Unit 6
The Bernoulli and Binomial Distributions

"If you believe in miracles, head for the Keno lounge"
- Jimmy the Greek

The Amherst Regional High School provides flu vaccinations to a random sample of 200 students. How many will develop the flu? A new treatment for stage IV melanoma is given to 75 cases. How many will survive two or more years? In a sample of 300 cases of uterine cancer, how many have a history of IUD use?

Counts. This unit is all about counts. The number of "events" in each of the scenarios above is modeled well using a Binomial probability distribution. When the number of trials is just one, the probability model is called a Bernoulli trial.

The **Bernoulli** and **Binomial** probability distributions are used to model the chance occurrence of "success/failure" outcomes. They are also the basis of logistic regression which is used to explore (through modeling) hypothesized predictors of "yes/no" (event/non-event) outcomes.

Cheers!

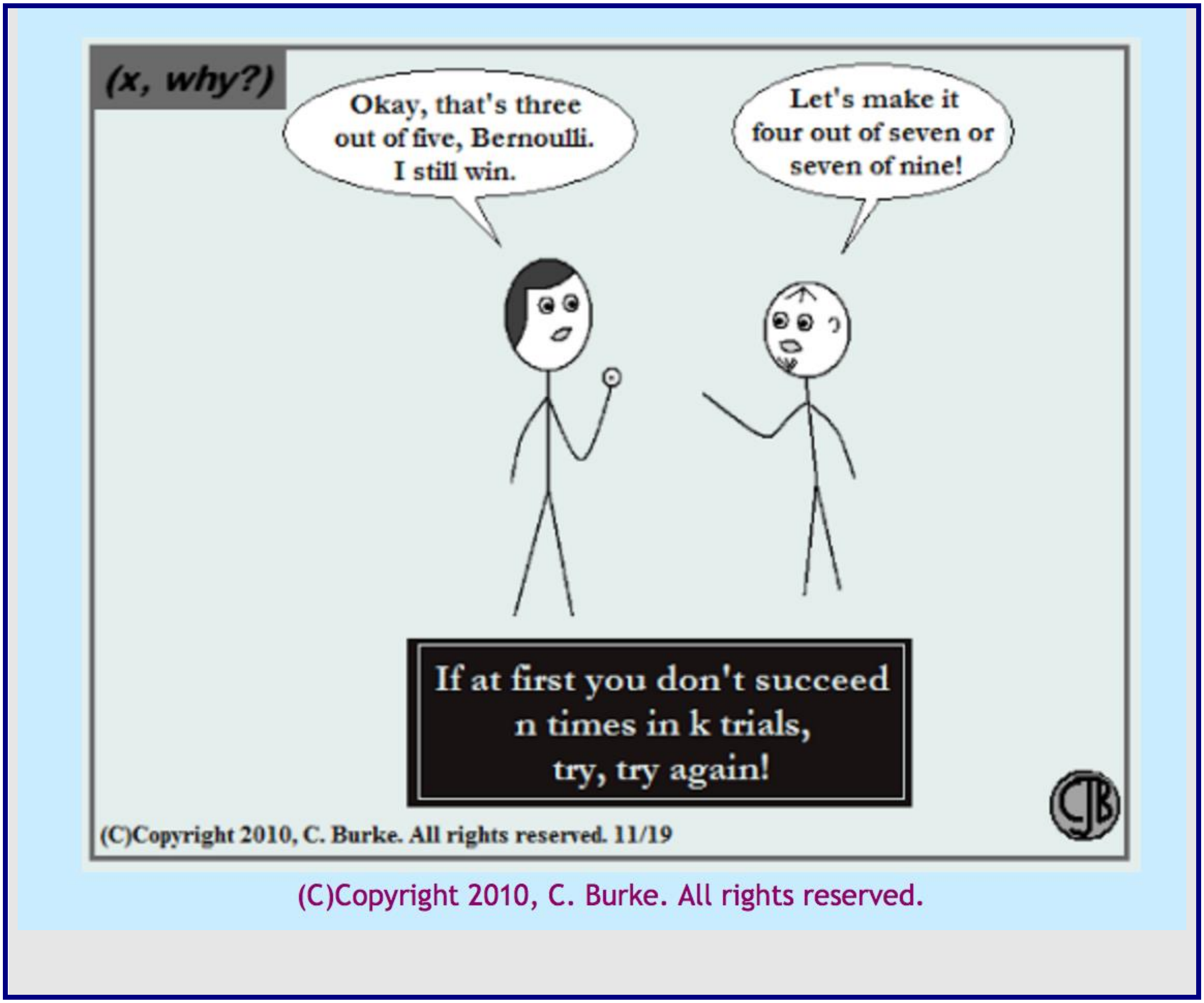
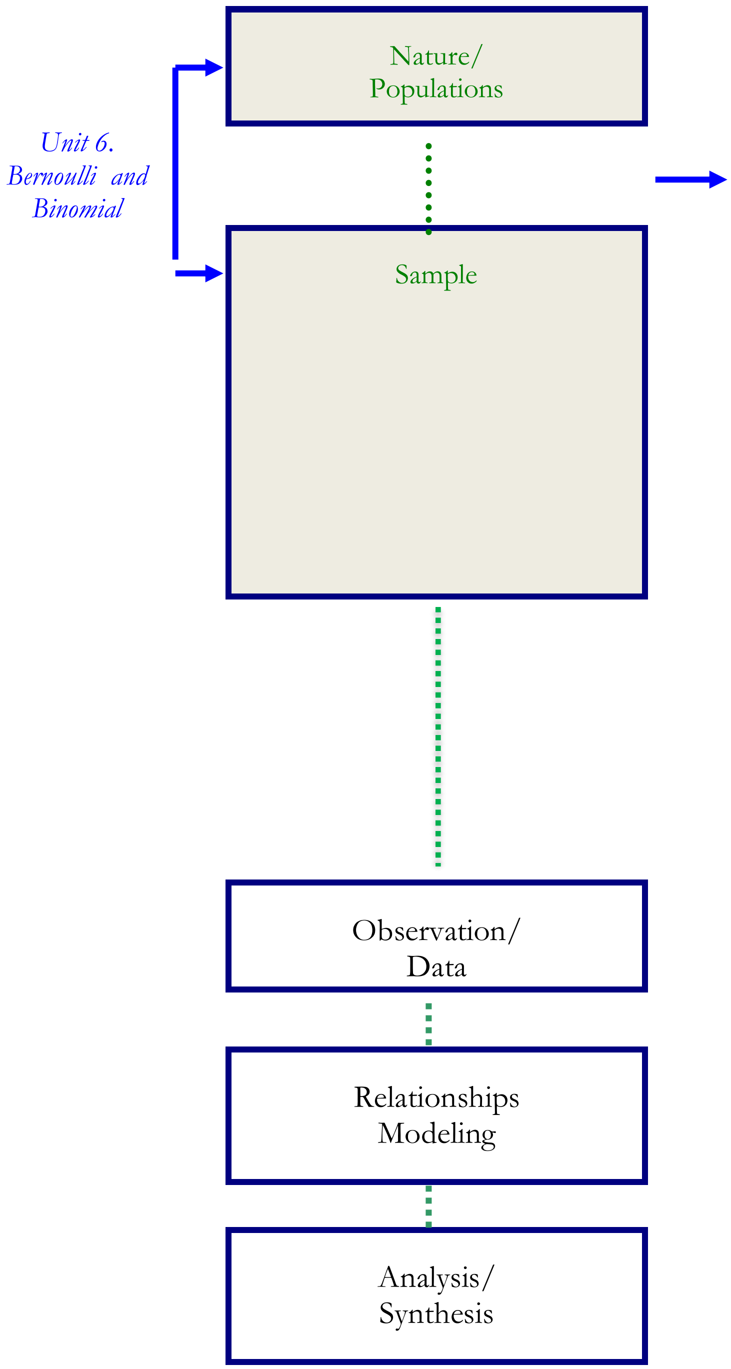

