

INTRODUCTION TO PROBABILITY DISTRIBUTION

KRISHAN K. PANDEY (Ph.D.)

Professor & Dean 'Office of Doctoral Studies'
O.P. Jindal Global University, Sonapat, Haryana-131001

[Website: www.krishanpandey.com](http://www.krishanpandey.com)

Basic Statistics

Consider the different possible orderings of boy (B) and girl (G) in four sequential births. There are $2*2*2*2=2^4 = 16$ possibilities, so the sample space is:

BBBB	BGBB	GBBB	GGBB
BBBG	BGBG	GBBG	GGBG
BBGB	BGGB	GBGB	GGGB
BBGG	BGGG	GBGG	GGGG

If girl and boy are each equally likely [$P(G)=P(B) = 1/2$], and the gender of each child is independent of that of the previous child, then **the probability of each of these 16 possibilities is:**
 $(1/2)(1/2)(1/2)(1/2) = 1/16$.

Random Variables

Now count the **number of girls** in each set of four sequential births:

BBBB	(0)	BGBB	(1)	GBBB	(1)	GGBB	(2)
BBBG	(1)	BGBG	(2)	GBBG	(2)	GGBG	(3)
BBGB	(1)	BGGB	(2)	GBGB	(2)	GGGB	(3)
BBGG	(2)	BGGG	(3)	GBGG	(3)	GGGG	(4)

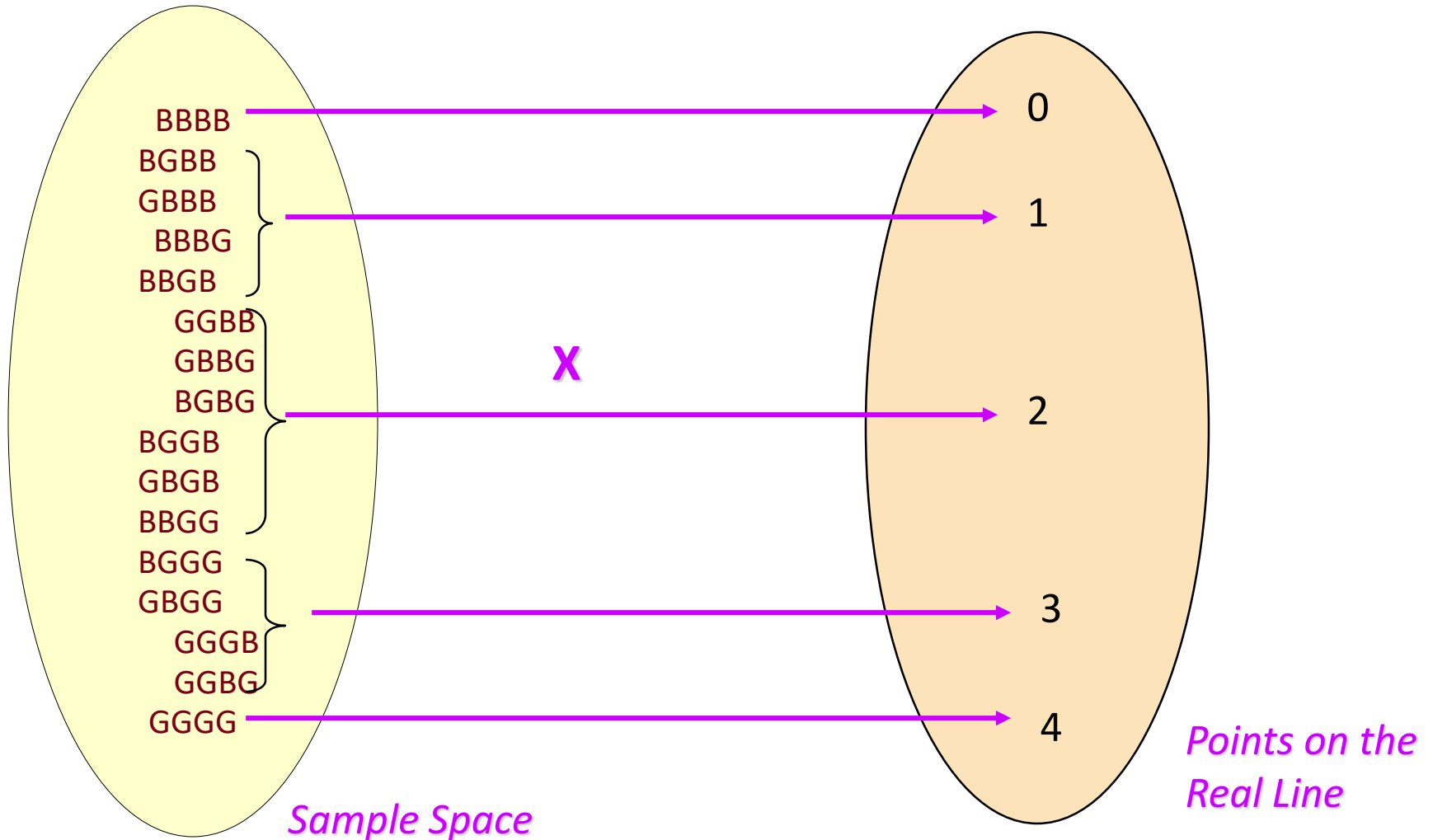
Notice that:

- each possible outcome is assigned a single numeric value,
- all outcomes are assigned a numeric value, and
- the value assigned varies over the outcomes.

