STAT 515 -- Chapter 4: Discrete Random Variables

Random Variable: A variable whose value is the numerical outcome of an experiment or random phenomenon.

Discrete Random Variable : A numerical r.v. that takes on a countable number of values (there are gaps in the range of possible values).

Examples:

1. **Number of phone calls received in a day by a company**
2. **Number of heads in 5 tosses of a coin**

Continuous Random Variable : A numerical r.v. that takes on an uncountable number of values (possible values lie in an unbroken interval).

Examples:

1. **Length of nails produced at a factory**
2. **Time in 100-meter dash for runners**

Other examples?

The probability distribution of a random variable is a graph, table, or formula which tells what values the r.v. can take and the probability that it takes each of those values.

Example 1: Roll 1 die. The r.v. X = number of dots showing.

x	1	2	3	4	5	6
P(x)	**1/6**	**1/6**	**1/6**	**1/6**	**1/6**	**1/6**

Example 2: Toss 2 coins. The r.v. X = number of heads showing.

x	0	1	2
P(x)	**¼**	**½**	**¼**

Graph for Example 2:

For <u>any</u> probability distribution:

(1) P(x) is between 0 and 1 for any value of x.
(2) $\sum_x P(x)$ = 1. That is, the sum of the probabilities for all possible x values is 1.

Example 3: P(x) = x / 10 for x = 1, 2, 3, 4.

Valid Probability Distribution?
Property 1?
Property 2?

Expected Value of a Discrete Random Variable

The expected value of a r.v. is its mean (i.e., the mean of its probability distribution).

For a discrete r.v. X, the expected value of X, denoted μ or E(X), is:

$$\mu = E(X) = \Sigma\, x\, P(x)$$

where Σ represents a summation over all values of x.

Recall Example 3:

$\mu =$

Here, the expected value of X is

Example 4: Suppose a raffle ticket costs \$1. Two tickets will win prizes: First prize = \$500 and second prize = \$300. Suppose 1500 tickets are sold. What is the expected profit for a ticket buyer?

x (profit)	
P(x)	

E(X) =

E(X) = -0.47 dollars, so on average, a ticket buyer will lose 47 cents.

The expected value does <u>not</u> have to be a possible value of the r.v. --- it's an <u>average</u> value.

<u>Variance of a Discrete Random Variable</u>

The variance σ^2 is the expected value of the squared deviations from the mean μ; that is, $\sigma^2 = E[(X - \mu)^2]$.

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

Shortcut formula:

$$\sigma^2 = [\Sigma\,x^2\,P(x)] - \mu^2$$

where Σ represents a summation over all values of x.

<u>Example 3</u>: Recall $\mu = 3$ for this r.v.

$\Sigma\,x^2\,P(x) =$

Thus $\sigma^2 =$

Note that the standard deviation σ of the r.v. is the square root of σ^2.

For Example 3, $\sigma =$

The Binomial Random Variable

Many experiments have responses with 2 possibilities (Yes/No, Pass/Fail).

Certain experiments called binomial experiments yield a type of r.v. called a binomial random variable.

Characteristics of a binomial experiment:

(1) The experiment consists of a number (denoted n) of identical trials.

(2) There are only two possible outcomes for each trial – denoted "Success" (S) or "Failure" (F)

(3) The probability of success (denoted p) is the same for each trial.
(Probability of failure = $q = 1 - p$.)

(4) The trials are independent.

Then the binomial r.v. (denoted X) is the number of successes in the n trials.

Example 1: A fair coin is flipped 5 times. Define "success" as "head". X = total number of heads.
Then X is

Example 2: A student randomly guesses answers on a multiple choice test with 3 questions, each with 4 possible answers. X = number of correct answers.
Then X is

What is the probability distribution for *X* in this case?

Outcome	***X***	**P(outcome)**

Probability Distribution of *X*

x	**P(*x*)**

<u>General Formula</u>: (Binomial Probability Distribution)
(n = number of trials, p = probability of success.)
The probability there will be exactly x successes is:

$$P(x) = \binom{n}{x} p^x q^{n-x} \qquad (x = 0, 1, 2, \ldots , n)$$

where

$$\binom{n}{x} = \text{“}n \text{ choose } x\text{”}$$

$$= \frac{n!}{x!\,(n-x)!}$$

Here, $0! = 1$, $1! = 1$, $2! = 2 \cdot 1 = 2$, $3! = 3 \cdot 2 \cdot 1 = 6$, etc.

