Lecture 4: Discrete Random Variables and Probability Distributions

**Definition:** \((E, S)\): Experiment and Sample Space.  
\(\mathcal{R}\): all real numbers.  
Random variable: \(X: \mathcal{S} \rightarrow \mathcal{R}\).  
\[X(s) = x\]

**Examples:**  
Toss an unfair coin.  
Toss an unfair coin until head.  
Toss two fair 4-sided dice.  
Randomly select a student on campus.  
1. gender  
2. undergraduate/graduate  
3. which year  
4. height

**Discrete Random Variable vs Continuous Random Variable**  
Bernoulli Random Variable

**Probability Distributions for d.r.v**  
Toss two fair dice (continued):
\[ S = \{(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4), (4,1), (4,2), (4,3), (4,4)\} \]

Random variable \( Y \): the sum of the outcomes

possible values \(2, 3, 4, 5, 6, 7, 8\)
probabilities

**Probability Distribution or Probability Mass Function**

For each possible value \( x \), \( p(x) \): the prob. of observing \( x \) when the experiment is performed

**Bernoulli r.v.**

\[
p(x) = \begin{cases} 
1 - \alpha & \text{if } x = 0 \\
\alpha & \text{if } x = 1 \\
0 & \text{otherwise}
\end{cases}
\]

• Bernoulli distribution
• \( \alpha \) called parameter
• Family of Bernoulli distributions

**Toss a coin until head (continued):**

\( X \): the number of tosses
Let $P(H) = p$,
$P(X = 1)=P(H) = p$
$P(X = 2)=P(TH) = P(T)P(H) = (1−p)p$
$P(X = 3)=P(TTH) = P(T)P(T)P(H) = (1−p)^2p$
In general, if $x$ is an positive integer,
$P(X = x)=

**Tire Example**

$X$: the number of tires on a randomly selected car that are underinflated. Its probability mass function is given as follows,

<table>
<thead>
<tr>
<th>$x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x)$</td>
<td>0.4</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3</td>
<td></td>
</tr>
</tbody>
</table>

Compute $P(2 \leq X \leq 4)$ and $P(X \neq 0)$.

**Another Representation of Probability Distribution**

**Tire Example continued**

Let’s calculate $P(X \leq x)$ for any given $x$

If $x < 0$, $P(X \leq x)=0$

If $x = 0$, $P(X \leq 0) = P(X = 0) = 0.4$

If $0 < x < 1$, $P(X \leq x) = P(X = 0) = 0.4$

If $x = 1$, $P(X \leq x)=P(X = 0)+P(X = 1)=0.4+0.1=0.5$

If $1 < x < 2$, $P(X \leq x)=0.5$

if $2 \leq x < 3$, $P(X \leq x)=0.6$

if $3 \leq x < 4$, $P(X \leq x)=0.7$

if $x \geq 4$, $P(X \leq x)=1$
Hence, \( P(X \leq x) \) is defined for any \( x \), it is called cumulative distribution function, denoted by \( F(x) \).

Step function

**Definition** The cumulative distribution function (cdf) \( F(x) \) of a discrete rv \( X \) with pmf \( p(x) \) is defined for every number \( x \) by

\[
F(x) = P(X \leq x) = \sum_{y:y\leq x} p(y)
\]

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

**Calculate Probability Using cdf \( F(x) \)**

For any two numbers \( a \) and \( b \) with \( a \leq b \),

\[
P(a \leq X \leq b) = F(b) - F(a-)
\]

where \( a- \) represents the largest possible value of \( X \) that is strictly less than \( a \).

- pmf and cdf are equivalent.

