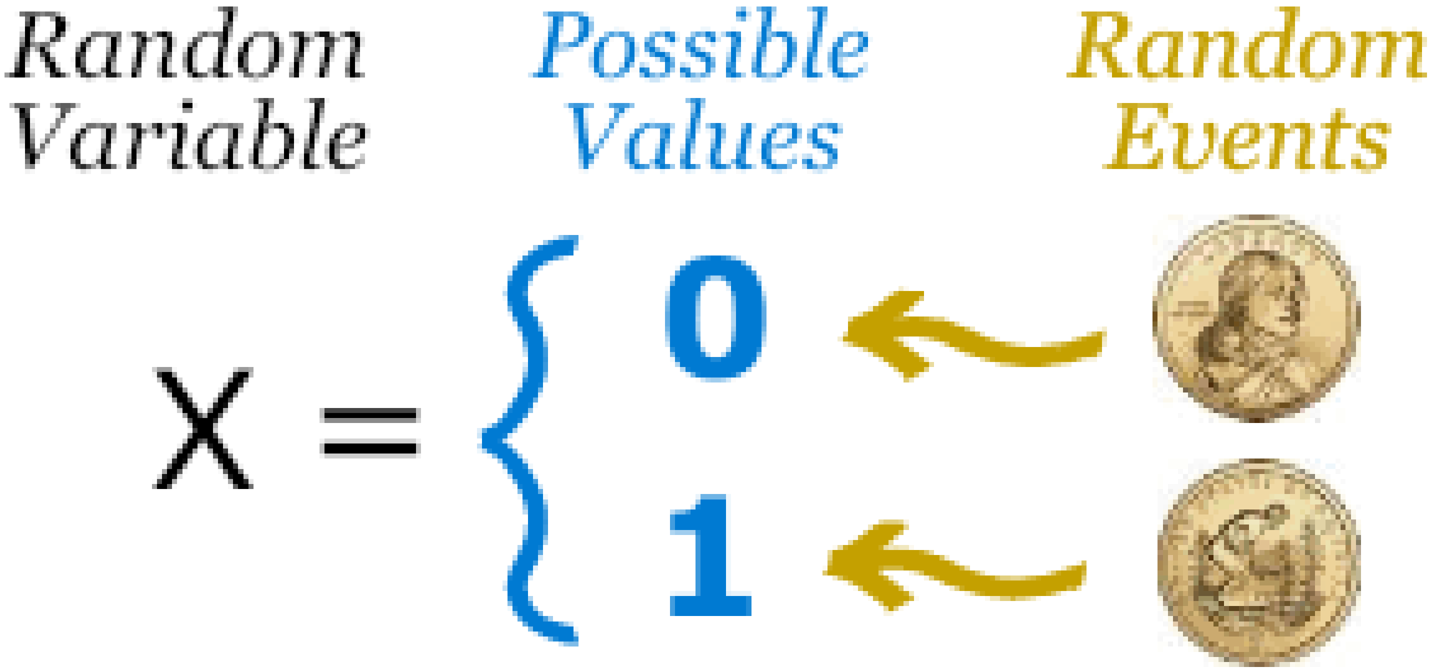
Discrete Random Variable



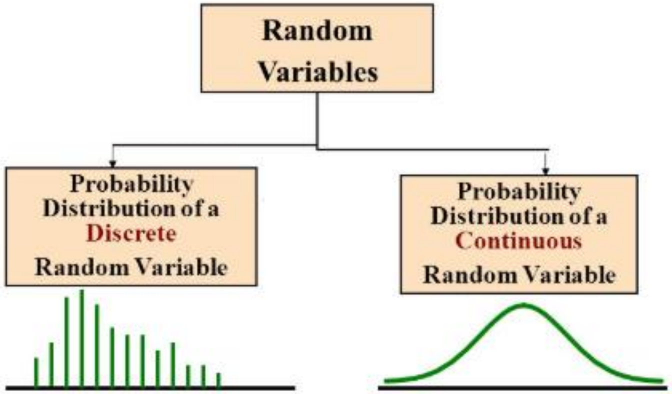
Random Variables

- A quantitative variable x is a **random variable** if the value that it assumes, corresponding to the outcome of an experiment is a chance or random event.
- **Random Variable (RV):** A numeric outcome that results from an experiment.
- For each element of an experiment's sample space, the random variable can take on exactly one value
- **Examples:**
  - ✓ x = SAT score for a randomly selected student
  - ✓ x = number of people in a room at a randomly selected time of day
  - ✓ x = number on the upper face of a randomly tossed die

