DS Bootcamp 2023.02 (Math)

1. **[5pt] Matrix Properties (True/False).**
    a. If an eigenvalue of an $n \times n$ matrix A is 0, then $\operatorname{rank}(A) < n$.
    b. $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$
    c. $(AB)CD = A(BC)D = AB(CD)$
    d. $(AB)^{-1} = B^{-1}A^{-1}$
    e. $(A^T)^{-1} = (A^{-1})^T$

2. **[4pt] Basis (True/False).**
    a. Are the following vectors linearly independent?

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},\ x_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},\ x_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

    b. If A is an $n \times n$ matrix and the columns of A form a basis for $\mathbb{R}^n$ then $\det(A) = 0$.

3. **[4pt] Orthogonal / Determinant / Norm (True/False).**
    a. If Q is an orthogonal matrix, so that $Q^TQ = I$, then $\det(Q)$ equals 1
    b. If $\det(A) = 3$ then $\det(2A) = 6$.
    c. Every norm is induced by an inner product such that $\|x\| \coloneqq \sqrt{\langle x, x \rangle}$.
    d. If a norm is on a vector space V and for all $x, y \in V$, then $\|x + y\| \leq \|x\| + \|y\|$.

4. **[4pt]** Find A^{-1}.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & -3 & 0 \end{bmatrix}$$

5. **[5pt]** Solve the least squares problem $Ax = b$. What is the $\hat{x}$?

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 12 \\ 6 \\ 18 \end{bmatrix}, \quad \hat{x} = ?$$

6. **[5pt]** Solve the following questions for the matrix.

$$A = \begin{bmatrix} 1 & a \\ a & 4 \end{bmatrix}$$

   a. If $a = 3$, A is positive definite?
   b. Find the range of values of a such that A is positive definite.

7. **[6pt]** For a subspace $U = \mathrm{span}[\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}] \subseteq \mathbb{R}^3$, and $x = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \in \mathbb{R}^3$

   a. Find the projection of x onto U, i.e. $\pi_U(x)$.
   b. Find the distance $d(x, U)$.

8. **[5pt]** If $f(x_1, x_2) = x_1^2 + x_1 x_2$, where $x_1 = t^2$ and $x_2 = \ln t$, what is $\frac{\partial f}{\partial t}$?

9. **[6pt]** Consider the function $f : \mathbb{R}^2 \to \mathbb{R}^2$,
given by $f(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} x_1 x_2 \\ e^{x_1} + x_2^2 \end{bmatrix}$, what is the Jacobian matrix of f?

10. **[6pt]** Find the singular value decomposition (SVD) of

$$A = \begin{bmatrix} 4 & 0 \\ 3 & -5 \end{bmatrix}.$$

DS Bootcamp 2023.02 (Statistics)

Please answer the questions below. There are 10 questions over 4 pages totaling 50 pts. You can write down **answers** only (no need for a step-by-step derivation).

1. (5pts) True or False. ($P(A), P(B) > 0$ for any events A and B).

   (a) (1pt) $P((A \cup B)^C) = 1 - P(A - B) - P(B - A) - P(A)P(B)$.

   (b) (1pt) $P((A \cap B)^C) = P(A^C) + P(B^C)$.

   (c) (1pt) If event A and B are disjoint, then A and B are dependent.

   (d) (1pt) If event A and B are independent, then $P(A \mid B) = P(B \mid A)$.

   (e) (1pt) If $P(A \mid B) < P(B \mid A)$, then $P(A) > P(B)$.

2. (5pts) True or False. PDF is short for Probability Density Function and $\mathbb{E}$ is expectation.

   (a) (1pt) PDF is a probability that a continuous random variable X takes some value.

   (b) (1pt) If a joint PDF $f_{X,Y}(x, y) = g_X(x)h_Y(y)$, then random variables X and Y are independent.

   (c) (1pt) If a binomial random variable X has a large sample size n, then sample distribution of X is close to a Poisson distribution in general.

   (d) (1pt) Sum of two Poisson random variables always follows Poisson distribution, even if they don't have the same mean (λ).

   (e) (1pt) If random variables $X_1, \cdots, X_n$ are independent, then $\mathbb{E}(\prod_{i=1}^n X_i) = \prod_{i=1}^n \mathbb{E}(X_i)$.

3. (5pts) In year of 2123, the human race finally found a planet where life existed. A race called the Na'vi has inhabited the planet, which consisted of two tribes, Omatikaya and Matkaina. Human researchers found that the height of Omatikaya tribe follows normal distribution of $\mathcal{N}(215, 3^2)$, and the height of Matkaina follows $\mathcal{N}(195, 5^2)$. Researchers assumed that the numbers of Omatikaya and Matkaina are the same. Solve the following questions.

   (You do not have to evaluate expressions made with numbers. For example, if the question is about the mean of the Na'vi's height, you can just write down $= \frac{215}{2} + \frac{195}{2} = 205$)

   (a) (2pts) Compute the variance of the Na'vi's height.

   Through further explorations, the human researchers found that the numbers of two tribes are different: Omatikaya tribe comprises 60% of Na'vi. Solve the following questions with this new information.

   (b) (1pt) Compute the mean of Na'vi's height.

   (c) (2pts) Compute the variance of Na'vi's height.

4. (5pt) True or False. (V stands for variance.)

   (a) (1pt) When a and b are constants, $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$ can be satisfied only if two random variable X and Y are **independent**.

