Class 13

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
Agenda:

Recap Chapter 6.4 - 6.6

Lecture Chapter 7.1- 7.3

Review Chapters 4.3 - 5.6
Recap Chapter 6.4 - 6.6
Assume that IQ scores are normally distributed with a mean $\mu$ of 100 and a standard deviation $\sigma$ of 16.

If a person is picked at random, what is the probability that his or her IQ is between 100 and 115?

i.e. $P(100 < x < 115)$ ?
6: Normal Probability Distributions
6.4 Applications of Normal Distributions

IQ scores normally distributed
\( \mu = 100 \) and \( \sigma = 16 \).

\[ P(100 < x < 115) \]

\[ z = \frac{x - \mu}{\sigma} \]

\[ z_1 = \frac{x_1 - \mu}{\sigma} = \frac{100 - 100}{16} = 0 \]

\[ z_2 = \frac{x_2 - \mu}{\sigma} = \frac{115 - 100}{16} = 0.94 \]

Figures from Johnson & Kuby, 2008.
6: Normal Probability Distributions

6.4 Applications of Normal Distributions

IQ scores normally distributed
\( \mu = 100 \) and \( \sigma = 16 \).

\[ P(100 < x < 115) \]

Now we can use the table.

\[ P(100 < x < 115) = P(0 < z < 0.94) \]

Figures from Johnson & Kuby, 2008.
6: Normal Probability Distributions
6.5 Notation

We can use the table in reverse.

Example:
Let $\alpha=0.05$. Find $P(z > z(0.05)) = 0.05$.

Same as finding $P(0 < z < z(0.05)) = 0.45$.

Figures from Johnson & Kuby, 2008.
6: Normal Probability Distributions

6.5 Notation

Example:

Same as finding

\[ P(0 < z < z(0.05)) = 0.45. \]

The table shows the area corresponding to the z-value of 0.05, which is 0.4500.

<table>
<thead>
<tr>
<th>z</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
<th>0.07</th>
<th>0.08</th>
<th>0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0000</td>
<td>0.0040</td>
<td>0.0080</td>
<td>0.0120</td>
<td>0.0160</td>
<td>0.0199</td>
<td>0.0239</td>
<td>0.0279</td>
<td>0.0319</td>
<td>0.0359</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0398</td>
<td>0.0438</td>
<td>0.0478</td>
<td>0.0517</td>
<td>0.0557</td>
<td>0.0596</td>
<td>0.0636</td>
<td>0.0675</td>
<td>0.0714</td>
<td>0.0753</td>
</tr>
<tr>
<td>0.2</td>
<td>0.0793</td>
<td>0.0832</td>
<td>0.0871</td>
<td>0.0910</td>
<td>0.0948</td>
<td>0.0987</td>
<td>0.1026</td>
<td>0.1064</td>
<td>0.1103</td>
<td>0.1141</td>
</tr>
<tr>
<td>0.3</td>
<td>0.1179</td>
<td>0.1217</td>
<td>0.1255</td>
<td>0.1293</td>
<td>0.1331</td>
<td>0.1368</td>
<td>0.1406</td>
<td>0.1443</td>
<td>0.1480</td>
<td>0.1517</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1554</td>
<td>0.1591</td>
<td>0.1628</td>
<td>0.1664</td>
<td>0.1700</td>
<td>0.1736</td>
<td>0.1772</td>
<td>0.1808</td>
<td>0.1844</td>
<td>0.1879</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\text{Second Decimal Place in } z &
\end{align*}
\]

\[
\begin{align*}
1.0 & \quad 0.00 & \quad 0.01 & \quad 0.02 & \quad 0.03 & \quad 0.04 & \quad 0.05 & \quad 0.06 & \quad 0.07 & \quad 0.08 & \quad 0.09 \\
1.5 & \quad 0.4332 & \quad 0.4345 & \quad 0.4357 & \quad 0.4370 & \quad 0.4382 & \quad 0.4394 & \quad 0.4406 & \quad 0.4418 & \quad 0.4429 & \quad 0.4441 \\
1.6 & \quad 0.4452 & \quad 0.4463 & \quad 0.4474 & \quad 0.4484 & \quad 0.4495 & \quad 0.4505 & \quad 0.4515 & \quad 0.4525 & \quad 0.4535 & \quad 0.4545 \\
1.7 & \quad 0.4554 & \quad 0.4564 & \quad 0.4573 & \quad 0.4582 & \quad 0.4591 & \quad 0.4599 & \quad 0.4608 & \quad 0.4616 & \quad 0.4625 & \quad 0.4633 \\
1.8 & \quad 0.4641 & \quad 0.4649 & \quad 0.4656 & \quad 0.4664 & \quad 0.4671 & \quad 0.4678 & \quad 0.4686 & \quad 0.4693 & \quad 0.4699 & \quad 0.4706 \\
1.9 & \quad 0.4713 & \quad 0.4719 & \quad 0.4726 & \quad 0.4732 & \quad 0.4738 & \quad 0.4744 & \quad 0.4750 & \quad 0.4756 & \quad 0.4761 & \quad 0.4767 \\
\end{align*}
\]

Figures from Johnson & Kuby, 2008.
Approximate binomial probabilities with normal areas. Use a normal with $\mu = np$, $\sigma^2 = np(1 - p)$

$\mu = (14)(.5) = 7$

$\sigma^2 = (14)(.5)(1 - .5) = 3.5$

\[
P(x) = \frac{n!}{x!(n-x)!} p^x (1 - p)^{n-x}
\]
6: Normal Probability Distributions

6.6 Normal Approximation of the Binomial Distribution

From the binomial formula

\[ P(x = 4) = \frac{14!}{4!(14 - 4)!} (0.5)^4 (1 - 0.5)^{14-4} \]

\[ P(x = 4) = 0.061 \]

\[ P(1.34 < z < 1.87) = 0.0594 \]

\[ P(0 < z < 1.87) \]

\[ P(0 < z < 1.34) \]

\[ P(1.34 < z < 1.87) \]

From the Normal Distribution

\[ P(3.5 < x < 4.5) \]

\[ \mu = 7, \sigma^2 = 3.5 \]

\[ z_1 = \frac{x_1 - \mu}{\sigma} = \frac{3.5 - 7}{\sqrt{3.5}} = -1.87 \]

\[ z_2 = \frac{x_2 - \mu}{\sigma} = \frac{4.5 - 7}{\sqrt{3.5}} = -1.34 \]

\[ n=14, p=1/2 \]

Rowe, D.B.
6: Normal Probability Distributions

Questions?

Homework: Chapter 6 # 7, 9, 13, 17, 29, 31, 33, 41, 43, 57, 59, 71, 89
Lecture Chapter 7.1- 7.3
Chapter 7: Sample Variability

Daniel B. Rowe, Ph.D.

Department of Mathematics, Statistics, and Computer Science
7: Sample Variability
7.2 Sampling Distributions

When we take a random sample \( x_1, \ldots, x_n \) from a population, one of the things that we do is compute the sample mean \( \bar{x} \). The value of \( \bar{x} \) is not \( \mu \). Each time we take a random sample of size \( n \), we get a different set of values \( x_1, \ldots, x_n \) and a different value for \( \bar{x} \).
Recall: When we take a sample of data $x_1, \ldots, x_n$ from a population, then compute an estimate of a parameter it is called a sample statistic. i.e. $\bar{x}$ for $\mu$.

**Sampling Distribution of a sample statistic:** The distribution of values for a sample statistic obtained from repeated samples, all of the same size and all drawn from the same population.
7: Sample Variability
7.2 Sampling Distributions

The relationship between the sample mean and the population mean.

Assume that we have a population of items with population mean $\mu$ and population standard deviation $\sigma$.

If we take a random sample of size $n$ and compute sample mean.

The collection of all possible means is called the sampling distribution of the sample mean.
7: Sample Variability
7.2 Sampling Distributions

\( N = 5 \) balls in bucket, select \( n = 1 \) \textit{with} replacement.

Population data values: 0, 2, 4, 6, 8.

<table>
<thead>
<tr>
<th>( x )</th>
<th>( P(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1/5</td>
</tr>
<tr>
<td>2</td>
<td>1/5</td>
</tr>
<tr>
<td>4</td>
<td>1/5</td>
</tr>
<tr>
<td>6</td>
<td>1/5</td>
</tr>
<tr>
<td>8</td>
<td>1/5</td>
</tr>
</tbody>
</table>

5 possible values
7: Sample Variability
7.2 Sampling Distributions

$N=5$ balls in bucket, select $n=1$ with replacement.
Population data values: 0, 2, 4, 6, 8.  
5 possible values

<table>
<thead>
<tr>
<th>$x$</th>
<th>$P(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1/5</td>
</tr>
<tr>
<td>2</td>
<td>1/5</td>
</tr>
<tr>
<td>4</td>
<td>1/5</td>
</tr>
<tr>
<td>6</td>
<td>1/5</td>
</tr>
<tr>
<td>8</td>
<td>1/5</td>
</tr>
</tbody>
</table>

$$\mu = \sum_{i=1}^{n} [x_i P(x_i)]$$

= 

Rowe, D.B.
7: Sample Variability
7.2 Sampling Distributions

$N=5$ balls in bucket, select $n=1$ with replacement.
Population data values: 0, 2, 4, 6, 8.

5 possible values

$\sigma^2 = \sum_{i=1}^{n} [(x_i - \mu)^2 P(x_i)]$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$P(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1/5</td>
</tr>
<tr>
<td>2</td>
<td>1/5</td>
</tr>
<tr>
<td>4</td>
<td>1/5</td>
</tr>
<tr>
<td>6</td>
<td>1/5</td>
</tr>
<tr>
<td>8</td>
<td>1/5</td>
</tr>
</tbody>
</table>
7: Sample Variability
7.2 Sampling Distributions

\[ N = 5 \] balls in bucket, select \( n = 2 \) with replacement.

Population data values:
0, 2, 4, 6, 8.

\[
\begin{align*}
(0,0) & \quad (2,0) & \quad (4,0) & \quad (6,0) & \quad (8,0) \\
(0,2) & \quad (2,2) & \quad (4,2) & \quad (6,2) & \quad (8,2) \\
(0,4) & \quad (2,4) & \quad (4,4) & \quad (6,4) & \quad (8,4) \\
(0,6) & \quad (2,6) & \quad (4,6) & \quad (6,6) & \quad (8,6) \\
(0,8) & \quad (2,8) & \quad (4,8) & \quad (6,8) & \quad (8,8)
\end{align*}
\]

25 possible samples
Example: There are \( N=5 \) items in the population. Population data values: 0, 2, 4, 6, 8. Take samples of size \( n=2 \) (with replacement).

There are 25 possible samples.

\[
\begin{array}{cccccc}
(0,0) & (2,0) & (4,0) & (6,0) & (8,0) \\
(0,2) & (2,2) & (4,2) & (6,2) & (8,2) \\
(0,4) & (2,4) & (4,4) & (6,4) & (8,4) \\
(0,6) & (2,6) & (4,6) & (6,6) & (8,6) \\
(0,8) & (2,8) & (4,8) & (6,8) & (8,8) \\
\end{array}
\]

Each sample has mean \( \bar{x} \).

\[
\begin{array}{ccccccc}
0 & 1 & 2 & 3 & 4 \\
1 & 2 & 3 & 4 & 5 \\
2 & 3 & 4 & 5 & 6 \\
3 & 4 & 5 & 6 & 7 \\
4 & 5 & 6 & 7 & 8 \\
\end{array}
\]
7: Sample Variability
7.2 Sampling Distributions

**Example:** $N=5$, values: 0, 2, 4, 6, 8, $n=2$ (with replacement). 25 possible samples.

Each possible sample is equally likely.

Prob. of each sample $= 1/25 = 0.04$

There are 25 possible samples.

$(0,0) \quad (2,0) \quad (4,0) \quad (6,0) \quad (8,0)$

$(0,2) \quad (2,2) \quad (4,2) \quad (6,2) \quad (8,2)$

$(0,4) \quad (2,4) \quad (4,4) \quad (6,4) \quad (8,4)$

$(0,6) \quad (2,6) \quad (4,6) \quad (6,6) \quad (8,6)$

$(0,8) \quad (2,8) \quad (4,8) \quad (6,8) \quad (8,8)$
7: Sample Variability
7.2 Sampling Distributions

Example: \( N=5 \), values: 0, 2, 4, 6, 8, \( n=2 \) (with replacement). 25 possible samples. Each possible sample is equally likely.

Prob. of each samples
mean = \( \frac{1}{25} = 0.04 \)

There are 25 possible samples.

\[
\begin{array}{cccccccc}
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
0 & (0,0) & (2,0) & (4,0) & (6,0) & (8,0) \\
1 & (0,2) & (2,2) & (4,2) & (6,2) & (8,2) \\
2 & (0,4) & (2,4) & (4,4) & (6,4) & (8,4) \\
3 & (0,6) & (2,6) & (4,6) & (6,6) & (8,6) \\
4 & (0,8) & (2,8) & (4,8) & (6,8) & (8,8) \\
\end{array}
\]
7: Sample Variability
7.2 Sampling Distributions

**Example:** \( N=5 \), values: 0, 2, 4, 6, 8, \( n=2 \) (with replacement).

25 possible samples.

Prob. of each samples mean = \( 1/25 = 0.04 \)

\[ \begin{array}{cccccc}
0 & 1 & 2 & 3 & 4 \\
1 & 2 & 3 & 4 & 5 \\
2 & 3 & 4 & 5 & 6 \\
3 & 4 & 5 & 6 & 7 \\
4 & 5 & 6 & 7 & 8 \\
\end{array} \]

\( \bar{x} = 0 \), occurs one time
\( \bar{x} = 1 \), occurs two times
\( \bar{x} = 2 \), occurs three times
\( \bar{x} = 3 \), occurs four times
\( \bar{x} = 4 \), occurs five times
\( \bar{x} = 5 \), occurs four times
\( \bar{x} = 6 \), occurs three times
\( \bar{x} = 7 \), occurs two times
\( \bar{x} = 8 \), occurs one time
7: Sample Variability
7.2 Sampling Distributions

Example: \( N=5 \), values: 0, 2, 4, 6, 8, \( n=2 \) (with replacement).
25 possible samples.

Prob. of each samples
mean = 1/25 = 0.04

\[
\begin{align*}
P(\bar{x} = 0) &= 1 / 25 \\
P(\bar{x} = 1) &= 2 / 25 \\
P(\bar{x} = 2) &= 3 / 25 \\
P(\bar{x} = 3) &= 4 / 25 \\
P(\bar{x} = 4) &= 5 / 25 \\
P(\bar{x} = 5) &= 4 / 25 \\
P(\bar{x} = 6) &= 3 / 25 \\
P(\bar{x} = 7) &= 2 / 25 \\
P(\bar{x} = 8) &= 1 / 25
\end{align*}
\]
Example: \(N=5\), values: 0, 2, 4, 6, 8, \(n=2\) (with replacement).

\[
\begin{align*}
P(\bar{x} = 0) &= 1/25 \\
P(\bar{x} = 1) &= 2/25 \\
P(\bar{x} = 2) &= 3/25 \\
P(\bar{x} = 3) &= 4/25 \\
P(\bar{x} = 4) &= 5/25 \\
P(\bar{x} = 5) &= 4/25 \\
P(\bar{x} = 6) &= 3/25 \\
P(\bar{x} = 7) &= 2/25 \\
P(\bar{x} = 8) &= 1/25
\end{align*}
\]

Represent this distribution function with a histogram.

Figure from Johnson & Kuby, 2008.
7: Sample Variability
7.2 Sampling Distributions

Don’t forget that the two values that we draw are random.

That is, we may know the sample space of possible outcomes but we do not know exactly which ones we will get!

**Random Sample:** A sample obtained in such a way that each possible sample of fixed size $n$ has an equal probability of being selected.
As the number of samples increases the empirical dist. turns into theoretical dist.
# 7: Sample Variability

## 7.2 Sampling Distributions

### The Sampling Distribution of Sample Means

<table>
<thead>
<tr>
<th>Statistical population being studied</th>
<th>Repeated sampling is needed to form the sampling distribution.</th>
<th>All possible samples of size $n$</th>
<th>One value of the sample statistic ($\bar{x}$ in this case) corresponding to the parameter of interest ($\mu$ in this case) is obtained from each sample</th>
<th>Then all of these values of the sample statistic, $\bar{x}$, are used to form the sampling distribution.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter of interest, $\mu$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As the number of samples increases the empirical dist. turns into theoretical dist.

portion of Figure from Johnson & Kuby, 2008.

Rowe, D.B.
Sample distribution of sample means (SDSM): If all possible random samples, each of size \( n \), are taken from any population with mean \( \mu \) and standard deviation \( \sigma \), then the sampling distribution of sample means will have the following:

1. A mean \( \mu_{\bar{x}} \) equal to \( \mu \)
2. A standard deviation \( \sigma_{\bar{x}} \) equal to \( \frac{\sigma}{\sqrt{n}} \)

Furthermore, if the sampled population has a normal distribution, then the sampling distribution of \( \bar{x} \) will also be normal for all samples of all sizes.
7: Sample Variability
7.3 The Sampling Distribution of Sample Means

Example: $N=5$, values: 0, 2, 4, 6, 8, $n=2$ (with replacement). Instead of drawing two values with replacement and computing the sample mean, we can think of this as drawing one of the sample means with replacement.

The probability for each sample mean is →

\[
\begin{align*}
P(\bar{x} = 0) &= 1 / 25 \\
P(\bar{x} = 1) &= 2 / 25 \\
P(\bar{x} = 2) &= 3 / 25 \\
P(\bar{x} = 3) &= 4 / 25 \\
P(\bar{x} = 4) &= 5 / 25 \\
P(\bar{x} = 5) &= 4 / 25 \\
P(\bar{x} = 6) &= 3 / 25 \\
P(\bar{x} = 7) &= 2 / 25 \\
P(\bar{x} = 8) &= 1 / 25
\end{align*}
\]
7: Sample Variability
7.3 The Sampling Distribution of Sample Means

Example: \( N=5 \), values: 0, 2, 4, 6, 8, \( n=2 \) (with replacement).

\[
\mu_{\bar{x}} = \sum_{\bar{x}} \bar{x} P(\bar{x})
\]

\[
\mu_{\bar{x}} = \begin{align*}
P(\bar{x} = 0) &= \frac{1}{25} \\
P(\bar{x} = 1) &= \frac{2}{25} \\
P(\bar{x} = 2) &= \frac{3}{25} \\
P(\bar{x} = 3) &= \frac{4}{25} \\
P(\bar{x} = 4) &= \frac{5}{25} \\
P(\bar{x} = 5) &= \frac{4}{25} \\
P(\bar{x} = 6) &= \frac{3}{25} \\
P(\bar{x} = 7) &= \frac{2}{25} \\
P(\bar{x} = 8) &= \frac{1}{25}
\end{align*}
\]
7: Sample Variability
7.3 The Sampling Distribution of Sample Means

Example: \( N=5 \), values: 0, 2, 4, 6, 8, \( n=2 \) (with replacement).

\[
\sigma^2_x = \sum_{\bar{x}} (\bar{x} - \mu_{\bar{x}})^2 P(\bar{x})
\]

\[
\sigma^2_x = \frac{2}{25}
\]

\[
P(\bar{x} = 0) = \frac{1}{25}
\]
\[
P(\bar{x} = 1) = \frac{2}{25}
\]
\[
P(\bar{x} = 2) = \frac{3}{25}
\]
\[
P(\bar{x} = 3) = \frac{4}{25}
\]
\[
P(\bar{x} = 4) = \frac{5}{25}
\]
\[
P(\bar{x} = 5) = \frac{4}{25}
\]
\[
P(\bar{x} = 6) = \frac{3}{25}
\]
\[
P(\bar{x} = 7) = \frac{2}{25}
\]
\[
P(\bar{x} = 8) = \frac{1}{25}
\]
7: Sample Variability
7.3 The Sampling Distribution of Sample Means

We have a couple of definitions.

**Standard error of the mean** ($\sigma_{\bar{x}}$): The standard deviation of the sampling distribution of sample means.

**Central Limit Theorem (CLT):** The sampling distribution of sample means will more closely resemble the normal distribution as the sample size increases.

The CLT is extremely important in Statistics.
7: Sample Variability
7.3 The Sampling Distribution of Sample Means

The CLT: Assume that we have a population (arbitrary distribution) with mean $\mu$ and standard deviation $\sigma$.

If we take random samples of size $n$ (with replacement), then for “large” $n$, the distribution of the sample means the $\bar{x}$‘s is approximately normally distributed with

$$\mu_{\bar{x}} = \mu, \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

where in general $n \geq 30$ is sufficiently “large,” but can be as small as15 or as big as 50 depending upon the shape of distribution!
7: Sample Variability

Questions?

Homework: Chapter 7 # 1, 6, 21, 23, 29, 35
Review Chapters 4.3 – 5.6
(Exam 3 Chapters)

Just the highlights!
Conditional probability an event will occur:  A conditional probability is the relative frequency with which an event can be expected to occur under the condition that that additional preexisting information is known about some other event.

\[ P(A \mid B) \], the “\( \mid \)” is spoken as “given”
4: Probability

4.3 Conditional Probability

Example: Draw card from deck. Let $A =$ red card, $B =$ heart.

$P(A) = ?$ vs $P(A|B) = ?$

$P(A) = 1/2$ vs. $P(A|B) = 1$

Conditional Probability Formula:

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

$P(A \text{ and } B) = 1/4$  \quad $P(B) = 1/4$

$$P(A \mid B) = 1$$

Figure from Johnson & Kuby, 2008.
4: Probability
4.4 Rules of Probability

**Union Example** \((A \text{ or } B)\): Rolling a single die. 
\(A=\text{event #1,2,3. } B=\text{event odd number.}\)

\((A \text{ or } B)\)

\[
P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)
\]

\begin{align*}
\text{S} & \quad \begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6
\end{array} \\
A & \begin{array}{cccccc}
\require{cancel} & \require{cancel} & \require{cancel} & 4 & 5 & \require{cancel}
\end{array} \\
B & \begin{array}{cccccc}
\require{cancel} & \require{cancel} & \require{cancel} & 4 & 5 & \require{cancel}
\end{array}
\end{align*}

double count 1 and 3
4: Probability
4.4 Rules of Probability

Intersection Example \((A \text{ and } B)\): Rolling a single die. 
\(A=\text{event #1,2,3. } B=\text{event odd number.}\) 
\((A \text{ and } B) =\{1,3\}\)

\[ P(A \text{ and } B) = P(A) \cdot P(B \mid A) \]
4: Probability
4.5 Mutually Exclusive Events

Mutually Exclusive Example: Rolling a single die. $A=$event #1,2. $B=$event #5,6.

\[ A \text{ and } B = \emptyset \]

\[ P(A \text{ and } B) = 0 \]

\[ S \]

no overlap

null set

Rowe, D.B.
4: Probability
4.6 Independent Events

**Independent events:** … the occurrence or nonoccurrence of one gives no information about … occurrence for the other.

\[ P(A) = P(A \mid B) = P(A \mid \text{not } B) \]

**Dependent events:** … occurrence of one event does have an effect on the probability of occurrence of the other event.

\[ P(A) \neq P(A \mid B) \]

**Special multiplication rule:**
Let \( A \) and \( B \) … independent events … in a sample space \( S \).

\[ P(A \text{ and } B) = P(A) \cdot P(B) \]
Random Variables: A variable that assumes a unique numerical value for each of the outcomes in the sample space of a probability experiment.

