Probability theory and Linear Algebra

Question 1: Some probability basics

(a) How many different σ-algebras is it possible to define on the set $\Omega = \{A, B, C\}$? Specifically write down every possibility.

(b) Prove that the pdf of the Binomial distribution indeed qualifies as a probability function.
Hint: You may use the Binomial theorem, which states that

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i} \,. \tag{1}$$

(c) Consider the following function

$$F : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto \begin{cases} 1, & x \geq d \\ \frac{1}{3}x^3 + \frac{1}{6}x + \frac{1}{2}, & -d \leq x \leq d \\ 0, & x \leq -d \end{cases}$$

for some $d \in \mathbb{R}_{>0}$.

(i) Determine a value for d so that F is a CDF.

(ii) Write down the corresponding probability (density) function and calculate $P(X \in A)$ for a random variable X with CDF F and

$$A = \big\{\{x \in \mathbb{R} \mid x > 0\} \cup \{x \in \mathbb{R} \mid x \leq -d\}\big\}^C.$$

(d) Consider a test to detect a disease that 0.1% of the population have. The test is 99% effective in detecting an infected person. However, the test gives a false positive result in 0.5% of cases. If a person tests positive for the disease what is the probability that they actually have it?

Solution:

(a) There are 5 possibilities for a σ-algebra on Ω:

(i) $\{\emptyset, \Omega\}$

(ii) $\{\emptyset, \{A\}, \{B, C\}, \Omega\}$

(iii) $\{\emptyset, \{B\}, \{A, C\}, \Omega\}$

(iv) $\{\emptyset, \{C\}, \{A, B\}, \Omega\}$

(v) $\{\emptyset, \{A\}, \{B\}, \{C\}, \{B, C\}, \{A, C\}, \{A, B\}, \Omega\} = \mathcal{P}(\Omega)$

(b) Given that the probability function of the Binomial distribution is $\binom{n}{x} p^x (1-p)^{n-x}$ with support $\{0, 1, \ldots, n\} \, \forall n \in \mathbb{N}$, this immediately follows from:

$$1 = (p + 1 - p)^n = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} \quad \forall n \in \mathbb{N} \, .$$

(c) (i) We require that

$$\int_{-\infty}^{\infty} F'(x)\mathrm{d}x = \int_{-d}^{d} F'(x)\mathrm{d}x \stackrel{!}{=} 1 \, ,$$

and, therefore,

$$\begin{aligned} \int_{-\infty}^{\infty} F'(x)\mathrm{d}x &= F(d) - F(-d) \\ &= \frac{1}{3}d^3 + \frac{1}{6}d + \frac{1}{2} - [\frac{1}{3}(-d)^3 + \frac{1}{6}(-d) + \frac{1}{2}] \\ &= \frac{2}{3}d^3 + \frac{1}{3}d \stackrel{!}{=} 1 \, . \end{aligned}$$

It follows that $d = 1$.

(ii) The corresponding density is given by

$$f(x) = \begin{cases} 0, & x > 1 \\ F'(x) = x^2 + \frac{1}{6}, & -1 \leq x \leq 1 \\ 0, & x \leq -1. \end{cases}$$

Meanwhile, we have that

$$\begin{aligned} P(X \in \left\{\{x \in \mathbb{R} \mid x > 0\} \cup \{x \in \mathbb{R} \mid x \leq -d\}\right\}^C) &= P(X \in \{x \in \mathbb{R} \mid x > 0\}^C \cap \{x \in \mathbb{R} \mid x \leq -d\}^C) \\ &= P(X \in \{x \in \mathbb{R} \mid x \leq 0\} \cap \{x \in \mathbb{R} \mid x > -d\}) \\ &= P(X \in]-d, 0]) = P(X \leq 0) - P(X \leq -d) \\ &= F(0) - F(-d) = \frac{1}{2} \, . \end{aligned}$$

(d) The first step in solving this problem is to recognize what we are trying to calculate and what quantities we have been given in the question. We can introduce a Bernoulli random variable and say that $D = 1$ when the person has the disease. We can then introduce a second Bernoulli random variable T and say that $T = 1$ when a person gets a positive test result. We then note that these two random variables are **not** independent. With these symbols in place we can now state clearly what it we are trying to calculate. We are trying to calculate the conditional probability $P(D = 1|T = 1)$. In addition, the question tells us that:

$$P(D = 1) = 0.001 \qquad P(T = 1|D = 1) = 0.990 \qquad P(T = 1|D = 0) = 0.005$$

From these quantities we can calculate the probability of getting a positive test result, $P(T = 1)$ using

$$\begin{aligned} P(T = 1) &= P(T = 1|D = 1)P(D = 1) + P(T = 1|D = 0)P(D = 0) \\ &= P(T = 1|D = 1)P(D = 1) + P(T = 1|D = 0)[1 - P(D = 1)] \\ &= 0.99 \times 0.1 + 0.005 \times (1 - 0.001) = 0.005985 \end{aligned}$$

We can now insert this result into Bayes theorem to get the desired conditional probability.