The count of the number of girls is a **random variable**:

A **random variable, X** , is a function that assigns a single, but variable, value to each element of a sample space.

Random Variables (Continued)



Random Variables (Continued)

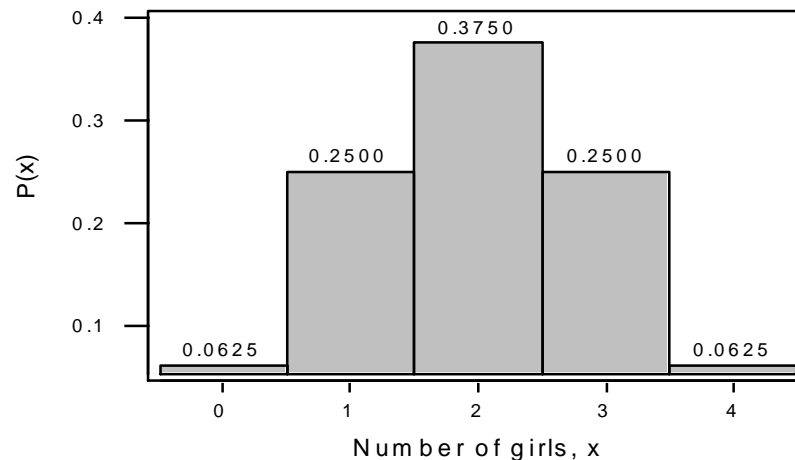
Since the random variable $X = 3$ when any of the four outcomes BGGG, GBGG, GGBG, or GGGB occurs,

$$P(X = 3) = P(\text{BGGG}) + P(\text{GBGG}) + P(\text{GGBG}) + P(\text{GGGB}) = 4/16$$

The **probability distribution** of a random variable is a table that lists the possible values of the random variables and their associated probabilities.

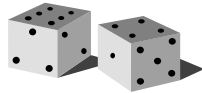
<u>x</u>	<u>P(x)</u>
0	1/16
1	4/16
2	6/16
3	4/16
4	1/16
	16/16=1

Probability Distribution of the Number of Girls in Four Births



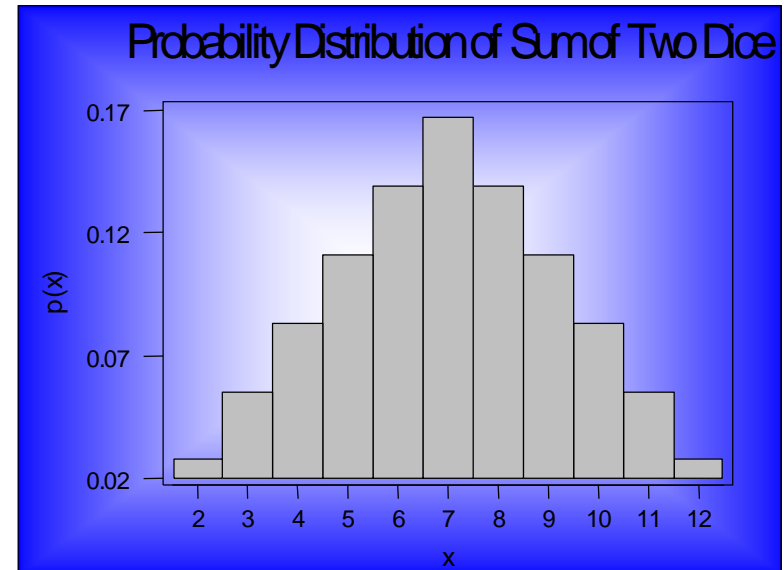
Example 1

Consider the experiment of **tossing two six-sided dice**. There are **36 possible outcomes**. Let the random variable **X** represent the **sum of the numbers on the two dice**:



	2	3	4	5	6	7
1,1	1,2	1,3	1,4	1,5	1,6	8
2,1	2,2	2,3	2,4	2,5	2,6	9
3,1	3,2	3,3	3,4	3,5	3,6	10
4,1	4,2	4,3	4,4	4,5	4,6	11
5,1	5,2	5,3	5,4	5,5	5,6	12
6,1	6,2	6,3	6,4	6,5	6,6	

