CALIFORNIA STATE UNIVERSITY, LONG BEACH

EE 381 – Probability and Statistics with Applications to Computing

# Project on Binomial and Poisson Distributions

## 0. Introduction and Background Material

### 0.1. Random experiments that can be described by well-known probability distributions

In this project you will **simulate the rolling of three dice *n* times**. Your random variable "$X$" is the number of "successes" in $n$ rolls. This is considered one experiment. You will repeat the experiment $N$ times and you will create the probability distribution of the variable "$X$".

As an **alternative method** to the simulation experiments, you will use the formula for the **Binomial distribution** to calculate the probability distribution for the random variable "$X$". This method involves only calculations using the binomial formula, and does not involve simulations.

Similarly, **another alternative** to the simulation experiments, is to use the formula for the **Poisson distribution**, which can approximate the Binomial under certain conditions.

### 0.2. Binomial distribution

Consider the following experiment: You toss a coin, with probability of success $p$ and probability of failure $q = 1 - p$. This toss is called a *Bernoulli trial*. You repeat tossing the coin $n$ times, i.e. you have $n$ Bernoulli trials. These $n$ Bernoulli trials are independent, since the outcome of each trial does not depend on the others. The question is: *what is the probability of getting exactly $x$ successes in $n$ independent Bernoulli trials*?

The answer can be calculated from the Binomial distribution: consider the random variable $X = \{$*number of successes in $n$ Bernoulli trials*$\}$. Then:

$$p(X = x) = \binom{n}{x} p^x q^{n-x}$$

The probability distribution of $X$ in called the *Binomial distribution*.

## 0.3. Poisson distribution

Consider the following experiment: You observe the occurrence of a particular event during a time interval that has duration one unit of time. You count how many times the event has occurred during this interval. The occurrences are independent of each other, and the event occurs at an average rate of $\lambda$ times per unit of time. The question is: *what is the probability of getting exactly $x$ occurrences during the observation interval (which has duration of one time unit)* ?

The answer can be calculated from the *Poisson distribution*: consider the random variable $X = \{$*number of occurrences during a unit time interval*$\}$. Then:

$$p(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

## 1. Experimental Bernoulli Trials

Consider the following experiment: You have three identical multi-sided unfair dice. The probability vector ($p$) for the dice has been provided to you.
One roll is considered "*success*" if you get: "one" for the first die; "two" for the second die; "three" for the third die.
You roll the three dice $n$=1000 times, and the number of successes in $n$ rolls, will be your random variable "$X$". This is considered one experiment. The goal is to create the PMF plot of "$X$".

- In order to generate the PMF plot repeat the experiment $N$=10,000 times, and record the values of "$X$" each time, i.e. the number of "successes" in $n$ rolls.
- Create the experimental **Probability Mass Function** plot, using the histogram of "$X$" as you did in previous projects.
- Include the PMF plot in your report, in addition to all other requirements. See Figure 1 for an example of a properly labeled PMF plot.

## 2. Calculations using the Binomial Distribution

In this problem you will use the theoretical formula for the Binomial distribution to calculate the probability $p$ of success in a single roll of the three dice.

- Use the Binomial formula to generate the **Probability Mass Function** plot of the random variable $X = \{number\ of\ successes\ in\ n\ Bernoulli\ trials\}$.
- Compare the plot you obtain using the Binomial formula, to the plot you obtained from the experiments in Problem 1.
- "Include the PMF plot in your report, in addition to all other requirements. The graph should be plotted in the same scale as the graph in Problem 1 so that they can be compared. The title should reflect the calculations for problem 2: Bernoulli Trials: PMF – Binomial Formula"

## 3. Approximation of Binomial by Poisson Distribution

Consider the case when the probability $p$ of success in a Bernoulli trial is small and the number of trials $n$ is large (in practice this means that $n \geq 50$ and $np \leq 5$). In that case you can use the Poisson distribution formula to approximate the probability of success in $n$ trials, as an alternative to the Binomial formula. The parameter $\lambda$ that is needed for the Poisson distribution is obtained from the equation $\lambda = np$