**Measure Location and Dispersion of Probability Distributions**

**Expected value (or mean value, or population mean)**
Suppose D is the collection of all possible values of X, and \( p(x) \) is the pmf.

\[
E(X) = \mu_x = \sum_{x \in D} x \cdot p(x)
\]

**Tire Example continued**

**Expected Value of a Function of X**
Let \( h(X) \) be any function depending on \( X \), then

\[
E(h(X)) = \mu_{h(X)} = \sum_{D} h(x) \cdot p(x)
\]

**Proposition**

\[
E(aX + b) = aE(X) + b
\]

or

\[
\mu_{aX+b} = a \cdot \mu_X + b
\]

**Tire Example continued**
Suppose the time used to inflate the tires is \( 0.5x^2 \), where \( x \) is the number of underinflated tire.

**Variance of \( X \)**
Let \( X \) have pmf \( p(x) \) and expected value \( \mu \). Then the variance of \( X \), denoted by \( V(X) \), or \( \sigma_X^2 \), or just \( \sigma^2 \), is

\[
V(X) = \sum_{D}(x - \mu)^2 \cdot p(x) = E((X - \mu)^2)
\]
The standard deviation (SD) of \( X \) is
\[
\sigma_X = \sqrt{\sigma_X^2}
\]

**Example**

**Proposition**
1. 
\[
V(h(X)) = \sigma_{h(X)}^2 = \sigma_D(h(X) - E(h(X)))^2 \cdot p(x)
\]
2. 
\[
V(aX + b) = \sigma_{aX+b}^2 = a^2 \cdot \sigma_X^2 \text{ and } \sigma_{aX+b} = |a| \cdot \sigma_X
\]
3. 
\[
V(X) = E(X^2) - (E(X))^2
\]

**Useful Probability Distributions**

**Binomial Probability Distribution**
1. A sequence of \( n \) trials
2. The trials are identical, each with two outcomes, denoted by \( S \) or \( F \).
3. The trials are independent.
4. \( P(S) = \) constant, denoted by \( p \)

**Binomial Experiment**

**Example:** Toss a coin 10 times, \( X \): the number of heads

**Approximate Binomial Experiment**

**Example:** Select 10 students on campus, \( X \): the number of female
student

**Definition:** Given a binomial experiment consisting of $n$ trials (exact or approximate), let

$$X = \text{the number of } S\text{'s among the } n \text{ trials.}$$

$X$ is a binomial random variable.

Possible values?

pmf $b(x; n, p)$?

**Example:** $n=4$

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b(x; 4, p)$</td>
<td>$b(0; 4, p)$</td>
<td>$b(1; 4, p)$</td>
<td>$b(2; 4, p)$</td>
<td>$b(3; 4, p)$</td>
<td>$b(4; 4, p)$</td>
</tr>
</tbody>
</table>

$b(0; 4, p) = P(X = 0) = P(FFFF) = P(F)P(F)P(F)P(F) = (1-p)^4$

$b(1; 4, p) = P(X = 1) = P(SFFF + P(FSSF) + P(FSFS) + P(FFFS) = (1-p)^3p + (1-p)^3p + (1-p)^3p + (1-p)^3p = 4(1-p)^3p$

Similarly,
\[ b(3; 4, p) = 4(1 - p)p^3 \]
\[ b(4; 4, p) = p^4 \]

**General Formula**

\[ X \sim \text{Bin}(n, p) \]

pmf:

<table>
<thead>
<tr>
<th>( x )</th>
<th>( p(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( \binom{n}{0} (1 - p)^n p^0 )</td>
</tr>
<tr>
<td>1</td>
<td>( \binom{n}{1} (1 - p)^{n-1} p^1 )</td>
</tr>
<tr>
<td>2</td>
<td>( \binom{n}{2} (1 - p)^{n-2} p^2 )</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>( i )</td>
<td>( \binom{n}{i} (1 - p)^{n-i} p^i )</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>( n )</td>
<td>( \binom{n}{n} (1 - p)^0 p^n )</td>
</tr>
</tbody>
</table>

**cdf of binomial distribution**
\[ P(X \leq x) = B(x; n, p) = \sum_{y=0}^{x} b(y; n, p) \]

For \( n = 5, 10, 15, 20, 25 \), and \( p = 0.01, 0.05, \ldots, 0.95, 0.99 \), the probabilities are given in Appendix Table A.1

