Specific Discrete Random Variables and Probability Models

Bernoulli Random Variable and Distribution

- The random variable is the number of successes in **one trial**
- There are only **2 disjoint and exhaustive outcomes of interest** in the trial (success, failure)
- The **probability of a "success"** is denoted by **p**

$\mathbf{X \sim Bernoulli\ (p)}$

$$\boxed{q = 1 - p}$$

$$\mathbf{pmf}: \ p(x) = \begin{cases} \mathbf{p} \ ; & \mathbf{x = 1} \\ \mathbf{1 - p} \ ; & \mathbf{x = 0} \\ \mathbf{0} \ ; & \mathbf{otherwise} \end{cases} \qquad \text{and} \quad 0 < p < 1$$

$\mathbf{Mean}: E(X) = p$

$\mathbf{Variance}: Var(X) = p(1 - p)$

Ex 1: Roll a regular fair die once and let X be the number of times number 4 is observed.

Are there only 2 disjoint and exhaustive outcomes of interest for each trial?

Binomial Random Variable and Distribution

- The random variable is the number of successes in **n trials.**
- There is a **fixed number of identical Bernoulli trials**. The total number of trials is denoted by **n**
- There are only **2 disjoint and exhaustive outcomes of interest** for each trial (success, failure)
- The trials are **independent**
- The **probability of a "success" is a constant** across the trials. The probability of success is denoted by **p**

$\mathbf{X \sim B\ (n,\ \ p)}$

$$\mathbf{pmf}: \ p(x) = P(X = x) = \binom{\mathbf{n}}{\mathbf{x}} \mathbf{p^x (1 - p)^{n - x}} \quad ; \quad x = 0, 1, 2, \dots, n \quad \text{and} \quad 0 < p < 1$$

$\mathbf{Mean}: E(X) = np$

$\mathbf{Variance}: Var(X) = np(1 - p)$

$$\boxed{\binom{\mathbf{n}}{\mathbf{x}} = {}^{\mathbf{n}}\mathbf{C_x} = \frac{\mathbf{n!}}{\mathbf{x!\ (n - x)!}} \quad \text{and} \quad q = 1 - p}$$

Ex 2: Suppose we roll a regular fair die three times and observe top face. Let X be the number of times we observe number 5. Is X a binomial random variable?

The support of $X =$

Does the experiment have a fixed number of trials (rolls)?

Are there only 2 disjoint and exhaustive outcomes of interest for each trial?

Are trials independent?

Does probability of a "success" remain constant across the trials?

In MATLAB: X has B(n,p)
When a is an integer from 0 to n,

Probability of observing at most a successes = P(X ≤ a) = binocdf(a, n, p)

Probability of observing more than a successes = P(X > a) = binocdf(a, n, p, 'upper')

Ex 3: Transmission Error
A binary communications channel introduces a bit error in a transmission with probability of 0.01. Let X be the number of errors in 10 independent transmissions.

(a) Write the pmf of X.

(b) Find the probability of 8 errors in 10 transmissions.

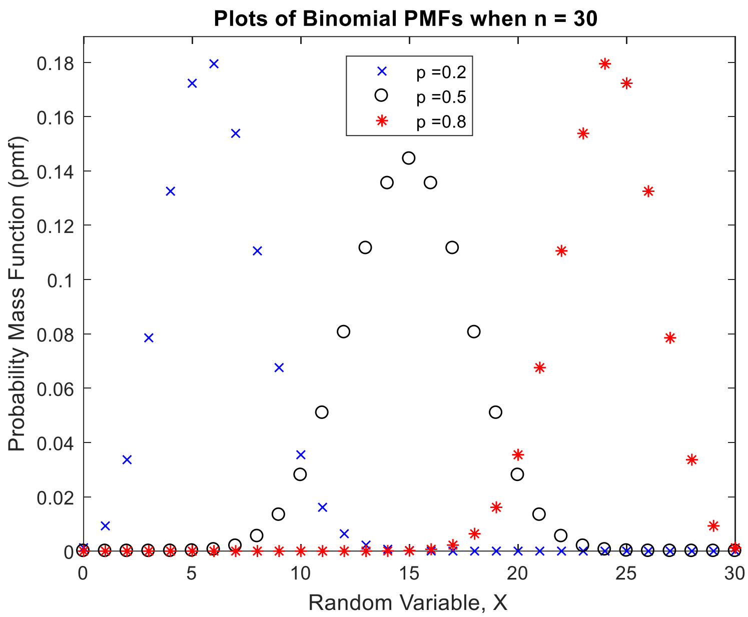
(c) Find the probability of one or fewer errors in 10 transmissions.

