CSE 575: Statistical Machine Learning
Self-Evaluation Test

Instructor: Prof. Jingrui He
January 11, 2019

First Name:			
Last Name:			
Email:			
ASU ID:			
Q	Topic	Max Score	Score
1	Probability	24	
2	Iterative Algorithms	21	
3	MLE	20	
4	Quadratic Optimization	17	
5	Distance Metric	18	
Total:		100	

- This exam book has **6** pages, including this cover page and a blank page at the end.
- Good luck!

1 Probability (24 points)

(7 points.) If A and B are **DISJOINT** events, and $P(B) > 0$, what is the value of $P(A|B)$?

Solution. 0

(7 points.) Suppose that the PDF of a random variable X is as follows:

$$f(x) = \begin{cases} \frac{4}{3}(1 - x^3), & \text{for } 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Then what is the value of $P(X < 0)$?

Solution. 0

(10 points.) Suppose that X is a random variable for which $E(X) = \mu$ and $Var(X) = \sigma^2$, and let c be an arbitrary constant. What is the value of $E[(X-c)^2]$? **Hint: What is the definition of variance?**

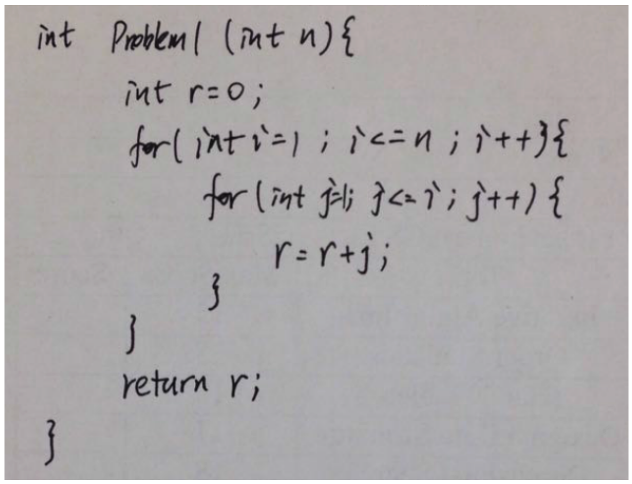
Solution.

$$E[(X-c)^2] = E[X^2] - 2cE[X] + c^2 = (\mu - c)^2 + \sigma^2$$

2 Iterative Algorithms and Big-O (21 points)

[7 points.] Write a program/function (**Not Pseudocode**) to calculate the following sum (c++ or any of your favorite programming language). Your program takes an integer n ($n \geq 1$) as the input; and returns r defined as $r = \sum_{i=1}^{n} \sum_{j=1}^{i} j$.

Solution:



```
int Problem1 (int n){
    int r=0;
    for( int i=1 ; i<=n ; i++){
        for (int j=1; j<=i; j++) {
            r=r+j;
        }
    }
    return r;
}
```

Figure 1: Code

(7 points.) What is the value r returned by your program? Express your answer as a function of n and use *closed-form* solution.

Solution: $r = \sum_{i=1}^{n} \sum_{j=1}^{i} j = \sum_{i=1}^{n} \frac{1}{2} i(i+1) = \frac{1}{6} n(n+1)(n+2)$.

(7 points.) Using $O()$ notation, give the worst-case running time of your program.

Solution: $O(n^3)$.

3 Maximum Likelihood Estimation (20 points)

[10 points.] Suppose we flip a coin, and observe either a head or a tail. The probability of observing a head in each trial is p $(0 \leq p \leq 1)$. If we flip the coin five times, and observe *(head, head, tail, tail, head)*, what is the maximum likelihood estimate of p? Justify your answer.

Solution: likelihood is $p^3(1-p)^2$. take the logorithm, and calculate its derivative, and set it as zero, we have

$$3/p - 2/(1-p) = 0$$

which gives $p = 0.6$.

[10 points.] Suppose we flip a coin, and observe either a head or a tail. The probability of observing a head in the first trial is p $(0 \leq p \leq 1)$. The probability of observing a head in the second trial is $2p$. The probability of observing a head in the third and fourth trials is $3p$, respectively. The probability of observing a head in the fifth trial is $5p$. If we flip the coin five times, and observe *(head, head, tail, tail, head)*, what is the maximum likelihood estimate of p? Justify your answer.

Solution: likelihood is $p * 2p(1-3p)^2 * 4p$. take the logorithm, and calculate its derivative, and set it as zero, we have

$$3/p - 6/(1-3p) = 0$$

which gives $p = 0.2$.

4 Quadratic Optimization (17 points)

[7 points.] Solve the following optimization problem

$$\text{argmin}_x \; 3x^2 - 12x + 8$$

Solution: $f(x) = 3(x-2)^2 - 4$. therefore $x_* = 2$, and $f_* = -4$

[10 points.] Solve the following optimization problem

$$\begin{gathered}\text{argmin}_x \; 3x^2 - 12x + 8 \\ \text{subject to } 6 < x < 10\end{gathered}$$

Solution: $f(x) = 3(x-2)^2 - 4$. therefore $x_* = 6$, and $f_* = 44$

5 Distance Metric (18 points)

[18 points]. Given two data points $x_1 = (0, 0)'$ and $x_2 = (1, 2)'$ in two-dimensional space, what is the L_2 distance between them? What is the L_1 distance between them? What is the L_∞ distance between them? Justify your answer.

Solution: $L_2 = \sqrt{5}$

$L_1 = 3$

$L_1 = 2$