<u>x</u>	<u>P(x)*</u>
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

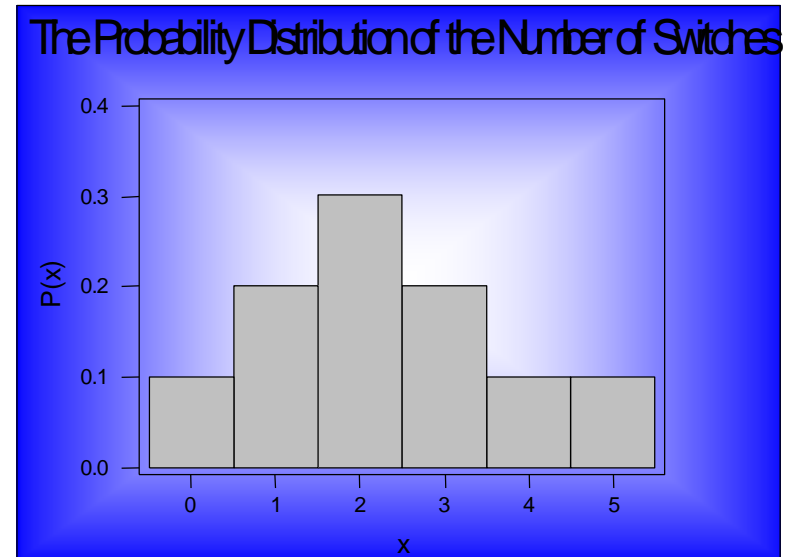


* Note that: $P(x) = (6 - \sqrt{(7-x)^2}) / 36$

Example 2

Probability Distribution of the Number of Switches

<u>x</u>	<u>P(x)</u>
0	0.1
1	0.2
2	0.3
3	0.2
4	0.1
5	<u>0.1</u>
	1



Probability of more than 2 switches:

$$P(X > 2) = P(3) + P(4) + P(5) = 0.2 + 0.1 + 0.1 = 0.4$$

Probability of at least 1 switch:

$$P(X \geq 1) = 1 - P(0) = 1 - 0.1 = .9$$

Discrete and Continuous Random Variables

A **discrete random variable**:

- has a countable number of possible values
- has discrete jumps (or gaps) between successive values
- has measurable probability associated with individual values
- counts

A **continuous random variable**:

- has an uncountably infinite number of possible values
- moves continuously from value to value
- has no measurable probability associated with each value
- measures (e.g.: height, weight, speed, value, duration,

length)

Rules of Discrete Probability Distributions

The probability distribution of a discrete random variable X **must** satisfy the following two conditions.

1. $P(x) \geq 0$ for all values of x .

2.
$$\sum_{\text{all } x} P(x) = 1$$

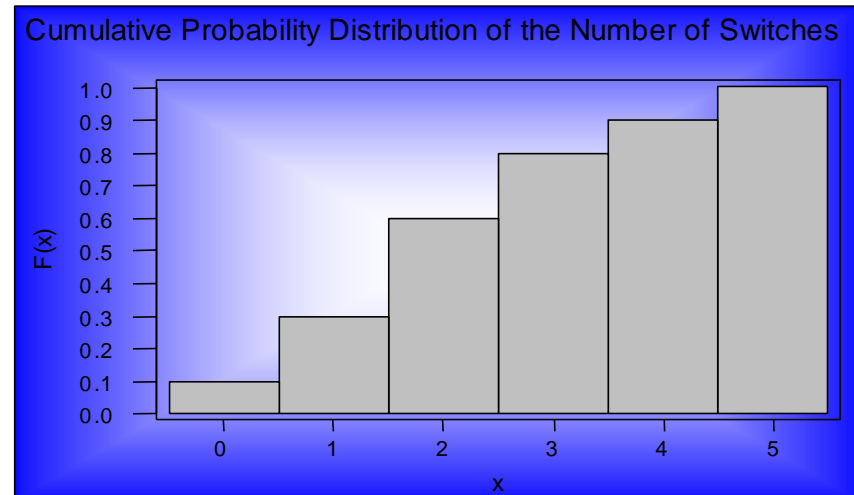
[Corollary: $0 \leq P(X) \leq 1$]

Cumulative Distribution Function

The **cumulative distribution function**, $F(x)$, of a discrete random variable X is:

$$F(x) = P(X \leq x) = \sum_{\text{all } i \leq x} P(i)$$

<u>x</u>	<u>$P(x)$</u>	<u>$F(x)$</u>
0	0.1	0.1
1	0.2	0.3
2	0.3	0.6
3	0.2	0.8
4	0.1	0.9
5	<u>0.1</u>	<u>1.0</u>
	1	

