2.4 THE BINOMIAL DISTRIBUTION:

<u>Bernoulli Trials:</u> A **Bernoulli trial** is a random experiment with two possible outcomes, traditionally labeled "success" and "failure". The probability of success is traditionally denoted by $\boldsymbol{p}$. The probability of failure $(\mathbf{1}-\boldsymbol{p})$ is often denoted by $\boldsymbol{q}$. A **Bernoulli random variable** is simply the indicator of success in a Bernoulli trial.
Let $\boldsymbol{X}$ be a random variable associated with a Bernoulli trial by defining it as follows:

$$\boldsymbol{X}(\text{success}) = \mathbf{1} \quad \text{and} \quad \boldsymbol{X}(\text{failure}) = \mathbf{0}$$

That is, the two outcomes, success and failure, are denoted by **one** and **zero**, respectively. The ***pmf*** of $\boldsymbol{X}$ can be written as:

$$\boldsymbol{f(x) = p^x(1-p)^{1-x} = p^x q^{1-x}; \quad x = 0, 1.}$$

and we say that $\boldsymbol{X}$ has a **Bernoulli distribution**.
The expected value of $\boldsymbol{X}$ is $\boldsymbol{\mu} = \boldsymbol{E}(\boldsymbol{X}) = \boldsymbol{p} =$ Mean, Variance is $\boldsymbol{\sigma^2} = \boldsymbol{Var}(\boldsymbol{X}) = \boldsymbol{pq}$ and standard deviation is $\boldsymbol{\sigma} = \sqrt{\boldsymbol{p}(\mathbf{1}-\boldsymbol{p})} = \sqrt{\boldsymbol{pq}}$.

<u>The Binomial Distribution:</u> If there are $\boldsymbol{n}$ stochastically independent Bernoulli trials with the same probability $\boldsymbol{p}$ of success, the probability distribution of the number of successes is called the **Binomial distribution**. A **Binomial random variable** is simply the count of the number of successes in $\boldsymbol{n}$ trials. To get exactly $\boldsymbol{x}$ successes, there must be $(\boldsymbol{n}-\boldsymbol{x})$ failures. There are $\binom{\boldsymbol{n}}{\boldsymbol{x}} = \frac{\boldsymbol{n}!}{\boldsymbol{x}!(\boldsymbol{n}-\boldsymbol{x})!}$ such outcomes (where by convention $\mathbf{0}! = \mathbf{1}$), and by independence each has probability $\boldsymbol{p^x(1-p)^{1-x}}$.

N.B. In this case, since success and failure are not equally likely, so the points in the sample space are not equally likely. Simple counting is not going to be adequate. Thus,

$$\boldsymbol{P}(\boldsymbol{x} \text{ successes in } \boldsymbol{n} \text{ independent Bernoulli trials}) =$$
$$\boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) = \binom{\boldsymbol{n}}{\boldsymbol{x}} \boldsymbol{p^x(1-p)^{n-x}}; \quad \boldsymbol{x = 0, 1, 2, \ldots, n.}$$

These probabilities are called binomial probabilities, and the random variable $\boldsymbol{X}$ is said to have a **binomial distribution**.
Note that the Binomial random variable is simply the sum of the Bernoulli random variables for each trial.

A binomial distribution will be denoted by the symbol $\boldsymbol{b}(\boldsymbol{n},\ \boldsymbol{p})$, and we say that the distribution of $\boldsymbol{X}$ is $\boldsymbol{b}(\boldsymbol{n},\ \boldsymbol{p})$. The constants $\boldsymbol{n}$ and $\boldsymbol{p}$ are called the **parameters** of the binomial distribution; they correspond to the number $\boldsymbol{n}$ of independent trials and the probability $\boldsymbol{p}$ of success on each trial. Thus, if we say that the distribution of $\boldsymbol{X}$ is $\boldsymbol{b}(\mathbf{12},\ \frac{\mathbf{1}}{\mathbf{4}})$, we mean that $\boldsymbol{X}$ is the number of successes in a random sample of size $\boldsymbol{n} = \mathbf{12}$ from a Bernoulli distribution with $\boldsymbol{p} = \frac{\mathbf{1}}{\mathbf{4}}$.

