List of assignments

| Till (midnight) | Assignments |
| --- | --- |
| 6.11 | 3 (Bayes), 9 (KNN)  10,11 (Expectation) |
| 13.11 | 4,5,6, 10,13,14,15,16,17(Statistics I) 18-20, 24,26,27 (Statistics II) |
| 20.11 | all others up to 27 |
| | |

---

1. Tossing a coin three times:  H – head, T – tail

- What is the sample space Ω ?
- What is the event space F for selection of at least two heads?

---

2.

Show that P(B) = Σi P(B|Ai) * P(Ai)
for Ai being disjoint sets, partitioning the whole sample space (Ω = U Ai ) and
P(Ai) >0 for all I

---

3. A pharmaceutical company developed a test for detecting the rare disease, which is carried by 0.5 % cases of the whole population . Let's assume that the test gives the positive results for 96% of the cases if the patient is ill, but it also gives the positive results in 5% of the healthy patient. What is the probability that a patient is ill if his test gave a positive result?

---

4. Tossing two coins: Ω = { HT , TH , HH , TT }

X is a random variable - number of heads

X(E) = 1 if E = {HT, TH}

0 if E= {TT}

2 if E = {HH}

- What are CDF and PMF functions?
- Draw it

---

5. What are the PDF and CDF functions for the uniform distribution defined for X =[0, a]?

---

6.Write a program that simulates the tossing of two coins, and estimate the CDF and PMF functions for the problem 4.

---

7. Derive the optimal algorithm f*(X) assuming L(f(X), Y) = |(f(X) -Y)|

---

---

8. Derive the optimal algorithm f*(X) assuming L(f(X), Y) = $(f(X) - Y)^2$

---

9. Let's assume that we use k-NN model to perform the classification of two type of objects (think of spam/non-spam) with k=2, k =3 and k=10 and N=100 training samples with 2 features each e.g. { X(1)=[0,1]T, X(2)=[-1,2]T, X(3)=[2,2]T, X(4)=[0,0.3]T, …, X(100)=[-100,12,5]T } and {Y(1)=[0], Y(2)=[1], Y(3)=[1], Y(4)=[0]]

- For which k value do we expect the smallest training error?
- For which k value do we expected the highest/smallest stability?
- How would we classify X=[0,0] if we use first 4 training samples for k=1, k=2, k=3? Calculate the majority vote result for each case.
- We increase the training sample to N=101. How we expect it affects the stability for different k values?
- Now instead of kNN model we use the linear regression. Discuss the stability issues.

---

10. Let g(X) =1 for some set A being a subset of sample space Ω:

What is E[g(X)] if X is discrete with a given PMF or continuous with a given PDF

---

11. What is the interpretation of E[g(X)] for g(X) =x

---

12. Show that Var[X] = $E[X^2] - E^2[X]$

---

13. Calculate the mean and the variance of the uniform distribution

---

14.Implement a function that returns a mean of a vector represented as a list of numbers.

---

15. Implement a function that returns Var[X] for vector X represented as a list of numbers

---

16.Implement a function that returns Euclidean distance between two vectors represented as a list of numbers

---

17. Implement a function that returns Manhattan distance between two vectors represented as a list of numbers

---

18. Tossing a coin: Ω = { H, T } and rolling strange die Ω = { 1 , 2 , 3}

If H we roll the die twice, if T we roll the die once

X number of heads   X = {0,1}

Y sum from die Y = {1,2,3,4,5,6}

- Calculate joint PMF
- Calculate marginal PMF based on joint ones.

---

19* Calculate joint and marginal CDF

---

20. Write a program that estimates the PMFs distribution from 18

---

21. Rolling a die Ω = { 1 , 2 , 3 , 4 , 5, 6 }

X is 1 if even number 0 otherwise.

Y is 1 if prime number 0 otherwise.

- Calculate joint PMF
- Calculate marginal PMF of X and of Y
- Calculate conditional PMF pY|X (r|X=1)
- Check if h(k) = pY|X (r|k) is a proper probability function with respect to k

---

22. Let X and Y have a joint PDF $f_{XY}$ (x,y) = x+y for $0 <x<1$, $0<y<1$

- Find conditional PDF $f_{Y|X}$ (y|x)
- Show that the integral of $f_{Y|X}$ (y|x) over all y values is equal to 1

---

23. Show that:

- Cov[X,Y] = E[XY] – E[X]E[Y]
- Var[X+Y] = Var[X] + Var[Y] + 2 Cov[X,Y]

---

24. Let X be uniform in (-1,1) and Y = $X^2$

- Check if X and Y are correlated
- Check if X and Y are independent

---

25. Rolling a die Ω = { 1 , 2 , 3 , 4 , 5, 6 }

X is 1 if even number 0 otherwise.

Y is 1 if prime number 0 otherwise.

- Calculate E[Y|1], E[Y|0]
- Var[Y|1], Var[Y|0]

---

26

- Implement a function that returns Cov[X,Y] for two vectors X,Y represented as lists of numbers

---

27.

- Implement a function that returns Cosine similarity for two vectors represented as lists of numbers
- Implement a function that returns Pearson correlation coefficient for two vectors represented as lists of numbers

---