$$P(D=1|T=1) = \frac{P(T=1|D=1)P(D=1)}{P(T=1)} = \frac{0.99 \times 0.001}{0.005985} \approx 0.165\,.$$

Question 2: Matrix rank and linear independence

(a) Calculate the rank of the following matrix:

$$\begin{bmatrix} 4 & -6 & 0 \\ -6 & 0 & 1 \\ 0 & 9 & -1 \\ 0 & 1 & 4 \end{bmatrix}$$

(b) Are the following vectors linearly independent?

$$v_1 = \begin{pmatrix} 3 \\ 0 \\ 1 \\ 2 \end{pmatrix} \; ; \; v_2 = \begin{pmatrix} 6 \\ 1 \\ 0 \\ 0 \end{pmatrix} \; ; \; v_3 = \begin{pmatrix} 12 \\ 1 \\ 2 \\ 4 \end{pmatrix} \; ; \; v_4 = \begin{pmatrix} 6 \\ 0 \\ 2 \\ 4 \end{pmatrix} \; ; \; v_5 = \begin{pmatrix} 9 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

If they are not, find the largest number of linearly independent vectors among them.

(c) Prove that if a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ is not square, then either the row vectors or the column vectors are linearly dependent.

Solution:

(a)

$$\begin{bmatrix} 4 & -6 & 0 \\ -6 & 0 & 1 \\ 0 & 9 & -1 \\ 0 & 1 & 4 \end{bmatrix} \overset{R1\cdot(-1)-R2-R3}{\Longrightarrow} \begin{bmatrix} 2 & -3 & 0 \\ -6 & 0 & 1 \\ 0 & 9 & -1 \\ 0 & 1 & 4 \end{bmatrix}$$

$$\overset{R2+3R1}{\Longrightarrow} \begin{bmatrix} 2 & -3 & 0 \\ 0 & -9 & 1 \\ 0 & 9 & -1 \\ 0 & 1 & 4 \end{bmatrix}$$

$$\overset{R4+R3\;;\;R5+\frac{1}{9}R2}{\Longrightarrow} \begin{bmatrix} 2 & -3 & 0 \\ 0 & -9 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 37/9 \end{bmatrix}$$

$$\overset{R4+R5\;;\;R5-R5}{\Longrightarrow} \begin{bmatrix} 2 & -3 & 0 \\ 0 & -9 & 1 \\ 0 & 0 & 37/9 \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore, the rank of the original matrix is 3 .

(b) This question is equivalent to asking for the rank of the matrix

$$\begin{bmatrix} 3 & 0 & 1 & 2 \\ 6 & 1 & 0 & 0 \\ 12 & 1 & 2 & 4 \\ 6 & 0 & 2 & 4 \\ 9 & 0 & 1 & 2 \end{bmatrix} \Longrightarrow \begin{bmatrix} 3 & 0 & 1 & 2 \\ 0 & 1 & -2 & -4 \\ 0 & 1 & -2 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & -4 \end{bmatrix} \Longrightarrow \begin{bmatrix} 3 & 0 & 1 & 2 \\ 0 & 1 & -2 & -4 \\ 0 & 0 & -2 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The rank of the matrix is 3. It follows that the maximum number of linearly independent vectors is also 3. They are the ones that correspond to the non-zero rows of the final matrix:

$$(3, 0, 1, 2)^\top;\ (6, 1, 0, 0)^\top;\ (9, 0, 1, 2)^\top .$$

(c) The maximum number of linearly independent row vectors is the rank of $\boldsymbol{A}$, while the maximum number of linearly independent column vectors is the rank of $\boldsymbol{A}^\top$. If $n < m$, then $\text{rank}(\boldsymbol{A})^\top = \text{rank}(\boldsymbol{A}) \leq n < m$. Therefore, the column vectors are linearly dependent. Similarly, if $m < n$, then the row vectors are linearly dependent.

Question 3: Matrixdecomposition

Consider the following matrices:

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad \text{und} \quad \boldsymbol{D} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 26 \end{pmatrix} .$$

You may assume it to be known that the matrix $\boldsymbol{D}$ is positive definite.

a) Determine the definitiveness of $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$.

b) Decompose the matrix $\boldsymbol{A}$ using Eigendecomposition.

c) Determine the entries of $\boldsymbol{L}$ in the Cholesky-decomposition $\boldsymbol{D} = \boldsymbol{L}\boldsymbol{L}^T$ of $\boldsymbol{D}$.