(x, why?)
Okay, that's three out of five, Bernoulli. I still win.
Let's make it four out of seven or seven of nine!
If at first you don't succeed n times in k trials, try, try again!
(C)Copyright 2010, C. Burke. All rights reserved. 11/19
(C)Copyright 2010, C. Burke. All rights reserved.

Table of Contents

1. Unit Roadmap



Data are the measurements of our observations of nature. This unit focuses on nominal data that are binary.

Previously, we learned that data can be of several types – nominal, ordinal, quantitative, etc. A **nominal** variable that is binary or dichotomous has exactly two possible values. Examples are vital status (alive/dead), exposure (yes/no), tumor remission(yes/no), etc.

The "frequentist" view of probability says that *probability is the relative frequency in an indefinitely large number of trials*. In this framework, a probability distribution model is a model of chance. It describes the way that probability is distributed among the possible values that a random variable can take on.

The **Bernoulli** and **Binomial** probability distribution models are often very good models of patterns of occurrence of binary ("yes/no") events that are of interest in public health; eg - mortality, disease, and exposure.

Nature — Population/Sample — Observation/Data — Relationships/Modeling — Analysis/Synthesis

2. Learning Objectives

When you have finished this unit, you should be able to:

- Explain the "frequentist" approach to probability;
- Define a discrete probability distribution;
- Explain statistical expectation for a discrete random variable;
- Define the Bernoulli probability distribution model;
- Explain factorials and combinatorials and how to "count the # ways" ;
- Define the Binomial probability distribution model; and
- Calculate binomial probabilities.

3. Introduction to Discrete Probability Distributions

Recall.

A variable is a quantity (or characteristic) that may vary from object to object *(typically, the object is a person on whom we are making an observation)*.

A random variable denotes a quantity (or characteristic) that varies from object to object according to some sort of random process. *Chance!*

Remember:
- We represent the name of the random variable with capital letters (X, Y, etc)
- We represent the values of the random variable with lower case letters (x, y, etc.)

