49 PM.png Download

(b) List the names of all 4 groups of KKT conditions and their corresponding mathematical equations or inequalities for this specific maximization problem [2pts]

$$\frac{\partial L}{\partial x} = y + 4x\lambda_1 + \lambda_2 = 0$$
$$\frac{\partial L}{\partial y} = x - 8 + 2y\lambda_1 = 0$$
$$\lambda_1 \left(2x^2 + y^2 - 20\right) = 0$$
$$\lambda_2(x - 1) = 0$$

Q7.3 Problem 3c
7 Points

Solve for 4 possibilities formed by each constraint being active or inactive. Do not forget to check the inactive constraints for each point. Candidate points must satisfy the inactive constraints.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

No files uploaded

Q7.4 Problem 3d
2 Points

List the candidate point(s) (there is at least 1). Please round answers to 3 decimal points and use that answer for calculations in further parts.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

No files uploaded

Q7.5 Problem 3e
2 Points

Find the **one** candidate point for which $f(x, y)$ is largest.

You may either type LaTeX into the provided box or upload a screenshot of your answer.

No files uploaded

Check if $L(x, y)$ is concave or convex at this point by using the Hessian in the second partial derivative test.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

No files uploaded

Q7.6 Problem 3f - BONUS FOR ALL
4 Points

Make a 3D plot of the objective function $f(x, y)$ and constraints g_1 and g_2 using Math3d. Mark the maximum candidate point and include a screenshot of your plot.

NOTE: Use an explicit surface for the objective function, implicit surfaces for the constraints, and a point for the maximum candidate point.

No files uploaded

Briefly explain why your plot makes sense in one sentence.

Q8 Problem 4.1: Discrete Example
10 Points

Mastermind Mahdi decides to give a challenge to his students for their MLE Final. He provides a spinner with 10 sections, each numbered 1 through 10. The students can change the sizes of each section, meaning that they can select the probability the spinner lands on a certain section. Mahdi then proposes that the students will get a 100 on their final if they can spin the spinner 10 times such that it doesn't land on section 1 during the first 9 spins and lands on section 1 on the 10th spin. If the probability of the spinner landing on section 1 is θ, what value of θ should the students select to most likely ensure they get a 100 on their final? Use your knowledge of Maximum Likelihood Estimation to get a 100 on the final.

NOTE: You must specify the log-likelihood function and use MLE to solve this problem for full credit. You may assume that the log-likelihood function is concave for this question.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 7.42.44 PM.png Download

4 Maximum Likelihood [10pts + 15pts Bonus for Undergrads]

4.1 Discrete Example [10pts]

Mastermind Mahdi decides to give a challenge to his students for their MLE Final. He provides a spinner with 10 sections, each numbered 1 through 10. The students can change the sizes of each section, meaning that they can select the probability the spinner lands on a certain section. Mahdi then proposes that the students will get a 100 on their final if they can spin the spinner 10 times such that it doesn't land on section 1 during the first 9 spins and lands on section 1 on the 10th spin. If the probability of the spinner landing on section 1 is θ, what value of θ should the students select to most likely ensure they get a 100 on their final? Use your knowledge of Maximum Likelihood Estimation to get a 100 on the final.

NOTE: You must specify the log-likelihood function and use MLE to solve this problem for full credit. You may assume that the log-likelihood function is concave for this question

Probability of $1 = \theta$

x	1	2	3	4	5	6	7	8	9	10
p	2	$\frac{1-\theta}{9}$	$\frac{1-\theta}{9}$	$\frac{1-\theta}{9}$	$\frac{1-\theta}{9}$	$\frac{1-\theta}{9}$	$\frac{1-\theta}{9}$	$\frac{1-\theta}{9}$	$\frac{1-\theta}{9}$	$\frac{1-\theta}{9}$

$$L(\theta) = \left(\frac{1-\theta}{a}\right)^9 (\theta)$$

$$\ln L(\theta) = 9 \ln \left(\frac{1-\theta}{a}\right) + \ln(\theta)$$

$$\ln L(\theta) = 9(\ln(1-\theta) - \ln(9)) + \ln(\theta)$$

$$\frac{d \ln L(\theta)}{d\theta} = \frac{-9\theta + 1 - \theta}{\theta - \theta^2}$$

$$1 = 10\theta$$

$$\theta = 0.1$$

Q9 Problem 4.2: Poisson Distribution - BONUS FOR UNDERGRADS

15 Points

The Poisson distribution is defined as:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} (k = 0, 1, 2, ...)$$

Q9.1 Problem 4.2a BONUS FOR UG

2 Points

Let $X_1 \sim Poisson(\lambda)$, and x_1 be an observed value of X_1. What is the likelihood given λ?