Solution:

a) **Definiteness:**

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

characteristic polynomial:

$$\begin{aligned} det(\boldsymbol{A} - \lambda \boldsymbol{I}) &= det \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 1^2 \stackrel{!}{=} 0 \\ (2-\lambda)^2 &= 1 \\ 2-\lambda &= \pm 1 \\ \Rightarrow \lambda_1 = 1 & \qquad \lambda_2 = 3 \end{aligned}$$

$\Rightarrow$ $\boldsymbol{A}$ is positive definite, since all eigenvalues are greater than zero!

$$\boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

characteristic polynomial:

$$\begin{aligned} det(\boldsymbol{B} - \lambda \boldsymbol{I}) &= det \begin{pmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 - 2^2 \stackrel{!}{=} 0 \\ (1-\lambda)^2 &= 4 \\ 1-\lambda &= \pm 2 \\ \Rightarrow \lambda_1 = 3 & \quad \lambda_2 = -1 \end{aligned}$$

$\Rightarrow$ $\boldsymbol{B}$ is indefinite, since one eigenvalue is smaller and one greater than zero!

$$\boldsymbol{C} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

$\boldsymbol{C}$ is not symmetric. Therefore, definiteness cannot be directly determined through eigenvalues. However, we can use the eigenvalues of the corresponding symmetric matrix! Simply put: we keep the diagonal, but take the pairwise average over the elements of the off-diagonal.

$$\boldsymbol{C}_S = \frac{1}{2}\left(\boldsymbol{C} + \boldsymbol{C}^\top\right).$$

$$\boldsymbol{C}_S = \frac{1}{2}\left[\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}\right] = \begin{pmatrix} 1 & 1.5 \\ 1.5 & 2 \end{pmatrix}$$

characteristic polynomial:

$$\begin{aligned} det(\boldsymbol{C}_S - \lambda \boldsymbol{I}) &= det \begin{pmatrix} 1-\lambda & 1.5 \\ 1.5 & 2-\lambda \end{pmatrix} \stackrel{!}{=} 0 \\ &\Rightarrow (1-\lambda)(2-\lambda) - 1.5^2 \\ &= \lambda^2 - 3\lambda - 0.25 \stackrel{!}{=} 0 \\ \Rightarrow \lambda_{1/2} &= \frac{3 \pm \sqrt{10}}{2} \\ \lambda_1 \approx -0.08 & \quad \lambda_2 \approx 3.08 \end{aligned}$$

$\Rightarrow$ $\boldsymbol{C}$ is indefinite, since one eigenvalue is smaller and one greater than zero!

b) **Eigendecomposition:**

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \qquad \text{with eigenvalue} \quad \lambda_1 = 1,\ \lambda_2 = 3$$

Calculating the eigenvectors:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \stackrel{!}{=} (\boldsymbol{A} - \lambda_1 \boldsymbol{I}) \cdot \boldsymbol{x} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 + x_2 \end{pmatrix}$$

$$\Rightarrow x_2 = -x_1 \Rightarrow \text{normed eigenvector: } \sqrt{1/2} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \stackrel{!}{=} (\boldsymbol{A} - \lambda_2 \boldsymbol{I}) \cdot \boldsymbol{x} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 + x_2 \\ x_1 + -x_2 \end{pmatrix}$$

$$\Rightarrow x_2 = x_1 \Rightarrow \text{normed eigenvector: } \sqrt{1/2} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\Rightarrow \boldsymbol{A} = \underbrace{\sqrt{1/2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}}_{\boldsymbol{P}} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}}_{\text{diag}(\lambda_1, \lambda_2)} \cdot \underbrace{\sqrt{1/2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}}_{\boldsymbol{P}^T}$$

c) **Cholesky-decomposition:**

$$\boldsymbol{D} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 26 \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{pmatrix} = \boldsymbol{L}\boldsymbol{L}^T$$

Calculating the Cholesky decomposition:

$$l_{ii} = (d_{ii} - \sum_{k=1}^{i-1} {l_{ik}}^2)^{\frac{1}{2}}$$

$$l_{ji} = \frac{1}{l_{ii}}(d_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik})$$

Specifically:

$$l_{11} = (d_{11} - \sum_{k=1}^{1-1} {l_{1k}}^2)^{\frac{1}{2}} = (1-0)^{\frac{1}{2}} = 1$$

$$l_{21} = \frac{1}{l_{11}}(d_{21} - \sum_{k=1}^{1-1} l_{2k} l_{1k}) = \frac{d_{21}}{l_{11}} = \frac{2}{1} = 2$$

$$l_{31} = \frac{1}{l_{11}}(d_{31} - \sum_{k=1}^{1-1} l_{3k} l_{1k}) = \frac{d_{31}}{l_{11}} = \frac{3}{1} = 3$$

$$l_{22} = (d_{22} - \sum_{k=1}^{2-1} {l_{2k}}^2)^{\frac{1}{2}} = (5 - 2^2)^{\frac{1}{2}} = 1$$

$$l_{32} = \frac{1}{l_{22}}(d_{32} - \sum_{k=1}^{2-1} l_{3k} l_{2k}) = \frac{1}{1}(7 - 3 \cdot 2) = 1$$

$$l_{33} = (d_{33} - \sum_{k=1}^{3-1} {l_{3k}}^2)^{\frac{1}{2}} = \left(d_{33} - ({l_{31}}^2 + {l_{32}}^2)^{\frac{1}{2}}\right) = (26 - 3^2 - 1^2)^{\frac{1}{2}} = 16^{\frac{1}{2}} = 4$$

$$\Longrightarrow \boldsymbol{L} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 4 \end{pmatrix} \qquad \text{und} \qquad \boldsymbol{D} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 26 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{pmatrix}.$$