A discrete random variable is one for which there is only a *finite* set of possible values. *Finite" is nice, because it means that, in principle, we can list them out.*

A probability distribution is a *model* that links the possible values of a random variable with the likelihood (chances) of their occurrence. The model might take the form of a table or an equation. *Tables versus equations will be clear in the examples that follow.*

How to define a discrete probability distribution. This is straightforward. We need two things::

1st - a listing of all the possible random variable values
Check to be sure that you got them all (that it is exhaustive); and

2nd - the listing of the associated likelihoods (probabilities) of the occurrence of each.
Each will be a number between 0 and 1. Or it will be an equation (the probability density function) that yields a number between 0 and 1.
Check to be sure that, when you add these all up, the total is 1 or 100%

How to define a continuous probability distribution. *We'll get to this in Unit 7, The Normal Distribution.*

Example of a Discrete Random Variable - Assigned sex at birth of a randomly selected student

- We'll use **capital X** as our placeholder for the random variable name:
  X = Assigned sex at birth of randomly selected student from the population of students at a University

- We'll use **small x** as our placeholder for a value of the random variable X:
  x = 0 if selected student male sex at birth
  x = 1 if selected student is female sex at birth

Example, continued - Probability Distribution of X= Sex at Birth of a randomly selected student

1st – Listing of all possible values Value of the Random Variable X is x =	2nd – Listing of associated likelihoods (probabilities) Probability density function: Pr [X = x] =
0 = male sex at birth 1 = female sex at birth ↑ ***Note that this roster exhausts all possibilities.***	Pr [X = 0] = 0.53 Pr [X = 1] = 0.47 ↑ ***Note that the sum of these individual probabilities, because the sum is taken over all possibilities, is 100% or 1.00.***

Some useful terminology -

1. For discrete random variables, a ***probability model*** is the set of assumptions used to assign probabilities to each outcome in the sample space.

   The ***sample space*** is the collection of all possible values of the random variable.

2. A ***probability distribution*** defines the relationship between the outcomes and their likelihood of occurrence.

3. ***Putting it all together …*** To ***define a probability distribution***, we make an assumption (the probability model) and use this to assign likelihoods.

4. Statistical Expectation for Discrete Random Variables

Statistical expectation was introduced in Appendix 2 of Unit 3 *Probability Basics*, pp 42-43. The following paragraph is excerpted from that appendix.

An Example to Motivate the Ideas:
Suppose you stop at a convenience store on your way home and play the lottery.

Suppose further that the specifics of the lottery game that you are playing are the following–

$1 is won with probability = 0.50
$5 is won with probability = 0.25
$10 is won with probability = 0.15
$25 is won with probability = 0.10 *(wow, no matter what, you'll win something!)*

Equipped with this information, you can calculate for yourself the average winning or, put another way, what the winning is *likely to be in the long run*.

Average winning ("what the winning is likely to be in the long run")

= [$1](probability of a $1 ticket)
+ [$5](probability of a $5 ticket)
+ [$10](probability of a $10 ticket)
+ [$25](probability of a $25 ticket)

= [$1](0.50) + [$5](0.25) +[$10](0.15) + [$25](0.10)

= $5.75

The solution to a statistical expectation is NOT necessarily a possible outcome. Bin this example, the statistical expectation of "winning" is $5.75 which is not a possible outcome. Taking a weighted sum of things will do that, sometimes.

Some other names for statistical expectation here:

- ♣ Expected winnings
- ♣ "Long range average"
- ♣ Statistical expectation!

The statistical expectation for a discrete random variable is a weighted sum of the possible outcomes. Each possible outcome is given a weight equal to its chances of occurrence (likelihood, probability").

[Statistical expectation = $5.75]

= [$1 winning] * (chances, or probability, of this winning = 0.50) +
[$5 winning] * (chances, or probability, of this winning = 0.25) +
[$10 winning] * (chances, or probability, of this winning = 0.15) +
[$25 winning]] * (chances, or probability, of this winning = 0.10)

= [$1](.50) + [$5](.25) + [$10](.15) + [$25](.10)

The *statistical expectation*, $5.75, is useful even if it does not correspond to a possible winning. Imagine you are in the budget office of the State of Massachusetts. This lottery game was your brain-child. The statistical expectation $5.75 has the following interpretation. It represents what the State of Massachusetts can expect to have to pay out *on average, in the long run* to each player. *Still with me? At this point a bell might go off in your head, saying: well then, if the state hopes to make money then it had better hike up the cost to purchase a lottery ticket. Stay tuned.*

We have what we need to define the statistical expectation of a discrete random variable. *Note – we'll have to modify this slightly when we come to random variables that are continuous. Stay tuned for Unit 7 (Normal Distribution).*

Statistical Expectation
Discrete Random Variable X

For a discrete random variable X (e.g. winning in lottery) having probability distribution as follows:

Value of X, x =	P[X = x] =
$ 1	0.50
$ 5	0.25
$10	0.15
$25	0.10

The **statistical expectation** of the random variable X is written as $E[X]=\mu$. When X is discrete, it is equal to the weighted sum of all the possible values x, using weights equal to associated probabilities of occurrence Pr[X=x]

$$E[X]=\mu=\sum_{\text{all possible} X=x}[x]P(X=x)$$

Example, continued - In the "likely winnings" example, μ = $5.75

We can calculate the statistical expectation of other things, too.

