

PPGEE2249 Aprendizado de Máquina  
Assignment 1  
Prof. Daniel Guerreiro e Silva

Thiago Tomás de Paula

November 8, 2025

► **Question 1**

$X$  is a continuous-valued random variable with uniform density in  $(-1, +1)$ .

▷ **Item a**

Draw its probability density function (PDF). What is the area under the curve? Justify your answer.

▷ **Item b**

Draw its cumulative distribution function (CDF).

▷ **Item c**

Calculate the probability of the event  $X \in (-0.2, 0.2)$ .

▷ **Item d**

Calculate the expected value  $E[X]$ , the second  $E[X^2]$  and the fourth moment  $E[X^4]$  of the random variable. Calculate its variance  $\sigma^2[X]$ , as well.

Let  $p : \mathbb{R} \rightarrow \mathbb{R}$  denote the PDF of  $X$  and  $c : \mathbb{R} \rightarrow \mathbb{R}$  its CDF. Since the random variable is of uniform density in  $(-1, +1)$ ,

$$p(x) = \begin{cases} k, & |x| < 1 \\ 0, & |x| \geq 1 \end{cases} \quad (1)$$

for some constant  $k > 0$ . By definition, the area under the PDF curve is equal to the probability  $P$  of  $X$  assuming a value in the associated range  $(a, b)$ ,  $P(a < X < b) = \int_a^b p(x) dx$ . For  $a = -\infty$  and  $b = \infty$ ,  $P(-\infty < X < \infty) = 1$ , i.e., the probability of  $X$  assuming *any* value is 100%. We can use this result (in general, an identity) to find the value of  $k$ :

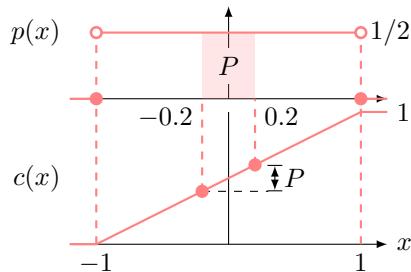
$$P(-\infty < X < \infty) = \int_{-\infty}^{\infty} p(x) dx = 1 \implies \int_{|x| \geq 1} p(x) dx + \int_{|x| < 1} p(x) dx = 1 \quad (2)$$

$$\int_{|x| \geq 1} 0 dx + \int_{|x| < 1} k dx = 1 \quad (3)$$

$$2k = 1 \quad \therefore \quad k = \frac{1}{2}. \quad (4)$$

With  $k$  evaluated, we can find any  $P$ , e.g.,  $P(-0.2 < X < 0.2) = [0.2 - (-0.2)] \cdot (1/2) = 0.2$ . A drawing of  $p(x)$  and  $c(x) = \int_{-\infty}^x p(z) dz$  is presented in fig. 1.

Figure 1: PDF and CDF of a uniform distribution in  $(-1, 1)$ .



Next, the variance and some of the moments of  $p$  is calculated. By definition,  $E[X^n] = \int_{\mathbb{R}} x^n p(x) dx$ . For this question's distribution,

$$E[X^n] = \int_{-1}^1 x^n \cdot \frac{1}{2} dx = \frac{1}{2} \cdot \frac{x^{n+1}}{n+1} \Big|_{x=-1}^{x=1} = \frac{1 + (-1)^n}{2(n+1)} = \begin{cases} 0, & n \text{ odd} \\ 1/(n+1), & n \text{ even} \end{cases} \quad (5)$$

The above expression gives, for example,  $E[X] = 0$ ,  $E[X^2] = 1/3$ , and  $E[X^4] = 1/5$ . As for the variance, the definition  $\sigma^2[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$  results in  $\sigma^2[X] = 1/3$ .

► **Question 2**

$X$  is a discrete-valued random variable with uniform distribution over the set  $\{-2, -1, 0, 1, 2\}$ . Draw its probability mass function (PMF) and calculate  $E[X]$  and  $\sigma^2[X]$ .

