Random Variable:

A random variable is a variable whose value is determined by the outcome of an experiment.

Example:

Flip a fair coin twice, and let *X* be the number of tails that occur.

$$S = \{HH, HT, TH, TT\}$$

X	0	1	2

Probability Distribution of a Random Variable:

It's a table of all the values of a random variable along with their associated probabilities.

Example: Using the previous random variable,

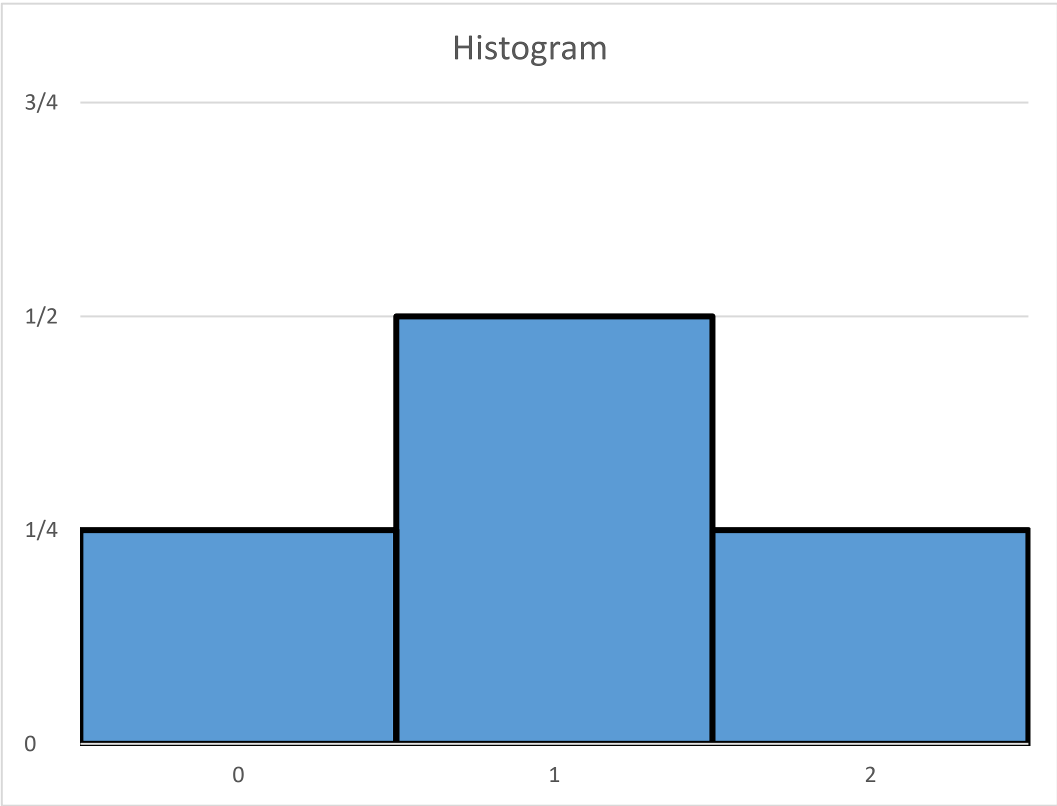
$$S = \{HH, HT, TH, TT\}$$

X	**0**	**1**	**2**
P(X)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Histogram of a Random Variable:

It's a special bar graph of the probability distribution where the bars are centered at the values of the random variable, adjacent bars touch, and the area of the bar is equal to the probability that the random variable takes on that particular value.