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

No files uploaded

Q9.2 Problem 4.2b BONUS FOR UG

3 Points

Now, assume we are given n such values. Let $(X_1, ..., X_n) \sim Poisson(\lambda)$ where $X_1, ..., X_n$ are i.i.d. random variables, and $x_1, ..., x_n$ be observed values of $X_1, ..., X_n$. What is the likelihood of this data given λ? You may leave your answer in product form.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

No files uploaded

Q9.3 Problem 4.2c BONUS FOR UG
10 Points

What is the maximum likelihood estimator of λ?

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

No files uploaded

Q10 Problem 5.1: Marginal Distribution

6 Points

Suppose the joint probability distribution of two binary random variables X and Y are given as follows. X are the rows, and Y are the columns.

	Y=0	Y=1
X=0	$\frac{5}{16}$	$\frac{1}{16}$
X=1	$\frac{1}{2}$	$\frac{1}{8}$

Q10.1 Problem 5.1a

3 Points

Show the marginal distribution of X and Y, respectively.

$\boldsymbol{P(X=0)=}$

3/8

$\boldsymbol{P(X=1)=}$

5/8

$\boldsymbol{P(Y=0)=}$

13/16

$\boldsymbol{P(Y=1)=}$

3/16

Q10.2 Problem 5.1b
3 Points

Find mutual information $I(X,Y)$ for the joint probability distribution in the previous question to at least 5 decimal places (please use base 2 to compute logarithm).

Provide your answer as a number up to 5 decimal places in the provided boxes. Gradescope accepts any of the following operators: `+`, `-`, `*`, `/`, `%`, `^`, `()`

$$\boldsymbol{I(X,Y)} =$$

0.00125

Show your work using LaTeX in the provided box or upload a screenshot of your work. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 7.58.22 PM.png Download

(b) Find mutual information $I(X,Y)$ for the joint probability distribution in the previous question to at least 5 decimal places (please use base 2 to compute logarithm) [3pts]

$$I(x,y) = \sum_{x,x} P(x,y) \cdot \log\left(\frac{P(x,y)}{P(x)\cdot P(y)}\right)$$

$$I(x,y) = P(0,0)\cdot\log_2\left(\frac{P(0,0)}{P(0)\cdot P(0)}\right) + P(0,1)\cdot\log_2\left(\frac{P(0,1)}{P(0)\cdot P(1)}\right)$$

$$+ P(1,0)\cdot\log_2\left(\frac{P(1,0)}{P(1)\cdot P(0)}\right) + P(1,1)\cdot\log_2\left(\frac{P(1,1)}{P(1)\cdot P(1)}\right)$$

$$= \left(\frac{5}{16}\right)\cdot\log_2\left(\frac{\frac{5}{16}}{\frac{3}{8}\cdot\frac{13}{16}}\right) + \frac{1}{16}\cdot\log_2\left(\frac{\frac{1}{16}}{\frac{3}{8}\cdot\frac{3}{16}}\right) + \frac{1}{2}\cdot\log_2\left(\frac{\frac{1}{2}}{\frac{5}{8}\cdot\frac{13}{16}}\right) + \frac{1}{8}\cdot\log_2\left(\frac{\frac{1}{8}}{\frac{5}{8}\cdot\frac{3}{16}}\right)$$

$$= 0.00125$$

Q11 Problem 5.2: Mutual Information and Entropy

19 Points

A recent study has shown symptomatic infections are responsible for higher transmission rates. Using the data collected from positively tested patients, we wish to determine which feature(s) have the greatest impact on whether or not some will present with symptoms. To do this, we will compute the entropies, conditional entropies, and mutual information of select features. Please use base 2 when computing logarithms.

