Bookmark this page

Probability Distribution: Probability Histograms

Learning Objective: Find the probability distribution of discrete random variables, and use it to find the probability of events of interest.

In the previous two examples and activity, we needed to specify the probability distributions ourselves, based on the physical circumstances of the situation. In some situations, as in the following example, the probability distribution may be specified with an algebraic formula. Such a formula must be consistent with the constraints imposed by the laws of probability, so that the probability of each outcome must be between 0 and 1, and the probabilities of all possible outcomes together must sum to 1.

Example: Formulas to Define Random Variables

A random variable X has a probability distribution of

P(X = x) = (x + 2) / 25 for x = 1, 2, 3, 4, 5.

Show the probability distribution in a table, and verify that the above requirements are satisfied.

Substituting x = 1, 2, 3, 4, and 5, respectively, into the formula for P(X = x), we have

X	1	2	3	4	5
P(X=x)	3/25	4/25	5/25	6/25	7/25

Clearly, each probability is between 0 and 1. Also, the probabilities sum to (3 + 4 + 5 + 6 + 7) / 25 = 25/25 = 1.

Scenario: Telemarking Sales

The number of sales that a telemarketing salesperson makes in an hour is a random variable X having the following probability distribution:

P(X = x) = (x + 2)(5-x) / 50 for x = 0, 1, 2, 3, 4

Did I Get This

1/1 point (graded)

Which of the following tables represents the probability distribution of X?

Table A

x	0	1	2	3	4
P(X=x)	1/5	1/5	1/5	1/5	1/5

Table B

x	0	1	2	3	4
P(X=x)	0/50	1/50	2/50	3/50	4/50

Table C

x	1	2	3	4	5
P(X=x)	10/50	12/50	12/50	10/50	6/50

Table D

x	0	1	2	3	4
P(X=x)	10/50	12/50	12/50	10/50	6/50

- ○ Table A
- ○ Table B
- ○ Table C
- ◉ Table D ✔

Answer

Correct:

This is the right probability distribution. In order to find the probability that X takes a certain value, you need to plug that value into the formula. So, for example, P(X = 1) = (1 + 2)(5 - 1) / 50 = 12/50. Note that this is a legitimate probability distribution since all the probabilities are between 0 and 1, and their sum is 1.

Submit

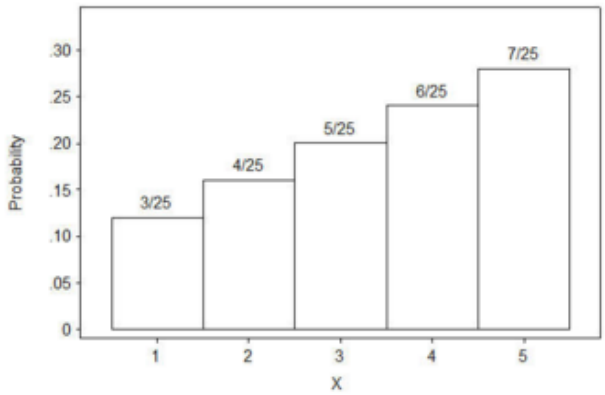
Probability Histograms

We learned to display the distribution of sample values for a quantitative variable with a histogram in which the horizontal axis represented the range of values in the sample. The vertical axis represented the frequency or relative frequency (sometimes given as a percentage) of sample values occurring in that interval. So the width of each rectangle in the histogram was an interval, or part of the possible values for the quantitative variable, and the height of each rectangle was the frequency (or relative frequency) for that interval.

Similarly, we can display the probability distribution of a random variable with a probability histogram. The horizontal axis represents the range of all possible values of the random variable, and the vertical axis represents the probabilities of those values.

Here is the probability histogram for the previous example:

X	1	2	3	4	5
P(X=x)	3/25	4/25	5/25	6/25	7/25

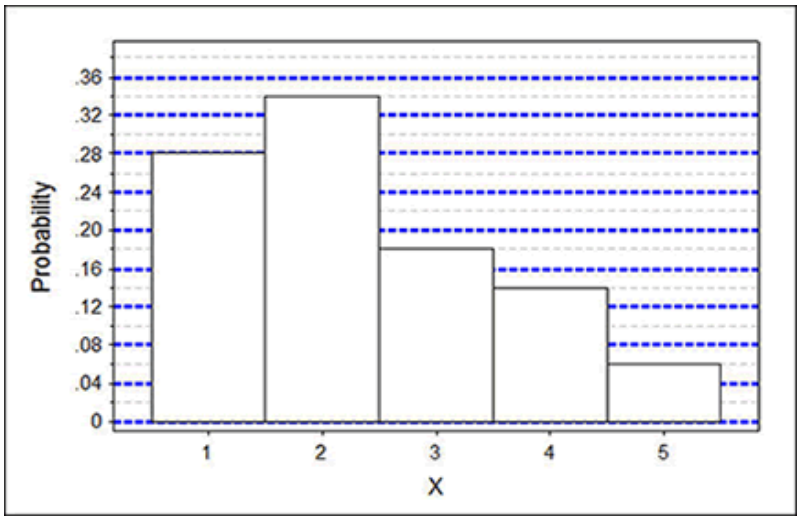


Area of a Probability Histogram

Notice that each rectangle in the histogram has a width of 1 unit. The height of each rectangle is the probability that it will occur. Thus, the area of each rectangle is base times height, which for these rectangles is 1 times its probability for each value of X. This means that the sum of the areas of all of the rectangles is the same as the sum of all of the probabilities. Therefore, the total area = 1.