Example: Using the previous random variable and distribution,



Expected Value of a Random Variable:

For *X* a random variable with probability distribution

X	x_1	x_2	x_3	$\cdots$	x_n
P(*X*)	p_1	p_2	p_3	$\cdots$	p_n

, the expected value of *X*, *E*(*X*), is defined as $E\left(X\right)=x_1p_1+x_2p_2+x_3p_3+\cdots+x_np_n$**.**

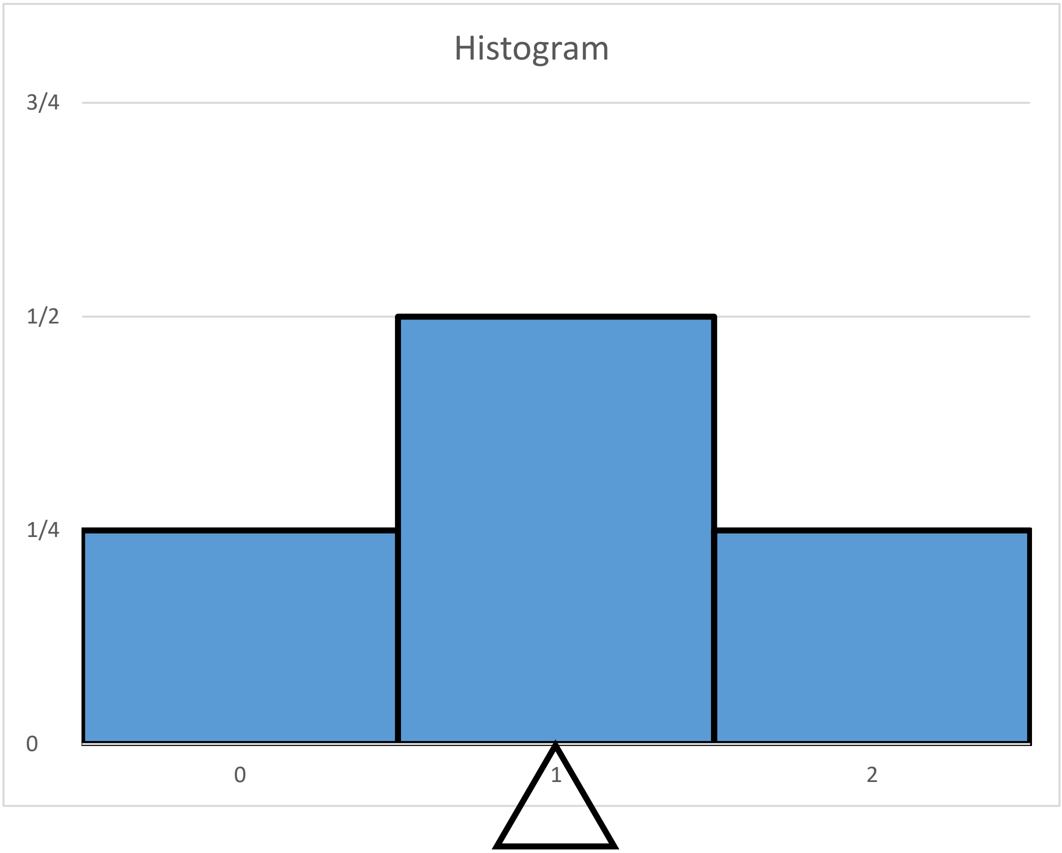
Example: Using the previous random variable and distribution,

X	**0**	**1**	**2**
P(X)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$$E\left(X\right)=0\cdot\frac{1}{4}+1\cdot\frac{1}{2}+2\cdot\frac{1}{4}=0+\frac{1}{2}+\frac{1}{2}=1$$

What does $E(X)$ mean or represent?

Geometrically, it's the balance position of the histogram:



More practically, it represents the long-term average value of *X*.

If the experiment is performed *n* times with *X* taking on the values $X_1, X_2, \ldots, X_n$**,**

then the average value of *X* would be $\frac{X_1 + X_2 + \ldots + X_n}{n}$**.**

For large values of *n*, the average value of *X* will be close to *E*(*X*), and the bigger *n* gets, the closer the average value of *X* will get to *E*(*X*).

So in our example, if the experiment of flipping a fair coin twice and recording the number of tails is done a bunch of times, the average number of tails per trial will be close to 1.

Examples:

1. In a game, you roll a fair die. If you roll a 1, you win $1; if you roll a 2, you win $2; if you roll a 3, you win $3. If you roll a 4, 5, or 6, you win nothing. Determine your expected winnings from one play of the game.