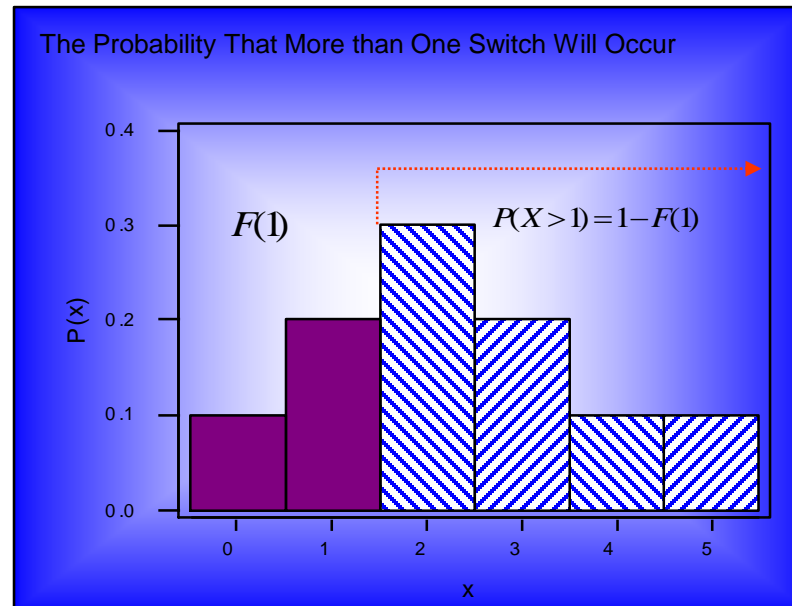


Using Cumulative Probability Distributions

The probability that **more than one switch** will occur:

x	$P(x)$	$F(x)$
0	0.1	0.1
1	0.2	0.3
2	0.3	0.6
3	0.2	0.8
4	0.1	0.9
5	0.1	1.0

1

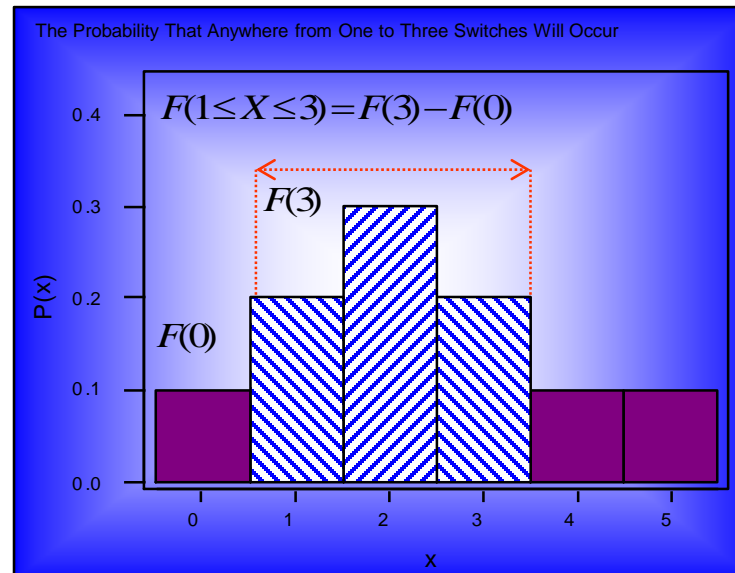


Note: $P(X > 1) = P(X \geq 2) = 1 - P(X \leq 1) = 1 - F(1) = 1 - 0.3 = 0.7$

Using Cumulative Probability Distributions (Figure 3-9)

The probability that anywhere **from one to three switches** will occur:

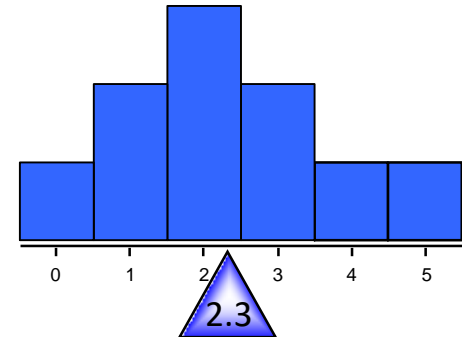
<u>x</u>	<u>$P(x)$</u>	<u>$F(x)$</u>
0	0.1	0.1
1	0.2	0.3
2	0.3	0.6
3	0.2	0.8
4	0.1	0.9
5	0.1	1.0
	<u>1</u>	



Note: $P(1 \leq X \leq 3) = P(X \leq 3) - P(X \leq 0) = F(3) - F(0) = 0.8 - 0.1 = 0.7$

3-2 Expected Values of Discrete Random Variables

The mean of a probability distribution is a measure of its centrality or location, as is the mean or average of a frequency distribution. It is a **weighted average**, with the values of the random variable weighted by their probabilities.



The mean is also known as the **expected value** (or **expectation**) of a random variable, because it is the value that is expected to occur, on average.

The **expected value** of a discrete random variable X is equal to the sum of each value of the random variable **multiplied** by its probability.