Let  $p : \mathbb{Z} \rightarrow \mathbb{R}$  denote the PMF of  $X$ . Since the random variable is uniform on  $\{x \in \mathbb{Z} : |x| < 3\}$ , we have

$$p(x) = \begin{cases} k, & |x| < 3 \\ 0, & |x| \geq 3 \end{cases} \quad (6)$$

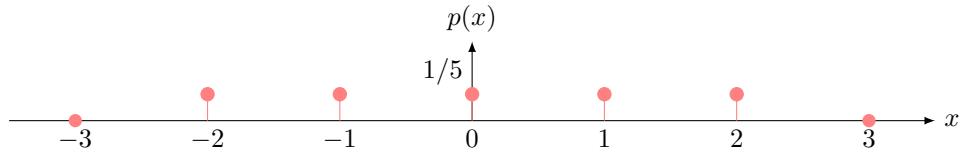
for some constant  $k > 0$ . Since the PMF must satisfy the identity  $\sum_{x \in \mathbb{Z}} p(x) = 1$ , we have

$$\sum_{x \in \mathbb{Z}} p(x) = \sum_{|x| \geq 3} p(x) + \sum_{|x| < 3} p(x) = 1 \quad (7)$$

$$0 + 5k = 1 \quad \therefore \quad k = \frac{1}{5} \quad (8)$$

The drawing in fig. 2 illustrates the PMF with  $k$  evaluated. As for  $E[X]$  and  $\sigma^2[X]$ , we can use the

Figure 2: PMF of an uniform distribution in  $\{-2, -1, 0, 1, 2\}$ .



definition for the  $n$ -th moment  $E[X^n] = \sum_{x \in \mathbb{Z}} x^n p(x)$  and the definition  $\sigma^2[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$ :

$$E[X] = \sum_{|x| < 3} \frac{x}{5} = \frac{1}{5}(-2 - 1 + 0 + 1 + 2) = 0 \quad (9)$$

$$E[X^2] = \sum_{|x| < 3} \frac{x^2}{5} = \frac{1}{5}[(-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2] = 2 \quad (10)$$

$$\sigma^2[X] = E[X^2] - E[X]^2 = 2 - 0 = 2 \quad (11)$$

► **Question 3**

Consider the normal random variables  $X_1 \sim N(-2, 2)$ ,  $X_2 \sim N(1, 4)$ , with  $\text{cov}(X_1, X_2) = -0.8$ . Calculate the joint pdf of  $X = (X_1, X_2)^T$ .

---

Let  $X_i \sim N(\mu_i, \sigma_i^2)$ ,  $1 \leq i \leq d$  be several univariate normal random variables. The joint PDF  $p : \mathbb{R}^{d \times 1} \rightarrow \mathbb{R}$  of  $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_d]^T$  is the multivariate gaussian

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \times \exp \left[ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right] \quad (12)$$

with  $\boldsymbol{\mu} = [\mu_1 \ \mu_2 \ \cdots \ \mu_d]^T$  and  $\Sigma_{ij} = \sigma_{ij} = \text{cov}(X_i, X_j)$  the covariance matrix of  $\mathbf{x}$ .

The question statement gives

$$\mu_1 = -2, \quad \sigma_1^2 = 2, \quad \mu_2 = 1, \quad \sigma_2^2 = 4, \quad \sigma_{12} = \sigma_{21} = -0.8 \quad (13)$$