(d) Find the probability of at least 2 errors in 10 transmissions.

Key Questions:

- How is a Binomial random variable similar to a Bernoulli random variable?

- How are they different?




```
%% Plot of Binomial Distributions
n = 30;
x = 0:n;
p1=0.2; p2=0.5 ; p3= 0.8;
y1 = binopdf(x,n,p1);
y2 = binopdf(x,n,p2);
y3 = binopdf(x,n,p3);
plot(x,y1,'xb',x,y2,'ok', x,y3,'*r')
xlabel('Random Variable, X')
ylabel ('Probability Mass Function (pmf)')
title(['Plots of Binomial PMFs when n = '
num2str(n) ' '])
legend(strcat('p = ',num2str(p1)),
strcat('p = ',num2str(p2)), strcat('p =
',num2str(p3)),'Location','north')
ylim([0,max(max([y1(:),y2(:),y3(:)]))+.01])
```

- **If p < 0.5, then**

- **If p = 0.5, then**

- **If p > 0.5, then**

Ex 4: A 10-digit binary string is randomly generated. What is the probability that the ten digits (each of which is either 0 or 1) sum to 7?

Geometric Random Variable and Distribution

A random variable which gives **the number of the Bernoulli trial on which the first success occurs** is a **geometric random variable** and has a **geometric distribution**.

- There is **no fixed number of trials**. It is unknown the maximum times experiment should be repeated.
- There are only **2 outcomes of interest** for each trial (success, failure).
- The trials are independent.
- The **probability of a "success" is a constant** across the trials. The probability of success is denoted by $\mathbf{p}$

$\mathbf{X \sim Geometric\ (p)}$

pmf : $p(x) = P(X = x) = \mathbf{p(1-p)^{x-1}}$; $x = 1, 2, \ldots$ and $0 < p < 1$

Mean : $E(X) = \frac{1}{p}$

Variance : $Var(X) = \dfrac{1-p}{p^2}$

$q = 1 - p$

Geometric Series for $|r| < 1$: $\displaystyle\sum_{\mathbf{k=0}}^{\infty} (\mathbf{ar^k}) = \frac{\mathbf{a}}{\mathbf{1-r}}$

Key Questions:

- Why does the support of the geometric random variable start at 1 (not zero)?
- How is a Binomial random variable similar to a Geometric random variable?
- How are they different?

In MATLAB: Y = X – 1 is the number of failures before 1^{st} success

When a is an integer from 0,

Probability of observing at most a failures before 1^{st} success = P(X ≤ a) = geocdf(a, p)

Probability of observing more than a failures before 1^{st} success = P(X > a) = geocdf(a, p, 'upper')

Ex5: Message Transmission in a feedback channel

Let X be the number of times a message needs to be transmitted until it arrives correctly at its destination.