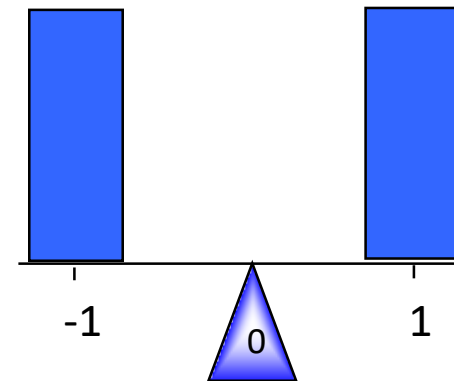
$$\mu = E(X) = \sum_{\text{all } x} xP(x)$$

<u>x</u>	<u>$P(x)$</u>	<u>$xP(x)$</u>
0	0.1	0.0
1	0.2	0.2
2	0.3	0.6
3	0.2	0.6
4	0.1	0.4
5	0.1	0.5
1.0		2.3 = $E(X) = \mu$

A Fair Game

Suppose you are playing a coin toss game in which you are paid \$1 if the coin turns up heads and you lose \$1 when the coin turns up tails. The expected value of this game is $E(X) = 0$. A game of chance with an expected payoff of 0 is called a *fair game*.

x	$P(x)$	$xP(x)$
-1	0.5	-0.50
1	0.5	0.50
1.0		0.00 = $E(X) = \mu$



Expected Value of a Function of a Discrete Random Variables

The **expected value of a function** of a discrete random variable X is:

$$E[h(X)] = \sum_{\text{all } x} h(x)P(x)$$

Example 3-3: Monthly sales of a certain product are believed to follow the given probability distribution. Suppose the company has a fixed monthly production cost of \$8000 and that each item brings \$2. Find the expected monthly profit $h(X)$, from product sales.

<i>Number of items, x</i>	<i>P(x)</i>	<i>xP(x)</i>	<i>h(x)</i>	<i>h(x)P(x)</i>
5000	0.2	1000	2000	400
6000	0.3	1800	4000	1200
7000	0.2	1400	6000	1200
8000	0.2	1600	8000	1600
9000	0.1	900	10000	1000
	1.0	6700		5400

$$E[h(X)] = \sum_{\text{all } x} h(x)P(x) = 5400$$

Note: $h(X) = 2X - 8000$ where $X = \#$ of items sold

The expected value of a **linear function** of a random variable is:

$$E(aX+b) = aE(X) + b$$

In this case: $E(2X-8000) = 2E(X) - 8000 = (2)(6700) - 8000 = 5400$

Variance and Standard Deviation of a Random Variable

The ***variance*** of a random variable is the expected squared deviation from the mean:

$$\begin{aligned}\sigma^2 &= V(X) = E[(X - \mu)^2] = \sum_{\text{all } x} (x - \mu)^2 P(x) \\ &= E(X^2) - [E(X)]^2 = \left[\sum_{\text{all } x} x^2 P(x) \right] - \left[\sum_{\text{all } x} x P(x) \right]^2\end{aligned}$$

The ***standard deviation*** of a random variable is the square root of its variance: $\sigma = SD(X) = \sqrt{V(X)}$

Variance and Standard Deviation of a Random Variable – using Example 2

Table 3-8

Number of Switches, x	$P(x)$	$xP(x)$	$(x-\mu)$	$(x-\mu)^2$	$P(x-\mu)^2$	$\underline{x^2P(x)}$
0	0.1	0.0	-2.3	5.29	0.529	0.0
1	0.2	0.2	-1.3	1.69	0.338	0.2
2	0.3	0.6	-0.3	0.09	0.027	1.2
3	0.2	0.6	0.7	0.49	0.098	1.8
4	0.1	0.4	1.7	2.89	0.289	1.6
5	0.1	<u>0.5</u>	2.7	7.29	<u>0.729</u>	<u>2.5</u>
		2.3			2.010	7.3

Recall: $\mu = 2.3$.

$$\begin{aligned} \sigma^2 &= V(X) = E[(X - \mu)^2] \\ &= \sum_{\text{all } x} (x - \mu)^2 P(x) = 2.01 \\ &= E(X^2) - [E(X)]^2 \\ &= \left[\sum_{\text{all } x} x^2 P(x) \right] - \left[\sum_{\text{all } x} x P(x) \right]^2 \\ &= 7.3 - 2.3^2 = 2.01 \end{aligned}$$

Variance of a Linear Function of a Random Variable

The variance of a **linear function** of a random variable is:

$$V(aX+b) = a^2V(X) = a^2\sigma^2$$

Example 3-3:

Number of items, x	P(x)	xP(x)	x ² P(x)
5000	0.2	1000	5000000
6000	0.3	1800	10800000
7000	0.2	1400	9800000
8000	0.2	1600	12800000
9000	0.1	900	8100000
	1.0	6700	46500000

$$\sigma^2 = V(X)$$

$$= E(X^2) - [E(X)]^2$$

$$= \left[\sum_{all\ x} x^2 P(x) \right] - \left[\sum_{all\ x} xP(x) \right]^2$$

$$= 46500000 - (6700^2) = 1610000$$

$$\sigma = SD(X) = \sqrt{1610000} = 1268.86$$

$$V(2X - 8000) = (2^2)V(X)$$

$$= (4)(1610000) = 6440000$$

$$\sigma_{(2x-8000)} = SD(2x - 8000)$$

$$= 2\sigma_x = (2)(1268.86) = 2537.72$$

Some Properties of Means and Variances of Random Variables

The **mean or expected value of the *sum*** of random variables is the sum of their means or expected values:

$$\mu_{(X+Y)} = E(X+Y) = E(X) + E(Y) = \mu_X + \mu_Y$$

For example: $E(X) = \$350$ and $E(Y) = \$200$

$$\longrightarrow E(X+Y) = \$350 + \$200 = \$550$$

The **variance of the *sum* of **independent**** random variables is the sum of their variances:

$$\sigma^2_{(X+Y)} = V(X+Y) = V(X) + V(Y) = \sigma^2_X + \sigma^2_Y$$

if and only if X and Y are ***independent***.