   (b) (1pt) If X and Y are **independent** random variables, then $V(aX + bY) = a^2V(X) + b^2V(Y)$ for constant a and b.

   (c) (1pt) If a and b are non-zero constants and X follows a normal distribution, $aX + b$ follows a normal distribution.

   (d) (1pt) Suppose X_n converges to X in distribution and Y_n converges to a constant c. Then X_nY_n converges to cX in distribution.

   (e) (1pt) If X_n converges to X in distribution, then X_n converges X in probability.

5. (2pt) Suppose X is a random variable with $\mathbb{E}(X) = 5$ and $V(X) = 2$. What is $\mathbb{E}(X^2)$?

6. (3pt) Consider a large freight elevator that can transport 1020 kg at maximum. Suppose a load of cargo containing 400 boxes must be transported via the elevator. The weight of each box follows a distribution with mean $\mu = 2.5$ kg and standard deviation $\sigma = 0.5$ kg.

   What is the probability that all 400 boxes can be safely loaded onto the freight elevator and transported? Please utilize the central limit theorem to approximate the probability, and express the answer in the form of $P(Z \leq x)$ using a standard normal distribution.

7. (6pts) True or False

   (a) (1pt) Bootstrap is a parametric method to approximately estimate the standard error and confidence interval.

   (b) (1pt) Under appropriate regularity conditions, MLE of θ, i.e., $\hat{\theta}_n$ converges to θ in quadratic mean, that is, $\mathbb{E}(\hat{\theta}_n - \theta)^2 \to 0$.

   (c) (1pt) Under appropriate regularity conditions, for MLE of θ, i.e., $\hat{\theta}_n$, asymptotic variance of MLE is variance of score function.

   (d) (1pt) When we construct a 95% confidence interval for a parameter θ, we can interpret it as with 95% frequency, the confidence intervals trap the true parameter θ.

   (e) (2pts) Under the null hypothesis, p-values follow a uniform(0,1) distribution.

8. (8pt) Consider a random sample $(X_1, X_2, X_3, X_4, X_5)$ from an unknown distribution. Let's denote observed values in the sample $\{X_i\}$ as x_i for $i = 1, 2, 3, 4, 5$, and assume that realized values are all different.

   Researchers want to estimate the variance of sample maximum and decide to use bootstrap. So they generated 100 bootstrap samples, where each bootstrap sample's size is 5. Then, they obtained sample maximum $\hat{\theta}_j^*$ from each bootstrap sample ($j = 1, 2, 3, ..., 100$).
   Please answer the following questions.

   (a) (1pt) Let j^{th} bootstrap sample be denoted as $(X_{1,j}^*, X_{2,j}^*, ..., X_{5,j}^*)$. Please calculate $P(X_{i,j}^* = x_3)$.

   (b) (2pt) Let's suppose we had x_5 as a sample maximum for the original sample. Please write an expression for $P(\hat{\theta}_j^* = x_5)$.

   (c) (2pt) One of researchers thinks that original sample size is too small, so he or she decides to use bootstrap but generates each bootstrap sample with sample size 10 (i.e. $(X_{1,j}^*, X_{2,j}^*, ..., X_{10,j}^*)$ where $j = 1, 2, ..., 100$). In this setting, please write an expression for $P(\hat{\theta}_j^* = x_5)$.

   (d) (3pt) Based on the above answers, please mark the right ones for three selections in the following sentence.

   In order to obtain a appropriate estimator of the variance of sample maximum, setting **((b), (c))** is more suitable because estimated variance in setting **((b), (c))** would **(underestimate, overestimate)** the true variance.

9. (6pts) We want to know whether we have a fair coin(fair coin should obtain exactly the same probability of having heads with that of having tails). Tested coin is tossed 100 times and landed heads 20 times. (Assume that each trial follows an i.i.d. Bernoulli distribution.)

   (a) (1pt) Compute the maximum likelihood estimate (MLE, $\hat{p}$) for the probability of heads (p).

   (b) (1pt) Compute the $\widehat{SE}(\hat{p})$. Note the variance of $X \sim \text{Bernoulli}(p)$ is $p(1-p)$.

   (c) (1pt) Please express H_0 (null hypothesis) that the coin is fair and H_1 (alternative hypothesis) that the coin isn't fair.

   (d) (2pt) Please compute the Wald test statistic for the above hypothesis test using $\hat{p}$.

   (e) (1pt) Would the null hypothesis be rejected at the significance level $\alpha = 0.05$?
   (for $Z \sim \mathcal{N}(0,1), P(Z \geq 1.96) \approx 0.025, P(Z \geq 1.65) \approx 0.05$)

10. (5pt) Consider waiting times for a bus. The times until the arrival of the first Gwanak-02 bus are $(X_1, X_2, X_3, X_4) = (2, 4, 1, 3)$ (in minutes). We assume that $X_i \sim \text{Exp}(\lambda)$ where the pdf is defined as $f(x;\lambda) = \lambda e^{-\lambda x}$ for $x \in [0, \infty)$. Researchers want to estimate population median. (Note that the mean and the variance of an exponential distribution are respectively $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$.)

   (a) (2pt) Compute (or express) $\hat{\lambda}$, the maximum likelihood estimate (MLE) of λ.

   (b) (3pt) Estimate (or express) population median using the answer in (a). (Hint: utilize the inverse of CDF for the exponential distribution.)