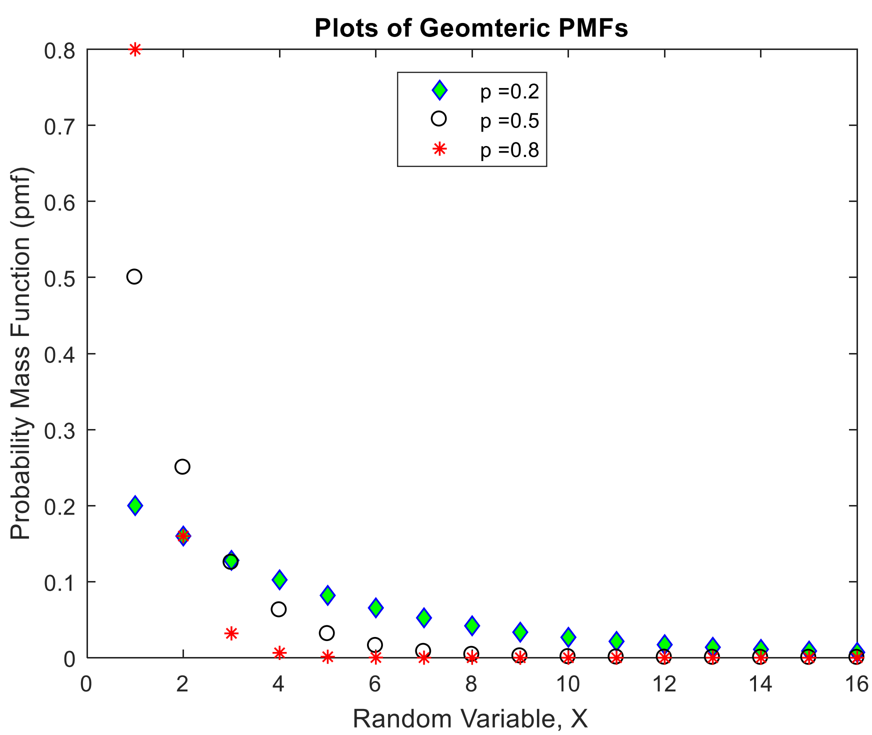
Let p be the probability a message arrives correctly at its destination.

(a) Write the pmf of X

(b) If $p = 0.8$, find the probability that a message needs to be transmitted at least 3 times until it arrives correctly at its destination.

(c) If $p = 0.8$, find the expected number of times a message needs to be transmitted until it arrives correctly at its destination.

(d) If $p = 0.8$, find the variance of number of times a message needs to be transmitted until it arrives correctly at its destination.




```
%% Plots of Geometric Distributions
x = 0:15;
p1=0.2; p2=0.5 ; p3= 0.8;
y1 = geopdf(x,p1);
y2 = geopdf(x,p2);
y3 = geopdf(x,p3);
plot(x+1,y1,'db',
'MarkerFaceColor',[0 1 0]) % Specify
fill color
hold on;
plot(x+1,y2,'ok', x+1,y3,'*r')
xlabel('Random Variable, X')
ylabel ('Probability Mass Function
(pmf)')
title('Plots of Geomteric PMFs')
legend(strcat('p = ',num2str(p1)),
strcat('p = ',num2str(p2)), strcat('p
= ',num2str(p3)),'Location','north')
```

Poisson Random Variable and Distribution

A random variable which represents **the number of occurrences (successes) over a certain amount of time or space** has a **Poisson distribution.** Trials are independent.

$\mathbf{X \sim P(\lambda)}$

$\mathbf{pmf}$: $p(x) = P(X = x) = \dfrac{\mathbf{e^{-\lambda}\lambda^{x}}}{\mathbf{x!}}$; $x = 0, 1, 2, \ldots$ and $\lambda > 0$

$\mathbf{Mean}$: $E(X) = \lambda$

$\mathbf{Variance}$: $Var(X) = \lambda$

Ex 6: Do the following random variables follow Poisson distribution? If yes, specify the value of λ:
A typist makes, on average, 3 typing errors per page.
If more than 4 errors appear on a given page, the typist must retype the whole page.