For example: $V(X) = 84$ and $V(Y) = 60$ $\longrightarrow V(X+Y) = 144$

3-3 Bernoulli Random Variable

- If an experiment consists of a single trial and the outcome of the trial can only be either a success* or a failure, then the trial is called a **Bernoulli trial**.
- The number of success X in one Bernoulli trial, which can be 1 or 0, is a **Bernoulli random variable**.
- **Note:** If p is the probability of success in a Bernoulli experiment, the $E(X) = p$ and $V(X) = p(1 - p)$.

* The terms *success* and *failure* are simply statistical terms, and do not have positive or negative implications. In a production setting, finding a defective product may be termed a “success,” although it is not a positive result.

The Binomial Random Variable

Consider a **Bernoulli Process** in which we have a sequence of n identical trials satisfying the following conditions:

1. Each trial has two possible outcomes, called **success** *and **failure**. The two outcomes are **mutually exclusive** and **exhaustive**.
2. The **probability of success**, denoted by p , remains **constant** from trial to trial. The **probability of failure** is denoted by q , where $q = 1-p$.
3. The n trials are **independent**. That is, the outcome of any trial does not affect the outcomes of the other trials.

A random variable, X , that counts the number of successes in n Bernoulli trials, where p is the probability of success* in any given trial, is said to follow the **binomial probability distribution** with parameters n (number of trials) and p (probability of success). We call X the **binomial random variable**.

* The terms *success* and *failure* are simply statistical terms, and do not have positive or negative implications. In a production setting, finding a defective product may be termed a “success,” although it is not a positive result.

Binomial Probabilities (Introduction)

Suppose we toss a single fair and balanced coin five times in succession, and let X represent the number of heads.

There are $2^5 = 32$ possible sequences of H and T (S and F) in the sample space for this experiment. Of these, there are 10 in which there are exactly 2 heads ($X=2$):

HHTTT HTHTH HTTHT HTTTH THHTT THTHT THTTH TTHHT TTHTH TTTHH

The probability of each of these 10 outcomes is $p^2q^3 = (1/2)^2(1/2)^3 = (1/32)$, so the probability of 2 heads in 5 tosses of a fair and balanced coin is:

$$P(X = 2) = 10 * (1/32) = (10/32) = .3125$$

10	(1/32)
Number of outcomes with 2 heads	Probability of each outcome with 2 heads

Binomial Probabilities (continued)

$$P(X=2) = 10 * (1/32) = (10/32) = .3125$$

Notice that this probability has two parts:

10	(1/32)
Number of outcomes with 2 heads	Probability of each outcome with 2 heads

In general:

1. The **probability of a given sequence** of x successes out of n trials with probability of success p and probability of failure q is equal to:
2. The **number of different sequences** of n trials that result in exactly x successes is equal to the number of choices of x elements out of a total of n elements. This number is denoted:

$$p^x q^{(n-x)}$$

$$nC_x = \binom{n}{x} = \frac{n!}{x!(n-x)!}$$

The Binomial Probability Distribution

The **binomial probability distribution**:

$$P(x) = \binom{n}{x} p^x q^{(n-x)} = \frac{n!}{x!(n-x)!} p^x q^{(n-x)}$$

where :

p is the probability of success in a single trial,

$q = 1-p$,

n is the number of trials, and

x is the number of successes.

Number of
successes, x

Probability $P(x)$

0	$\frac{n!}{0!(n-0)!} p^0 q^{(n-0)}$
1	$\frac{n!}{1!(n-1)!} p^1 q^{(n-1)}$
2	$\frac{n!}{2!(n-2)!} p^2 q^{(n-2)}$
3	$\frac{n!}{3!(n-3)!} p^3 q^{(n-3)}$
⋮	⋮
n	$\frac{n!}{n!(n-n)!} p^n q^{(n-n)}$
	1.00

The Cumulative Binomial Probability Table (Table 1, Appendix C)

n=5

x	p												
	0.01	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99
0	.951	.774	.590	.328	.168	.078	.031	.010	.002	.000	.000	.000	.000
1	.999	.977	.919	.737	.528	.337	.187	.087	.031	.007	.000	.000	.000
2	1.000	.999	.991	.942	.837	.683	.500	.317	.163	.058	.009	.001	.000
3	1.000	1.000	1.000	.993	.969	.913	.813	.663	.472	.263	.081	.023	.001
4	1.000	1.000	1.000	1.000	.998	.990	.969	.922	.832	.672	.410	.226	.049

