V. Discrete Probability Distributions

A. Basic Definitions

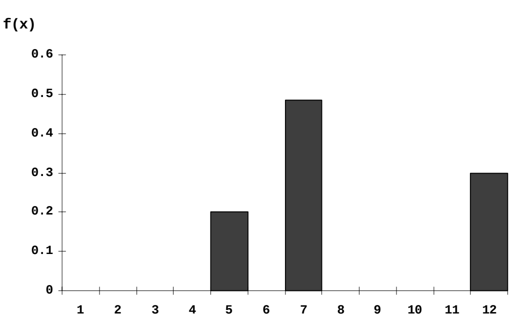
1. Random Variable - a numerical description of the outcome of an experiment.
2. Discrete Random Variable - a numerical description of the outcome of an experiment that can yield only a finite number of values or an infinite sequence such as 0, 1, 2, ...
3. Continuous Random Variable - a numerical description of the outcome of an experiment whose outcome can assume any numerical value in an interval or collection of intervals.

4. Probability Distribution - description of how probabilities are allocated to potential values of a random variable

   Example: If we have a random variable x that can take on a value of 5 with probability 0.20, a value of 7 with probability 0.50, and a value of 12 with probability 0.30, the probability distribution of x is

x	f(x)
5	0.20
7	0.50
12	0.30
	1.00

Note that the area under the graph of f(x) represents probability.



What is the area under the curve at x = 5?

Area is height*width, and height under the curve is f(x), so the area under the curve at x=5 is

(height)(width) = f(x)(1) = (0.20)(1) = 0.20

which is also P(x=5) for this probability distribution.

What is the area under the curve at x = 9?

(height)(width) = f(x)(1) = (0.00)(1) = 0.00

which is also P(x=9) for this probability distribution.

Thus f(x) = P(X=x) for a discrete probability distribution.

5. Probability Distribution Function - convenient mathematical means of allocating/assigning probabilities to potential values of a random variable. For a random variable x, it is denoted as f(x) which is equal to P(x) for a discrete variable. Note that the following conditions must be met:

$f(x) \geq 0$ **for all x** **and** $\Sigma f(x) = 1.0$

The mean and variance of a discrete random variable can be expressed as:

$$E(X) = \mu = \sum x f(x)$$

and

$$V(X) = E(X-\mu)^2 = \sigma^2 = \sum (x-\mu)^2 f(x)$$

Example: for the previous random variable x with the probability distribution

x	f(x)
5	0.20
7	0.50
12	0.30
	1.00

The mean (expected value) is

$$E(X) = \mu = 5(0.20) + 7(0.50) + 12(0.30) = 8.1$$

and the variance is

$$V(X) = E(X-\mu)^2 = \sigma^2 = \sum (x-\mu)^2 f(x)$$
$$= .2(5 - 8.1)^2 + .5(7 - 8.1)^2 + .3(12 - 8.1)^2$$
$$= 1.922 + 0.605 + 4.563 = 7.09$$

B. Some Important Discrete Probability Distributions

1. Discrete Uniform Probability Function - expresses the probabilities of outcomes for a discrete random variable x for which all possible outcomes are equally likely. Given by

$$f(x) = \begin{cases} \frac{1}{n} \text{ if } x \in S \\ 0 \text{ otherwise} \end{cases}$$

where n = the number of possible unique values that the random variable x may assume.

Example: Suppose that random variable x can take on any integer value between 0 and 10 with equal probability. What is the probability that x equals 9?

$$f(9) = \frac{1}{11} = 0.090909$$

What is the probability that x is at least 9 and no more than 11?

$P(9 \le x \le 11) = f(9) + f(10) + f(11) = 0.0909 + 0.0909 + 0.00 = .1818$

What is the probability that x is more than 9 and no more than 11?

$P(9 < x \le 11) = f(10) + f(11) = 0.0909 + 0.00 = 0.09090$

2. Binomial Probability Function - mathematical expression of the probability of the number of 'successes' (x) in n identical trials.