Example, continued – Now suppose that, what we really want to know is about profit! That is, suppose we want to know how much we can expect to win or lose, by taking into account the cost of the purchase of the lottery ticket.

Suppose a lottery ticket costs $15 to purchase.

So, really, a lottery ticked purchaser can expect to win a -$9.25. Put another way, he or she can expect to lose $9.25. Here's how it works.

[Statistical expectation of net profit = -$9.25]

= [$1 winning - $15 cost] * (percent of the time this winning occurs=0.50) +
[$5 winning - $15 cost] * (percent of the time this winning occurs =0.25) +
[$10 winning - $15 cost] * (percent of the time this winning occurs = 0.15) +
[$25 winning - $15 cost] * (percent of the time this winning occurs = 0.10)

= [-$14](.50) + [-$10](.25) + [-$5](.15) + [+$10](.10)

Statistical Expectation
Discrete Random Variable Y = [X-15]

Value of Y, y =	P[Y=y] =
$ 1 - $15 = -$14	0.50
$ 5 - $15 = -$10	0.25
$10 - $15 = -$5	0.15
$25 - $15 = +$10	0.10

How much can you expect to lose? In the long run, you can expect to lose an amount that is equal to the **statistical expectation of Y**, denoted $E[Y]=\mu_Y$

$$\mu_Y = \sum_{\text{all possible } Y=y} [y]\, P(Y=y) = -\$9.25$$

It seems that, on average you can expect to lose $9.25. Maybe you don't want to buy a lottery ticket on your way home.….

5. The Variance of a Random Variable is a Statistical Expectation

Example, continued - One play of the Massachusetts State Lottery.

- **The random variable X** is the "winnings". X has possible values (think sampling frame) x=\$1, \$5, \$10, and \$25.
- **The statistical expectation of X** is μ = \$5.75. Recall that this figure is what the state of Massachusetts can expect to pay out, on average, in the long run.
- **What about the variability in X?** In learning about population variance σ^2 for the first time, we understood this to be a measure of the variability of individual values in a population.

The population variance σ^2 of a random variable X is also a statistical expectation! It is the statistical expectation of the quantity $[X - \mu]^2$

Discrete Random Variables
Variance σ^2 = Statistical Expectation of $[X\text{-}\mu]^2$
$= E[X\text{-}\mu]^2$

For a discrete random variable X (e.g. winning in lottery) having probability distribution as follows:

Value of $[X\text{-}\mu]^2$ =		P[X = x] =
$[1 - 5.75]^2$ =	22.56	0.50
$[5 - 5.75]^2$ =	0.56	0.25
$[10 - 5.75]^2$ =	18.06	0.15
$[25 - 5.75]^2$ =	370.56	0.10

The **variance of a random variable X** is the **statistical expectation of the random variable $[X\text{-}\mu]^2$** and is written as $Var[X]=\sigma^2$ When *X is discrete*, it is calculated as the weighted sum of all the possible values $[x\text{-}\mu]^2$, using weights equal to associated probabilities of occurrence Pr[X=x]

$$\sigma^2 = E\left[(X-\mu)^2\right] = \sum_{\text{all possible } X=x} [(x-\mu)^2]\, P(X=x)$$

Example, continued - In the "likely winnings" example, σ^2 = 51.19 dollars *squared*.

6. The Bernoulli Distribution

The Bernoulli Distribution is an example of a discrete probability distribution (perhaps the simplest – think coin toss). It is often the probability model that is used for the analysis of proportions and rates. In BIOSTATS 640, we'll learn about another distribution that is used for the analysis of rates, the Poisson distribution.

Example – The fair coin toss.

- We'll use **capital Z** as our placeholder for the random variable name here **(there's a reason for Z as you'll see in the coming pages)**:

  Z = Face of a single coin toss

- We'll use **small z** as our placeholder for a value of the random variable Z:
  z = 1 if "heads"
  z = 0 if "tails"

- We'll use **π and (1-π)** as our placeholder for the associated probabilities
  π = Pr[Z=1] **eg – This is the probability of "heads" and is equal to .5 when the coin is fair**
  $(1-\pi)$ = Pr[Z=0]

Bernoulli Distribution (π) ("Bernoulli Trial")

A random variable Z is said to have a **Bernoulli Distribution** if it takes on the value 1 with probability π and takes on the value 0 with probability $(1-\pi)$.

Value of Z =	P[Z = z] =
1	π
0	$(1-\pi)$

(1) **μ = Mean = E[Z] = Statistical Expectation of Z**
$\mu = \pi$

(2) **σ^2 = Variance = Var[Z] = E[$(Z-\mu)^2$] = Statistical Expectation of $(Z-\mu)^2$**
$\sigma^2 = \pi(1-\pi)$

A Bernoulli Distribution is used to model the outcome of a SINGLE "event" trial
Eg – mortality, MI, etc.

