Math 3339
Section 21155 - SR 117 - MW 1 - 2:30pm

Bekki George: bekki@math.uh.edu

University of Houston

Sections 4.3-4.5
Office Hours: Mondays 11am - 12:30pm, Tuesdays 3-4pm  
(also available by appointment)  
Office: 206 PGH  
Course webpage: www.casa.uh.edu
Empirical probability (experimental probability) of an event is the ratio of the number of outcomes in which a specified event occurs to the total number of trials, not in a theoretical sample space but in an actual experiment.

1. A die is rolled 60 times with the following results recorded:

<table>
<thead>
<tr>
<th>Outcome</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>10</td>
<td>6</td>
<td>12</td>
<td>9</td>
<td>8</td>
<td>15</td>
</tr>
</tbody>
</table>

The empirical probability of getting a 3 is:

a) $\frac{1}{12}$

b) $\frac{1}{10}$

c) $\frac{1}{6}$

d) $\frac{1}{5}$

e) none of these
Expected Values

The **expected value** or mean of the distribution of a random variable \( X \) is given by:

\[
E[X] = \mu = \sum_{x} x \cdot f(x) = \sum_{i=1}^{n} x_i \cdot p_i
\]

Example: From Monday's example where we had number of defectives per lot, find the expected number of defective items.

<table>
<thead>
<tr>
<th>Lot</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. Defective</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

\[
f(x) = p(x) = \frac{1}{2}, \frac{1}{6}, \frac{1}{3}
\]

\[
E[X] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{3} = \frac{5}{6}
\]
Expected Values

Another Example: Consider the table below which gives the number of years required to obtain a Bachelor’s degree for graduates of high school A, and the number of students who needed each:

<table>
<thead>
<tr>
<th>X: Years</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Students</td>
<td>17</td>
<td>23</td>
<td>38</td>
<td>19</td>
</tr>
</tbody>
</table>

= 97 students

How would compute the average number of years required by graduates of high school A?

\[ P(X = x) = \frac{1}{97}, \frac{23}{97}, \frac{38}{97}, \frac{19}{97} \]

\[ E[X] = 3 \left( \frac{1}{97} \right) + 4 \left( \frac{23}{97} \right) + 5 \left( \frac{38}{97} \right) + 6 \left( \frac{19}{97} \right) \approx 4.61 \]

\[ \frac{3 \cdot 17 + 4 \cdot 23 + 5 \cdot 38 + 6 \cdot 19}{97} \]
Variance

The variance of a rv $X$ is

$$\sigma^2 = \text{Var}[X] = E[(X - \mu)^2] = E[X^2] - E[X]^2$$
Expected Values and Variance

Properties of Expected values and Variance

1. $E[c] = c$ for any constant $c \in \mathbb{R}$

2. $E[aX + b] = aE[X] + b$

3. $E[aX + bY] = aE[X] + bE[Y]$

4. $E[h(x)] = \sum_x h(x) \cdot f(x)$

5. $Var[aX + b] = a^2 Var[X]$  \[ E[X+b] = E[X] + b \]

6. $Var[aX + bY] = a^2 Var[X] + b^2 Var[Y]$  \[ E[X^2] = x_1^2 \cdot p_1 + x_2^2 \cdot p_2 + ... \]

7. $Var[aX + bY] = \sqrt{a^2 Var[X] + b^2 Var[Y]}$  \[ a \cdot sd[X] + b \cdot sd[Y] \]
Expected Values and Variance

Example: Let $X$ have pmf given by

<table>
<thead>
<tr>
<th>$x$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(x)$</td>
<td>0.4</td>
<td>0.2</td>
<td>0.3</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Determine $E[X]$, $E[X^2]$, $Var[X]$ and the standard deviation of $X$.

$E[X] = 1(.4) + 2(.2) + 3(.3) + 4(.1) = 2.1$

$E[X^2] = 1(.4) + 4(.2) + 9(.3) + 16(.1) = 5.5$

$Var[X] = 5.5 - (2.1)^2 = 1.09$

$sd[X] = \sqrt{1.09} = 1.044$
Expected Values and Variance

Example: Determine the expected value and variance of the rv $Y$ defined by $Y = 5X - 1$, where $X$ is given in the previous problem.

\[
\begin{array}{cccccc}
X & 1 & 2 & 3 & 4 \\
Y & 4 & 9 & 14 & 19 \\
\end{array}
\]

\[
E[Y] = E[5X-1] = 5E[X] - 1 = 5(2.1) - 1 = 9.5
\]

\[
Var[Y] = Var[5X-1] = 25 \cdot Var[X] = 25(1.09) = 27.25
\]
Suppose \( X \) is a rv with \( E[X] = 3.7 \) and \( Var[X] = 2.25 \) and \( Y = 2X - 3 \). Find:

2. \( E[Y] = \) \[ E[2X-3] = 2 \cdot E[X] - 3 \]
   
   a) 7.4
   b) 4.4
   c) 3.7
   d) 2.5
   e) none of these

3. \( Var[Y] = \)
   
   a) 9
   b) 4.5
   c) 3.7
   d) 3
   e) none of these
The Binomial Probability Distribution

Definition: A Bernoulli Trial is a random experiment with the following characteristics:

1. The outcome can be classified as either success or failure (where these are mutually exclusive and exhaustive).

2. The probability of success is $p$, so the probability of failure is $q = 1 - p$. e.g. a coin is flipped (heads or tails), someone is pulled over for speeding (ticket or warning), etc.

Suppose that a coin is flipped. Let $X$ be the random variable that indicates that heads was flipped (i.e.). Here heads represents “success” and tails represents “failure” so that $X$ is a Bernoulli random variable.
The Binomial Probability Distribution

Suppose that we flip a coin 10 times. This is a sequence of Bernoulli trials. We are interested in calculating the probability of obtaining a certain number of heads. Let $X_i$ indicate heads on the $i$-th flip.

\[
X(\text{Heads}) = 1 \quad X(\text{Tail}) = 0
\]

Define $Y = X_1 + X_2 + \ldots + X_{10}$. What does $Y$ represent?

What is the probability that $Y = 0$?

\[
P(\text{no heads}) = \frac{1}{2^{10}} = \frac{1}{1024}
\]

What is the probability that $Y = 1$?

\[
P(\text{one heads}) = \left(\frac{1}{2} \cdot \frac{1}{2^9}\right)10 = \frac{10}{1024}
\]

What is the probability that $Y = 2$?

\[
P(Y = 2) = \frac{10 \cdot \binom{2}{2}}{2^{10}} = 10 \cdot \binom{2}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^8
\]

What is the probability that $Y = n$?

\[
10 \binom{n}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^{10-n}
\]
The Binomial Probability Distribution

Here $Y$ is the sum of 10 independent Bernoulli trials. We call this type of random variable a **Binomial random variable**.

A random variable $X$ is a Binomial random variable if the following conditions are satisfied:

1. $X$ represents the number of successes on $n$ Bernoulli trials.
2. The probability of success for each trial is $p$.
3. The trials are mutually independent.

If $X$ is a binomial random variable with probability $p$ of success on each of $n$ trials, we write $X \sim \text{Binomial}(n,p)$.

If $X \sim \text{Binomial}(n,p)$, then $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ where $x = 0, 1, 2, ...$

$$P(X = 3) = \binom{n}{3} p^3 (1-p)^{n-3}$$  

**Formula**
The Binomial Probability Distribution

R commands:

\[ P(X = x) = \text{dbinom}(x, n, p) \]

\[ P(X \leq x) = \text{pbinom}(x, n, p) \]

\[ P(X > x) = 1 - \text{pbinom}(x, n, p) \]

\[ 1 - P(X \leq x) \]

\[ P(X \geq 3) = 1 - P(X \leq 2) \]

\[ P(X > 3) = 1 - P(X \leq 3) \]
The Binomial Probability Distribution

The expected value of a Binomial rv is \( E[X] = np \) and the variance is \( \sigma^2 = np(1 - p) \).

