Topic 6 - Probability Mass Function

Statistics for Managers

June 3, 1999

Random Variable

Let S be a discrete sample space with the set of elementary events denoted by E = {e_i, i = 1, 2, 3…}. A random variable is a function Y(e_i) that assigns a real value to each elementary event, e_i. The random variable is denoted by Y.

The set of all possible values of Y is the set {y_i = Y(e_i), i = 1, 2, 3…}. The set of all possible values of Y is a finite or countably infinite set, and Y is said to be a discrete random variable.

Let y be the set of values of Y. The subset of all elementary events that are assigned the y, is the compound event {e_i: Y(e_i) = y}. Usually the probabilities of all the elementary events are known, therefore the probability of this compound event can be readily computed by summing over all the elementary events.

Example 1. Number of heads

Toss a coin twice. Let Y denote the number of heads.

Denote (Tail, Tail) to be the elementary event that the first toss is tail and the second toss is tail. Denote the other elementary events accordingly.

Random Variable

Compound Event	Elementary Events
(Y=0)	`(Tail, Tail)`
(Y=1)	`(Tail, Head), (Head, Tail)`
(Y=2)	`(Head, Head)`

Probability Mass Function

Let Y be a discrete random variable. A probability mass function is f(y) = P{e_i: Y(e_i) = y}. It is a function with values between 0 and 1 and whose sum is 1 over all values of y.

Example 2. Number heads in two coin tosses

Toss a balanced coin once. Let Y be the number of heads that occurs. Find the probability mass function of Y.

Random Variable

Number of Heads	Elementary Events
(Y=0)	(Tail)
(Y=1)	(Head)

Probability Mass Function

y	f(y)
0	½
1	½

Example 3. Number heads in two coin tosses

Toss a balanced coin twice. Let Y denote the number of heads. Find the probability mass function of Y.

Denote (Tail, Tail) to be the elementary event that the first toss is tail and the second toss is tail. Denote the other elementary events accordingly.

Random Variable

Number of Heads (y)	Elementary Events
0	(Tail, Tail)
1	(Tail, Head) (Head, Tail)
2	(Head, Head)

Probability Mass Function

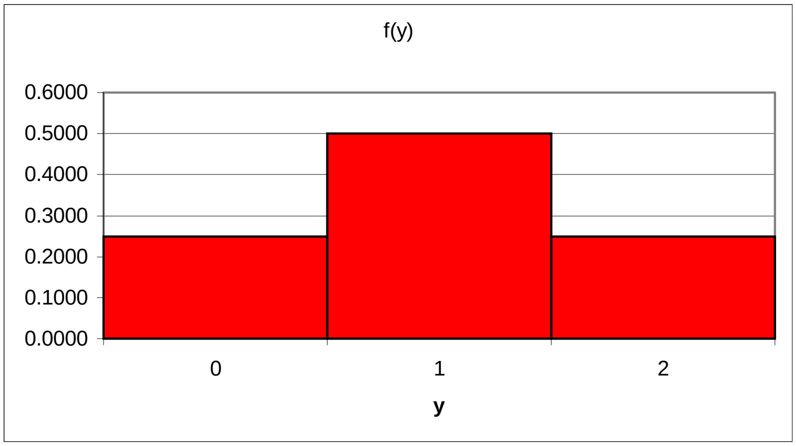
y	f(y)
0	¼
1	½
2	¼

Histogram

A <u>histogram</u> is a graph of the probability mass function. The total area under a histogram is one.

For a discrete random variable, the probability that Y is equal to y is the area in a histogram corresponding to a value y.

Example 4. Histogram of number heads in two coin tosses



Example 5. Sum of numbers on two dice

Toss a pair of dice; win dollars equal to the sum of numbers on the two dice. Let Y denote the winnings after playing the game once. Find the probability mass function of Y.