The expected value of $\boldsymbol{X}$ is $\boldsymbol{\mu} = \boldsymbol{E}(\boldsymbol{X}) = \boldsymbol{M}'(\mathbf{0}) = \boldsymbol{np} =$ Mean and
Variance is $\boldsymbol{\sigma^2} = \boldsymbol{Var}(\boldsymbol{X}) = \boldsymbol{M}''(\mathbf{0}) - [\boldsymbol{M}'(\mathbf{0})]^2 = \boldsymbol{np}(\mathbf{1}-\boldsymbol{p}) = \boldsymbol{npq}$

The ***cdf*** of **binomial distribution** is $\boldsymbol{F(x) = P(X \leq x)}$

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

Examples: 2.4-1 to 2.4-9 (See yourself)
Exercises: 2.4-1 to 2.4-6 (Try yourself)

2.6 THE POISSON DISTRIBUTION:

Poisson Process: A Poisson Process is a model for a series of discrete event where the ***average time*** between events is known, but the exact timing of events is random. The arrival of an event is independent of the event before (waiting time between events is memory less). For example, suppose we own a website which our content delivery network (**CDN**) tells us goes down on average once per $\mathbf{60}$ days, but one failure doesn't affect the probability of the next. All we know is the average time between failures.

Definition: Let the number of occurrences of some event in a given continuous interval be counted. Then we have an **approximate Poisson process** with parameter $\boldsymbol{\lambda > 0}$ if the following conditions are satisfied:

- **(a)** The numbers of occurrences in non-overlapping subintervals are independent.
- **(b)** The probability of exactly one occurrence in a sufficiently short subinterval of length $\boldsymbol{h}$ is approximately $\boldsymbol{\lambda h}$.
- **(c)** The probability of two or more occurrences in a sufficiently short subinterval is essentially zero.

Let $\boldsymbol{X}$ denote the number of occurrences in an interval of length $\mathbf{1}$ (where "length $\mathbf{1}$" represents one unit of the quantity under consideration). We would like to find an approximation for $\boldsymbol{P(X = x)}$, where $\boldsymbol{x}$ is a nonnegative integer. To achieve this, we partition the unit interval into $\boldsymbol{n}$ subintervals of equal length $\frac{1}{n}$. If $\boldsymbol{n}$ is sufficiently large (i.e., much larger than $\boldsymbol{x}$), we shall approximate the probability that there are $\boldsymbol{x}$ occurrences in this unit interval by finding the probability that exactly $\boldsymbol{x}$ of these $\boldsymbol{n}$ subintervals each has one occurrence. The probability of one occurrence in any one subinterval of length $\frac{1}{n}$ is approximately $\lambda(\frac{1}{n})$, by condition **(b)**. The probability of two or more occurrences in any one subinterval is essentially **zero**, by condition **(c)**. So, for each subinterval, there is exactly one occurrence with a probability of approximately $\lambda(\frac{1}{n})$. Consider the **occurrence** or **nonoccurrence** in each subinterval as a **Bernoulli trial**. By condition **(a)**, we have a sequence of $\boldsymbol{n}$ Bernoulli trials with probability $\boldsymbol{p}$ approximately equal to $\lambda(\frac{1}{n})$. Thus, an approximation for $\boldsymbol{P(X = x)}$ is given by the binomial probability:

$$\binom{n}{x}\left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

The ***pmf*** of the **Poisson distribution** for the discrete random variable $\boldsymbol{X}$ is

$$P(X = x) = f(x) = \frac{\lambda^x e^{-\lambda}}{x!}; \quad \text{where } \lambda > 0 \text{ and } x = 0, 1, 2, 3, \ldots$$

The expected value of $\boldsymbol{X}$ is $\boldsymbol{\mu} = \boldsymbol{E}(\boldsymbol{X}) = \boldsymbol{M}'(\boldsymbol{0}) = \boldsymbol{\lambda} =$ Mean and Variance is $\boldsymbol{\sigma^2} = \boldsymbol{Var}(\boldsymbol{X}) = \boldsymbol{M}''(\boldsymbol{0}) - [\boldsymbol{M}'(\boldsymbol{0})]^2 = (\boldsymbol{\lambda^2} + \boldsymbol{\lambda}) - \boldsymbol{\lambda^2} = \boldsymbol{\lambda}$

The ***cdf*** of **Poisson distribution** is $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{P}(\boldsymbol{X} \leq \boldsymbol{x})$

$$= \sum_{\boldsymbol{k}=\boldsymbol{0}}^{\boldsymbol{x}} \frac{\boldsymbol{\lambda^k e^{-\lambda}}}{\boldsymbol{k}!}$$

Examples: 2.6-1 to 2.6-5 (See yourself)
Exercises: 2.6-1 to 2.6-9 (Try yourself)