- Use the parameter $\lambda$ and the Poisson distribution formula to create a plot of the **probability distribution function** approximating the probability distribution of the random variable $X = \{number\ of\ successes\ in\ n\ Bernoulli\ trials\}$.
- Compare the plot you obtained from the Poisson formula to the plot you obtained from the experiments in Problem 1.
- Include the PMF plot in your report, in addition to all other requirements. The graph should be plotted in the same scale as the graph in Problem 1 so that they can be compared. The title should reflect the calculations for problem 3: "Bernoulli Trials: PMF – Poisson Approximation"

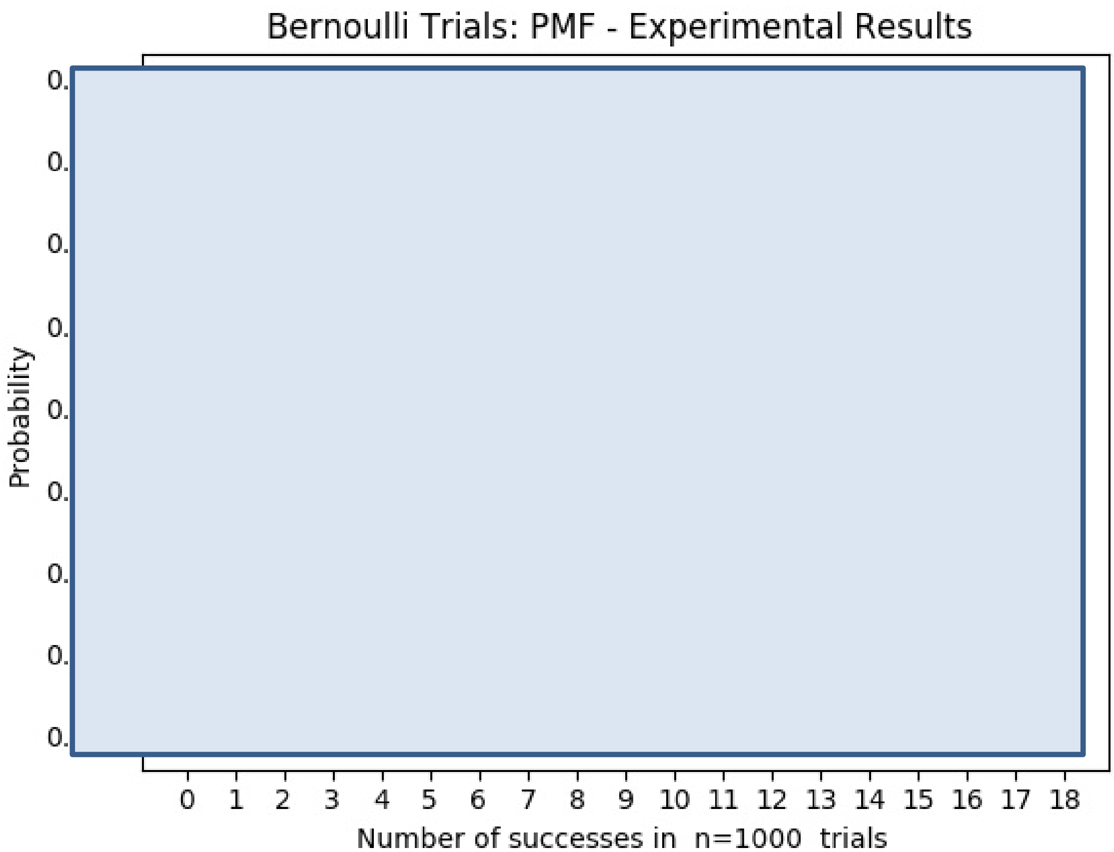


**Figure 1. Example of an appropriately labeled PMF plot.**