Two Types of Random Variables

- **Random variables can be discrete or continuous.**

- **Discrete Random Variable**: An RV that can take on only a finite or countably infinite set of outcomes
- **Continuous Random Variable**: An RV that can take on any value along a continuum
- Random Variables are denoted by upper case letters (*Y*)
- Individual outcomes for an RV are denoted by lower case letters (*y*)



Discrete Random Variable Examples

Experiment	Random Variable	Possible Values
Make 100 Sales Calls	# Sales	0, 1, 2, ..., 100
Inspect 70 Radios	# Defective	0, 1, 2, ..., 70
Answer 33 Questions	# Correct	0, 1, 2, ..., 33
Count Cars at Toll Between 11:00 & 1:00	# Cars Arriving	0, 1, 2, ..., ∞

Continuous Random Variable

Random variables that can assume values corresponding to any of the points contained in one or more intervals (i.e., values that are infinite and uncountable) are called **continuous**.

Continuous Random Variable Examples

Experiment	Random Variable	Possible Values
Weigh 100 People	Weight	45.1, 78, ...
Measure Part Life	Hours	900, 875.9, ...
Amount spent on food	$ amount	54.12, 42, ...
Measure Time Between Arrivals	Inter-Arrival Time	0, 1.3, 2.78, ...

Probability Distributions for Discrete Random Variables

- The **probability distribution for a discrete random variable *x*** resembles the relative frequency distributions we constructed previously. It is a graph, table or formula that gives the possible values of *X* and the probability *p(X=x)* associated with each value.
- The probability distribution for a discrete variable X must satisfy the following two conditions:

1) $0 \leq p(X = x) \leq 1, and$
2) $\sum p(X=x)=1$

Additional Examples of Discrete Random Variables

1- The number of houses in a certain block.
2- The number of Female students in our class.
3- The number of kids in a given family.
4- The number of cars sold at a dealership during a month.
5- The number boys in families with 4 children.
6- The number of fish caught on a fishing trip.
7- The number of complaints received at the office of an airline on a given day.
8- The number of customers who visit a pharmacy during any given hour.
9- The number of heads obtained in three tosses of a coin.
10- The number of typos in a textbook consists of 100 pages.
11- The number of visitors to PETRA in a day.
12- The number of episodes of otitis media in the first 2 years of life (A common disease of the middle ear in early childhood).