Example: Suppose probability of “red” in a roulette wheel spin is 18/38. In 5 spins of the wheel, what is the probability of exactly 4 red outcomes?

- **The mean (expected value) of a binomial r.v. is $\mu = np$.**
- **The variance of a binomial r.v. is $\sigma^2 = npq$.**
- **The standard deviation of a binomial r.v. is $\sigma =$**

Example: What is the mean number of red outcomes that we would expect in 5 spins of a roulette wheel?

$\mu = np =$

What is the standard deviation of this binomial r.v.?

<u>Using Binomial Tables</u>

Since hand calculations of binomial probabilities are tedious, Table II gives "cumulative probabilities" for certain values of *n* and *p*.

Example:
Suppose *X* is a binomial r.v. with $n = 10$, $p = 0.40$.
Table II (page 785) gives:

Probability of 5 or fewer successes: $P(X \leq 5) =$

Probability of 8 or fewer successes: $P(X \leq 8) =$

What about …

… the probability of exactly 5 successes?

… the probability of more than 5 successes?

… the probability of 5 or more successes?

… the probability of 6, 7, or 8 successes?

Why doesn’t the table give P($X \leq 10$)?

Hypergeometric Random Variables

Consider the situation of selecting several items from a group, where each item is either a "success" or a "failure" and counting how many of the selected items are successes.

If the probabilities of "success" do not change from one selection to the next, and the selections are independent, then the count of overall successes is a ___________ r.v.

But suppose we are selecting (without replacement) several elements from a relatively small group. In this case, is the count of successes a ________________ r.v.?

Fact: If we select n elements, without replacement, from a set of N elements, r of which are Successes and $(N - r)$ of which are Failures, then the number X of the selected elements that are Successes follows a hypergeometric distribution.

Probability distribution for X
(if X is Hypergeometric with parameters n, r, and N)

Mean of Hypergeometric probability distribution:

Variance of Hypergeometric probability distribution:

Example: There are 12 faculty members in a statistics department, and 5 are specialists in Bayesian statistics. Suppose we must randomly select 4 of the statistics faculty to form a committee. What is the expected number of Bayesian specialists on the committee?

What is the probability of having exactly two Bayesian specialists on the committee?

What is the probability of having more than two Bayesian specialists on the committee?

Poisson Random Variables

The Poisson distribution is a common distribution used to model "count" data:

- **Number of telephone calls received per hour**
- **Number of claims received per day by an insurance company**
- **Number of accidents per month at an intersection**

The mean number of events for a Poisson distribution is denoted λ.

Which values can a Poisson r.v. take?

Probability distribution for X
(if X is Poisson with mean λ)

$$P(x) = \frac{\lambda^{x} e^{-\lambda}}{x!} \qquad \text{(for } x = 0, 1, 2, \ldots)$$

Mean of Poisson probability distribution: λ

Variance of Poisson probability distribution: λ

Example: A call center averages 10 calls per hour. Assume *X* (the number of calls in an hour) follows a Poisson distribution. What is the probability that the call center receives exactly 3 calls in the next hour?

What is the probability the call center will receive 2 or more calls in the next hour?

Calculating Poisson probabilities by hand can be tedious. Table III gives cumulative probabilities for a Poisson r.v., P($X \leq k$) for various values of k and λ.

Example 1: X is Poisson with $\lambda = 1$. Then

P($X \leq 1$) =

P($X \geq 3$) =

P($X = 2$) =

Example 2: X is Poisson with $\lambda = 6$. Then

… probability that X is 5 or more?

… probability that X is 7, 8, or 9?

Linear Transformations, Sums, and Differences of Random Variables

In general, the expected value is a "linear operator". This means:

If:

Then:

In particular:

Also:

Proof:

Hence:

Example: Suppose *X* = July daily temperature (in degrees Fahrenheit) has mean 92 and standard deviation 2. If *Y* = July daily temperature (in degrees Celsius), then find the mean and std. deviation of *Y*.

For two r.v.'s X and Y,

If X and Y are independent, then:

Example 1: A business undertakes a venture in Atlanta and a venture in Chicago. Assume X (revenue in \$ from Atlanta venture) and Y (revenue in \$ from Chicago venture) are independent. X has expected value 50,000 and variance 1,000,000. Y has expected value 40,000 and variance 1,000,000.

Find expected total revenue:

Find standard deviation of total revenue:

Example 2: Suppose the Atlanta venture has expected cost \$35,000 and cost variance = 500,000.

Find expected profit:

Find standard deviation of profit: