5.
Introduction to Probability

Topics

Please Quiet Cell Phones and Pagers

Thank you.

1. Why We Need Probability

So far we've simply described the data at hand

- **histograms, frequency tables, plots;**
- **means, medians; and**
- **variance, SD, SE, MAD, MADM**

And we've considered the data as a sample from a population

- **We asked if the sample is representative of its source population.**
- **We learned what simple random sampling is is.**
- **We learned methods for obtaining probability samples. Thus,**

We already have some intuition for probability.

We need probability for inferential statistics.

Suppose we relax our previous view of the data as a "given". Reasonably, we might now ask

- **Where did the data come from? (See, *Unit 4 – Populations and Samples*)**

- **If the population source is known, what are the chances of obtaining a particular value? A particular sample of values?**

<u>Chance</u> - The notion of "chance" is described using concepts of probability and events. We recognize this in such familiar questions as:

- **What are the chances that a diseased person will obtain a test result that indicates the same? (*Sensitivity*)**

- **If the experimental treatment has no effect, how likely is the observed discrepancy between the average response of the controls and the average response of the treated? (*Clinical Trial)***

2. Definition Probability Model

Setting -

- **The source population is assumed known**
- **The sample is assumed to be a simple random sample**

Question -

> **If the available sample is representative of the source population, what are the "chances" of obtaining the observed values?**

This is a "frequentist" approach to probability. It is not the only approach.

Alternative approaches -

- **Bayesian - "This is a fair coin. So it will lands "heads" with probability 1/2**
- **Frequentist – "In 100 tosses, this coin landed heads 48 times"**
- **Subjective - "This is my lucky coin"**

In this unit, we consider the frequentist approach

Probabilities and probability distributions are nothing more than extensions of the ideas of relative frequency and histograms, respectively:

> **Ignoring certain mathematical details, a probability distribution consists of:**
>
> 1. **The possible values a random value can assume, together with**
> 2. **The probabilities with which these values are assumed.**

Example -

- **Suppose the universe of all university students is known to include men and women in the ratio 53:47.**
- **Consider the random variable, X = gender of an individual student**
  For convenience, we'll set

  X = 0 when the student is "male"
  X = 1 when the student if "female"

- **We have what we need to define a probability distribution:**

Possible value of X is represented as x	**Probability [X = x]**
0 = male **1 = female** ↑ **Enumeration of all possible outcomes must be "exhaustive"**	**0.53** **0.47** ↑ **Therefore, check that these probabilities add up to 100% or a total of 1.00.**

More Formally

Probability can be defined as

- **the chance of observing a particular outcome, or**
- **the likelihood of an event.**

- **The concept of probability assumes a <u>stochastic</u> or <u>random</u> process: i.e., the outcome is not predetermined – there is an element of chance.**

- **In discussing probabilities, we assign a <u>numerical weight</u> or "probability" to each outcome which measures the likelihood of its occurrence.**

<u>Notation</u> -
The probability of outcome O_i is denoted $P(O_i)$

- **The probability of each outcome is between 0 and 1, inclusive:**

$$0 <= P(O_i) <= 1 \text{ for all } i$$

- **The probabilities of all possible elementary outcomes sum to 1**

$$\sum_{\text{all possible outcomes } O_i} P(O_i) = 1 \qquad \textbf{(\textit{something happens})}$$

- **An event E might be a one or several outcomes, O. If an event, E, is certain,**

$$P(E) = 1$$

- **If an event, E, is impossible,**

$$P(E) = 0$$

Some More Formal Language

1. **A *probability model* is the set of assumptions used to assign probabilities to each outcome in the sample space.**

   The *sample space* is the universe, or collection, of all possible outcomes.

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

3. **To *define a probability distribution*, we make an assumption (the probability model) and use this to assign a probability to each outcome.**

4. **E.g. all outcomes are equally likely (*uniform probability model)***

3. The "Equally Likely" Setting
Introduction to Probability Calculations

<u>An "Equally Likely" Setting</u> –

Rolling a die.

There are 6 possible outcomes: {1, 2, 3, 4, 5, 6}. The probability of each is:

P(1) = 1/6
P(2) = 1/6
. . .
<u>P(6) = 1/6</u>
Sum 1

<u>An "Equally Likely" Setting</u> –

Tossing a coin. There are 2 possible outcomes" {H, T}.

Probability Distribution:

O_i	$P(O_i)$
H	.5
T	.5
Sum	1

An "Equally Likely" Setting –

The set of all possible samples of size n that can be taken, with replacement, from a population of size N. E.g., for N=3, n=2:

Sample Space:
S = { (1,1), (1,2), (1,3), (2,2), (2,1), (2,3), (3,1), (3,2), (3,3) }

Probability Model:
Assumption: equally likely outcomes, with $N^n = 3^2 = 9$ outcomes

Probability Distribution:

O_i	$P(O_i)$
(1,1)	1/9
(1,2)	1/9
…	…
(3,3)	1/9
Sum	1

An "Equally Likely" Setting –

Toss 2 coins
Set of all possible outcomes: S = {HH, HT, TH, TT}

Probability Distribution:

O_i	$P(O_i)$
HH	.25
HT	.25
TH	.25
TT	.25
Sum	1

<u>Elementary Outcomes "O" versus (Composite) Events "E"</u> –

Recall that we can also define composite events of interest, and compute their probabilities. Such composite events are each composed of a set of elementary outcomes from the sample space:

Event, E	Qualifying Outcomes, O	$P(E_i)$
E_1: 2 heads	{HH}	.25
E_2: Just 1 head	{HT, TH}	.50
E_3: 0 heads	{TT}	.25
E_4: Both the same	{HH, TT}	.50
E_5: At least 1 head	{HH, HT, TH}	.75

The probability of each composite event "E" is determined by summing the probabilities of each of the elementary outcomes "O" that make up the event.

<u>Calculation of a Composite Event in An "Equally Likely" Setting</u> –

Recall the set of all possible samples of size n=2 that can be taken, with replacement, from a population of size N=3:

Sample Space:
S = { (1,1), (1,2), (1,3), (2,2), (2,1), (2,3), (3,1), (3,2), (3,3) }

Probability Model:
Each outcome is equally likely and is observed with probability 1/9

Probability Distribution:

O_i	$P(O_i)$
(1,1)	1/9
(1,2)	1/9
…	…
(3,3)	1/9
Sum	1

Suppose we are interested in the event that subject "2" is in our sample. Thus,

E: subject 2 is in the sample.

S={ (1,1), <u>(1,2)</u>, (1,3), <u>(2,1)</u>, <u>(2,2)</u>, <u>(2,3)</u>, (3,1), <u>(3,2)</u>, (3,3) }

By the "underlines", we notice that the event of interest occurs in 5 of the samples

Pr{ E } = Pr { <u>(1,2)</u>, <u>(2,1)</u>, <u>(2,2)</u>, <u>(2,3)</u>, <u>(3,2)</u> } = 1/9 + 1/9 + 1/9 + 1/9 + 1/9 = 5/9 = .56

We have what we need to define

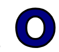
1. **Sample Space**

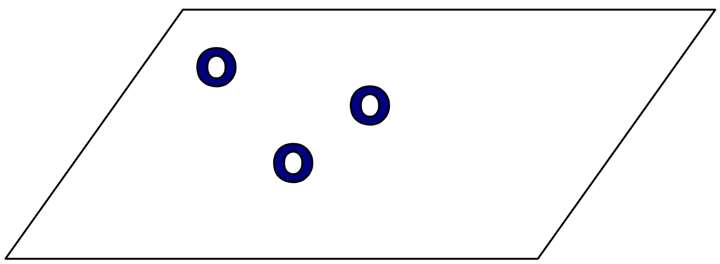
2. **Elementary Outcomes or Sample Points**

3. **Events**

3a. Sample Space, Elementary Outcomes, Event

<u>Population or Sample Space</u> -

The population or sample space is defined as the set of all possible outcomes  **of a random variable.**



Example –

The sexes of the first and second of a pair of twins.
The population or sample space of all possible outcomes “O” is therefore:

{ boy, boy }
{ boy, girl }
{ girl, boy }
{ girl, girl },

The first sex is that of the first individual in the pair while the second sex is that of the second individual.

***NOTE*: This random variable is a little different than the random variables described so far. Here, specification of one outcome requires two pieces of information (the sex of the first individual and the sex of the second individual) instead of one piece of information. It is an example of a bivariate random variable.**

<u>Elementary Outcome (Sample Point), O</u> –

One sample point corresponds to each possible outcome "O" of a random variable.

Example –

For the twins example, a single sample point is:

{ boy, girl }

There are three other single sample points: { girl, boy }, { girl, girl }, and { boy, boy }.

<u>Event</u> -

An event "E" is a collection of individual outcomes "O".

Notice that the individual outcomes or sample points are denoted O_1, O_2, …
while events are denoted E_1, E_2, …

Example –

Consider again the random variable defined as the sexes of the first and second individuals in a pair of twins. Consider, in particular, the event defined "boy". The set of outcomes defined by this event is the collection of all possible ways in which a pair of twins can include a boy. There are three such ways:

{ boy, boy }
{ boy, girl }
{ girl, boy }

Probability -

The probability of an event is the relative frequency with which at least one outcome of the event occurs. If an event is denoted by E, then the probability that the event E occurs is written as P(E).

Example –

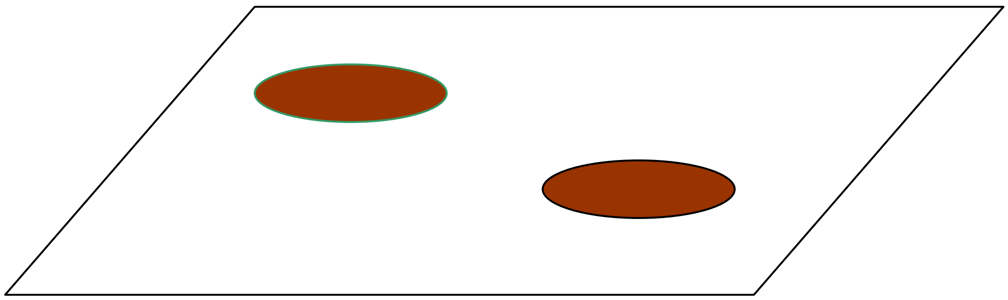
For the twins example, we might be interested in three events: E_1 = "two boys", E_2 = "two girls", E3 = "one boy, one girl". Assume that the chances of each sex are 1/2 for both the first and second twin and that the sex of the second twin is not determined in any way by the sex of the first twin. A table summarizing the individual probabilities of these events is:

Event	**Description**	**Probability of Event**
E_1	"two boys" {boy,boy}	Prob{1^{st}=boy}Prob{2^{nd}=boy} = (1/2) x (1/2) = 1/4 = 0.25
E_2	"two girls" {girl,girl}	Prob{1^{st}=girl}Prob{2^{nd}=girl} = (1/2) x (1/2) = 1/4 = 0.25
E_3	"one boy, one girl" {girl,boy} OR {boy,girl}	Prob{1^{st}=girl}Prob{2^{nd}=boy} + Prob{1^{st}=boy}Prob{2^{nd}=girl} = (1/2) x (1/2) = 1/4 = 0.25 + (1/2) x (1/2) = 1/4 = 0.25 = 0.50

3b. Types of Events

<u>Mutually Exclusive ("nothing in common")</u> -

Two events are mutually exclusive if they have no outcome in common.

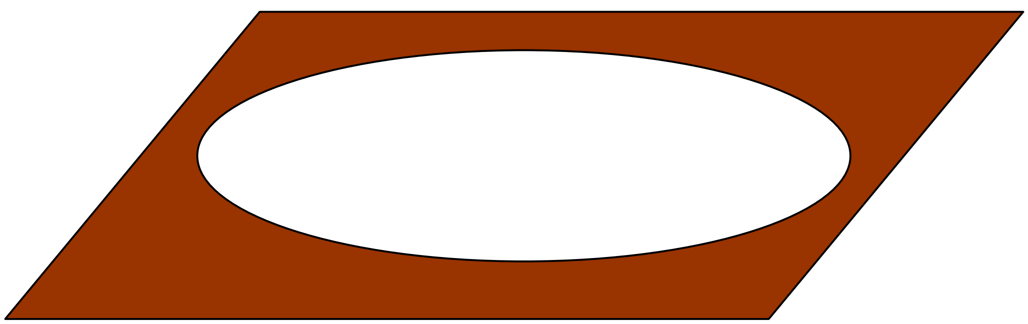


Example –

For the twins example, the events E1 = "two boys" and E2 = "two girls" are mutually exclusive.

<u>Complement ("opposite")</u> -

The complement of an event E is the event consisting of all outcomes in the population or sample space that are *<u>not</u>* contained in the event E. The complement of the event E is denoted using a superscript c, as in E^c .



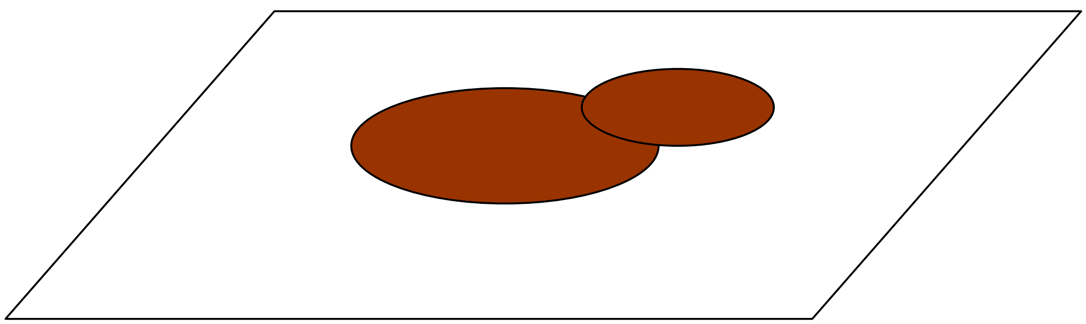
Example –

For the twins example, consider the event E_1 = "two boys". The complement of the event E_1 is

$$E_1^{\,c} = \{ \text{boy, girl} \}, \{ \text{girl, boy} \}, \{ \text{girl, girl} \}$$

<u>Union, A or B ("either or")</u> -

The union of two events, say A and B, is another event which contains those outcomes which are contained either in A or in B. The notation used is $A \cup B$.



Example –

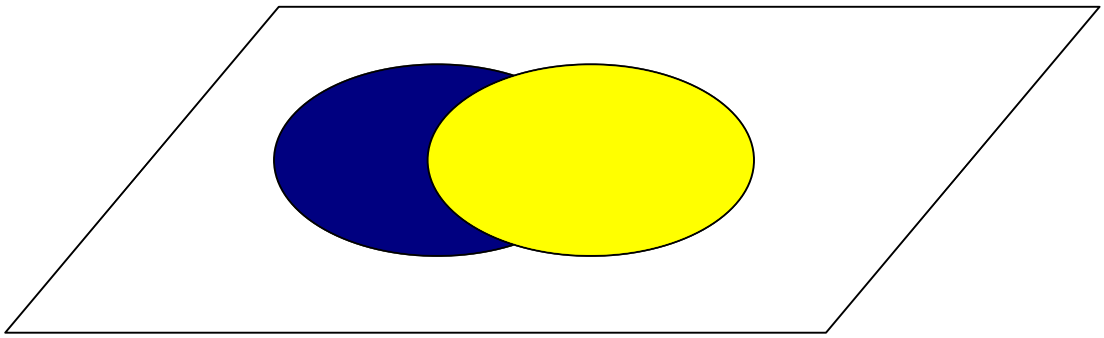
For the twins example suppose event "A" is defined { boy, girl } and event "B" is defined { girl, boy }. The union of events A and B is:

$$A \cup B = \{ \text{boy, girl} \}, \{ \text{girl, boy} \}$$

<u>Intersection, A and B ("only the common part")</u> -

Note: Still need to insert correct graphics here to show the green that is intersection …

The intersection of two events, say A and B, is another event which contains only those outcomes which are contained in both A and in B. The notation used is A $\cap$ B.



Example –

For the twins example, consider next the events E_1 defined as "having a boy" and E_2 defined as "having a girl". Thus,

E_1 = { boy, boy }, { boy, girl }, { girl, boy }
E_2 = { girl, girl }, { boy, girl }, { girl, boy }

These two events do share some common outcomes.
The intersection of E_1 and E_2 is "have a boy and a girl":

$E_1 \cap E_2$ = { girl, boy }, { boy, girl }

4. Conditional Probability

Recall the twins example –
The sex of the second twin was not determined in any way by the sex of the first twin. This is what is meant by independence.

The idea of conditional probability -
The occurrence of a first event alters, at least in part, the occurrence of a second event.

<u>Independence</u> -

Two events A and B are independent if the chances, or likelihood, of one event is in no way related to the likelihood of the other event. When A and B are independent:

$$P(A \text{ and } B) = P(A)\, P(B)$$

Example –

Event A = "a woman is hypertensive"
Event B = "her mother-in-law is hypertensive".

The assumption of independence seems reasonable since the two women are not genetically related. If the probability of being hypertensive is 0.07 for each woman, then the probability that BOTH the woman and her mother-in-law are hypertensive is:

$$P(A \text{ and } B) = P(A) \times P(B) = 0.07 \times 0.07 = 0.0049$$

<u>Dependence</u> -

Two events are dependent if they are not independent.

The probability of both occurring depends on the outcome of at least one of the two.

Two events A and B are dependent if the probability of one event is related to the probability of the other:

$$P(A \text{ and } B) \neq P(A)\, P(B)$$

Example -

An offspring of a person with Huntington's Chorea has a 50% chance of contracting Huntington's Chorea.

Let Event A = "parent has Huntington's Chorea"
Let Event B = "offspring has Huntington's Chorea".

Without knowing anything about the parent's family background, suppose the chances that the parent has Huntington's Chorea is 0.0002.

Suppose further that, if the parent does not have Huntington's Chorea, the chances that the offspring has Huntington's Chorea is 0.000199.

However, if the parent has Huntington's Chorea, the chances that the offspring has Huntington's Chorea jumps to 0.50. Thus, the chances that both have Huntington's Chorea is not simply 0.0002 x 0.000199. That is,

$$P(A \text{ and } B) \neq P(A) \times P(B)$$

It is possible to calculate P(A and B), but this requires knowing how to relate P(A and B) to conditional probabilities. This is explained below.

BE CAREFUL !!

The concepts of *mutually exclusive* and *independence* are distinct and in no sense equivalent. To see this:

For A and B *independent*:
Pr(A and B) = P(A) P(B)

For A and B *mutually exclusive*:
Pr(A and B) = Pr (empty event) = 0

Conditional Probability ("what happened first is assumed") -

Conditional probability refers to the probability of an event, given that another event is known to have occurred. This measure is useful when it is of interest to assess how dependent two events are, relative to each other.

The *conditional probability* that event B has occurred given that event A has occurred is denoted P(B|A) and is defined

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

provided that $P(A) \neq 0$

***Hint* - When thinking about conditional probabilities, think of the two events A and B being spaced apart, either in time or space.**

Example - Huntington's Chorea

- **The conditional probability that an offspring has Huntington's Chorea given a parent has Huntington's Chorea is 0.50.**
- **It is also known that the parent has Huntington's Chorea with probability 0.0002.**
- **Consider**

A = event that parent has Huntington's Chorea
B = event that offspring has Huntington's Chorea

- **Thus, we know**

Pr (A) = 0.0002
Pr (B|A) = 0.5

- **With these two "knowns" we can solve for the probability that both parent and child will have Huntington's Chorea. This is P(A and B)**

- $P(B|A) = \dfrac{P(A \text{ and } B)}{P(A)}$ **is the same as saying** $P(A \text{ and } B) = P(A)\,P(B|A)$

- **Probability both parent and child have Huntington's Chorea**

= P(A and B)

= P(A) P(B|A)

= 0.0002 x 0.5

= 0.0001

This type of staged probability calculation is often how probabilities of sequences of events are calculated.

4a. Theorem of Total Probabilities
A tool for calculating probabilities of sequences of events

The previous example sets the stage.

Now we'll play a game. The game has two steps.

<u>Game</u> -

Step 1: **Choose one of two games to play: G_1 or G_2**
G_1 is chosen with probability = 0.85
G_2 is chosen with probability = 0.15 (*notice that probabilities sum to 1)*

Step 2: **Given that you choose a game, play it and hope to win**
G_1 yields win with conditional probability $P(\text{win}|G_1) = 0.01$
G_2 yields win with conditional probability $P(\text{win}|G_2) = 0.10$

<u>What is the overall probability of a win, Pr(win)?</u>

Hint – Think of all the distinct and mutually ways in which a win could occur and sum their associated probabilities. There are only two such scenarios, and each has two steps.

$$\Pr(\text{win}) = \underbrace{\Pr[G_1 \text{ chosen}]}_{\text{step 1}} \; \underbrace{\Pr[\text{win}|G_1]}_{\text{step 2}} + \underbrace{\Pr[G_2 \text{ chosen}]}_{\text{step 1}} \; \underbrace{\Pr[\text{win}|G_2]}_{\text{step 2}}$$

scenario #1 **scenario #2**

$$= (.85)\ (.01) + (.15)\ (.10)$$

$$= 0.0235$$

This intuition has a name –
The Theorem of Total Probabilities.

Theorem of Total Probabilities

Suppose that a sample space S can be partitioned (carved up into bins) so that S is actually a union that looks like

$$S = G_1 \cup G_2 \cup \ldots \cup G_K$$

If you are interested in the overall probability that an event "E" has occurred, this is calculated

$$P[E] = P[G_1]P[E|G_1] + P[G_2]P[E|G_2] + \ldots + P[G_K]P[E|G_K]$$

provided the conditional probabilities are known.

Example – The lottery game just discussed.

G_1 = Game #1
G_2 = Game #2
E = Event of a win.

<u>Applications of the Theorem of Total Probabilities</u> –

We'll see this again in this course and also in BioEpi640

- ♣ **Diagnostic Testing**
- ♣ **Survival Analysis**

4b. Bayes Rule

At this point, we have two handy tools:

1. P(A and B) = P(A) P(B|A) = P(B) P(A|B)

This provides us with three ways of determining a joint probability

2. $P[E] = P[G_1]P[E|G_1] + P[G_2]P[E|G_2] + ... + P[G_K]P[E|G_K]$

This provides us with a means of calculating an overall probability when things happen in a "sequence" kind of way.

Putting these together provides us with a third useful tool, that of Bayes Rule

Bayes Rule

Suppose that a sample space S can be partitioned (carved up into bins) so that S is actually a union that looks like

$$S = G_1 \cup G_2 \cup \dots \cup G_K$$

If you are interested in calculating $P(G_i|E)$, this is calculated

$$P[G_i|E] = \frac{P(E|G_i)P(G_i)}{P[G_1]P[E|G_1] + P[G_2]P[E|G_2] + ... + P[G_K]P[E|G_K]}$$

provided the conditional probabilities are known.

5. Introduction to the Concept of Expected Value

From playing the lottery, we have a feel for "what are the likely winnings".

Consider the following example –

IF $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

THEN "likely winning" = [$1](0.50) + [$5](0.25) +[$10](0.15) + [$25](0.10)

= $5.75

Other names for this intuition are

- **Expected winnings**
- **"Long range average"**
- **Statistical expectation**

Statistical Expectation for a Discrete Random Variable is the Same Idea.

