Probability Mass Function

(Formally defining something mentioned previously)

- a probability mass function (pmf) is a ***function*** that gives the probability that a discrete random variable is exactly equal to some value.
- The probability mass function is often the primary means of defining a discrete probability distribution
- It is conventional to present the probability mass function in the form of a table.

Binomial Example

(Revision from Last Class)
Suppose a die is tossed 5 times. What is the probability of getting exactly 2 fours?
Solution: This is a binomial experiment in which the number of trials is equal to 5, the number of successes is equal to 2, and the probability of success on a single trial is 1/6 or about 0.167. Therefore, the binomial probability is:

$$P(X = 2) = ^5C_2 \times (1/6)^2 \times (5/6)^3 = 0.161$$

Probability Tables

In the sulis workspace there are two important tables used for this part of the course. This class will feature a demonstration on how to read those tables.

- The Cumulative Binomial Tables (Murdoch Barnes Tables 1)
- The Cumulative Poisson Tables (Murdoch Barnes Tables 2)

Please get a copy of each as soon as possible.

Probability Tables

- For some value r the tables record the probability of $P(X \geq r)$.
- The Student is required to locate the appropriate column based on the parameter values for the distribution in question.
- A copy of the Murdoch Barnes Tables will be furnished to the student in the End of Year Exam. The Tables are not required for the first mid-term exam.
- Knowledge of the sample space, partitioning of the sample points, and the complement rule are advised.

Binomial Distribution : Using Tables

It is estimated by a particular bank that 25% of credit card customers pay only the minimum amount due on their monthly credit card bill and do not pay the total amount due. 50 credit card customers are randomly selected.

1. (3 marks) What is the probability that 9 or more of the selected customers pay only the minimum amount due?
2. (3 marks) What is the probability that less than 6 of the selected customers pay only the minimum amount due?
3. (3 marks) What is the probability that more than 5 but less than 10 of the selected customers pay only the minimum amount due?

Binomial Expected Value and Variance

If the random variable X has a binomial distribution with parameters n and p, we write

$$X \sim B(n,p)$$

Expectation and Variance If $X \sim B(n,p)$, then:

- Expected Value of X : E(X) = np
- Variance of X : Var(X) = np(1-p)

Suppose n=3 and p=0.5 (like our coin flipping example for tutorial 1) Then E(X) = 1.5 and V(X) = 0.75.
Remark: Referring to the expected value and variance may be used to validate the assumption of a binomial distribution.

The Geometric Distribution

- The Geometric distribution is related to the Binomial distribution in that both are based on independent trials in which the probability of success is constant and equal to p.
- However, a Geometric random variable is the number of trials until the ?rst failure, whereas a Binomial random variable is the number of successes in n trials.
- The Geometric distributions is often used in IT security applications.

The Geometric Distribution

Suppose that a random experiment has two possible outcomes, success with probability p and failure with probability 1p . The experiment is repeated until a success happens. The number of trials before the success is a random variable X computed as follows

$$P(X = k) = (1 - p)^{(k-1)} \times p$$

(i.e. The probability that first success is on the k-th trial)

The Geometric Distribution: Notation

If X has a geometric distribution with parameter p, we write

$$X \sim Geo(p)$$

Expectation and Variance If $X \sim Geo(p)$, then:

- Expected Value of X : E(X) = 1/p
- Variance of X : Var(X) = $(1-p)/p^2$.

Poisson Experiment

A Poisson experiment is a statistical experiment that has the following properties:

- The experiment results in outcomes that can be classified as successes or failures.
- The average number of successes (m) that occurs in a specified region is known.
- The probability that a success will occur is proportional to the size of the region.
- The probability that a success will occur in an extremely small region is virtually zero.

Note that the specified region could take many forms. For instance, it could be a length, an area, a volume, a period of time, etc.

Poisson Distribution

A Poisson random variable is the number of successes that result from a Poisson experiment.
The probability distribution of a Poisson random variable is called a Poisson distribution.
Given the mean number of successes (m) that occur in a specified region, we can compute the Poisson probability based on the following formula:

The Poisson Probability Distribution

- The number of occurrences in a unit period (or space)
- The expected number of occurrences is m

Poisson Formulae

The probability that there will be k occurrences in a unit time period is denoted $P(X = k)$, and is computed as follows.

$$P(X = k) = \frac{m^k e^{-m}}{k!}$$

Poisson Formulae

Given that there is on average 2 occurrences per hour, what is the probability of no occurrences in the next hour?

i.e. Compute $P(X = 0)$ given that $m = 2$

$$P(X = 0) = \frac{2^0 e^{-2}}{0!}$$

- $2^0 = 1$
- $0! = 1$

The equation reduces to

$$P(X = 0) = e^{-2} = 0.1353$$

Poisson Formulae

What is the probability of one occurrences in the next hour?

i.e. Compute $P(X=1)$ given that $m=2$

$$P(X=1) = \frac{2^1 e^{-2}}{1!}$$

- $2^1 = 2$
- $1! = 1$

The equation reduces to

$$P(X=1) = 2 \times e^{-2} = 0.2706$$