Mean (μ) and Variance (σ^2) of a Bernoulli Distribution

Mean of Z = μ = π

μ = the mean of Z (the statistical expectation of Z) is represented as E[Z].

E[Z] = π because the following is true:

$$\mu = \text{E[Z]} = \sum_{\text{All possible z}} [\text{z}]\text{Probability}[\text{Z=z}]$$

$$= [0]\Pr[\text{Z=0}] + [1]\Pr[\text{Z=1}]$$

$$= [0](1-\pi) + [1](\pi)$$

$$= \pi$$

Variance of Z = σ^2 = (π) (1-π)

The variance of Z is Var[Z] = E[(Z – (EZ)2].

Var[Z] = π(1-π) because the following is true:

$$\text{Var[Z]} = \text{E}[(\text{Z}-\pi)^2] = \sum_{\text{All possible z}} [(\text{z}-\pi)^2]\text{Probability}[\text{Z=z}]$$

$$= [(0-\pi)^2]\Pr[\text{Z=0}] + [(1-\pi)^2]\Pr[\text{Z=1}]$$

$$= [\pi^2](1-\pi) + [(1-\pi)^2](\pi)$$

$$= \pi(1-\pi)[\pi + (1-\pi)]$$

$$= \pi(1-\pi)$$

7. Introduction to Factorials and Combinatorials

Why are we doing this?
From 1 Trial to Many Trials -
When we do a SINGLE trial of event/non-event occurrence, this is a *Bernoulli* trial.
When we do SEVERAL trials of event/non-event occurrence, this is a *Binomial* random variable. We need to understand factorials and combinatorials in order to understand the *Binomial distribution*.

Example -
Birthday party: 5 Guests, 2 Hats - ***who gets to wear a hat?***
A birthday party has 5 guests (n=5): Bill, Ed, Sarah, John, and Alice. There are just 2 hats (x=2).
From among 5 guests, how many ways are there to choose 2 guests to wear a hat? This is an example of:

- Sampling 2 items without replacement from a collection of 5

- "5 choose 2" which is written more generally as "n choose x"

  "5 choose 2" is written with either of two special notations: ${}_5C_2$ or $\binom{5}{2}$ tip – "C" is for "choose"

- "n choose x" is written with the either of two special notations: ${}_nC_x$ or $\binom{n}{x}$

The answer is: there are 10 ways (hat):

1	**Bill**	**Ed**	**Sarah**	**John**	**Alice**
2	**Bill**	**Ed**	**Sarah**	**John**	**Alice**
3	**Bill**	**Ed**	**Sarah**	**John**	**Alice**
4	**Bill**	**Ed**	**Sarah**	**John**	**Alice**
5	**Bill**	**Ed**	**Sarah**	**John**	**Alice**
6	**Bill**	**Ed**	**Sarah**	**John**	**Alice**
7	**Bill**	**Ed**	**Sarah**	**John**	**Alice**
8	**Bill**	**Ed**	**Sarah**	**John**	**Alice**
9	**Bill**	**Ed**	**Sarah**	**John**	**Alice**
10	**Bill**	**Ed**	**Sarah**	**John**	**Alice**

- **Oh, how annoying you say.** Is there a way we can obtain # ways = 10 without having to write them all out?

- Yes, and it makes use of 2 tools: **factorial and combinatorial**

Preliminary – Introduction to the factorial

The **"factorial"** is just a shorthand that saves us from having to write out in longhand multiplications of the form (3)(2)(1) or (5)(4)(3)(2)(1) or (10)(9)(8)(7)(6)(5)(4)(3)(2)(1) … well, you get the idea.

- Notation: "n factorial" is written n!
- Definition: n! = (n)(n-1)(n-2) … (3)(2)(1)
- Example - 3! = (3)(2)(1) = 6
- Example - 8! = (8)(7)(6)(5)(4)(3)(2)(1) = 40,320
- Definition: 0! = 1
  Dear reader, I know. 0! =1 seems weird. It turns out to be the device we need for the idea of factorials to make sense

Factorial

n! = (n)(n-1)(n-2) … (2)(1)

0! = 1 *by convention (otherwise, we get into a pickle)*

Motivating the Combinatorial – Example, continued

A **"combinatorial"** is the name given to the solution to the question **"how many ways can you choose x from n**, when sampling **without replacement**?" The solution itself makes use of factorials.

Five guests are at a birthday party. How many ways can you choose 2 to wear a hat?
Important – Choosing "2 to wear a hat" is actually saying "2 wear a hat and the remaining 3 do NOT wear a hat".

