AP Statistics Chapter 6 – Discrete, Binomial & Geometric Random Variables

6.1: Discrete Random Variables

Random Variable
A random variable is a variable whose value is a numerical outcome of a random phenomenon.

Discrete Random Variable
A discrete random variable *X* has a *countable* number of possible values. Generally, these values are limited to integers (whole numbers). The probability distribution of X lists the values and their probabilities.

Value of X	x_1	x_2	x_3	…	x_k
Probability	p_1	p_2	p_3	…	p_k

The probabilities p_i must satisfy two requirements:

1. Every probability p_i is a number between 0 and 1.
2. $p_1 + p_2 + \ldots + p_k = 1$

Find the probability of any event by adding the probabilities p_i of the particular values x_i that make up the event.

Continuous Random Variable
A continuous random variable *X* takes all values in an interval of numbers and is *measurable*.

Mean (Expected Value) of A Discrete Random Variable
Suppose that X is a discrete random variable whose distribution is

Value of X	x_1	x_2	x_3	…	x_k
Probability	p_1	p_2	p_3	…	p_k

To find the **mean** of X, multiply each possible value by its probability, then add all the products:

$$\mu_x = E(x) = \sum x_i \cdot p_i = x_1 \cdot p_i + x_2 \cdot p_2 + \cdots + x_k \cdot p_k$$

6.3: The Binomial Distributions

A **binomial probability distribution** occurs when the following requirements are met.

1. Each observation falls into one of just two categories – call them "success" or "failure."
2. The procedure has a fixed number of trials – we call this value *n*.
3. The observations must be *independent* – result of one does not affect another.
4. The probability of success – call it *p* - remains the same for each observation.

Notation for binomial probability distribution

n denotes the number of fixed trials

k denotes the number of successes in the *n* trials

p denotes the probability of success

1 – p denotes the probability of failure

Binomial Probability Formula

$$P(X=k)=\frac{n!}{k!(n-k)!}(p)^k(1-p)^{n-k}$$

How to use the TI-83/4 to compute binomial probabilities *

There are two binomial probability functions on the TI-83/84, *binompdf* and *binomcdf*

binompdf is a *probability distribution function* and determines $P(X=k)$

binomcdf is a *cumulative distribution function* and determines $P(X\le k)$

*Both functions are found in the DISTR menu (2nd-VARS)

Probability	Calculator Command	Example (assume n = 4, p = .8)
$P(X=k)$	*binompdf(n, p, k)*	$P(X=3)$ = ***binompdf(4, .8, 3)***
$P(X\le k)$	*binomcdf(n, p, k)*	$P(X\le 3)$ = ***binomcdf(4, .8, 3)***
$P(X<k)$	*binomcdf(n, p, k - 1)*	$P(X<3)$ = ***binomcdf(4, .8, 2)***
$P(X>k)$	*1 – binomcdf(n, p, k)*	$P(X>3)$ = **1 – *binomcdf(4, .8, 3)***
$P(X\ge k)$	*1 – binomcdf(n, p, k - 1)*	$P(X\ge 3)$ = **1 – *binomcdf(4, .8, 2)***

Mean (expected value) of a Binomial Random Variable

Formula: $\mu=np$ Meaning: Expected number of successes in *n* trials (think *average*)

Example: *Suppose you are a 80% free throw shooter. You are going to shoot 4 free throws.*

For *n* = 4, *p* = .8, $\mu=(4)(.8)=3.2$, which means we expect 3.2 makes out of 4 shots, on average

6.3: The Geometric Distributions

A **geometric probability distribution** occurs when the following requirements are met.

1. Each observation falls into one of just two categories – call them "success" or "failure."
2. The observations must be *independent* – result of one does not affect another.
3. The probability of success – call it *p* - remains the same for each observation.
4. The variable of interest is the number of trials required to obtain the first success.*

* As such, the geometric is also called a "waiting-time" distribution

Notation for geometric probability distribution

n denotes the number of trials required to obtain the first success

p denotes the probability of success

1 – p denotes the probability of failure

Geometric Probability Formula
$P(X = n) = (1 - p)^{n-1}(p)$

How to use the TI-83/4 to compute geometric probabilities *

There are two geometric probability functions on the TI-83/84, *geometpdf* and *geometcdf*

geometpdf is a *probability distribution function* and determines $P(X = n)$

geometcdf is a *cumulative distribution function* and determines $P(X \le n)$

*Both functions are found in the DISTR menu (2^{nd}-VARS)

Probability	Calculator Command	Example (assume p = .8, n = 3)
$P(X = n)$	*geometpdf (p, n)*	$P(X = 3)$ = ***geometpdf(.8, 3)***
$P(X \le n)$	*geometcdf(p, n)*	$P(X \le 3)$ = ***geometcdf(.8, 3)***
$P(X < n)$	*geometcdf(p, n-1)*	$P(X < 3)$ = ***geometcdf(.8, 2)***
$P(X > n)$	*1 – geometcdf(p, n)*	$P(X > 3)$ = **1 –** ***geometcdf(.8, 3)***
$P(X \ge n)$	*1 – geometcdf(p, n-1)*	$P(X \ge 3)$ = **1 –** ***geometcdf(.8, 2)***

Mean (expected value) of a Geometric Random Variable

Formula: $\mu = \frac{1}{p}$ Meaning: Expected number of *n* trials to achieve first success (*average*)

Example: *Suppose you are a 80% free throw shooter. You are going to shoot until you make.*

For *p* = .8, $\mu = \frac{1}{.8} = 1.25$, which means we expect to take 1.25 shots, on average, to make first