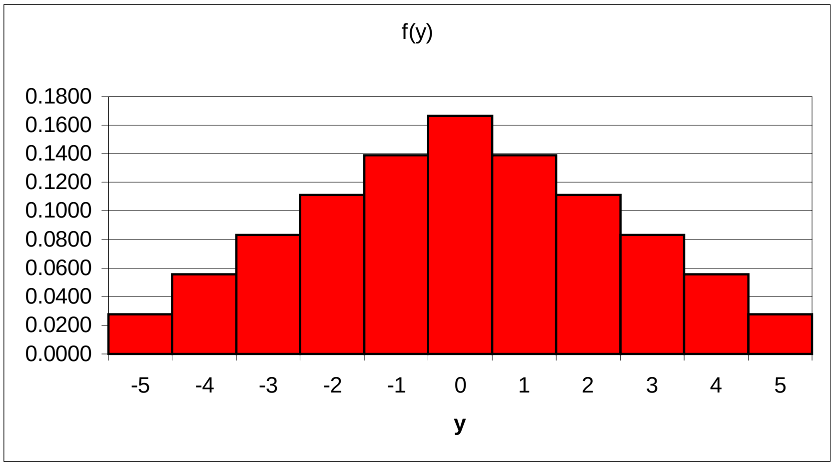
Random Variable

Winnings (y)	Elementary Events
2	(1,1)
3	(1,2) (2,1)
4	(1,3) (2,2) (3,1)
5	(1,4) (2,3) (3,2) (4,1)
6	(1,5) (2,4) (3,3) (4,2) (5,1)
7	(1,6) (2,5) (3,4) (4,3) (5,2) (6,1)
8	(2,6) (3,5) (4,4) (5,3) (6,2)
9	(3,6) (4,5) (5,4) (6,3)
10	(4,6) (5,5) (6,4)
11	(5,6) (6,5)
12	(6,6)

Probability Mass Function

y	f(y)
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

Histogram



f(y)
0.1800
0.1600
0.1400
0.1200
0.1000
0.0800
0.0600
0.0400
0.0200
0.0000
-5
-4
-3
-2
-1
0
1
2
3
4
5
y

Homework 1 Number of defective parts

Take a production run of 100 machine parts, with 25 defective and 75 non-defective. Randomly draw two parts, replacing them after each draw. Let Y denote the number of defective parts drawn. Create a table for the random variable and the compound events. Find the probability mass function of Y. Draw the histogram.

Homework 2 Number of defective without replacement

Take a carton of 12 machine parts, with 3 defective and 9 non-defective. Randomly draw two parts for inspection, but **do not** replace them after each draw. Let Y denote the number of defective parts drawn. Create a table for the random variable and the compound events. Find the probability mass function of Y. Draw the histogram.

Example 6. Waiting until a part fails inspection

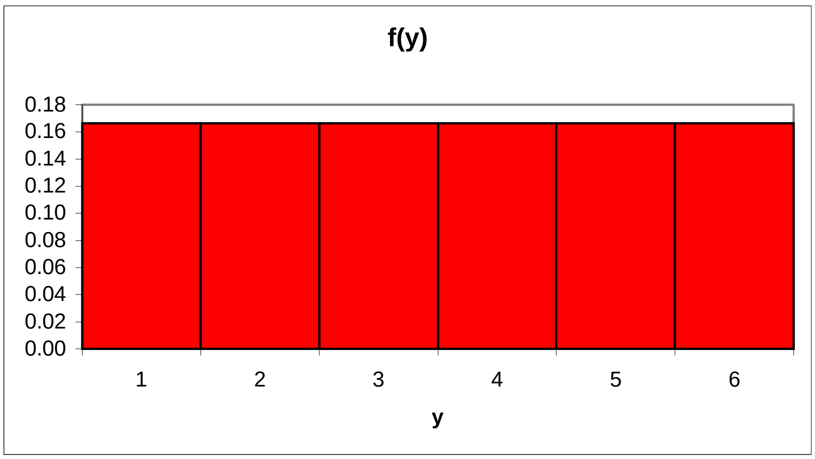
Take a carton of six machine parts, with one defective and five non-defective. Randomly draw parts for inspection, but **do not** replace them after each draw, until a defective part fails inspection. Let Y denote the number of parts drawn. Find the probability mass function of Y.

Let "F" denote the event that a part fails inspection, and let "P" denote the event a part passes inspection.