Q11.1 Problem 5.2a

3 Points

Find entropy $H(Y)$ to at least 3 decimal places.

Provide your answer as a number up to 3 decimal places in the provided boxes. Gradescope accepts any of the following operators: `+`, `-`, `*`, `/`, `%`, `^`, `()`

$\boldsymbol{H(Y) =}$

0.881

Show your work using LaTeX in the provided box or upload a screenshot of your work. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 8.02.12 PM.png Download

(a) Find entropy $H(Y)$ to at least 3 decimal places. [3pts]

$$H(Y) = \frac{-3}{10} \cdot \log_2\left(\frac{3}{10}\right) - \frac{7}{10} \cdot \log_2\left(\frac{7}{10}\right)$$
$$= 0.881$$

Q11.2 Problem 5.2b
8 Points

Find conditional entropy $H(Y|X_1)$ and $H(Y|X_3)$ to at least 3 decimal places.

Provide your answer as a number up to 3 decimal places in the provided boxes.
Gradescope accepts any of the following operators: `+`, `-`, `*`, `/`, `%`, `^`, `()`

$$\boldsymbol{H(Y|X_1) =}$$

0.761

Show your work using LaTeX in the provided box or upload a screenshot of your work.
Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 8.20.15 PM.png Download

(b) Find conditional entropy $H(Y|X_1)$ and $H(Y|X_3)$ to at least 3 decimal places. [8pts]

	H	M	L	total
T	2	2	3	7
F	1	0	2	3

$$H(y \mid X_1) = \sum p(x,y) \log_2 \frac{p(x)}{p(x,y)}$$

$$0.2 \log \frac{0.3}{0.2} + 0.1 \log \frac{0.3}{0.1} + 0.2 \log \frac{0.2}{0.2} + 0 + 0.3 \log \frac{0.5}{0.3} + 0.2 \log \frac{0.5}{0.2}$$
$$= 0.761$$

	T	F	total
T	6	1	7
F	1	2	3

$$H(y \mid X_3) = \sum p(x,y) \log_2 \frac{p(x)}{p(x,y)}$$

$$0.6 \log \frac{0.7}{0.6} + 0.1 \log \frac{0.7}{0.1} + 0.1 \log \frac{0.3}{0.1} + 0.2 \log \frac{0.3}{0.2}$$
$$= 0.690$$

Provide your answer as a number up to 3 decimal places in the provided boxes.
Gradescope accepts any of the following operators: `+`, `-`, `*`, `/`, `%`, `^`, `()`

$H(Y|X_3) =$

0.690

Show your work using LaTeX in the provided box or upload a screenshot of your work. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 8.20.15 PM.png Download

(b) Find conditional entropy $H(Y|X_1)$ and $H(Y|X_3)$ to at least 3 decimal places. [8pts]

	H	M	L	total
T	2	2	3	7
F	1	0	2	3

$$H(y \mid X_1) = \sum p(x,y) \log_2 \frac{p(x)}{p(x,y)}$$

$$0.2 \log \frac{0.3}{0.2} + 0.1 \log \frac{0.3}{0.1} + 0.2 \log \frac{0.2}{0.2} + 0 + 0.3 \log \frac{0.5}{0.3} + 0.2 \log \frac{0.5}{0.2}$$
$$= 0.761$$

	T	F	total
T	6	1	7
F	1	2	3

$$H(y \mid X_3) = \sum p(x,y) \log_2 \frac{p(x)}{p(x,y)}$$

$$0.6 \log \frac{0.7}{0.6} + 0.1 \log \frac{0.7}{0.1} + 0.1 \log \frac{0.3}{0.1} + 0.2 \log \frac{0.3}{0.2}$$
$$= 0.690$$

Q11.3 Problem 5.2c
4 Points

Find mutual information $I(X_1, Y)$ and $I(X_3, Y)$ to at least 3 decimal places.