Let *X* be your winnings from one play of the game.

X	**$1**	**$2**	**$3**	**$0**
P(*X*)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$

$$E(X) = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 0\cdot\frac{1}{2} = \$1$$

Interpret the expected value.

If you play this game a large number of times, your average winnings will be close to $1 per play of the game.

2. In a raffle, there is 1 prize of $1,000, 2 prizes of $500, 4 prizes of $250, and 10 prizes of $100. One-thousand tickets are sold at $10 each. The prizes are awarded by randomly selecting the tickets. Determine your expected net winnings from one ticket.

Let *X* be your net winnings from one ticket.

X	**$990**	**$490**	**$240**	**$90**	**-$10**
P(*X*)	$\frac{1}{1,000}$	$\frac{2}{1,000}$	$\frac{4}{1,000}$	$\frac{10}{1,000}$	$\frac{983}{1,000}$

$$E(X) = 990 \cdot \frac{1}{1,000} + 490 \cdot \frac{2}{1,000} + 240 \cdot \frac{4}{1,000} + 90 \cdot \frac{10}{1,000} + (-10) \cdot \frac{983}{1,000}$$

$$= \frac{-6,000}{1,000} = -\$6$$

Interpret the expected value.

If you play this raffle a large number of times, on average you will lose close to $6 per play.

Another option:

Let *Y* be your gross winnings from one ticket.

Y	**\$1,000**	**\$500**	**\$250**	**\$100**	**\$0**
P(*Y*)	$\frac{1}{1{,}000}$	$\frac{2}{1{,}000}$	$\frac{4}{1{,}000}$	$\frac{10}{1{,}000}$	$\frac{983}{1{,}000}$

$$E(X) = E(Y) - \$10$$

$$= 1{,}000 \cdot \frac{1}{1{,}000} + 500 \cdot \frac{2}{1{,}000} + 250 \cdot \frac{4}{1{,}000} + 100 \cdot \frac{10}{1{,}000} + 0 \cdot \frac{983}{1{,}000} - 10$$

$$= 1 + 1 + 1 + 1 + 0 - 10 = -\$6$$

3. A 25-year-old can purchase a $10,000 one-year life insurance policy for $100. The probability of a person dying at age 25 is .002. Determine the expected profit to the insurance company from the sale of one policy.

Let *X* be the profit from the sale of one policy for the insurance company.

X	**-$9,900**	**$100**
P(*X*)	**.002**	.998

$$E(X) = (-9{,}900)(.002) + (100)(.998)$$
$$= -19.8 + 99.8 = \$80$$

Interpret the expected value.

If the insurance company sells a large number of these policies, then on average they will make close to $80 per policy.

4. What's the expected number of flips of a fair coin needed to see both heads and tails?

Let X be the number of flips to get both heads and tails.

X	2	3	4	5	…
P(X)	HT,TH $\frac{1}{2}$	HHT,TTH $\frac{1}{4}$	HHHT,TTTH $\frac{1}{8}$	HHHHT,TTTTH $\frac{1}{16}$	…

$$E(X) = 2\cdot\frac{1}{2} + 3\cdot\frac{1}{4} + 4\cdot\frac{1}{8} + 5\cdot\frac{1}{16} + 6\cdot\frac{1}{32} + 7\cdot\frac{1}{64} + \cdots$$