(a) Y is the number of errors per page:

(b) X is the number of errors per 5 pages:

(c) W is the number of retyped pages out of next 20 pages typed:

In MATLAB: X has P(λ)
When a is an integer from 0,

Probability of observing at most a successes = P(X ≤ a) = poisscdf(a, λ)

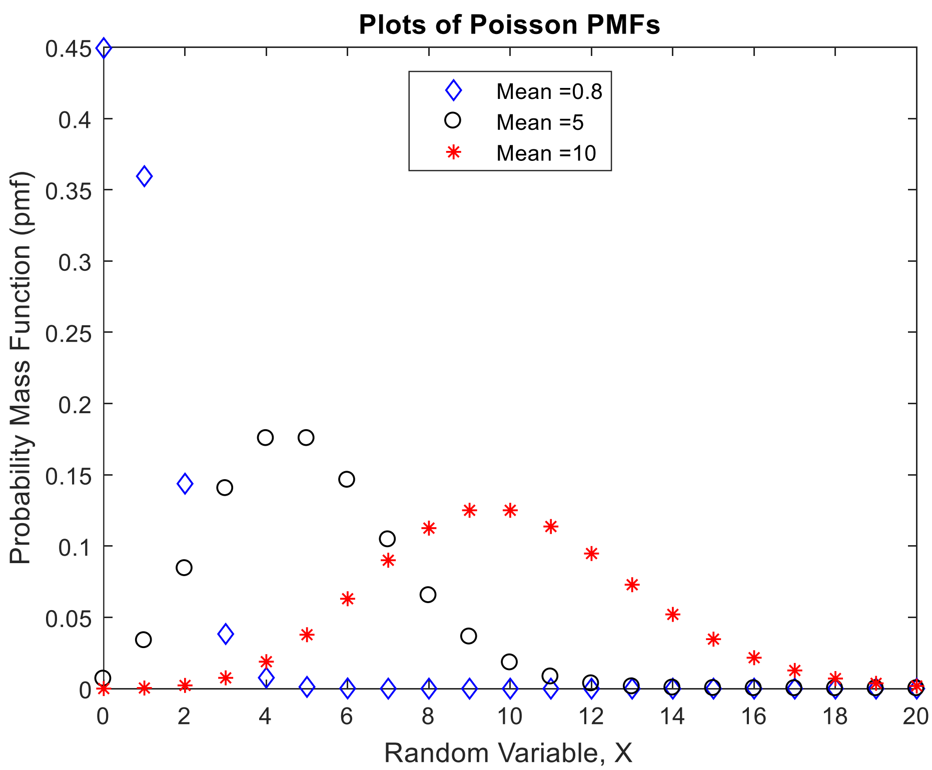
Probability of observing more than a successes = P(X > a) = poisscdf(a, λ, 'upper')

Ex 7: Queries at a Call Center
The number of queries arriving in *t* seconds at a call center is a Poisson random variable. Assume that the average arrival rate is four queries per minute.

Find the probability of the following events:
(a) 3 queries arrive in 10 seconds;

(b) more than 4 queries arrive in 10 seconds;

(c) fewer than 5 queries arrive in 2 minutes.




```
%% Plots of Poisson
Distributions
x = 0:20;
Mean1=0.8 ; Mean2 = 5; Mean3
=10;
y1 = poisspdf(x,Mean1);
y2 = poisspdf(x,Mean2);
y3 = poisspdf(x,Mean3);
figure(1);
plot(x,y1,'db',x,y2,'ok',
x,y3,'*r')
xlabel('Random Variable, X')
ylabel ('Probability Mass
Function (pmf)')
title('Plots of Poisson PMFs')
legend(strcat('Mean =
',num2str(Mean1)),
strcat('Mean =
',num2str(Mean2)),
strcat('Mean =
',num2str(Mean3)),'Location',
'north')
```

Statistical Inference using the Binomial Distribution
Under normal operating conditions 1.5% of the transistors produced in a factory are defective.
An inspector takes a random sample of forty transistors and finds that two are defective.

a) What is the probability that exactly two transistors will be defective from a random sample of forty under normal operating conditions?

b) What is the probability that more than two transistors will be defective from a random sample of forty if conditions are normal?