

Homework Assignment # 1
Due: Wednesday, September 16, 2015, 11:59 p.m.
Total marks: 90

Question 1. [5 MARKS]

Let (Ω, \mathcal{F}, P) be a probability space. Using only the set operations and axioms of probability, show that for any two sets, not necessarily disjoint, $A \subseteq \Omega$ and $B \subseteq \Omega$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Question 2. [5 MARKS]

Prove the following expression or provide a counterexample if it does not hold

$$P(A) = P(A|B) + P(A|B^c)$$

Question 3. [15 MARKS]

Two players perform a series of coin tosses. Player one wins a toss if the coin turns heads and player two wins a toss if it turns tails. The game is played until one player wins n times. However, the game is interrupted when player one had m wins and player two had l wins, where $0 \leq m < n$ and $0 \leq l < n$.

(a) [5 MARKS] Assuming $n = 8, m = 4$, and $l = 6$, what is the probability that player one would win the game if the game was to be continued later?

(b) [10 MARKS] Derive the general expression or write an algorithm that player one will win the game if the game is to be continued later. Your expression should be a function of m, l , and n . If you are providing an algorithm, implement it and submit your code along with your pseudo-code. You may not simulate the game as your solution.

Question 4. [10 MARKS]

Let X_1, \dots, X_m be independent multivariate Gaussian random variables, with $X_i \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, with $\boldsymbol{\mu}_i \in \mathbb{R}^d$ and $\boldsymbol{\Sigma}_i \in \mathbb{R}^{d \times d}$ for dimension $d \in \mathbb{N}$. Define $X = a_1 X_1 + a_2 X_2 + \dots + a_m X_m$ as a convex combination, $a_i \geq 0$ and $\sum_{i=1}^m a_i = 1$.

(a) [3 MARKS] Write the expected value $E[X]$ in terms of the givens $a_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i$. Show all you steps. What is the dimension of $E[X]$?

(b) [7 MARKS] Write the expected value $\text{Cov}[X]$ in terms of the givens $a_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i$. Show all you steps. What is the dimension of $\text{Cov}[X]$? Finally, briefly explain how the result for $\text{Cov}[X]$ would be different if the variables X_1 and X_2 are not independent and have covariance $\text{Cov}[X_1, X_2] = \mathbf{\Lambda}$.

Question 5. [15 MARKS]

Suppose you have three coins. Coin A has a probability of heads of 0.75, Coin B has a probability of heads of 0.5, and Coin C has a probability of heads of 0.25.

(a) [5 MARKS] Suppose you flip all three coins at once, and let X be the number of heads you see

(which will be between 0 and 3). What is the expected value of X , $E[X]$?

(b) [10 MARKS] Suppose instead you put all three coins in your pocket, select one at random, and then flip that coin 5 times. You notice that 3 of the 5 flips result in heads while the other 2 are tails. What is the probability that you chose Coin C?

Question 6. [10 MARKS]

Let X , Y , and Z be discrete random variables defined as functions on the same probability space (Ω, \mathcal{F}, P) . Prove or disprove the following expression

$$P_{X|Y}(X = x|Y = y) = \sum_{z \in \Omega} P_{X|YZ}(X = x|Y = y, Z = z)P_{Z|Y}(Z = z|Y = y)$$

Question 7. [5 MARKS]

Consider rolling two dice 24 times. A player wins if two sixes occur at least once. Otherwise, the house wins. Derive the probability that the player wins.

Question 8. [5 MARKS]

Consider a measurable space (Ω, \mathcal{F}) , where $\Omega = [0, 1]$ and $\mathcal{F} = B(\Omega)$. Let P be a set function on (Ω, \mathcal{F}) as follows

$$P(A) = \begin{cases} 1/2 & \text{if } 0 \in A \text{ or } 1 \in A, \text{ but not both} \\ 1 & \text{if } 0 \in A \text{ and } 1 \in A \\ 0 & \text{otherwise} \end{cases}$$

Is P a probability distribution? Show all your work.

Question 9. [10 MARKS]

This question uses Figure 1.5(B) in the notes, for the three random variables X , Y and Z .

(a) [4 MARKS] Derive $P(Z = 1|X = 1)$; show your steps.

(b) [6 MARKS] Derive $P(Z = 1)$; show your steps.

Question 10. [10 MARKS]

This question involves some simple simulations, to better visualize random variables and get some intuition for sampling, which is a central theme in machine learning. Use the attached code called `simulate.py`. This code is a simple script for sampling and plotting with python; play with some of the parameters to see what it is doing. Calling `simulate.py` runs with default parameters; `simulate.py 1 100` simulates 100 samples from a 1d Gaussian.

(a) [5 MARKS] Run the code for 10, 100 and 1000 samples with `dim=1` and `σ = 1.0`. Next run the code for 10, 100 and 1000 samples with `dim=1` and `σ = 10.0`. What do you notice about the sample mean?

(b) [5 MARKS] The current covariance for `dim=3` is

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

What does that mean about the multivariate Gaussian (i.e., about X , Y and Z)?

Change the covariance to

$$\Sigma = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

What happens?

Bonus. [20 MARKS]

High dimensional spaces.

(c) [10 MARKS] Show that in a high dimensional space, most of the volume of a cube is concentrated in corners, which themselves become very long “spikes”. Hints: compute the ratio of the volume of a hypersphere of radius a to the volume of a hypercube of side $2a$ and also the ratio of the distance from the center of the hypercube to one of the corners divided by the perpendicular distance to one of the edges.

(d) [10 MARKS] Show that for points which are uniformly distributed inside a sphere in d dimensions where d is large, almost all of the points are concentrated in a thin shell close to the surface. Hints: compute the fraction of the volume of the sphere which lies at values of the radius between $a - \epsilon$ and $0 < \epsilon < a$; evaluate this fraction for $\epsilon = 0.01a$ and also for $\epsilon = 0.5a$ for $d \in \{2, 3, 10, 100\}$.

Homework policies:

Your assignment must be typed; for example, in Latex, Microsoft Word, Lyx, etc. Images may be scanned and inserted into the document if it is too complicated to draw them properly. Submit a single pdf document or, if you are attaching your code, submit your code together with the typed (single) document as one .zip file.

All code (if applicable) should be turned in when you submit your assignment. Use Matlab, Python, R, Java or C.

Policy for late submission assignments: Unless there are legitimate circumstances, late assignments will be accepted up to 5 days after the due date and graded using the following rule:

on time: your score 1
 1 day late: your score 0.9
 2 days late: your score 0.7
 3 days late: your score 0.5
 4 days late: your score 0.3
 5 days late: your score 0.1

For example, this means that if you submit 3 days late and get 80 points for your answers, your total number of points will be $80 \times 0.5 = 40$ points.

All assignments are individual, except when collaboration is explicitly allowed. All the sources used for problem solution must be acknowledged, e.g. web sites, books, research papers, personal communication with people, etc. Academic honesty is taken seriously; for detailed information see Indiana University Code of Student Rights, Responsibilities, and Conduct.

Good luck!