The Probability Distribution of a Discrete Random Variable
by
W H Laverty
(modified)

A mathematical description of the possible values of the random variable together with the probabilities of those values

The probability distribution of a discrete random variable is describe by its :

probability function $p(x)$.

$p(x)$ = the probability that X takes on the value x.

This can be given in either a **tabular form** or in the form of **an equation**.

It can also be displayed in a **graph**.

Comments:

Every probability function must satisfy:

1. The probability assigned to each value of the random variable must be between 0 and 1, inclusive:

$$0 \le p(x) \le 1$$

2. The sum of the probabilities assigned to all the values of the random variable must equal 1:

$$\sum_x p(x) = 1$$

3. $$P[a \le X \le b] = \sum_{x=a}^{b} p(x)$$
$$= p(a) + p(a+1) + \cdots + p(b)$$

Example 1

- ***Discrete***
  – A die is rolled and X = number of spots showing on the upper face.

x	1	2	3	4	5	6
p(x)	1/6	1/6	1/6	1/6	1/6	1/6

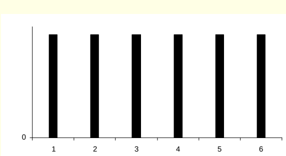
formula

– $p(x) = 1/6$ if x = 1, 2, 3, 4, 5, 6

Graphs

To plot a graph of *p(x)*, draw bars of height *p(x)* above each value of *x*.

Rolling a die



Example 2

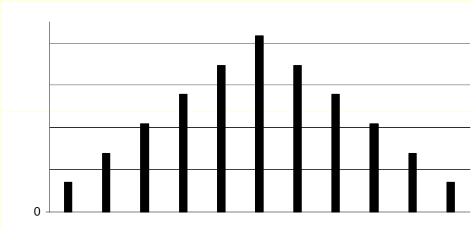
– Two dice are rolled and X = Total number of spots showing on the two upper faces.

x	2	3	4	5	6	7	8	9	10	11	12
p(x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

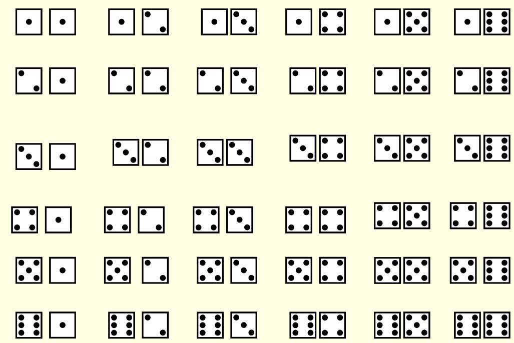
Formula:

$$p(x) = \begin{cases} \dfrac{x-1}{36} & x = 2,3,4,5,6 \\ \dfrac{13-x}{26} & x = 7,8,9,10,11,12 \end{cases}$$

Rolling two dice



36 possible outcome for rolling two dice



Mean, Variance and Standard Deviation of a Discrete Probability Distribution

Mean and Variance (standard deviation) of a Discrete Probability Distribution

- Describe the center and spread of a probability distribution
- The mean (denoted by greek letter μ (mu)), measures the centre of the distribution.
- The variance (σ^2) and the standard deviation (σ) measure the spread of the distribution.

  σ is the greek letter for s.

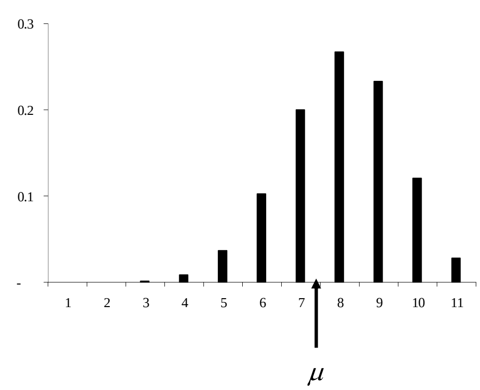
Mean of a Discrete Random Variable

- The mean, μ, of a discrete random variable *x* is found by multiplying each possible value of *x* by its own probability and then adding all the products together:

$$\mu = \sum_x [xp(x)]$$
$$= x_1 p(x_1) + x_2 p(x_2) + \ldots + x_k p(x_k)$$

Notes:

- The mean is a ***weighted average*** of the values of *X*.
- The mean is the ***long-run average*** value of the random variable.
- The mean is ***centre of gravity*** of the probability distribution of the random variable



Variance and Standard Deviation

Variance of a Discrete Random Variable: Variance, σ^2, of a discrete random variable x is found by multiplying each possible value of the squared deviation from the mean, $(x - \mu)^2$, by its own probability and then adding all the products together:

$$\sigma^2 = \sum_x \left[(x-\mu)^2 p(x)\right]$$
$$= \sum_x \left[x^2 p(x)\right] - \left\{\sum_x \left[xp(x)\right]\right\}^2$$
$$= \sum_x \left[x^2 p(x)\right] - \mu^2$$

Standard Deviation of a Discrete Random Variable: The positive square root of the variance:

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

Example 3

In baseball the number of individuals, X, on base when a home run is hit ranges in value from 0 to 3. The probability distribution is known and is given below:

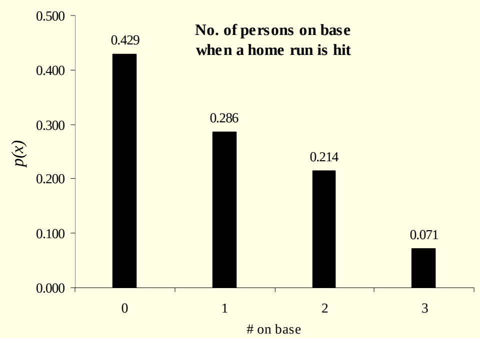
x	0	1	2	3
$p(x)$	6/14	4/14	3/14	1/14

Note:

- This chart implies the only values x takes on are 0, 1, 2, and 3.
- If the random variable X is observed repeatedly the probabilities, $p(x)$, represents the proportion times the value x appears in that sequence.

$$P(\text{the random variable } X \text{ equals } 2) = p(2) = \frac{3}{14}$$
$$P(\text{the random variable } X \text{ is at least } 2) = p(2) + p(3) = \frac{3}{14} + \frac{1}{14} = \frac{4}{14}$$

A Bar Graph



Example – contd.

The number of individuals, X, on base when a home run is hit ranges in value from 0 to 3.

x	$p(x)$	$xp(x)$	x^2	$x^2 p(x)$
0	0.429	0.000	0	0.000
1	0.286	0.286	1	0.286
2	0.214	0.429	4	0.857
3	0.071	0.214	9	0.643
Total	1.000	0.929		1.786
	$\sum p(x)$	$\sum xp(x)$		$\sum x^2 p(x)$

- Computing the mean:

$$\mu = \sum_x \left[xp(x)\right] = 0.929$$

Note:

- 0.929 is the long-run average value of the random variable
- 0.929 is the centre of gravity value of the probability distribution of the random variable

- Computing the variance:

$$\sigma^2 = \sum_x \left[(x-\mu)^2 p(x)\right]$$
$$= \sum_x \left[x^2 p(x)\right] - \left\{\sum_x \left[xp(x)\right]\right\}^2$$
$$= 1.786 - \{.929\}^2 = 0.923$$

- Computing the standard deviation:

$$\sigma = \sqrt{\sigma^2}$$
$$= \sqrt{0.923} = 0.961$$

The Binomial distribution

An important discrete distribution

Situation - in which the binomial distribution arises

- We have a random experiment that has two outcomes
  - Success (***S***) and failure (***F***)
  - $p = P[\boldsymbol{S}]$, $q = 1 - p = P[\boldsymbol{F}]$,
- The random experiment is repeated *n* times independently
- *X* = the number of times ***S*** occurs in the *n* repititions
- Then *X* has a binomial distribution