Example:
Let $x =$ the number of heads when we flip a coin twice.

$x = \{0, 1, 2\}$ ← numerical values for

$\{TT, TH, HT, HH\}$ ← outcomes in sample space
5: Probability Distributions (Discrete Variables)
5.3 Probability Distributions of a Discrete Random Variable

**Probability Function:** A rule $P(x)$ that assigns probabilities to the values of the random variable $x$.

**Example:**
Let $x$ = # of heads when we flip a coin twice.

$x = \{0, 1, 2\}$

$P(x) = \frac{2!}{x!(2-x)!} \left( \frac{1}{2} \right)^x \left( \frac{1}{2} \right)^{2-x}$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$P(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>1</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{4}$</td>
</tr>
</tbody>
</table>
5: Probability Distributions (Discrete Variables)
5.4 Mean and Variance of a Discrete Random Variable

\[ \mu = \sum_{i=1}^{n} [x_i P(x_i)] = x_1 P(x_1) + x_2 P(x_2) + \ldots + x_n P(x_n) \]

For the discrete distribution we just discussed it is:

\[ \mu = (0) \cdot P(0) + (1) \cdot P(1) + (2) \cdot P(2) \]

\[ \mu = (0) \cdot (1/4) + (1) \cdot (1/2) + (2) \cdot (1/4) \]

\[ \mu = 0 + 1/2 + 1/2 \]

\[ \mu = 1 \]

\[ \begin{array}{c|c}
  x & P(x) \\
  \hline
  0 & 1/4 \\
  x_1 & 1/2 \\
  x_2 & 1/4 \\
  x_3 & 2 \\
\end{array} \]
5: Probability Distributions (Discrete Variables)

5.4 Mean and Variance of a Discrete Random Variable

For the discrete distribution we just discussed it is:

\[ \sigma^2 = \sum_{i=1}^{n} [(x_i - \mu)^2 P(x_i)] = (x_1 - \mu)^2 P(x_1) + (x_2 - \mu)^2 P(x_2) + \ldots + (x_n - \mu)^2 P(x_n) \]

\[ \sigma^2 = (0-1)^2 \cdot P(0) + (1-1)^2 \cdot P(1) + (2-1)^2 \cdot P(2) \]

\[ \sigma^2 = (-1)^2 \cdot (1/4) + (0)^2 \cdot (1/2) + (1)^2 \cdot (1/4) \]

\[ \sigma^2 = 1/4 + 0 + 1/4 \]

\[ \sigma^2 = 1/2 \]
5: Probability Distributions (Discrete Variables)

5.4 Mean and Variance of a Discrete Random Variable

\[ \sigma^2 = \sum_{i=1}^{n} [(x_i - \mu)^2 P(x_i)] = (x_1 - \mu)^2 P(x_1) + (x_2 - \mu)^2 P(x_2) + \ldots + (x_n - \mu)^2 P(x_n) \]

For the discrete distribution we just discussed it is:

\[ \sigma = \sqrt{\sigma^2} = \sqrt{1/2} \approx 0.7071 \]

\[ \sigma^2 = 1/2 \]

\[ \sigma = 1/\sqrt{2} \approx 0.7071 \]
5: Probability Distributions (Discrete Variables)

5.5 The Binomial Probability Distribution

Experiment with only two outcomes is called a Binomial exp. Call one outcome *Success* and the other *Failure*. Each performance of expt. is called a trial and are independent.

\[
P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}
\]

\[x = 0, ..., n\]

\[n = \text{number of trials or times we repeat the experiment.}\]
\[x = \text{the number of successes out of } n \text{ trials.}\]
\[p = \text{the probability of success on an individual trial.}\]
5: Probability Distributions (Discrete Variables)
5.5 The Binomial Probability Distribution

Flip coin ten times.

