Name:

|  |  |  |
| --- | --- | --- |
| Problem | Points | Score |
| 1(a) | 10 |  |
| 1(b) | 10 |  |
| 1(c) | 10 |  |
| 2(a) | 10 |  |
| 2(b) | 10 |  |
| 3(a) | 10 |  |
| 3(b) | 10 |  |
| 3(c) | 10 |  |
| 3(d) | 10 |  |
| 4 | 10 |  |
| Total | 100 |  |

Notes:

1. The exam is closed books and notes except for one double-sided sheet of notes.
2. Please indicate clearly your answer to the problem.
3. If I can’t read or follow your solution, it is wrong and no partial credit will be awarded.

**Problem No. 1**: A biased four-sided die has faces labeled 1, 2, 3, and 4. Let the random variable *X* denote the outcome of a roll of the die. Extensive testing of the die shows that the PMF of *X* is given by:



(10) (a) Find the CDF of *X*.

(10) (b) What is the probability that a number less than 3 appears on a roll of the die?

(10) (c) What is the probability of obtaining a number whose value is at least 3 on a roll of the die?

**Problem No. 2**: The life of a light bulb in months is denoted by a random variable *X* with the PDF *fX*(*x*) = 0.5*e*− 0.5*x*, *x* ≥ 0.

(10) (a) Find the expected value of X.

(10) (b) Find the conditional expected value of *X*, given that *X* ≤ 1.5.

**Problem No. 3**: A random variable *X* is uniformly distributed between 3 and 15. Find the following:

(10) (a) The expected value of *X.*

(10) (b) The variance of *X.*

(10) (c) The probability that *X* lies between 5 and 10.

(10) (d) The probability that *X* is less than 6.

**(10) Problem No. 4**: Recall that when you estimated the PMF of the speech signal, the smoothness of the plot varied with the bin size. If the bin size was set too small, some bins would be empty while the adjacent bins were populated. This did not seem to match your intuition.

Suppose I ask you to estimate the probability of a word occurring in English from a file containing the text from 1 billion web pages. You begin by generating a list of each word that occurs at least once in the file, and then you count how many times each of these words occurs. You plot a histogram, normalized by the total number of words in the file. This is a very reasonable estimate of the PMF.

However, you observe that the word “Bernoulli” never occurs in your sample. Yet, you know from this course and your other science courses that this word does occur frequently in science documents. Explain how you would modify the PMF you estimated to give this word a non-zero probability of occurring. (Hint: remember that probabilities in a PMF must sum to 1.)