Provide your answer as a number up to 3 decimal places in the provided boxes.
Gradescope accepts any of the following operators: `+`, `-`, `*`, `/`, `%`, `^`, `()`

$$\boldsymbol{I(X_1, Y) =}$$

0.120

Show your work using LaTeX in the provided box or upload a screenshot of your work.
Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 8.21.17 PM.png Download

(c) Find mutual information $I(X_1, Y)$ and $I(X_3, Y)$ to at least 3 decimal places and determine which one (X_1 or X_3) is more informative. [4pts]

$$\begin{aligned} I(X_1, Y) &= H(Y) - H(Y \mid X_1) \\ &= 0.881 - 0.761 \\ &= 0.120 \\ I(X_3, Y) &= H(Y) - H(Y \mid X_3) \\ &= 0.881 - 0.690 \\ &= 0.191 \end{aligned}$$

X_3 is more informative

Provide your answer as a number up to 3 decimal places in the provided boxes.
Gradescope accepts any of the following operators: `+`, `-`, `*`, `/`, `%`, `^`, `()`

$$\boldsymbol{I(X_3, Y) =}$$

0.191

Show your work using LaTeX in the provided box or upload a screenshot of your work.
Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 8.21.17 PM.png Download

(c) Find mutual information $I(X_1, Y)$ and $I(X_3, Y)$ to at least 3 decimal places and determine which one (X_1 or X_3) is more informative. [4pts]

$$\begin{aligned} I(X_1, Y) &= H(Y) - H(Y \mid X_1) \\ &= 0.881 - 0.761 \\ &= 0.120 \\ I(X_3, Y) &= H(Y) - H(Y \mid X_3) \\ &= 0.881 - 0.690 \\ &= 0.191 \end{aligned}$$

X_3 is more informative

Which one (X_1 or X_3) is more informative?

- ○ X_1
- ◉ X_3

Q11.4 Problem 5.2d
4 Points

Find joint entropy $H(Y, X_2)$ to at least 3 decimal places.

Provide your answer as a number up to 3 decimal places in the provided boxes. Gradescope accepts any of the following operators: `+`, `-`, `*`, `/`, `%`, `^`, `()`

$$\boldsymbol{H(Y, X_2) =}$$

1.846

Show your work using LaTeX in the provided box or upload a screenshot of your work. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-09 at 8.27.51 PM.png Download

(d) Find joint entropy $H(Y, X_2)$ to at least 3 decimal places. [4pts]

$$H(Y, X_2) = 0.4 \log \frac{1}{0.4} + 0.3 \log \frac{1}{0.3} 0.1 \log \frac{1}{0.1} + 0.2 \log \frac{1}{0.2} = 1.846$$

Q12 Problem 5.3: Entropy Proofs

10 Points

Q12.1 Problem 5.3a
3 Points

Write the discrete case mathematical definition for $H(X|Y)$ and $H(X)$.

Provide your equation using LaTeX in the box below or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

$\boldsymbol{H(X|Y) =}$

▾ Screenshot 2023-02-09 at 8.30.54 PM.png Download

5.3 Entropy Proofs [10pts]

(a) Write the discrete case mathematical definition for $H(X|Y)$ and $H(X)$. [3pts]

$$H(x,y) = \sum_{x \in x} p(x_i)\, H(x \mid y = x_i) \sum_{x \in x, y \in y} p(x_i, y_i) \log \frac{P(x_i)}{p(x_i, y_i)}$$

Entropy measures. uncertainty of x

$$H(x) = -p_+ \log_2(p_+) - p \log_2(p_-)$$

$\boldsymbol{H(X) =}$

▾ Screenshot 2023-02-09 at 8.30.54 PM.png Download

5.3 Entropy Proofs [10pts]

(a) Write the discrete case mathematical definition for $H(X|Y)$ and $H(X)$. [3pts]

$$H(x,y) = \sum_{x \in x} p(x_i)\, H(x \mid y = x_i) \sum_{x \in x, y \in y} p(x_i, y_i) \log \frac{P(x_i)}{p(x_i, y_i)}$$

Entropy measures. uncertainty of x

$$H(x) = -p_+ \log_2(p_+) - p \log_2(p_-)$$

Q12.2 Problem 5.3b
7 Points

Using the mathematical definition of $H(X)$ and $H(X|Y)$ from part (a), prove that $I(X,Y) = 0$ if X and Y are statistically independent.

