# HOMEWORK 0

>>NAME HERE<<
>>ID HERE<<

**Instructions:** This is a background self-test on the type of mathematics and computer science we will encounter in class. If you find many questions intimidating, we suggest you drop 760 and take it again in the future when you are more prepared. This template can be generated by compiling the `hw0.tex` file, available under the "Files" tab on Canvas; please use it as a template to develop your homework. Submit your homework on time as a single pdf file to Gradescope. There is no need to submit the LaTeX source or any code. Please check Piazza for updates about the homework.

## 1 Vectors and Matrices [6 pts]

Consider the matrix $X$ and the vectors $\mathbf{y}$ and $\mathbf{z}$ below:

$$\mathbf{X} = \begin{pmatrix} 6 & 7 \\ 8 & 9 \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \qquad \mathbf{z} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$$

1. Compute $\mathbf{y}^\top \mathbf{X} \mathbf{z}$
   Solution goes here.

2. Is $\mathbf{X}$ invertible? If so, give the inverse, and if no, explain why not.
   Solution goes here.

## 2 Calculus [3 pts]

1. If $y = e^x + \tan(z)x^{6z} - \ln(\frac{7x+z}{x^4})$, what is the partial derivative of $y$ with respect to $x$?
   Solution goes here.

## 3 Probability and Statistics [10 pts]

Consider a sequence of data $S = (0, 1, 1, 0, 1, 1, 1)$ created by flipping a coin $x$ seven times, where 0 denotes that the coin turned up heads and 1 denotes that it turned up tails.

1. (2.5 pts) What is the probability of observing this data, assuming it was generated by flipping a biased coin with $p(x = 1) = 0.7$? Solution goes here.

2. (2.5 pts) Note that the probability of this data sample could be greater if the value of $p(x = 1)$ was not 0.7, but instead some other value. What is the value that maximizes the probability of $S$? Please justify your answer.
   Solution goes here.

3. (5 pts) Consider the following joint probability table where both $A$ and $B$ are binary random variables:

   | A | B | $P(A, B)$ |
   |---|---|-----------|
   | 0 | 0 | 0.4 |
   | 0 | 1 | 0.3 |
   | 1 | 0 | 0.2 |
   | 1 | 1 | 0.1 |

   (a) What is $P(A = 0|B = 1)$?
       Solution goes here.

  (b) What is $P(A = 0 \vee B = 0)$?
     Solution goes here.

# 4   Big-O Notation [6 pts]

For each pair $(f, g)$ of functions below, list which of the following are true: $f(n) = O(g(n))$, $g(n) = O(f(n))$, both, or neither. Briefly justify your answers.

1. $f(n) = \frac{n}{2}$, $g(n) = \log_2(n)$.
   Solution goes here.

2. $f(n) = \ln(n)$, $g(n) = \log_2(n)$.
   Solution goes here.

3. $f(n) = n^{100}$, $g(n) = 100^n$.
   Solution goes here.

# 5   Probability and Random Variables

## 5.1   Probability [12.5 pts]

State true or false. Here $\Omega$ denotes the sample space and $A^c$ denotes the complement of the event $A$.

1. For any $A, B \subseteq \Omega$, $P(A|B)P(B) = P(B|A)P(A)$.
   Solution goes here.

2. For any $A, B \subseteq \Omega$, $P(A \cup B) = P(A) + P(B) - P(A|B)$.
   Solution goes here.

3. For any $A, B, C \subseteq \Omega$ such that $P(B \cup C) > 0$, $\frac{P(A \cup B \cup C)}{P(B \cup C)} \geq P(A|B \cup C)P(B \cup C)$.
   Solution goes here.

4. For any $A, B \subseteq \Omega$ such that $P(B) > 0, P(A^c) > 0$, $P(B|A^C) + P(B|A) = 1$.
   Solution goes here.

5. For any $n$ events $\{A_i\}_{i=1}^n$, if $P(\bigcap_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$, then $\{A_i\}_{i=1}^n$ are mutually independent.
   Solution goes here.

## 5.2   Discrete and Continuous Distributions [12.5 pts]

Match the distribution name to its probability density / mass function. Below, $|\mathbf{x}| = k$.

(f) $f(\mathbf{x}; \boldsymbol{\Sigma}, \boldsymbol{\mu}) = \frac{1}{\sqrt{(2\pi)^k \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)$

(g) $f(x; n, \alpha) = \binom{n}{x}\alpha^x(1 - \alpha)^{n-x}$ for $x \in \{0, \ldots, n\}$; 0 otherwise

(a) Laplace Solution goes here.

(b) Multinomial Solution goes here.

(h) $f(x; b, \mu) = \frac{1}{2b} \exp\left(-\frac{|x - \mu|}{b}\right)$

(i) $f(\mathbf{x}; n, \boldsymbol{\alpha}) = \frac{n!}{\prod_{i=1}^k x_i!}\prod_{i=1}^k \alpha_i^{x_i}$ for $x_i \in \{0, \ldots, n\}$ and $\sum_{i=1}^k x_i = n$; 0 otherwise

(c) Poisson Solution goes here.