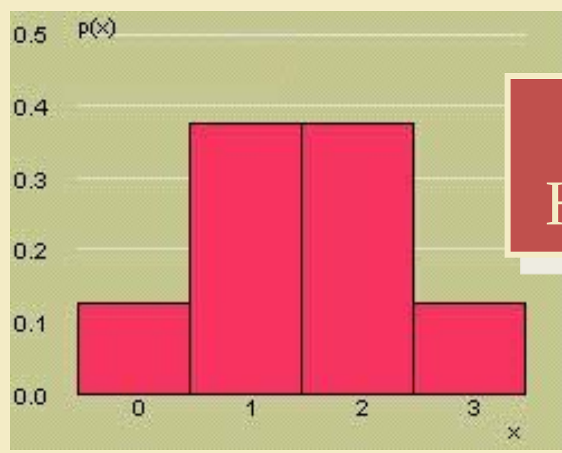
Example

- Toss a fair coin three times and define ***X* = number of heads.**

		x
HHH	1/8	3
HHT	1/8	2
HTH	1/8	2
THH	1/8	2
HTT	1/8	1
THT	1/8	1
TTH	1/8	1
TTT	1/8	0

$P(X = 0) = 1/8$
$P(X = 1) = 3/8$
$P(X = 2) = 3/8$
$P(X = 3) = 1/8$

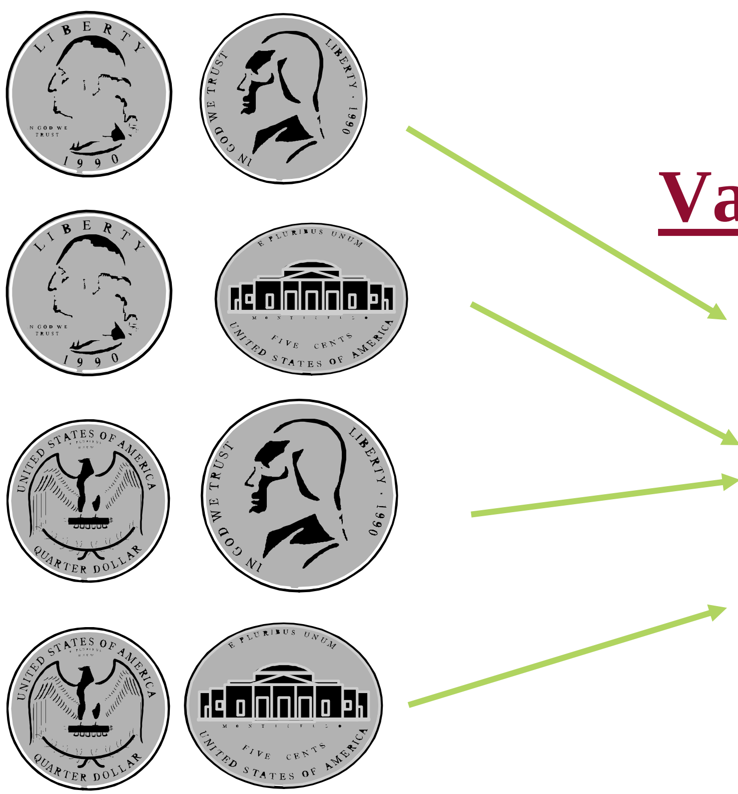
x	$p(x)$
0	1/8
1	3/8
2	3/8
3	1/8



Probability Histogram for x

Discrete Probability Distribution Example

Experiment: Toss 2 coins. Count number of tails.



Probability Distribution

Values, x	Probabilities, $p(x)$
0	1/4 = .25
1	2/4 = .50
2	1/4 = .25

Visualizing Discrete Probability Distributions

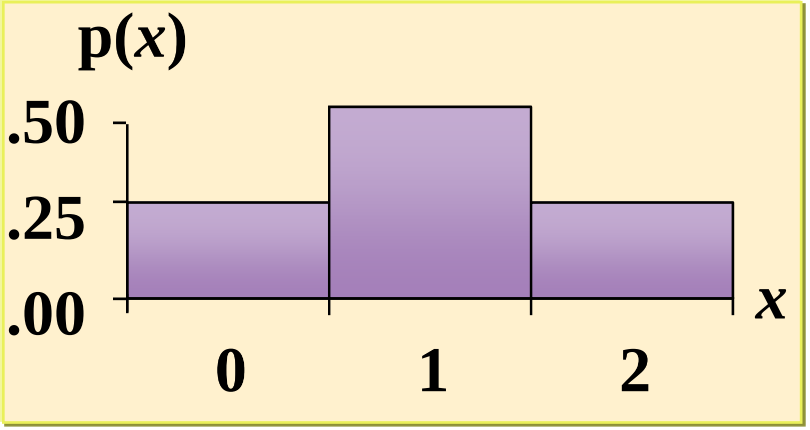
Listing

{ (0, .25), (1, .50), (2, .25) }

Table

# Tails	f(*x*) Count	p(*x*)
0	1	.25
1	2	.50
2	1	.25

Graph



Formula

$$p(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

Example

In the random experiment of rolling two dice once, then the number of the outcomes = n(S) = 6*6=36, and if we define X as follows:

X=the sum of the two numbers comes up

Then we will have the probability distribution for the discrete random variable X (X takes on value from 2 to 12) as follows:

$$
\begin{aligned}
S = \{&(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),\\
&(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6),\\
&(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),\\
&(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6),\\
&(5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6),\\
&(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}.
\end{aligned}
$$

[redacted]

X	2	3	4	5	6	7	8	9	10	11	12
# of Outcomes	1	2	3	4	5	6	5	4	3	2	1
P(x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Example

- Below is the probability distribution table for the random variable X whose values are the possible numbers of defective computers purchased by pharmacy in Jordan

x	0	1	2	3	4
$P(X = x)$	0.16	0.53	0.20	0.08	0.03

- Use this table to answer the questions that follow:

a) What is the probability a randomly selected pharmacy has exactly 2 defective computers?

b) What is the probability a randomly selected pharmacy has less than 2 defective computer?

Continued

Solution

a) $P(X=2)=0.2$

b) P(selected pharmacy has less than 2 defective computers)$=P(X<2)=P(X=0 \text{ or } X=1)$

$= P(X=0)+P(X=1)=0.16+0.53=0.69$

c) What is the probability a randomly selected pharmacy has 2 or fewer defective computers?

Answer:

$=P(X\leq2)$

$=P(X=0)+P(X=1)+P(X=2)$

$=0.16+0.53+0.2$

$=0.89$

Continued

d) What is the probability a randomly selected pharmacy has more than 2 defective computers?

Answer:

P(selected pharmacy has more than 2 defective computers)

$P(X>2)=P(X=3 \text{ or } 4)=P(X=3)+P(X=4)=0.08+0.03=0.11$

Rule: For a constant k, we have $P(X>k)+P(X \leq k) =1$

e) What is the probability a randomly selected pharmacy has more than 1 and less than or equal to 3 defective computers?

Answer:

P(selected pharmacy has more than 1 and less than or equal to 3 defective computers)

$=P(1< X \leq 3)$

$=P(X=2)+P(X=3)=0.2+0.08=0.28$

- **Probability distributions** can be used to describe the population, just as we described sample as follows:

✓ **Shape:** Symmetric, skewed, mound-shaped…

✓ **Outliers:** unusual or unlikely measurements in the data

✓ **Center (location) and spread:** the location is determined by the mean (**μ**) and the spread is determined by the standard deviation (**σ**).

The Mean and the Standard Deviation of a Discrete Random Variable

- **Mean (Expected Value)**

The mean of a discrete random variable is a weighted average of the possible values that the random variable can take.

Unlike the simple mean of a group of observations, which gives each observation equal weight, the mean of a random variable weights each outcome x according to its probability (X=x).

This common symbol for the mean (also known as the expected value is X of μ.

- Let X be a discrete random variable with probability distribution P(X=x) and let k be the number of possible values for X , then the mean (expected value) is given as follows:

$$\mu = \mathrm{E}\,(\mathrm{X}) = \sum_{i=1}^{k} xi * P(X = xi) \qquad \mu = E(x) = \Sigma x\, p(x)$$

- E(X) is not the value of the random variable X that you expect to observe if you perform the experiment once. E(X) is a long run average, if you perform the experiment many times and observe the random variable X each time, then the average of more and more values of the random variable X.

Variance and Standard Deviation

- The variance of a discrete random variable X is a weighted average of squared deviation about the mean. The standard deviation of a discrete random variable X is the positive square root for the variance. The common symbol for the variance σ^2 is and for the standard deviation is σ.
- Let X be a discrete random variable with probability distribution P(X=x) and mean (expected value) μ **=E(X)**, then the variance and standard deviation for the discrete random variable X can be calculated as follows:

$$\sigma^2 = E[(x-\mu)^2] = \Sigma\,(x-\mu)^2\,p(x)$$

$$\sigma = \sqrt{\sigma^2}$$

Summary Measures

1. Expected Value (Mean of probability distribution)
   - Weighted average of all possible values
   - $\mu = E(x) = \Sigma x\, p(x)$
2. Variance
   - Weighted average of squared deviation about mean
   - $\sigma^2 = E[(x - \mu)^2] = \Sigma\, (x - \mu)^2\, p(x)$
3. Standard Deviation
   - $\sigma = \sqrt{\sigma^2}$