- We've seen that one "way" that satisfies "2 hats and 3 non-hats" occurs when the first 2 guests (Bob and Ed) wear a hat and the remaining 3 do not wear a hat:

BOB ED SARAH JOHN ALICE

- Another "way" that satisfies "2 hats and 3 non-hats" occurs when the last 2 guests (John and Alice) wear a hat and the first 3 do not wear a hat:

BOB ED SARAH JOHN ALICE

- So now - what is the total number of outcomes that satisfy "2 hats and 3 non-hats"? The answer is a combinatorial. Here, it is solved as "five choose 2" and is equal to:

$$\text{"5 choose 2" ways} = {}_5C_2 = \binom{5}{2} = \frac{5!}{2!\,3!} = \frac{(5)(4)(3)(2)(1)}{(2)(1)(3)(2)(1)} = 10$$

- *Check:* All 10 are shown on page 14.

The Combinatorial

- Question: **"How many ways can we choose "x" from "n?"** Another wording of this question is the following: **"What is the number of combinations of n items that can be formed by taking them (*by handfuls*) x at a time?"**

- Notation: Recall. There are two notations for "n choose x": ${}_nC_x$ and $\binom{n}{x}$

Combinatorial

The number of ways to choose x items from n (without regard to order) is "n choose x":

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

Note: "choosing "x" is the same as leaving behind (n-x).

$$ \binom{n}{x} = \binom{n}{n-x} $$

Similarly: "choosing "0" is the same as leaving behind all n.

$$ \binom{n}{0} = \binom{n}{n} = 1 $$

HOMEWORK DUE Friday October 28, 2022
Question #1 of 5

Suppose that my 2022 BIOSTATS 540 class has just 10 students.

a. I wish to pair up students to work on homework together. How many pairs of 2 students could I form?

b. Next, I wish to form project groups of size 5. How many groups of 5 students could I form?

8. The Binomial Distribution

When is the Binomial Distribution Used?
The binomial distribution is used to answer questions of the form: "What is the probability that, in n independent success/failure trials with probability of success equal to π (also very often denoted as p), the result is x events of success?"

What are n, p = π, and X in the Binomial Distribution?

n = number of independent trials (eg – the number vaccinated for flu, n=100)
p = π = Probability individual trial yields "event" (eg – pr[single individual contracts flu] = .04)
x = number of events of success that is obtained (eg – x=13 "events of flu in subsequent season")

Example (details of the solution to a similar, but not identical, question are on page 22)
What is the probability that, among n=100 vaccinated for flu, with subsequent probability of flu p = π = .04, that x=13 will suffer flu?
Answer: Pr[X=13] for Binomial (n=100, p=π=.04) = .0001
using online calculator https://istats.shinyapps.io/BinomialDist/ and setting n=100, x=13 and p =π = .04

The solution for binomial distribution probabilities makes use of the combinatorial and factorial tools introduced on pp 14-17.

X ~ Binomial Distribution (n, p=π)
n = number of independent Bernoulli trials
p = π = Probability that trial outcome is event of interest ("success")

A random variable X is said to follow a **Binomial (n, p=π)** distribution if it is the sum of n independent Bernoulli Distribution trials each with probability of "success" = p=π.

Value of X =	P[X = x] =
0	$(1-\pi)^n$
1	$n\,\pi\,(1-\pi)^{n-1}$
…	…
x	$\binom{n}{x}\pi^x(1-\pi)^{n-x}$
…	…
n	π^n

(1) **μ = Mean = E[X] = Statistical Expectation of X**
$\mu = np = n\pi$

(2) **σ^2 = Variance = Var[X] = E[$(X-\mu)^2$] = Statistical Expectation of $(X-\mu)^2$**
$\sigma^2 = np(1-p) = n\pi\,(1-\pi)$

Binomial Probability Distribution
How to Calculate Probabilities by Hand

It can be done! The **binomial formula** is the binomial distribution probability that you use to calculate a binomial probabilities by hand.

Question: What is the probability of X = x

In words: *What is the probability of* *<u>exactly</u>* *x events of "success" in n independent Bernoulli trials, each with probability of success = π*

Solution:

$$\Pr[X=x] = \binom{n}{x} \pi^{x} (1-\pi)^{n-x} = \left[\frac{n!}{x!\,(n-x)!} \right] \pi^{x} (1-\pi)^{n-x}$$

Another Look
Binomial Distribution (n, π)
X = sum of n ***independent*** Bernoulli(π) Trials Z

The n Bernoulli trials are $Z_1\ Z_2 \ldots Z_n$

- Each Z_i has possible values of 1 ("success") or 0 ("failure")

- Pr [$Z_i = 1$] = π and
  Pr [$Z_i = 0$] = $(1-\pi)$ for i=1, 2, …, n

The Binomial random variable is $X = Z_1 + Z_2 + \ldots + Z_n$. X is distributed Binomial(n, π)

$$X = \sum_{i=1}^{i=n} Z_i$$

For X ~ Binomial (n, π), the probability that X =x is given by the binomial formula:

$$\text{Probability}[X=x] = \left[\frac{n!}{x!\,(n-x)!}\right]\pi^x (1-\pi)^{n-x},$$

where X has possible values x = 0, 1, 2, …, n

Hack!! E [X] and the variance Var [X] is obtained by working with the fact that $\mathbf{X = Z_1 + Z_2 + \ldots + Z_n}$

E [X] is actually is $E[\sum_{i=1}^{n} Z_i] = n\pi$

Var [X] is actually $Var[\sum_{i=1}^{n} Z_i] = n\pi(1-\pi)$

9. Worked Examples: Binomial Probabilities

Setting: A roulette wheel lands on each of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 with probability = .10. Write down the expression for, and then solve the calculation of each of the following.

