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 "three sixes" 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 λ 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 roll three fair dice n=1000 times. This is considered one experiment, or one Bernoulli Trial. If you get "*three sixes*" in a roll, it is considered "*success*". The number of successes in n rolls, will be your random variable "X". The goal is to create the probability mass function plot of "X".

- In order to generate the histogram 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.
- **SUBMIT** the PMF plot and your code in a Word file. Use 16 bins to plot the results. All plots should be properly labeled. See Figure 1 for an example. Note: Do not replicate the "scroll" in Figure 1. The scroll is used in the figure in order to hide the graph data.

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. Success is defined as the number of "*three sixes*" in $n = 1000$ trials.

- Use the Binomial formula to generate the **Probability Mass Function** plot of the random variable $X = \{\textit{number of successes in } n \textit{ Bernoulli trials}\}$.
- Compare the plot you obtain using the Binomial formula, to the plot you obtained from the experiments in Problem 1.
- **SUBMIT** the PMF plot and your code in a Word file. The graph should be plotted in the same scale as the graph in Problem 1 so that they can be compared.

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 λ that is needed for the Poisson distribution is obtained from the equation $\lambda = np$

- Use the parameter λ and the Poisson distribution formula to create a plot of the **probability distribution function** approximating the probability distribution of the random variable $X = \{\textit{number of successes in } n \textit{ Bernoulli trials}\}$.
- Compare the plot you obtained from the Poisson formula to the plot you obtained from the experiments in Problem 1.
- **SUBMIT** the PMF plot and your code in a Word file. The graph should be plotted in the same scale as the graph in Problem 1, so that they can be compared.

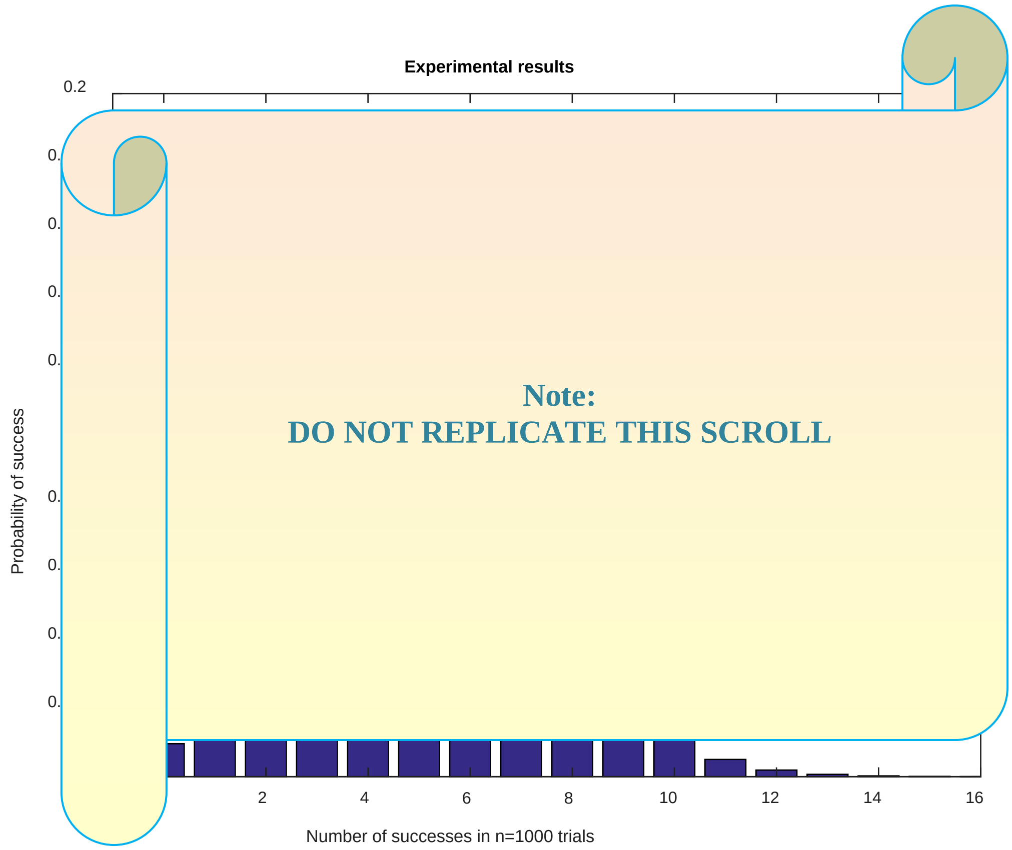

Figure 1. Example of an appropriately labeled histogram.