NOTE: you must provide a mathematical proof and cannot use the visualization shown in class found here

Start from: $I(X,Y) = H(X) - H(X|Y)$

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

No files uploaded

Q13 Problem 6: Programming
0 Points

See the Gradescope submission for Assignment 1 Programming to submit your .py and .pkl files to complete this section.

Problem 6: Programming is worth 2 points.

◉ Click here to confirm your understanding of the directions and/or notes above

Q14 Problem 7: Bonus for All

15 Points

Q14.1 Problem 7a BONUS FOR ALL

5 Points

Let X, Y be two statistically independent $N(0,1)$ random variables, and P, Q be random variables defined as:

$$P = 7X + 2XY^2$$
$$Q = X$$

Calculate the variance $Var(P+Q)$.

HINT: The following equality may be useful:
$Var(XY) = E[X^2Y^2] - [E(XY)]^2$
HINT: $E[Y^4] = \int_{-\infty}^{\infty} y^4 f_Y(y) dy$ where $f_Y(y)$ is the probability density function of Y.

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

No files uploaded

Q14.2 Problem 7b BONUS FOR ALL
5 Points

Suppose that X and Y have joint pdf given by:

$$f_{X,Y}(x,y) = \begin{cases} \frac{x}{2}e^{-y} & 0 \le x \le 2, y \ge 0 \\ 0 & otherwise \end{cases}$$

What are the marginal probability density functions for X and Y?

You may either type LaTeX into the provided box or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-10 at 5.10.46 PM.png Download

(b) Suppose that X and Y have joint pdf given by:

$$f_{X,Y}(x,y) = \begin{cases} \frac{x}{2}e^{-y} & 0 \le x \le 2, y \ge 0 \\ 0 & otherwise \end{cases}$$

What are the marginal probability density functions for X and Y? [5 pts]

$$f_x(x) = \int_0^\infty \frac{x}{2}e^{-y}dy = \frac{x}{2}$$

$$f_y(y) = \int_0^2 \frac{x}{2}e^{-1}dx = e^{-y}$$

$$\int_{x=0}^2 \int_{y:0}^\infty \frac{x}{2}e^{-y}dydx = 1$$

Q14.3 Problem 7c BONUS FOR ALL
5 Points

A person decides to toss a biased coin with $P(heads) = 0.3$ repeatedly until he gets a head. He will make at most 6 tosses. Let the random variable Y denote the number of heads. Find the probability distribution of Y. Then, find the variance of Y.

Provide your answer as a number up to 3 decimal places in the provided boxes. Gradescope accepts any of the following operators: `+`, `-`, `*`, `/`, `%`, `^`, `( )`

0.882

Provide your equation using LaTeX in the box below or upload a screenshot of your answer. Blurry screenshots or screenshots of handwritten work will result in 0 credit.

Screenshot 2023-02-10 at 3.33.22 PM.png Download

(c) A person decides to toss a biased coin with $P(heads) = 0.3$ repeatedly until he gets a head. He will make at most 6 tosses. Let the random variable Y denote the number of heads. Find the probability distribution of Y. Then, find the variance of Y. [5 pts]

$$
\begin{aligned}
&P(\text{ heads }) = 0.3 \\
&P(H) = 0.3 \\
&P(TH) = 0.7 \times 0.3 = 0.21 \\
&P(TTH) = 0.7^2 \times 0.3 = 0.147 \\
&P(TTTH) = 0.73 \times 0.3 = 0.1029 \\
&P(TTTTH) = 0.7^2 \times 0.3 = 0.07203 \\
&P(TTTTTH) = 0.7^5 \times 0.3 = 0.050421 \\
&P(TTTTTT) = 0.7^6 = 0.117649
\end{aligned}
$$

$$
\begin{aligned}
E(Y) &= 0.3 + 0.21 + 0.147 + 0.1019 + 0.07203 + 0.050421 \\
&= 0.882
\end{aligned}
$$