Cumulative Binomial Probability Distribution and Binomial Probability Distribution of H, the Number of Heads Appearing in Five Tosses of a Fair Coin	h	F(h)	P(h)
	0	0.031	0.031
	1	0.187	0.156
	2	0.500	0.313
	3	0.813	0.313
	4	0.969	0.156
	5	1.000	0.031
			1.000

Deriving Individual Probabilities from Cumulative Probabilities

$$F(x) = P(X \leq x) = \sum_{\text{all } i \leq x} P(i)$$

$$P(X) = F(x) - F(x - 1)$$

For example:

$$P(3) = F(3) - F(2)$$

$$= .813 - .500$$

$$= .313$$

Calculating Binomial Probabilities - Example

60% of Brooke shares are owned by LeBow. A random sample of 15 shares is chosen. What is the probability that **at most three** of them will be found to be owned by LeBow?

	p		
	.50	.60	.70
0	.000	.000	.000
1	.000	.000	.000
2	.004	.000	.000
3	.018	.002	.000
4	.059	.009	.001
...

$$F(x) = P(X \leq x) = \sum_{\text{all } i \leq x} P(i)$$

$$F(3) = P(X \leq 3) = .002$$

Mean, Variance, and Standard Deviation of the Binomial Distribution

Mean of a binomial distribution:

$$\mu = E(X) = np$$

Variance of a binomial distribution:

$$\sigma^2 = V(X) = npq$$

Standard deviation of a binomial distribution:

$$\sigma = SD(X) = \sqrt{npq}$$

For example, if H counts the number of heads in five tosses of a fair coin :

$$\mu_H = E(H) = (5)(.5) = 2.5$$

$$\sigma_H^2 = V(H) = (5)(.5)(.5) = 1.25$$

$$\sigma_H = SD(H) = \sqrt{1.25} = 1.118$$

The Poisson Distribution

The **Poisson probability distribution** is useful for determining the probability of a number of occurrences over a given period of time or within a given area or volume. That is, the Poisson random variable counts occurrences over a continuous interval of time or space. It can also be used to calculate approximate binomial probabilities when the probability of success is small ($p \leq 0.05$) and the number of trials is large ($n \geq 20$).

Poisson Distribution:

$$P(x) = \frac{\mu^x e^{-\mu}}{x!} \quad \text{for } x = 1, 2, 3, \dots$$

where μ is the mean of the distribution (which also happens to be the variance) and e is the base of natural logarithms ($e=2.71828\dots$).

The Poisson Distribution - Example

Example :

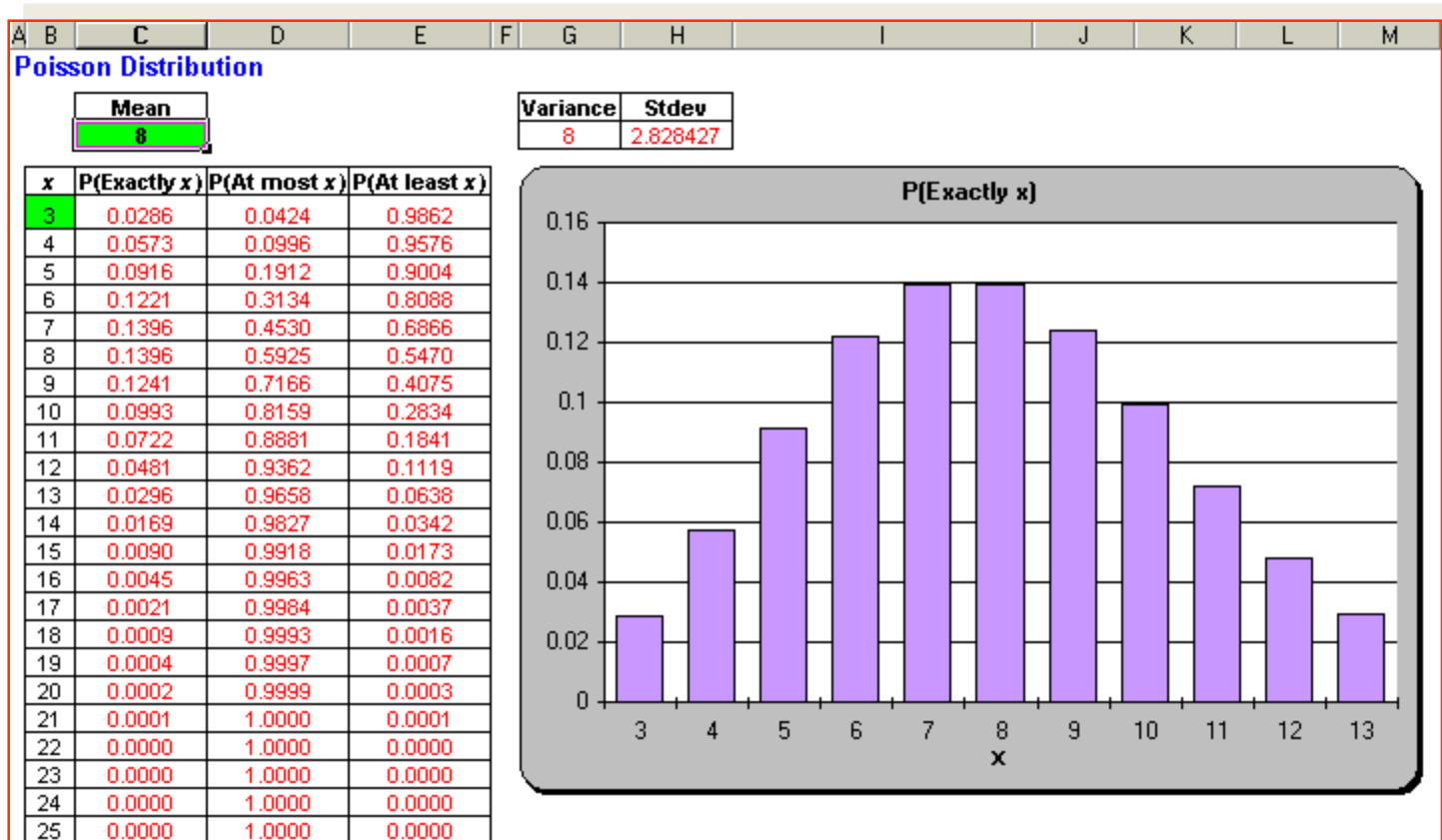
Telephone manufacturers now offer 1000 different choices for a telephone (as combinations of color, type, options, portability, etc.). A company is opening a large regional office, and each of its 200 managers is allowed to order his or her own choice of a telephone. Assuming independence of choices and that each of the 1000 choices is equally likely, what is the probability that a particular choice will be made by none, one, two, or three of the managers?

$$\begin{aligned}P(0) &= \frac{.2^0 e^{-.2}}{0!} = 0.8187 \\P(1) &= \frac{.2^1 e^{-.2}}{1!} = 0.1637 \\P(2) &= \frac{.2^2 e^{-.2}}{2!} = 0.0164 \\P(3) &= \frac{.2^3 e^{-.2}}{3!} = 0.0011\end{aligned}$$

$$n = 200 \quad \mu = np = (200)(0.001) = 0.2$$

$$p = 1/1000 = 0.001$$

Calculating Poisson Distribution Probabilities using the Template



The Poisson Distribution (continued)

- **Poisson assumptions:**

- ✓ The probability that an event will occur in a short interval of time or space is proportional to the size of the interval.
- ✓ In a very small interval, the probability that two events will occur is close to zero.
- ✓ The probability that any number of events will occur in a given interval is independent of where the interval begins.
- ✓ The probability of any number of events occurring over a given interval is independent of the number of events that occurred prior to the interval.

Discrete and Continuous Random Variables - Revisited

- A **discrete random variable**:

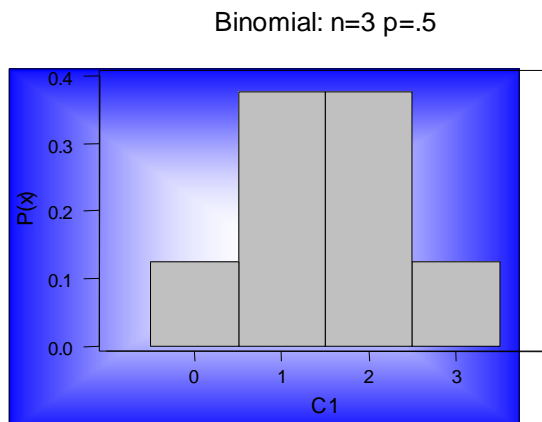
- counts occurrences
- has a countable number of possible values
- has discrete jumps between successive values
- **has measurable probability associated with individual values**
- **probability is height**

For example:

Binomial

n=3 p=.5

<u>x</u>	<u>P(x)</u>
0	0.125
1	0.375
2	0.375
3	<u>0.125</u>
	1.000



- A **continuous random variable**:

- measures (e.g.: height, weight, speed, value, duration, length)
- has an uncountably infinite number of possible values
- moves continuously from value to value
- **has no measurable probability associated with individual values**
- **probability is area**

For example:

In this case, the shaded area represents the probability that the task takes between 2 and 3 minutes.



Continuous Random Variables

A **continuous random variable** is a random variable that can take on any value in an interval of numbers.

The probabilities associated with a continuous random variable X are determined by the **probability density function** of the random variable. The function, denoted $f(x)$, has the following properties.

1. $f(x) \geq 0$ for all x .
2. The probability that X will be between two numbers a and b is equal to the area under $f(x)$ between a and b .
3. The total area under the curve of $f(x)$ is equal to 1.00.

The **cumulative distribution function** of a continuous random variable:

$F(x) = P(X \leq x)$ = Area under $f(x)$ between the smallest possible value of X (often $-\infty$) and the point x .