Example: Suppose that at a 4-way stop in a certain subdivision, only 12\% of drivers come to a complete stop. What is the probability that among 8 drivers, at least 6 of them will run the stop sign?

\[ n = 8 \]

\[ P(X \geq 6) = 1 - P(X \leq 5) \]

\[ P = 0.88 \]

\[ = 1 - \text{pbinom}(5, 8, 0.88) = 0.939 \]

What is the expected number of drivers who will run the stop sign?

\[ E[X] = n \cdot p = 8 \cdot 0.88 = 7.04 \]
The Binomial Probability Distribution

Suppose $X \sim Binomial(12, 0.3)$, find the following:

$$P(2 \leq X < 5) = P(X = 2) + P(X = 3) + P(X = 4) = \text{dbinom}(2, 12, 0.3) + \text{dbinom}(3, 12, 0.3) + \text{dbinom}(4, 12, 0.3) + \text{dbinom}(5, 12, 0.3) + \text{dbinom}(6, 12, 0.3) + \text{dbinom}(7, 12, 0.3) + \text{dbinom}(8, 12, 0.3) + \text{dbinom}(9, 12, 0.3) + \text{dbinom}(10, 12, 0.3) + \text{dbinom}(11, 12, 0.3) + \text{dbinom}(12, 12, 0.3)$$

$$P(X > 5) = 1 - P(X \leq 5) = 1 - \text{pbet}(4, 12, 0.3) - \text{pbet}(1, 12, 0.3) = 0.639$$

$$= 1 - \text{pbet}(5, 12, 0.3) = 0.118$$
Ex: Suppose that 20% of all copies of a particular textbook fail a certain binding strength test. Let \( X \) denote the number among 15 randomly selected copies that fail the test.

Is \( X \) a binomial random variable? \( \text{Yes} \) trials indep.

Determine the probability that exactly 3 fail the test.

\[
P(3 \text{ fail}) = P(X = 3) = \text{dbinom}(3, 15, 0.2)
\]

Determine the probability that at most 3 fail the test.

\[
P(X \leq 3) = \text{pbinom}(3, 15, 0.2)
\]

Determine the probability that at least 3 fail the test.

\[
P(X \geq 3) = 1 - P(X \leq 2) = 1 - \text{pbinom}(2, 15, 0.2)
\]

How many textbooks do we expect to fail the test? \( 15 \times 0.2 = 3 \)

What is the standard deviation of \( X \)?

\[
\sqrt{np(1-p)} = \sqrt{15 \times 0.2 \times 0.8} > \sqrt{15 \times 2 \times 0.8}
\]

\( [1] 1.549193 \)
Ex: Each year a company selects a number of employees for a management training program. On average, 70% of those sent complete the program. Out of the seven people sent, what is the probability that

\[ p = 0.7 \quad n = 7 \]

(a) Exactly five complete the program?

\[ P(X = 5) = \text{dbinom}(5, 7, 0.7) \]

\[ > \text{dbinom}(5,7,.7) \]
\[ [1] 0.3176523 \]

(b) Five or more complete the program?

\[ P(X \geq 5) = 1 - P(X \leq 4) \]

\[ 1 - \text{pbinom}(4, 7, 0.7) \]
Ex: Suppose it is known that 80% of the people exposed to the flu virus will contract the flu. Out of a family of five exposed to the virus, what is the probability that:
   a. No one will contract the flu?
   
   b. All will contract the flu?
   
   c. Exactly two will get the flu?
   
   d. At least two will get the flu?
4. Seventy percent of all trucks undergoing brake inspection at a certain facility pass the inspection. Consider a group of 15 trucks. What is the probability that between 10 and 12 trucks inclusively pass the inspection?
   a. 0.5948
   b. 0.3887
   c. 0.2186
   d. 0.5008
   e. none of these

   \[ \text{a. } \text{pbinom}(12, 15, 0.7) - \text{pbinom}(10, 15, 0.7) \]
   \[ \text{b. } \text{pbinom}(13, 15, 0.7) - \text{pbinom}(9, 15, 0.7) \]
   \[ \text{c. } \text{pbinom}(12, 15, 0.7) - \text{pbinom}(9, 15, 0.7) \]
   \[ \text{d. } \text{pbinom}(12, 15, 0.7) \]
   \[ \text{e. none of these} \]

5. Find the mean and standard deviation for the rv $X$ as described in the problem above.
   a. Mean = 10.5, standard deviation = 3.15
   b. Mean = 4.5, standard deviation = 3.15
   c. Mean = 10.5, standard deviation = 1.77
   d. Mean = 4.5, standard deviation = 1.77
   e. None of these