Characteristics of a binomial experiment are:

- experiment consists of a sequence of n identical trials
- there are two possible outcomes on each trial
- the probability of 'success' (denoted p) is constant across trials
- the trials are independent

Suppose we have a binomial experiment where the probability of successes in any trial is 0.70 (so the probability of failure in any trial is 0.30). What is the probability we would achieve exactly two successes in three trials?

- **first, recognize that the trials are independent, so we can multiply their probabilities directly:**

  pr(s, s, f) = pr(s)pr(s) pr(f)

  = (0.70)(0.70)(0.30) = 0.1470 = $p^x(1 - p)^{n-x}$ (why?)

- **second, recognize that the order doesn't matter with regard to the probability - (s, s, f), (s, f, s), and (f, s, s) have equivalent probabilities**

- **third, recognize that there are 3 ways to have two successes and one failure in three trials** $\binom{n}{x} = \binom{3}{2} = \frac{3!}{2!(3-2)!} = \frac{3 \cdot 2 \cdot 1}{2 \cdot 1(1)} = 3$

…so the probability of exactly x = two successes in n = three trials of a binomial experiment (where the probability of successes in any trial is p = 0.70) is 3(0.1470) = 0.4410,

i.e. the binomial probability distribution function is

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

where f(x) = probability of x successes in n trials

n = number of trials

x = number of successes

p = probability of success

1 - p = probability of *failure*

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

Example: What is the probability of getting exactly three heads in seven flips of a fair coin?

$$f(3) = \binom{7}{3} 0.50^3 (1 - 0.50)^{(7-3)}$$

$$= \frac{7!}{3!(7-3)!}(0.125)(0.0625)$$

$$= 0.2734$$

Example: What is the probability of getting at least one head in seven flips of a fair coin?

The sample space is S = {0, 1, 2, 3, 4, 5, 6, 7}

so $P(x \geq 1) = f(1) + f(2) + f(3) + f(4) + f(5) + f(6) + f(7)$

Is there an easier way?

$P(x \geq 1) = 1 - P(\overline{x \geq 1}) = 1 - P(x < 1) = 1 - P(x = 0) = 1 - f(0)$

and

$$f(0) = \binom{7}{0} 0.50^0 (1 - 0.50)^{(7-0)}$$
$$= \frac{7!}{0!(7-0)!}(1.00)(0.0078125)$$
$$= 0.0078125$$

so $P(x \geq 1) = 1 - 0.0078125 = 0.9921875$

Example: What is the probability of getting no more than one head in seven flips of a fair coin?

$P(x \leq 1) = P(x = 0) + P(x = 1) = 0.0078125 + P(x = 1)$

and

$$f(1) = \binom{7}{1} 0.50^1 (1 - 0.50)^{(7-1)}$$
$$= \frac{7!}{1!(7-1)!}(0.50)(0.015625)$$
$$= 0.0546875$$

so $P(x \leq 1) = 0.0078125 + 0.0546875 = 0.0625$

Note that Binomial Probabilities Tables (available in many textbooks) can be used for certain values of p, n, and x. Also note that

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

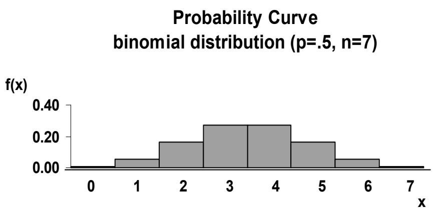
and

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

for the binomial distribution.

The probability curve and distribution for the previous problem (a binomially distributed random variable with a probability of success of p = 0.50 and n = 7 trials) look like this:

x	f(x)
0	0.007813
1	0.054688
2	0.164063
3	0.273438
4	0.273438
5	0.164063
6	0.054688
7	0.007813



3. Poisson Probability Function - mathematical expression of the probability for the number of random, independent occurrences (x) of a relatively rare event over some interval or continuum.