(d) Dirichlet Solution goes here.

(j) $f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$ for $x \in (0, +\infty)$; 0 otherwise

(e) Gamma Solution goes here.

(k) $f(\mathbf{x}; \boldsymbol{\alpha}) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \prod_{i=1}^k x_i^{\alpha_i-1}$ for $x_i \in (0, 1)$ and $\sum_{i=1}^k x_i = 1$; 0 otherwise

(l) $f(x; \lambda) = \lambda^x \frac{e^{-\lambda}}{x!}$ for all $x \in Z^+$; 0 otherwise

### 5.3 Mean and Variance [10 pts]

1. Consider a random variable which follows a Binomial distribution: $X \sim \text{Binomial}(n, p)$.

   (a) What is the mean of the random variable?
   Solution goes here.

   (b) What is the variance of the random variable?
   Solution goes here.

2. Let $X$ be a random variable and $\mathbb{E}[X] = 1, \text{Var}(X) = 1$. Compute the following values:

   (a) $\mathbb{E}[3X]$
   Solution goes here.

   (b) $\text{Var}(3X)$
   Solution goes here.

   (c) $\text{Var}(X + 3)$
   Solution goes here.

### 5.4 Mutual and Conditional Independence [12 pts]

1. (3 pts) If $X$ and $Y$ are independent random variables, show that $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

   Solution goes here.

2. (3 pts) If $X$ and $Y$ are independent random variables, show that $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.
   Hint: $\text{Var}(X + Y) = \text{Var}(X) + 2\text{Cov}(X, Y) + \text{Var}(Y)$

   Solution goes here.

3. (6 pts) If we roll two dice that behave independently of each other, will the result of the first die tell us something about the result of the second die?

   Solution goes here.

   If, however, the first die's result is a 1, and someone tells you about a third event — that the sum of the two results is even — then given this information is the result of the second die independent of the first die?

   Solution goes here.

### 5.5 Central Limit Theorem [3 pts]

Provide one line explanation. No calculation needed.

1. Let $X_i \sim \mathcal{N}(0, 1)$ and $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, then the distribution of $\bar{X}$ satisfies

$$\sqrt{n}\bar{X} \stackrel{n \to \infty}{\Longrightarrow} \mathcal{N}(0, 1)$$

   Solution goes here.

## 6 Linear algebra

### 6.1 Norms [3 pts]

Draw the regions corresponding to vectors $\mathbf{x} \in \mathbb{R}^2$ with the following norms:

1. $||\mathbf{x}||_1 \leq 1$ (Recall that $||\mathbf{x}||_1 = \sum_i |x_i|$)

2. $||\mathbf{x}||_2 \leq 1$ (Recall that $||\mathbf{x}||_2 = \sqrt{\sum_i x_i^2}$)

3. $||\mathbf{x}||_\infty \leq 1$ (Recall that $||\mathbf{x}||_\infty = \max_i |x_i|$)

   Solution figure goes here.

## 6.2 Geometry [10 pts]

Prove the following. Provide all steps.

1. The smallest Euclidean distance from the origin to some point $\mathbf{x}$ in the hyperplane $\mathbf{w}^\top \mathbf{x} + b = 0$ is $\frac{|b|}{||\mathbf{w}||_2}$. You may assume $\mathbf{w} \neq 0$.
   Solution goes here.

2. The Euclidean distance between two parallel hyperplane $\mathbf{w}^\top \mathbf{x} + b_1 = 0$ and $\mathbf{w}^\top \mathbf{x} + b_2 = 0$ is $\frac{|b_1 - b_2|}{||\mathbf{w}||_2}$ (Hint: you can use the result from the last question to help you prove this one).
   Solution goes here.

# 7 Programming Skills [12 pts]

Sampling from a distribution. For each question, submit a scatter plot (you will have 3 plots in total). Make sure the axes for all plots have the same ranges.

1. Draw 100 items $\mathbf{x} = [x_1, x_2]^\top$ from a 2-dimensional Gaussian distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with mean $\boldsymbol{\mu} = (0, 0)^T$ and identity covariance matrix $\boldsymbol{\Sigma} = \mathbf{I}$, i.e., $p(\mathbf{x}) = \frac{1}{2\pi} \exp\left(-\frac{||\mathbf{x}||^2}{2}\right)$, and make a scatter plot ($x_1$ vs. $x_2$).

   Solution figure goes here.

2. Make a scatter plot by drawing 100 items from $N(\boldsymbol{\mu} + (1, -1)^\top, 2\boldsymbol{\Sigma})$.

   Solution figure goes here.

3. Make a scatter plot by drawing 100 items from a mixture distribution $0.3N\left((1, 0)^\top, \begin{pmatrix} 1 & 0.2 \\ 0.2 & 1 \end{pmatrix}\right) + 0.7N\left((-1, 0)^\top, \begin{pmatrix} 1 & -0.2 \\ -0.2 & 1 \end{pmatrix}\right)$.

   Solution figure goes here.