$$= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots$$

$$+ \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots$$

$$+ \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots$$

$$+ \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots$$

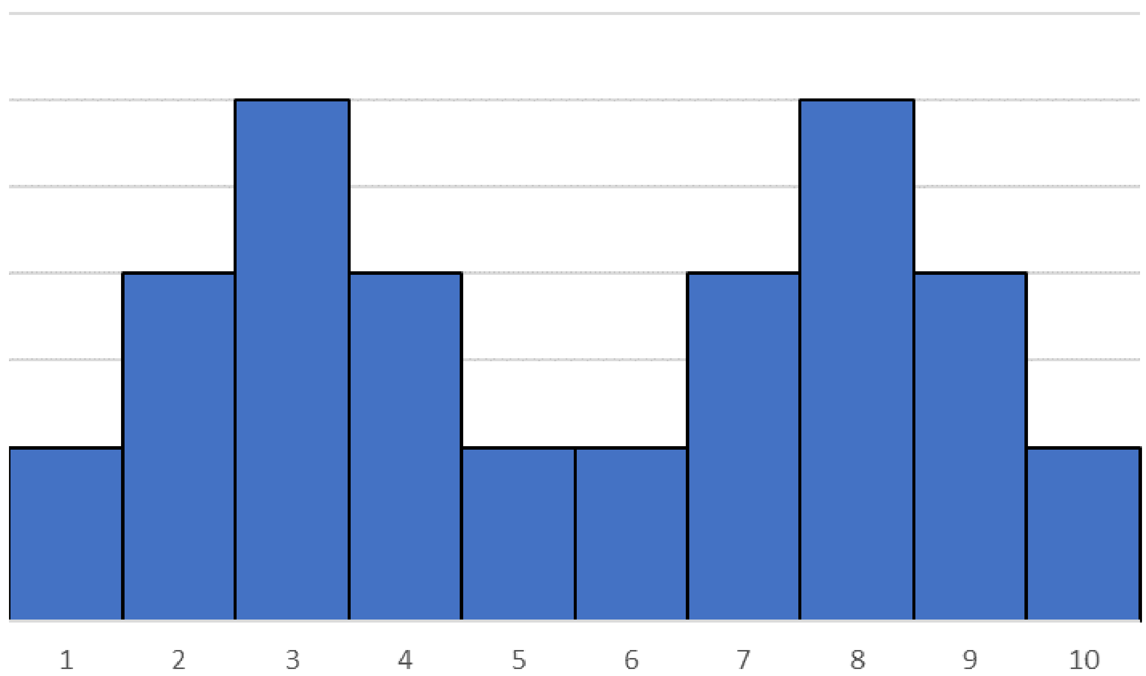
$$\vdots$$

If $S = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots$**, then** $2S = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 1 + S$**.**

So $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 1$**.**

So $E(X) = 1 + 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 1 + 1 + 1 = 3$**.**

One last example: Determine the expected value of the random variable with the following histogram.



The horizontal balance position of the histogram is 5.5, so the expected value of the random variable is 5.5.

Continuous Random Variables:

A continuous random variable can take on enough values to fill an entire interval.

It has a special curve associated with it called its probability density curve.

In general, a probability density curve must be on or above the horizontal axis, and the area of the region between the curve and the horizontal axis must be equal to 1.