Characteristics of a Poisson experiment are:

- the probability of an occurrence is the same over any two intervals of equal length
- the occurrence or nonoccurrence in any interval is independent of the occurrence or nonoccurrence in any other interval

The Poisson probability distribution function is

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

where f(x) = probability of x occurrences over an interval of a defined length

λ = expected or mean number of occurrences over an interval of a defined length

x = number of occurrences over an interval of a defined length

e = 2.71828 (Euler's constant)

Must all be defined over an interval of the same length (and measured in the same units!)

Example: Suppose that the express line in a grocery store averages 1 customer every 5 minutes. What is the probability of exactly 8 customers going through the express line over the next hour?

P(x = 8) where x = number of customers in an hour

so

$$\mu = 1\left(\frac{60}{5}\right) = 12.0$$

and

$$f(8) = \frac{12^8}{8!}e^{-12} = 0.0655$$

What is the probability of at least 3 customers going through the express line over the next hour?

$$P(x \geq 3) = 1 - P(x \leq 2) = 1 - [\, f(x = 0) + f(x = 1) + f(x = 2)] =$$
$$1 - [0.0000 + 0.0001 + 0.0004] = 1 - 0.0005 = 0.9995$$

Note that Poisson Probabilities Tables (available in many textbooks) can be used for certain values of μ and x. Also note that

$$E(X) = \mu$$

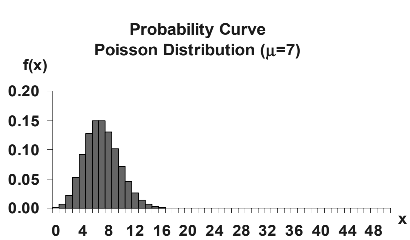
and

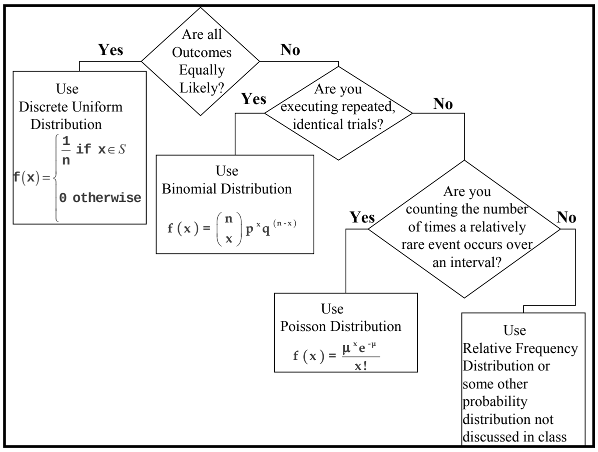
$$V(X) = \sigma^2 = \mu$$

for the Poisson distribution.

x	f(x)
0	0.000006144
1	0.000073731
2	0.000442383
3	0.001769533
4	0.005308599
5	0.012740639
6	0.025481277
7	0.043682190
8	0.065523285
9	0.087364380
10	0.104837256
11	0.114367916
12	0.114367916
13	0.105570384
14	0.090488900
15	0.072391120
16	0.054293340
17	0.038324711
18	0.025549807
19	0.016136720
20	0.009682032
21	0.005532590
22	0.003017776
23	0.001574492
24	0.000787246
25	0.000377878
⋮	⋮

The probability curve and distribution for the previous problem (a Poisson distributed random variable with a mean of μ = 7 customers per hour) look like this:





Are all Outcomes Equally Likely?
Yes
No
Use Discrete Uniform Distribution
$f(x) = \begin{cases} \frac{1}{n} \text{ if } x \in S \\ 0 \text{ otherwise} \end{cases}$
Are you executing repeated, identical trials?
Yes
No
Use Binomial Distribution
$f(x) = \binom{n}{x} p^x q^{(n-x)}$
Are you counting the number of times a relatively rare event occurs over an interval?
Yes
No
Use Poisson Distribution
$f(x) = \frac{\mu^x e^{-\mu}}{x!}$
Use Relative Frequency Distribution or some other probability distribution not discussed in class