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 random variable X has ***statistical expectation E[X]***$=\mu$

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

Example –

In the "likely winnings" example, μ = $5.75

6. Example – The Bernoulli Distribution

The Bernoulli Distribution is an example of a discrete probability distribution. It is an appropriate tool in the analysis of proportions and rates.

<u>Recall the coin toss.</u>

"50-50 chance of heads" can be re-cast as a random variable. Let

Z = random variable representing outcome of one toss, with

Z = 1 if "heads"
0 if "tails"

π= Probability [coin lands "heads" }. Thus,

7. = Pr [Z = 1]

<u>We have what we need to define a probability distribution.</u>

Enumeration of all possible outcomes - outcomes are mutually exclusive - outcomes are exhaust all possibilities	**1** **0**
Assiociated probabilities of each - each probability is between 0 and 1 - sum of probabilities totals 1	**<u>Outcome</u>** **<u>Pr[outcome]</u>** **0** **$(1 - \pi)$** **1** **π**

<u>In epidemiology, the Bernoulli might be a model for the description of ONE individual (N=1):</u>
This person is in one of two states. He or she is either in a state of:

1) **"event" with probability π**
2) **"non event" with probability $(1-\pi)$**

The description of the likelihood of being either in the "event" state or the "non-event" state is given by the <u>Bernoulli distribution</u>

Bernoulli Distribution

Suppose Z can take on only two values, 1 or 0, and suppose:

Probability [Z = 1] = π
Probability [Z = 0] = $(1-\pi)$

This gives us the following expression for the likelihood of Z=z.

Probability [Z = z] = $\pi^{z} \ (1-\pi)^{1-z}$ for z=0 or 1.

Expected value of Z is E[Z] = π
Variance of Z is Var[Z] = $\pi\,(1-\pi)$

<u>Example</u>: Z is the result of tossing a coin once. If it lands "heads" with probability = .5, then π = .5.
Later we'll see that individual Bernoulli distributions are the basis of describing patterns of disease occurrence in a logistic regression analysis.

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

Mean of Z = μ = π

The mean of Z is represented as E[Z].

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

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

$$= [0]\Pr[Z = 0] + [1] + \Pr[Z = 1]$$

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

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:

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

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

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

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

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

7. Example – Binomial Distribution

<u>From the Bernoulli to the Binomial….</u>

The outcome of a Binomial can be thought of as the <u>net</u> number of successes in a set of independent Binomial trials.

We'd like to know the probability of X=x successes in N separate Bernoulli trials, but we don't care about the order of the successes and failures among the N separate trials.

E.g.

- **What is the probability that 2 of 6 graduate students are female?**
- **What is the probability that of 100 infected persons, 4 will die within a year?**

<u>Steps in Calculating a Binomial Probability</u>
N = # of independent Binomial trials
π = **common probability of "event" accompanying each of the N trials**

$\pi^x (1-\pi)^{N-x}$ = **Probability of one "representative" sequence that yields a net of "x" events and "N-x" non-events.**

$\binom{N}{x}$ = **Count of the number of such sequences that yields a net of "x" events and "N-x" non-events. Recall that this is a combinatorial.**

<u>Probability [N trials yields x events] = (Count of sequences) (Pr[one sequence])</u>

$$= \binom{N}{x} \pi^x (1-\pi)^{N-x}$$

Binomial Distribution

Suppose each individual Z_i can take on only two values, 1 or 0, and suppose:

Probability [$Z_i = 1$] = π for every individual
Probability [$Z_i = 0$] = $(1-\pi)$ for every individual

Now consider the sample of size N. <u>Think of this as N trials</u>:

Among N trials, what are the chances of x events ? ($\sum_{i=1}^{N} Z_i = x$)?

The answer is the product of 2 terms.

