MATH 3339 - 03 15951
Statistics for the Sciences
Discrete Random Variables

Wendy Wang
wwang60@central.uh.edu

Lecture 13 - 3339
Outline

1. Discrete Random Variable Examples
2. Chebyshev’s inequality
3. Expected Values of Discrete Variables
4. Bernoulli Random Variables
Fill in all of the proper bubbles.

Make sure your ID number is correct.

Make sure the filled in circles are very dark.

This is popper number 08.
Discrete Random Variables

Def: The **probability mass function** (pmf) of a discrete random variable (r.v.) is defined for every number $x_i$ by $f(x_i) = P(X = x_i)$.

Example: Six lots of components are ready to be shipped by a certain supplier. The number of defective components in each lot is as follows:

<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>

One of these lots is to be randomly selected for shipment to a particular customer. Let $X =$ number of defectives in the selected lot. What are the possible values of $X$? Determine the values of the pmf.
Discrete Random Variables

Properties of $f$:
- $f(x) \geq 0$ for all $x \in \mathbb{R}$
- $\sum_i f(x_i) = 1$
- $P(X \in A) = \sum_{x \in A} f(x)$, where $A \subset \mathbb{R}$ is a discrete set.

Suppose we toss a fair coin 10 times. Let $X =$ number of heads in the 10 tosses. What are the possible values of $X$?

How many heads do we expect to get in the 10 tosses?

What is $f(0)$? $f(1)$?
Discrete Random Variables

Example: Suppose you are given the following distribution table:

<table>
<thead>
<tr>
<th>X</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(X)</td>
<td>0.15</td>
<td>0.05</td>
<td>0.10</td>
<td>0.10</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Find the following:

\[ P(X = 4) \]

\[ P(X < 2) \]

\[ P(2 < X \leq 5) \]

\[ P(X > 3) \]
The Cumulative Distribution Function: Discrete R.V.

Def: The **cumulative distribution function** (cdf) $F(x)$ of a discrete r.v $X$ with pmf $f(x)$ is defined for every number $x$ by $F(x) = P(X \leq x)$.

For any number $x$, $F(x)$ is the probability that the observed value of $X$ will be at most $x$. 

Discrete Random Variables

Example: A store carries flash drives with either 1 GB, 2 GB, 4 GB, 8 GB, or 16 GB of memory. The table below gives the distribution of the rv $X = \text{the amount of memory in a purchased drive}$.

<table>
<thead>
<tr>
<th>$X$</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X)$</td>
<td>0.05</td>
<td>0.10</td>
<td>0.35</td>
<td>0.40</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Determine the values of $F$ for the possible values of $X$.

What is the value of $F(2.5)$?

Sketch a graph of $F$. 
Discrete Random Variables

Example: Suppose the random variable $X$ takes on possible values $x = 0, 1, 2, 3$ and has pmf given by $f(x) = \frac{x+1}{k}$, determine the value of $k$. 

Wendy Wang  wwang60@central.uh.edu

MATH 3339  Lecture 13 - 3339  9/25
Chebyshev’s Inequality

- Chebyshev’s inequality places a universal restriction on the probabilities of deviations of random variables from their means.

- If $X$ is a random variable with mean $\mu$ and standard deviation $\sigma$ and if $k$ is a positive constant, then

  \[ P(|X - \mu| > k\sigma) \leq \frac{1}{k^2} \]

- That is: For any random variable, if we are $k$ standard deviations away from the mean, then no more than $\frac{1}{k^2} \times 100\%$ is beyond $k$ standard deviations from the mean. Or at least $(1 - \frac{1}{k^2}) \times 100\%$ is within $k$ standard deviations from the mean.
Chebyshev’s Inequality Application

Let $X =$ The number of traffic accidents daily in a small city. The following table is the probability distribution for $X$. With, $\mu = 2$ and $\sigma = 1.18$.

<table>
<thead>
<tr>
<th>$X$</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.10</td>
</tr>
<tr>
<td>1</td>
<td>0.20</td>
</tr>
<tr>
<td>2</td>
<td>0.45</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
</tr>
<tr>
<td>4</td>
<td>0.05</td>
</tr>
<tr>
<td>5</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Confirm Chebyshev’s Inequality for $k = 2$. 
Expected Values of Discrete Variables
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 a previous 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>
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>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>

How would compute the average number of years required by graduates of high school A?
Expected Values and Variance

Properties of Expected values

- \( E[c] = c \) for any constant \( c \in \mathbb{R} \)
- \( E[aX + b] = aE[X] + b \)
- \( E[aX + bY] = aE[X] + bE[Y] \)
- \( E[h(x)] = \sum_x h(x) \cdot f(x) \)
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

- $E[c] = c$ for any constant $c \in \mathbb{R}$
- $E[aX + b] = aE[X] + b$
- $E[aX + bY] = aE[X] + bE[Y]$
- $E[h(x)] = \sum_x h(x) \cdot f(x)$
- $Var[aX + b] = a^2 Var[X]$
- $Var[aX + bY] = a^2 Var[X] + b^2 Var[Y]$
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 $X$. 
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.
Expected Values and Variance

Example: A store carries flash drives with either 1 GB, 2 GB, 4 GB, 8 GB, or 16 GB of memory. The table below gives the distribution of the rv $X$ = the amount of memory in a purchased drive.

<table>
<thead>
<tr>
<th>X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(X)</td>
<td>0.05</td>
<td>0.10</td>
<td>0.35</td>
<td>0.40</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Determine $E[X]$ and $Var[X]$ using R.
Finding Expected Value and Standard Deviation in R

Wendy Wang wwang60@central.uh.edu
MATH 3339
Lecture 13 - 3339 21/25
Special cases for Discrete Random Variables:

- Bernoulli Random Variable
- Binomial Random Variable
- Hypergeometric Distribution
- Poisson Distribution
- Jointly Distribution Variables
The Bernoulli Probability Distribution

Definition: A Bernoulli random variable is a random experiment (a Bernoulli Trial) 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. $X = 1$). Here heads represents “success” and tails represents “failure” so that $X$ is a Bernoulli random variable.
Probability Function for Bernoulli Variable

\[ f(x) = P(X = x) = \begin{cases} 
p, & \text{if } x = 1 \\
1 - p, & \text{if } x = 0 \\
0, & \text{if } x \neq 0, 1 
\end{cases} \]

A compact way of writing this is:

\[ f(x) = P(X = x) = p^x (1 - p)^{1-x} \]
Mean and Variance of a Bernoulli Distribution

If $X$ has the Bernoulli distribution with probability of success $p$, the **mean** and **variance** of $X$ are

$$
\mu_X = E[X] = p
$$

$$
\sigma_X^2 = Var[X] = p(1 - p)
$$

Then the standard deviation is the square root of the variance.