Example 1 - What is the probability of "5 or 6" exactly 3 times in 20 spins?

Solution

Step 1: Define **"event of success"**. Answer: An event of success is an outcome of either "5" or "6"

Step 2: Solve for probability of event, **p or** π. Answer: Probability [event] = π = .20

Step 3: Solve for **n=** number of independent Bernoulli trials. Answer: "20 spins" says that the number of trials is n = 20

Thus, X is distributed Binomial(n=20, π=.20)

Step 4: Put it all together. Answer: Want Pr [X = 3]

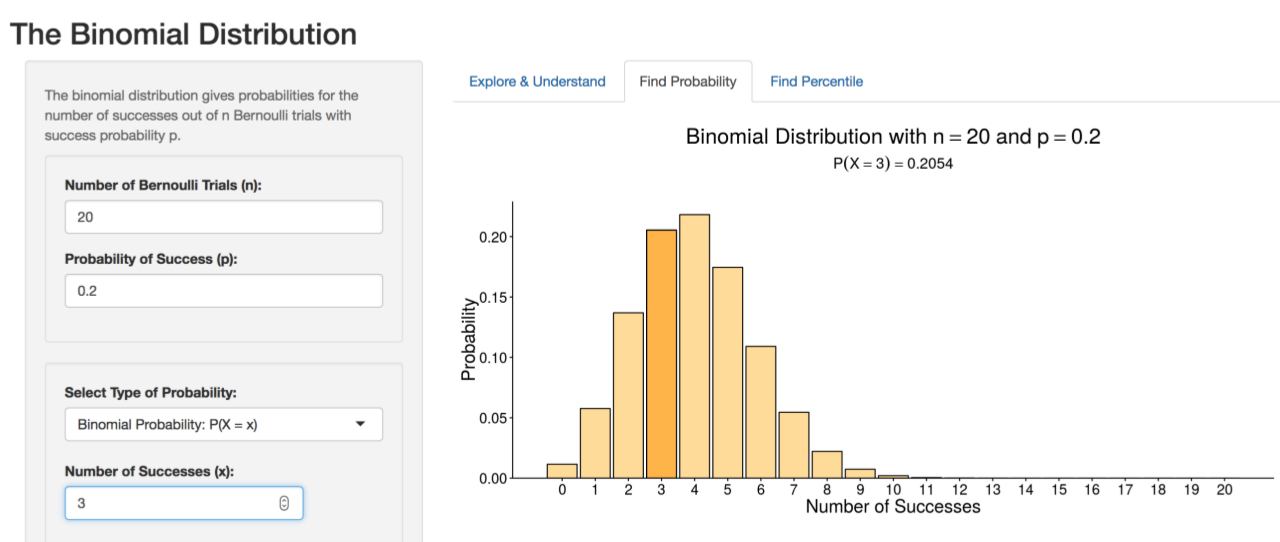
$$\Pr[X=3]=\binom{20}{3}[.20]^3[1-.20]^{20-3}$$

$$=\binom{20}{3}[.20]^3[.80]^{17}$$

$$=.2054$$

Artofstat.com

Artofstat.com > Online WebApps > Binomial Distribution > At top, click on tab: Find Probability



R using function dbinom( ) for exact probabilities

```
# Example 1 -  # Binomial(n, prob): Probability of exactly k events, Pr[X = k]
# Binomial(n=20, prob=.20), Prob[X=3] is:
# dbinom(x=,size=ntrials,prob=)

 # Plain with tags - Any order of input is allowed
dbinom(x=3, size=20, prob=0.20)
## [1] 0.2053641

 # Plain WITHOUT tags -  Inputs MUST be ordered x,size,prob
dbinom(3, 20, 0.20)
## [1] 0.2053641

# Fancy using paste("EXPRESSION" and round(STUFF, digits) to limit the number of digits shown to 4
paste("Pr [ Binom(20,.20) = 3] = ",round(dbinom(x=3,size=20,prob=0.20),4))
## [1] "Pr [ Binom(20,.20) = 3] =  0.2054"
```

Example 2 - What is the probability of "digit greater than 6" at most 3 times in 20 spins?