<table>
<thead>
<tr>
<th>x</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>n(x)</td>
<td>1</td>
<td>10</td>
<td>45</td>
<td>120</td>
<td>210</td>
<td>252</td>
<td>210</td>
<td>120</td>
<td>45</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>P(x)</td>
<td>(\frac{1}{1024})</td>
<td>(\frac{10}{1024})</td>
<td>(\frac{45}{1024})</td>
<td>(\frac{120}{1024})</td>
<td>(\frac{210}{1024})</td>
<td>(\frac{252}{1024})</td>
<td>(\frac{210}{1024})</td>
<td>(\frac{120}{1024})</td>
<td>(\frac{45}{1024})</td>
<td>(\frac{10}{1024})</td>
<td>(\frac{1}{1024})</td>
</tr>
</tbody>
</table>

\[n(x) = \frac{n!}{x!(n-x)!}\]

\[P(x) = \frac{n!}{x!(n-x)!} p^x (1 - p)^{n-x}\]

Note:
1. \(0 \leq P(x) \leq 1\)
2. \(\sum P(x) = 1\)
5: Probability Distributions (Discrete Variables)

5.5 The Binomial Probability Distribution

The binomial probability distribution is given by:

\[ P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \]

Binomial Probabilities \([\binom{n}{x} \cdot p^x \cdot q^{n-x}]\) (continued)

<table>
<thead>
<tr>
<th>(n)</th>
<th>(x)</th>
<th>0.01</th>
<th>0.05</th>
<th>0.10</th>
<th>0.20</th>
<th>0.30</th>
<th>0.40</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0</td>
<td>.904</td>
<td>.599</td>
<td>.349</td>
<td>.107</td>
<td>.028</td>
<td>.006</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>.091</td>
<td>.315</td>
<td>.387</td>
<td>.268</td>
<td>.121</td>
<td>.040</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.004</td>
<td>.075</td>
<td>.194</td>
<td>.302</td>
<td>.233</td>
<td>.121</td>
<td>.044</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0+</td>
<td>.010</td>
<td>.057</td>
<td>.201</td>
<td>.267</td>
<td>.215</td>
<td>.117</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0+</td>
<td>.001</td>
<td>.011</td>
<td>.088</td>
<td>.200</td>
<td>.251</td>
<td>.205</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0+</td>
<td>0+</td>
<td>.001</td>
<td>.026</td>
<td>.103</td>
<td>.201</td>
<td>.246</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>.006</td>
<td>.037</td>
<td>.111</td>
<td>.205</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>.001</td>
<td>.009</td>
<td>.042</td>
<td>.117</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>.001</td>
<td>.011</td>
<td>.044</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>.002</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>0+</td>
<td>.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(x)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P(x))</td>
<td>(\frac{1}{1024})</td>
<td>(\frac{10}{1024})</td>
<td>(\frac{45}{1024})</td>
<td>(\frac{120}{1024})</td>
<td>(\frac{210}{1024})</td>
<td>(\frac{252}{1024})</td>
<td>(\frac{210}{1024})</td>
<td>(\frac{120}{1024})</td>
<td>(\frac{45}{1024})</td>
<td>(\frac{10}{1024})</td>
<td>(\frac{1}{1024})</td>
</tr>
</tbody>
</table>
The formula for the mean $\mu$ and variance $\sigma^2$ of Binomial is

$$
\mu = \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}
$$

$$
= np
$$

(5.7)

$$
\sigma^2 = \sum_{x=0}^{n} (x - \mu)^2 \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}
$$

$$
= np(1-p) \quad \longrightarrow \quad \sigma = \sqrt{np(1-p)}
$$

(5.8)
5: Probability Distributions (Discrete Variables)

5.6 Mean and Standard Deviation of the Binomial Distribution

Example:
Before, using \( \mu = \sum_{x=0}^{n} [xP(x)] \), we found \( \mu = 1 \).

Now using \( \mu = np \), we get \( \mu = (2) \cdot (1/2) = 1 \).

Before, using \( \sigma^2 = \sum_{x=0}^{n} [(x-\mu)^2 P(x)] \), we found \( \sigma^2 = 1/2 \).

Now using \( \sigma^2 = np(1-p) \),
we get \( \sigma^2 = (2) \cdot (1/2) \cdot (1/2) = 1/2 \).