Mean : $\mu = \sum xp(x)$

Variance : $\sigma^2 = \sum (x-\mu)^2\, p(x)$

Standard deviation : $\sigma = \sqrt{\sigma^2}$

Summary Measures Calculation Table

x	$p(x)$	$x\,p(x)$	$x - \mu$	$(x - \mu)^2$	$(x - \mu)^2 p(x)$
Total		$\Sigma x\,p(x)$			$\Sigma(x - \mu)^2\,p(x)$

Thinking Challenge

You toss 2 coins. You're interested in the number of tails. What are the **expected value**, **variance**, and **standard deviation** of this random variable, number of tails?

Expected Value & Variance Solution*

x	$p(x)$	$x\,p(x)$	$x - \mu$	$(x - \mu)^2$	$(x - \mu)^2 p(x)$
0	**.25**	**0**	**–1.00**	**1.00**	**.25**
1	**.50**	**.50**	**0**	**0**	**0**
2	**.25**	**.50**	**1.00**	**1.00**	**.25**
		$\mu = 1.0$			$\sigma^2 = .50$
					$\sigma = .71$

Example

- Toss a fair coin 3 times and record *x* the number of heads.

x	$p(x)$	$xp(x)$	$(x-\mu)^2p(x)$
0	1/8	0	$(-1.5)^2(1/8)$
1	3/8	3/8	$(-0.5)^2(3/8)$
2	3/8	6/8	$(0.5)^2(3/8)$
3	1/8	3/8	$(1.5)^2(1/8)$

$$\mu = \sum xp(x) = \frac{12}{8} = 1.5$$

$$\sigma^2 = \sum (x-\mu)^2 p(x)$$

$$\sigma^2 = .28125 + .09375 + .09375 + .28125 = .75$$

$$\sigma = \sqrt{.75} = .688$$

Example

- A university medical research centre finds out that treatment of skin cancer by the use of chemotherapy has a success rate of 70%. Suppose that 5 patients are treated with chemotherapy. The probability distribution of X successful cures of the five patients is given in the table below:

x	0	1	2	3	4	5
$P(X = x)$	0.002	0.029	0.132	0.309	0.360	0.168

1) Find μ?
2) Find σ?

Continued

Answer:

1) $\mu = E(x) = \Sigma x$

$P(x)$=(0)(0.002)+(1)(0.029)+…+(4)(0.360)+(5)(0.168)

= 3.5 patient

2) $\sigma^2 = E[(x - \mu)^2] = \Sigma\ (x - \mu)^2\ p(x)$

$=(0-3.3)^2(0.002)+(1-3.5)^2(0.029)+\ldots+(5-3.5)^2(0.168)$

=1.05

$\sigma = \sqrt{\sigma^2} = \sqrt{1.5} = 1.025$

Continued

Conclusion

- The value of $\mu = 3.5$ is the centre of the probability distribution. In other words, if the five cancer patients receive chemotherapy treatment, we expect that the number of them who are cured to be near 3.5.
- The standard deviation, which is 1.025 in this case, measures the spread of the probability distribution, that is, on the average, the number of patients receive chemotherapy treatment who are cured will be less than or more than 3.5 by 1.025.

Example

Suppose that the number of cars, X, that pass through a car wash between 4:00 P.M and 5:00 P.M. on any sunny Friday has the probability distribution:

x	4	5	6	7	8	9
$P(X = x)$	0.083	0.083	0.250	0.250	0.167	0.167

Let Y=g(X)=2X-1 represents the amount of money in JD paid to the attendant by the manager. Find the attendant's expected earning for this particular time period?

- Solution

The attendant can expect to receive:

$E(Y)=E(g(X))=E(2X\text{-}1)=\sum_{i=4}^{9}(2x-1)P(X=x_i)$

$=(7)(0.083)+(9)(0.083)+(11)(0.250)+(13)(0.250)+$

$(15)(0167)+(17)(0.167) = 12.672$ JD