**Mean and Variance**

\[ E(X) = np \]
\[ V(X) = n(1 - p)p \]

**Examples 3.48**

**Poisson Probability Distribution**

\( X \): the number of events in a specific time period (or in a specific region)

Examples:
- the number of phone calls at an office
- the number of accidents at an intersection
- the number of certain animals found in a square mile area

\[ X \] is a Poisson r.v., its pmf (Poisson distribution) is

\[ p(x; \lambda) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & x = 0, 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases} \]

**Cumulative Distribution function (cdf)**
\[ F(x; \lambda) = \sum_{y=0}^{y=x} \frac{\lambda^y e^{-\lambda}}{y!} \]

cdf of Poisson distributions are given in Table A.2

<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>\ldots</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p(x; \lambda) )</td>
<td>( \frac{\lambda^0 e^{-\lambda}}{0!} )</td>
<td>( \frac{\lambda^1 e^{-\lambda}}{1!} )</td>
<td>( \frac{\lambda^2 e^{-\lambda}}{2!} )</td>
<td>( \frac{\lambda^3 e^{-\lambda}}{3!} )</td>
<td>( \frac{\lambda^4 e^{-\lambda}}{4!} )</td>
<td>( \frac{\lambda^5 e^{-\lambda}}{5!} )</td>
<td>( \frac{\lambda^6 e^{-\lambda}}{6!} )</td>
<td>( \frac{\lambda^7 e^{-\lambda}}{7!} )</td>
<td>( \frac{\lambda^8 e^{-\lambda}}{8!} )</td>
<td>\ldots</td>
</tr>
</tbody>
</table>

**Proposition**

\[
\sum_{x=0}^{x=+\infty} p(x; \lambda) = \frac{\lambda^0 e^{-\lambda}}{0!} + \frac{\lambda^1 e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!} + \frac{\lambda^3 e^{-\lambda}}{3!} + \frac{\lambda^4 e^{-\lambda}}{4!} + \ldots = 1
\]

\[
E(X) = \sum_{x=0}^{+\infty} x \cdot p(x; \lambda)
\]

\[
= 0 \cdot \frac{\lambda^0 e^{-\lambda}}{0!} + 1 \cdot \frac{\lambda^1 e^{-\lambda}}{1!} + 2 \cdot \frac{\lambda^2 e^{-\lambda}}{2!} + 3 \cdot \frac{\lambda^3 e^{-\lambda}}{3!} + 4 \cdot \frac{\lambda^4 e^{-\lambda}}{4!} + \ldots = \lambda
\]

\[
E(X^2) = \sum_{x=0}^{+\infty} x^2 \cdot p(x; \lambda)
\]

\[
= 0^2 \cdot \frac{\lambda^0 e^{-\lambda}}{0!} + 1^2 \cdot \frac{\lambda^1 e^{-\lambda}}{1!} + 2^2 \cdot \frac{\lambda^2 e^{-\lambda}}{2!} + 3^2 \cdot \frac{\lambda^3 e^{-\lambda}}{3!} + 4^2 \cdot \frac{\lambda^4 e^{-\lambda}}{4!} + \ldots = \lambda^2 + \lambda
\]

What are the mean and variance of \( X \)?

**Connection with Binomial Distributions**

Suppose that in the binomial pmf \( b(x; n, p) \), we let \( n \to \infty \) and \( p \to 0 \) in such a way that \( np \) approaches a value \( \lambda > 0 \). Then
\[ b(x; n, p) \to p(x; \lambda) \]

In any binomial experiment in which \( n \) is large and \( p \) is small,

\[ b(x; n, p) \approx p(x; \lambda) \text{ where } \lambda = np, \]

when

\[ n \geq 100, p \leq 0.01 \text{ and } np \leq 20 \]

**Examples 3.81 and 3.82**