Random Variable

No. of Draws (y)	Elementary Event	Probability
1	(F)	1/6
2	(P, F)	(5X1)/(6X5)=1/6
3	(P, P, F)	(5X4X1)/(6X5X4)=1/6
4	(P, P, P, F)	(5X4X3X1)/(6X5X4X3)=1/6
5	(P, P, P, P, F)	(5X4X3X2X1)/(6X5X4X3X2)=1/6
6	(P, P, P, P, P, F)	(5X4X3X2X1X1)/(6X5X4X3X2X1)=1/6

Histogram



Example 7. Number of tosses until the six appears

Repeatedly toss a die until the number six comes up. Let Y denote the number of tosses. Find the probability mass function of Y.

Let "F" denote the six appearing and let "P" the six not appearing. Imagine that a machine produces a defective part one time in six. Then Y is the number of the first defective part.

In calculating the probabilities of the elementary events, we have used either equally likely events or the classical definition. (The classical definition uses ratio of the number ways an event can occur by the total number of ways any event can occur). Here we must use the special multiplicative rule of probability that determines the probability of the intersection of independent events as the product of the probabilities of the individual events.

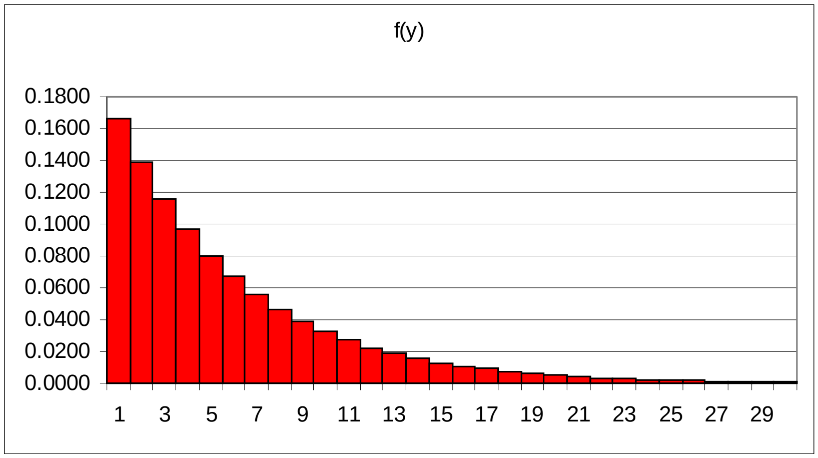
Random Variable

No. of Draws (y)	Elementary Event	Probability
1	(F)	1/6 = 0.1667
2	(P, F)	(5/6)X(1/6) = 0.1389
3	(P, P, F)	(5/6)X(5/6)X(1/6) = 0.1157
4	(P, P, P, F)	(5/6)X(5/6)X(5/6)X(1/6) = 0.0965
5	(P, P, P, P, F)	(5/6)X(5/6)X(5/6)X(5/6)X(1/6) = 0.0804
6	(P, P, P, P, P, F)	(5/6)X(5/6)X(5/6)X(5/6)X(5/6)X(1/6) = 0.0670
…	…	…

Formula for the Probability Mass Function:

$$f(y) = \left(\frac{5}{6}\right)^{y-1}\left(\frac{1}{6}\right)$$
$$y = 1,2,\ldots$$

Histogram



Note that this is an example of a countable infinite sample space. It is a special case of the geometric probability mass function.

The Geometric Distribution

Let p be the probability that a manufactured part is defective. Let Y be the number of parts manufactured until the first defective part is produced. The probability mass function of Y is called the Geometric Distribution:

$$f(y) = p(1-p)^{y-1}$$
$$y = 1,2,\ldots$$

The Cumulative Distribution Function (CDF)

Let f(y) be the probability mass function of a random variable Y. The cumulative distribution function F(y) is

$$F(y) = P(Y \le y) = \sum_{x \le y} f(x)$$

The CDF F(y) is the area in the histogram up to y.