Scenario: 2000 U.S. Census

Based upon data collected in the 2000 United States Census and an expanded number of households, the following histogram was constructed. It shows the distribution of people per household.



Learn By Doing

1/1 point (graded)

What does the random variable X represent?

- The number of children in each household
- The number of people in each household ✔
- The number of households contacted by the Census
- The area of each rectangle in the histogram
- The probabilities for each number of people per household

Answer

Correct: The variable X (on the horizontal axis) is described above as people per household.

Submit

Learn By Doing (1/1 point)

Using the completed probability distribution above, explain why it is indeed a legitimate probability distribution.

Your Answer:

total = 1 and each P(x) is within 0 and 1

Our Answer:
To be a probability distribution, two conditions must be satisfied. 1. Each of the probabilities must be between 0 and 1. Certainly 0.28, 0.34, 0.18, 0.14, and 0.06 are all between 0 and 1. 2. The sum of all of the probabilities must equal 1. Indeed, 0.28 + 0.34 + 0.18 + 0.14 + 0.06 = 1.

Resubmit Hint Reset

Learn By Doing (1/1 point)

Using the completed probability distribution above, show that the sum of the areas of all of the rectangles is one.

Your Answer:

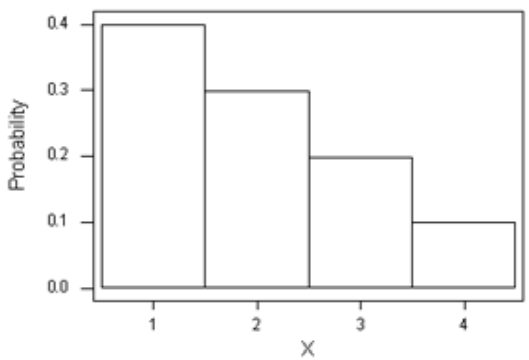
Sum of all areas = sum each base * height. Each base = 1, and each height = the probability, which we already know all sum up to 1.

Our Answer:
The sum of the areas = 1(.28) + 1(.34) + 1(.18) + 1(.14) + 1(.06) = 1. This make sense, since 1(.28) + 1(.34) + 1(.18) + 1(.14) + 1(.06) = 0.28 + 0.34 + 0.18 + 0.14 + 0.06, which we already know is equal to 1.

Resubmit Hint Reset

Did I Get This?

The probability distribution of the random variable X is represented by the following histogram.



Did I Get This

1/1 point (graded)

Fill in the probability distribution table values consistent with the histogram above:

x	1	cell 2	cell 3	4
P(X=x)	cell 1	.3	.2	cell 4

cell 1:

- 0
- 0.1
- 0.2
- 0.3
- 0.4 ✔
- 1
- 2
- 3
- 4

Answer

Correct: Indeed, P(X=1) = 0.4.

Submit

Did I Get This

1/1 point (graded)

Fill in the probability distribution table values consistent with the histogram above:

x	1	cell 2	cell 3	4
P(X=x)	cell 1	.3	.2	cell 4

cell 2:

- 0
- 0.1
- 0.2
- 0.3
- 0.4
- 1
- 2 ✔
- 3
- 4

Answer
Correct: Indeed, the second possible value that X can take is 2.

Submit

Did I Get This

1/1 point (graded)
Fill in the probability distribution table values consistent with the histogram above:

x	1	cell 2	cell 3	4
P(X=x)	cell 1	.3	.2	cell 4

cell 3:

- 0
- 0.1
- 0.2
- 0.3
- 0.4
- 1
- 2
- 3 ✔
- 4

Answer
Correct: Indeed the third possible value that X can take is 3.

Submit

Did I Get This

1/1 point (graded)
Fill in the probability distribution table values consistent with the histogram above:

x	1	cell 2	cell 3	4
P(X=x)	cell 1	.3	.2	cell 4

cell 4:

- 0
- 0.1 ✔
- 0.2
- 0.3
- 0.4
- 1
- 2
- 3
- 4

Answer
Correct: Indeed P(X = 4) = 0.1.

Submit

Did I Get This

1/1 point (graded)
The probability distribution represented in the histogram and table above may be also expressed by which formula?

- P(X = x) = x for x = 1, 2, 3, 4

- P(X = x) = 1/4 for x = 1, 2, 3, 4
- P(X = x) = x/10 for x = 1, 2, 3, 4
- P(X = x) = (5 - x) / 10 for x = 0, 1, 2, 3
- P(X = x) = (5 - x) / 10 for x = 1, 2, 3, 4 ✔

Answer

Correct:

Using this formula, we'll get P(X = 1) = (5 - 1) / 10 = 0.4, P(X = 2) = (5 - 2) / 10 = 0.3, P(X = 3) = (5 - 3) / 10 = 0.2, P(X = 4) = (5 - 4) / 10 = 0.1, which is exactly the one represented by the histogram above.

Submit