which leads to

$$\boldsymbol{\mu} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad (14)$$

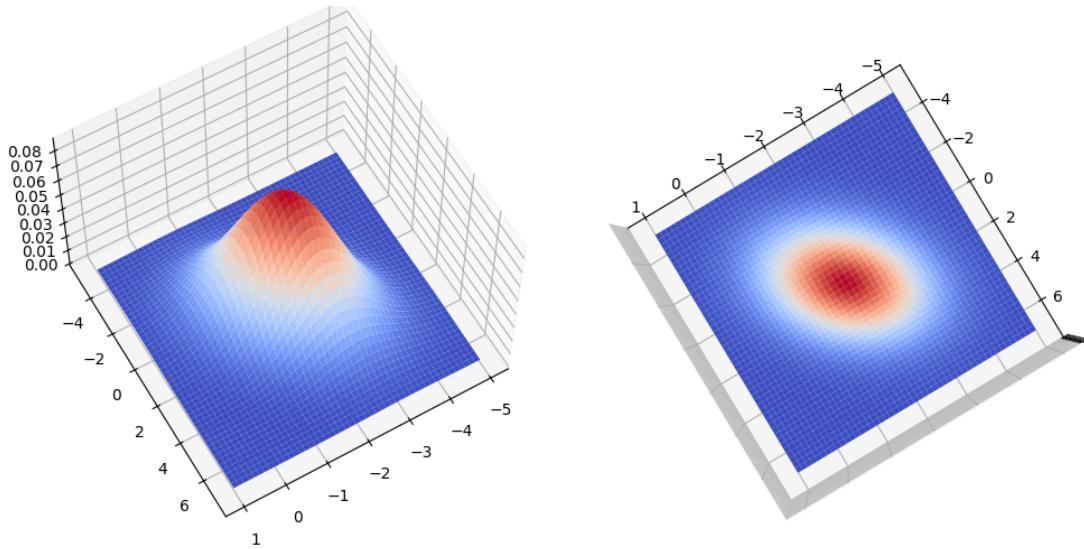
$$\Sigma = \begin{bmatrix} 2 & -0.8 \\ -0.8 & 4 \end{bmatrix}, \quad |\Sigma| = 8 - 0.36 = 7.64, \quad \Sigma^{-1} = \frac{1}{7.64} \begin{bmatrix} 4 & 0.8 \\ 0.8 & 1 \end{bmatrix} \quad (15)$$

and, making all substitutions in (12),

$$p(\mathbf{x}) = \frac{1}{2\pi\sqrt{7.36}} \times \exp \left[ -\frac{1}{2} \begin{bmatrix} x_1 + 2 \\ x_2 - 1 \end{bmatrix}^T \frac{1}{7.64} \begin{bmatrix} 4 & 0.8 \\ 0.8 & 1 \end{bmatrix} \begin{bmatrix} x_1 + 2 \\ x_2 - 1 \end{bmatrix} \right]. \quad (16)$$

fig. 3 shows a plot of the multivariate distribution given by the expression above.

Figure 3: Joint PDF of  $X_1 \sim N(-2, 2)$  and  $X_2 \sim N(1, 4)$ .



**► Question 4**

Assume you want to design a classifier which receives univariate data as input and must choose between class +1 or class -1. The data associated with class “+1” follow a gaussian density with mean 2 and unit variance. On the other hand, the data associated with class “-1” follow a uniform density between -2 and 2. The prior probabilities are  $P(C_{+1}) = 0.6$  and  $P(C_{-1}) = 0.4$ .

**▷ Item a**

Apply Bayes' methodology to obtain the optimal classification model and show, using a diagram, the model's architecture (with clear indication of the discriminant function(s)).

**▷ Item b**

Calculate the model's decision boundary. Analytically calculate the classifier error probability (the integrals can be numerically solved).

**▷ Item c**

Recalculate the model's decision boundary if the Maximum-Likelihood criterion is adopted, this time. Analytically calculate the classifier error probability (the integrals can be numerically solved).

**▷ Item d**

Compare the performances of both methodologies.

---

**► Question 5**

Consider the dataset of 3000 bivariate, labeled samples in `data.csv` file. Firstly, split the dataset into training (70%) and test set (30%).

**▷ Item a**

Assume the samples labeled with “+1” are drawn by a bivariate gaussian density with parameters  $\mu_{+1}$ ,  $\Sigma_{+1}$ , while the samples labeled with “−1” are drawn by another gaussian density with parameters  $\mu_{-1}$ ,  $\Sigma_{-1}$ . The prior probabilities  $P(C_{+1})$  and  $P(C_{-1})$  are unknown as well. Estimate the missing parameters (with the training set), present the discriminant functions of the Bayes classifier (MAP criterion) and evaluate its performance(accuracy, precision, recall) over the test set. Comment all your results.

**▷ Item b**

Calculate the decision boundary and plot it on a graph with the samples of the test set. Comment your results.

---

**► Question 6**

Do your own implementation and train a multivariate linear regression model for a given problem (suggestions: Kaggle, UCI Machine Learning Repository). You may use a validation set to train models with different subsets of features and select the best one. Then, use a test set to report and comment the final model results.

---