Example

- A coin is tosses $n = 20$ times
  - *X* = the number of heads
  - Success (***S***) = {head}, failure (***F***) = {tail
  - $p = P[\boldsymbol{S}] = 0.50$, $q = 1 - p = P[\boldsymbol{F}] = 0.50$
- An eye operation has 85 % chance of success. It is performed $n = 100$ times
  - *X* = the number of Sucesses (***S***)
  - $p = P[\boldsymbol{S}] = 0.85$, $q = 1 - p = P[\boldsymbol{F}] = 0.15$

The Binomial distribution

1. We have an experiment with two outcomes – *Success(S)* and *Failure(F).*
2. Let *p* denote the probability of *S (Success).*
3. In this case *q=1-p* denotes the probability of *Failure(F).*
4. This experiment is repeated *n* times independently.
5. *X* denote the number of successes occuring in the *n* repititions.

The possible values of *X* are

0, 1, 2, 3, 4, … , $(n - 2)$, $(n - 1)$, n

and *p(x)* for any of the above values of *x* is given by:

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{n}{x} p^x q^{n-x}$$

X is said to have the ***Binomial distribution*** with parameters *n* and *p.*

Summary:

X is said to have the ***Binomial distribution*** with parameters *n* and *p.*

1. *X* is the number of successes occurring in the *n* repetitions of a Success-Failure Experiment.
2. The probability of success is *p.*
3. The probability function

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

Example:

1. A coin is tossed n = 5 times. X is the number of heads occurring in the 5 tosses of the coin. In this case p = ½ and

$$p(x)=\binom{5}{x}\left(\tfrac{1}{2}\right)^x\left(\tfrac{1}{2}\right)^{5-x}=\binom{5}{x}\left(\tfrac{1}{2}\right)^5=\binom{5}{x}\left(\tfrac{1}{32}\right)$$

x	0	1	2	3	4	5
p(x)	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

Note: $$\binom{5}{x}=\frac{5!}{x!(5-x)!}$$

$$\binom{5}{0}=\frac{5!}{0!(5-0)!}=1 \qquad \binom{5}{3}=\frac{5!}{3!2!}=\frac{5(4)}{2(1)}=10$$

$$\binom{5}{1}=\frac{5!}{1!(5-1)!}=\frac{5!}{4!}=5 \qquad \binom{5}{4}=\frac{5!}{4!1!}=5$$

$$\binom{5}{2}=\frac{5!}{2!3!}=\frac{5(4)}{2(1)}=10 \qquad \binom{5}{5}=\frac{5!}{0!5!}=1$$

Computing the summary parameters for the distribution – μ, σ², σ

x	*p (x)*	*xp(x)*	*x ²*	*x ² p(x)*
0	0.03125	0.000	0	0.000
1	0.15625	0.156	1	0.156
2	0.31250	0.625	4	1.250
3	0.31250	0.938	9	2.813
4	0.15625	0.625	16	2.500
5	0.03125	0.156	25	0.781
Total	1.000	2.500		7.500
	$\sum p(x)$	$\sum xp(x)$		$\sum x^2 p(x)$

- Computing the mean:

$$\mu=\sum_x [xp(x)]=2.5$$

- Computing the variance:

$$\sigma^2=\sum_x\left[(x-\mu)^2 p(x)\right]$$
$$=\sum_x\left[x^2 p(x)\right]-\left\{\sum_x[xp(x)]\right\}^2$$
$$=7.5-\{2.5\}^2=1.25$$

- Computing the standard deviation:

$$\sigma=\sqrt{\sigma^2}$$
$$=\sqrt{1.25}=1.118$$

Example:

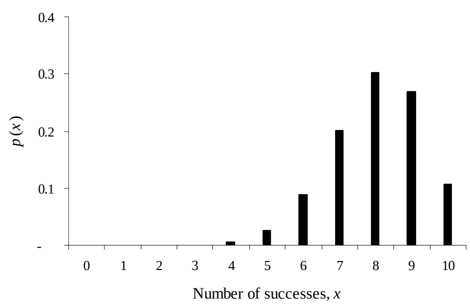
- A surgeon performs a difficult operation n = 10 times.
- X is the number of times that the operation is a success.
- The success rate for the operation is 80%. In this case p = 0.80 and
- X has a Binomial distribution with n = 10 and p = 0.80.

$$p(x)=\binom{10}{x}(0.80)^x(0.20)^{10-x}$$

Computing $p(x)$ for x = 0, 1, 2, 3, … , 10

x	0	1	2	3	4	5
p (x)	0.0000	0.0000	0.0001	0.0008	0.0055	0.0264
x	6	7	8	9	10	
p (x)	0.0881	0.2013	0.3020	0.2684	0.1074	

The Graph



Computing the summary parameters for the distribution – μ, σ^2, σ

x	$p(x)$	$xp(x)$	x^2	$x^2 p(x)$
0	0.0000	0.000	0	0.000
1	0.0000	0.000	1	0.000
2	0.0001	0.000	4	0.000
3	0.0008	0.002	9	0.007
4	0.0055	0.022	16	0.088
5	0.0264	0.132	25	0.661
6	0.0881	0.528	36	3.171
7	0.2013	1.409	49	9.865
8	0.3020	2.416	64	19.327
9	0.2684	2.416	81	21.743
10	0.1074	1.074	100	10.737
Total	1.000	8.000		65.600
		$\sum xp(x)$		$\sum x^2 p(x)$

- Computing the mean:

$$\mu = \sum_x [xp(x)] = 8.0$$

- Computing the variance:

$$\sigma^2 = \sum_x \left[(x-\mu)^2 p(x)\right]$$
$$= \sum_x \left[x^2 p(x)\right] - \left\{\sum_x [xp(x)]\right\}^2$$
$$= 65.6 - \{8.0\}^2 = 1.60$$

- Computing the standard deviation:

$$\sigma = \sqrt{\sigma^2}$$
$$= \sqrt{1.60} = 1.265$$

Mean, Variance and standard deviation of Binomial Random Variables

Mean,Variance & Standard Deviation of the Binomial Ditribution

- The mean, variance and standard deviation of the binomial distribution can be found by using the following three formulas:

1. $\mu = np$
2. $\sigma^2 = npq = np(1-p)$
3. $\sigma = \sqrt{npq} = \sqrt{np(1-p)}$

Example:

Find the mean and standard deviation of the binomial distribution when $n = 20$ and $p = 0.75$

Solutions:

1) $n = 20$, $p = 0.75$, $q = 1 - 0.75 = 0.25$

$$\mu = np = (20)(0.75) = 15$$

$$\sigma = \sqrt{npq} = \sqrt{(20)(075)(0.25)} = \sqrt{3.75} \approx 1.936$$

2) These values can also be calculated using the probability function:

$$p(x) = \binom{20}{x}(0.75)^x(0.25)^{20-x} \text{ for } x = 0, 1, 2, \ldots, 20$$

- Computing the mean:

$$\mu = \sum_x \left[x p(x) \right] = 15.0$$

- Computing the variance:

$$\begin{aligned} \sigma^2 &= \sum_x \left[(x-\mu)^2 p(x) \right] \\ &= \sum_x \left[x^2 p(x) \right] - \left\{ \sum_x \left[x p(x) \right] \right\}^2 \\ &= 228.75 - \{15.0\}^2 = 3.75 \end{aligned}$$

- Computing the standard deviation:

$$\begin{aligned} \sigma &= \sqrt{\sigma^2} \\ &= \sqrt{3.75} = 1.936 \end{aligned}$$