Probability Distributions

A **random variable** is a variable that assumes a unique numerical (quantitative) value for each of the outcomes in the sample space.

Example:

1)

2)

Random variables can be subdivided into two classes.

Discrete Random Variables: A quantitative variable that can assume a "countable" number. (Ie. Integers. Example 1)

Continuous Random Variables: A quantitative random variable that can assume an "uncountable" numbers of values (Ie. Real numbers. Example 2)

Example: Consider tossing two coins. Let x be the number of heads observed.

•

The distribution of probabilities associated with each of the values of the values of the random variable is called a **probability distribution**.

Example:

1) The probability distribution for tossing two coins:

2) The probability distribution for rolling a die:

What if we want to get a formula for P(x)?

A **probability function** is a rule that assigns probabilities to each of the random variables.

Every probability function must satisfy two basic properties of probabilities

1)

2)

Example: What is the P(x) for rolling a die?

Is $P(x) = \frac{x}{10}$ for $x = 1, 2, 3, 4$ a probability function?

Mean and variance of a discrete probability distribution

The mean of a discrete random variable, called the **expected value** and denoted by μ is given by

$$\boxed{\mu = \sum_x x P(x)}$$

•

The **variance of a discrete random variable** denoted by σ^2 is given by

$$\sigma^2 = \sum_x (x - \mu)^2 P(x)$$

The **standard deviation of a discrete random variable** is

$$\sigma = \sqrt{\sigma^2}$$

Computing σ^2

Example: Bellow is the sample space for rolling a black die and a white die

1,1	2,1	3,1	4,1	5,1	6,1	
1,2	2,2	3,2	4,2	5,2	6,2	(white, black)
1,3	2,3	3,3	4,3	5,3	6,3	
1,4	2,4	3,4	4,4	5,4	6,4	
1,5	2,5	3,5	4,5	5,5	6,5	
1,6	2,6	3,6	4,6	5,6	6,6	

Let x be the sum of both dice.

Summary of experiment:

Example: A small taxi company has four cabs. Let x be the number of cabs on the road on Friday night.

Probability Distribution

x	P(x)
0	0.10
1	0.15
2	0.25
3	0.35
4	0.15

a) How many cabs do we expect to find on the road at 9pm on Friday?

b) What is the standard deviation of the distribution?

c) What is the probability that the small taxi company will have more cabs than expected on the road at 9pm next Friday night?

Binomial Probability Distribution

A binomial probability experiment is an experiment that is made up of repeated trials that has the following properties

1. There are n repeated identical independent trials
2. Each trial has two possible outcomes (success or failure)
3. P(success) = p,   P(failure) = q,   p+q=1
4. The binomial random variable x is the count of the number of successful trials that occur, x may take on any integer value from 0 to n

Because the outcomes are independent it is easy to compute the probability for any particular sequence of events.

For example, the probability of getting two success followed by a failure would be

$$\begin{aligned} P(S_1 \cap S_2 \cap F_3) &= P(S_1) \cdot P(S_2) \cdot P(F_3) \\ &= p \cdot p \cdot (1-p) = p^2 q \end{aligned}$$

Example: A four question multiple choice test (with three possible answers each) is a binomial experiment made up of four trials when the answers are obtained by random guessing. Let's check to make sure al four properties are satisfied:

1)

2)

3)

4

.

Example: Consider the binomial experiment in which there are n = 5 trials. Let x be the number of successes in the five trials.

The probability distribution for a binomial experiment is called a binomial distribution.

<u>Formula for Binomial Distribution</u>

The probability of getting x successes in n trials of a binomial experiment with proportion of success p is

Example: Calculate the probability of getting at least 3 successes in a binomial experiment based on nine trials with p = 0.15

Example: 80% of all industrial accidents can be prevented by paying strict attention to safety regulations. Assuming accidents are independent events, what is the probability that at least two of the next six industrial accidents could have been prevented.

Example: A farmer guarantees that none of his cartons of a dozen eggs will contain more than one bad egg. If a carton contains more than one bad egg, he will replace the whole dozen and allow the customer to keep the original eggs. If the probability that an individual egg is bad is 0.05, what is the probability that the manager will have to replace a given carton of eggs? What assumptions did you have to make to solve this problem?

What number of bad eggs in a carton should the farmer allow before replacing the carton if he only wants to replace 1% or less cartons.

Expected value and variance of the binomial distribution

The formulas $$E(x) = \mu = n \cdot p$$

and $$\sigma^2 = \sum x^2 \, P(x) - \mu^2$$

can be simplified in the case of the binomial distribution.

$$\boxed{\begin{gathered} E(x) = \mu = n \cdot p \\ \sigma^2 = n \cdot p \cdot q \end{gathered}}$$

•

In the farmer example: