

**CALIFORNIA STATE UNIVERSITY, HAYWARD
DEPARTMENT OF STATISTICS**

Statistics 3601 Introductory Statistics for Scientists and Engineers

Homework #2

Additional Problems:

1. In the roll of two fair dice, let Z be the product of the two up faces.
 - a. Find the probability mass function for Z .
 - b. Compute the $E[Z]$ and explain its meaning.
2. The table below lists the probability mass function for X .

| | | | | | | |
|--------|------|------|------|------|------|------|
| x | 1.20 | 2.60 | 3.50 | 4.10 | 4.50 | |
| $p(x)$ | 0.24 | 0.36 | 0.21 | 0.15 | 0.04 | 1.00 |

- a. Find $E[X]$.
- b. Find $V(X)$.
- c. Find $F(x)$, the distribution function for X .
- d. Using $F(x)$, find $P(X \leq 4.0)$.

Problems for Simulation:

In testing blood samples for the presence of HIV (or other viruses) a method known as “pooled sampling” is often used. For this method a number of blood samples are mixed together into a single “pooled” sample. This pooled sample is then tested for the presence on HIV. A question arises: If your blood sample was pooled with other blood samples and the pooled sample tests positive for HIV, what is the probability that your blood contains HIV?

We will assume the accuracy of the test is as follows:

- If a blood sample contains HIV, then there is a 0.98 probability that the test will detect it.
- If a blood sample does not contain HIV, then there is a 0.005 probability that the test mistakenly detects it.

We will assume that the people submitting blood samples are from a high-risk group. In particular we assume:

- Of those that submit blood samples 1.5% have HIV.

Write a computer program that simulates the disease status of people from this group submitting blood samples, the samples being pooled, and the pooled sample being tested. The program should ask for two input variables:

n the number of samples that are in the single pool.

$reps$ the total number of samples that will be simulated (note that for convenience the value entered for $reps$ should be a multiple of n).

Use this program to approximate the probability that a person has HIV given that the pooled sample tests positive for HIV. Run the program several times for each of the following cases:

- a. $n = 5$ for $reps = 100, 1000, 10000$
- b. $n = 10$ for $reps = 100, 1000, 10000$
- c. $n = 20$ for $reps = 100, 1000, 10000$