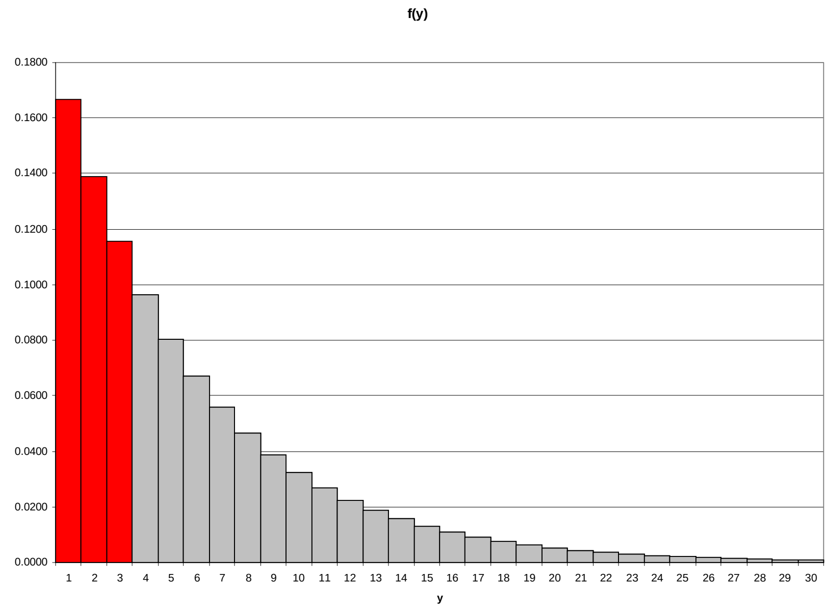
The probability mass function may be recovered from the CDF:

$$f(y) = F(y) - F(y-1)$$

Example 8. Number of tosses until the six appears

The probability that the number of tosses is 3 or less is the cumulative distribution function evaluated at y = 3:

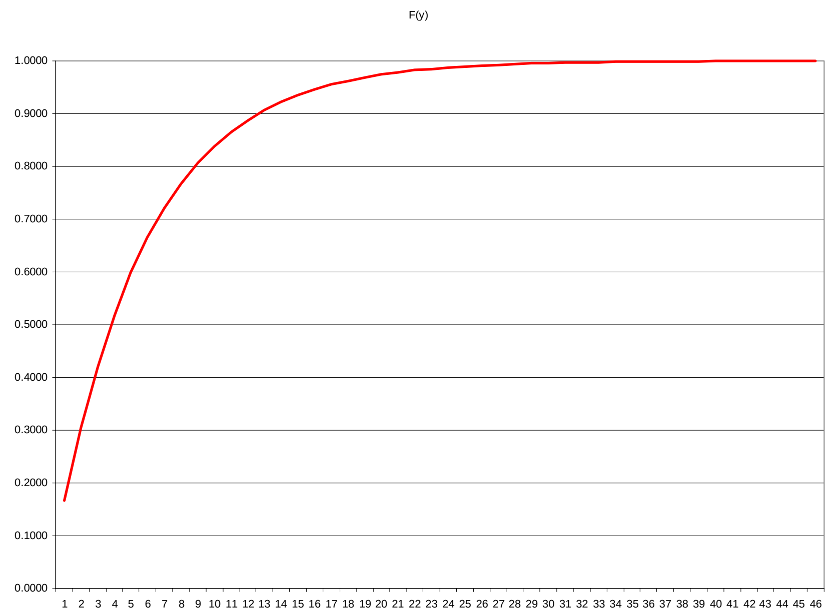
$$F(3) = f(1) + f(2) + f(3) = 0.1667 + 0.1389 + 0.1157 = 0.4213$$



Cumulative distribution function of number of tosses until the six appears

y	f(y)	F(y)
1	0.1667	0.1667
2	0.1389	0.3056
3	0.1157	0.4213
4	0.0965	0.5177
5	0.0804	0.5981
6	0.0670	0.6651
7	0.0558	0.7209
8	0.0465	0.7674
9	0.0388	0.8062
10	0.0323	0.8385
11	0.0269	0.8654
12	0.0224	0.8878
13	0.0187	0.9065
14	0.0156	0.9221
15	0.0130	0.9351
16	0.0108	0.9459
17	0.0090	0.9549
18	0.0075	0.9624
∞	0.0000	1.0000

Graph of the cumulative distribution function



The probability of exactly three tosses, f(3), may be obtained from the CDF:

f(3) = F(3) – F(2) = 0.4213 – 0.3056 = 0.1157

Homework 3 Sum of numbers on two dice

Toss a pair of dice; win dollars equal to the sum of numbers on the two dice. Let Y denote the winnings after playing the game once. Construct the table of the values of the CDF. Graph the CDF.