1st term: # ways to choose x from a collection of N
2nd term: Probability [$(Z_1=1) \ldots (Z_x=1)\ (Z_{x+1}=0) \ldots (Z_N=0)$]

This gives us the following expression for the likelihood of $\sum_{i=1}^{N} Z_i = x$:

$$\text{Probability}\left[\sum_{i=1}^{N} Z_i = x\right] = \binom{N}{x} \pi^x (1-\pi)^{N-x} \quad \text{for } x=0, \cdots, N$$

Expected value is $E\left[\sum_{i=1}^{N} Z_i = x\right] = N\pi$

Variance is $\text{Var}\left[\sum_{i=1}^{N} Z_i = x\right] = N\pi(1-\pi)$

$\binom{N}{x}$ **= # ways to choose X from N =** $\dfrac{N!}{x!(N-x)!}$

where N! = N(N-1)(N-2)(N-3) ⋯ (4)(3)(2)(1) and is called the <u>factorial.</u>

<u>The Binomial is a description of a SAMPLE (Size = N):</u>
Some experience the event. The rest do not.

Your Turn

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 the calculation of the following.

#1. The probability of "5 or 6" exactly 3 times in 20 spins.

#2. The probability of "digit greater than 6" at most 3 times in 20 spins.

#1. **"Event" is outcome of either "5" or "6"**
Pr[event] = π = .20
N = 20
X is distributed Binomial(N=20, π=.20)

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

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

$$= .2054$$

#2. **"Event" is outcome of either "7" or "8" or "9"**
Pr[event] = π = .30
N = 20
X is distributed Binomial(N=20, π=.30)

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

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

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

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

$$= .10709$$

Appendix
Some Elementary Laws of Probability

A. Definitions:

1) **one sample point corresponds to each possible outcome of a random variable**

2) **the sample space or population consists of all sample points**

3) **a group of events is said to be exhaustive if their union is the entire sample space or population. For the variable SEX, the events "male" and "female" exhaust all possibilities.**

4) **two events A and B are said to be mutually exclusive or disjoint if their intersection is the empty set. One cannot be simultaneously "male" and "female".**

5) **two events A and B are said to be complementary if they are both mutually exclusive and exhaustive**

6) **The events E1, E2, ..., En are said to partition the sample space or population if:**

   (i) Ei is contained in the sample space
   (ii) The event (Ei and Ej) = empty set for all i = j;
   (iii) The event (E1 or E2 or ... or En) is the entire sample space or population.

   In words: E1, E2, ..., En are said to partition the sample space if they are pairwise mutually exclusive and together exhaustive

7) **If the events E1, E2, ..., En partition the sample space such that P(E1) = P(E2) = ... P(En):**

   (i) P(Ei) = 1/n, for all i=1,...,n
   (ii) The events E1, E2, ..., En are equally likely

8) **For any event E in the sample space: 0 <= P(E) <= 1**

9) **P(empty event) = 0. The empty event is also called the null set.**

10) **P(sample space) = P(population) = 1**

11) **P(E) + P(Ec) = 1**

B. Addition of Probabilities -

1) **If events A and B are mutually exclusive:**
   (i) P(A or B) = P(A) + P(B)
   (ii) P(A and B) = 0

2) **More generally:**
   P(A or B) = P(A) + P(B) - P(A and B)

3) **If events E1, ..., En are all pairwise mutually exclusive:**
   P(E1 or ... or En) = P(E1) + ... + P(En)

C. Conditional Probabilities -

1) **P(B|A) = P(A and B) / P(A)**

2) **If A and B are independent:**
   P(B|A) = P(B)

3) **If A and B are mutually exclusive:**
   P(B|A) = 0

4) **P(B|A) + P(Bc|A) = 1**

5) **If P(B|A) = P(B|Ac):**
   Then the events A and B are independent

D. Theorem of Total Probabilities -

Let E1, ..., Ek be mutually exclusive events that partition the sample space. The unconditional probability of the event A can then be written as a weighted average of the conditional probabilities of the event A given the Ei; i=1, ..., k:

P(A) = P(A|E1)P(E1) + P(A|E2)P(E2) + ... + P(A|Ek)P(Ek)

E. Bayes Rule -

If the sample space is partitioned into k disjoint events E1, ..., Ek, then for any event A:

$$P(Ej|A) = \frac{P(A|Ej)\,P(Ej)}{P(A|E1)P(E1) + P(A|E2)P(E2) + \ldots + P(A|Ek)P(Ek)}$$