Solution

From the "setting", define **"event of success"**. Answer: An event of success is "7", "8" or "9"

Solve for probability of event. Answer: Probability [event] = π = .30

Solve for n= number of independent Bernoulli trials. Answer: "20 spins" says that the number of trials is n = 20

Thus, X is distributed Binomial(n=20, π=.30)

"At most 3 times" says we want Pr [$X \leq 3$]

Translation: **"At most 3 times"** is the same as saying "3 times or 2 times or 1 time or 0 times.
This is the same as saying **"less than or equal to 3 times"**

$$\Pr[X \leq 3] = \Pr[X=0] + \Pr[X=1] + \Pr[X=2] + \Pr[X=3]$$

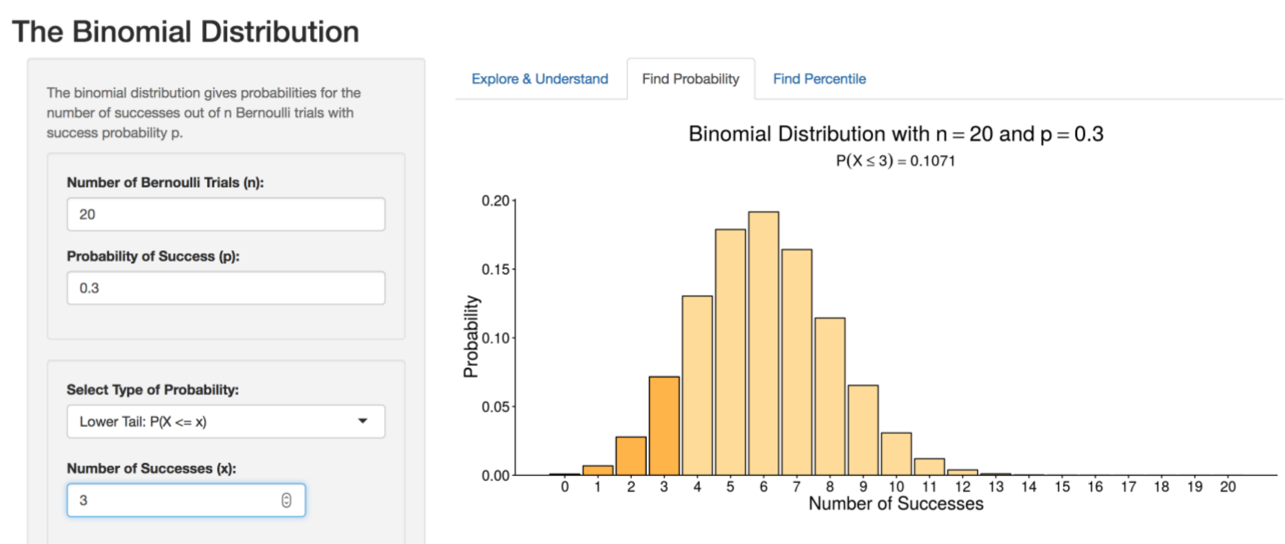
$$= \sum_{x=0}^{3} \left\{ \binom{20}{x} \right\} [.30]^x [.70]^{20-x}$$

$$= \binom{20}{0}[.30]^0[.70]^{20} + \binom{20}{1}[.30]^1[.70]^{19} + \binom{20}{2}[.30]^2[.70]^{18} + \binom{20}{3}[.30]^3[.70]^{17}$$

$$= .10709$$

Artofstat.com

Artofstat.com > Online WebApps > Binomial Distribution > At top, click on tab: Find Probability



R using function pbinom( ) for cumulative probabilities

```
# Example 2 -  # Binomial(n, prob): Probability of k or fewer events, Pr[X <= k]
# Binomial(n=20, prob=.30), Prob[X<=3] is:
# pbinom(q=,size=ntrials,prob=)

 # Plain with tags - Any order of input is allowed
pbinom(q=3, size=20, prob=0.30)
## [1] 0.1070868

 # Plain WITHOUT tags - Inputs MUST be ordered q, size, prob
pbinom(q=3, size=20, prob=0.30)
## [1] 0.1070868


# Fancy
paste("Pr [ Binom(20,.30) <= 3] = ",round(pbinom(q=3,size=20,prob=0.30),4))
## [1] "Pr [ Binom(20,.30) <= 3] =  0.1071"
```

HOMEWORK DUE Friday October 28, 2022
Question #2 of 5

A die will be rolled six times. What are the chances that, over all six rolls, the die lands neither ace (one dot showing) nor deuce (two dots showing) exactly 2 times?

HOMEWORK DUE Friday October 28, 2022
Question #3 of 5

Suppose that, in the general population, there is a 2% chance that a child will be born with a genetic anomaly. What is the probability that no congenital anomaly will be found among four random births?

HOMEWORK DUE Friday October 28, 2022
Question #4 of 5

Suppose it is known that, for a given couple, there is a 25% chance that a child of theirs will have a particular recessive disease. If they have three children, what are the chances that at least one of them will be affected?

HOMEWORK DUE Friday October 28, 2022
Question #5 of 5

Dear class – This is a harder question. Just give it a try and don't hesitate to look at the solutions to get you on your way! – cb

Suppose a quiz contains 20 true/false questions. You know the correct answer to the first 10 questions. You have no idea of the correct answer to questions 11 through 20 and decide to answer each using the coin toss method. Calculate the probability of obtaining a total quiz score of at least 85%.