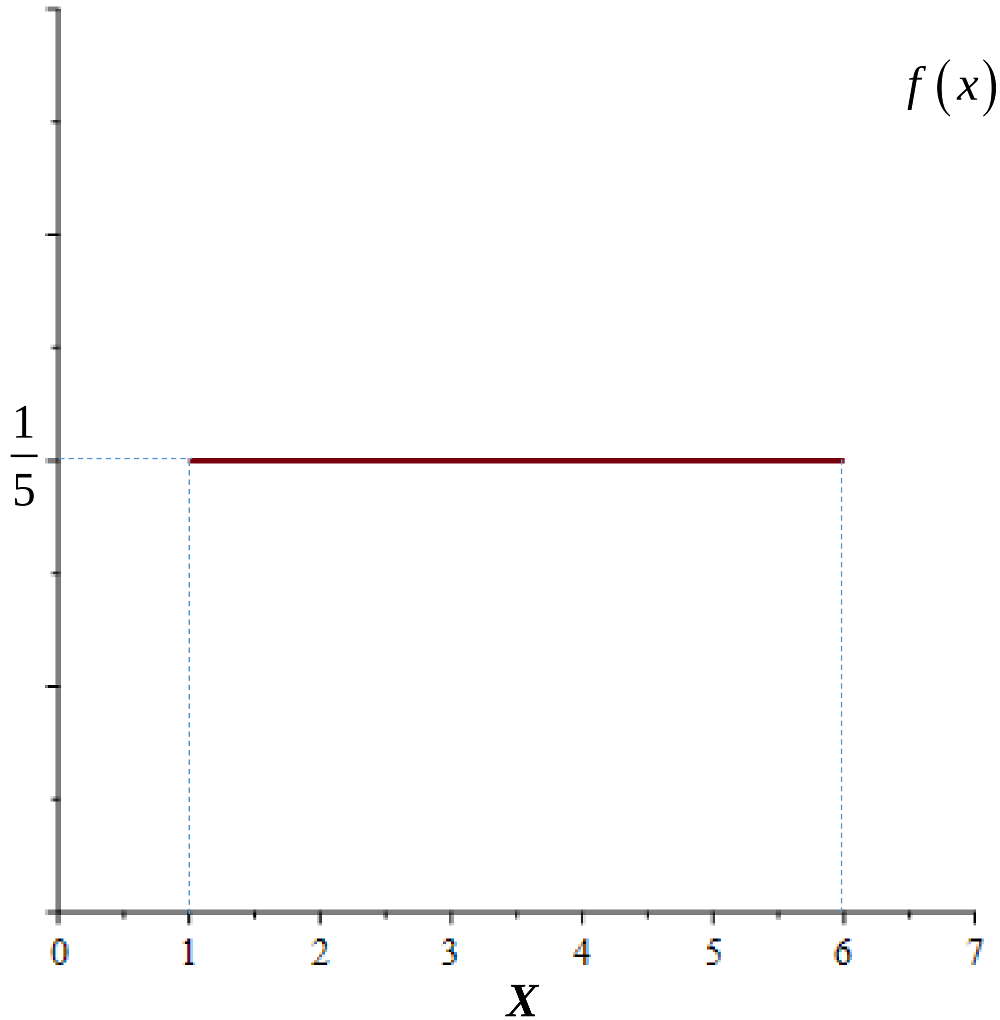
The probability that a continuous random variable will take on a value between *a* and *b* is equal to the area under its probability density curve from *a* to *b*.

Examples:

1. **Verify that the following curve can be a probability density curve for the random variable *X* that takes on values between 1 and 6.**

$$f(x)=\begin{cases}\frac{1}{5}; 1\le x\le 6\\ 0;\ \text{otherwise}\end{cases}$$

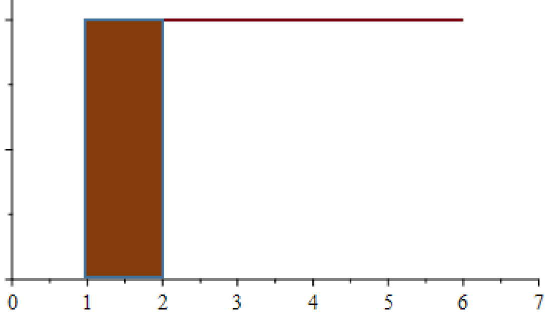
The region between the curve and the horizontal axis is a rectangle. The area of a rectangle is the product of its width and height. For this rectangle we get $5\cdot\frac{1}{5}=1$. So this curve can be a probability density curve.



Use the probability density curve to find the following probabilities:

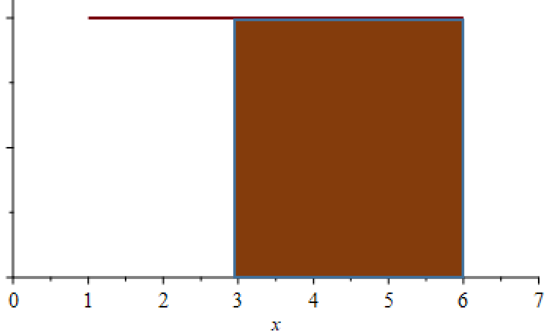
$P(1 \le X \le 2)$

$1 \cdot \frac{1}{5} = \boxed{\frac{1}{5}}$



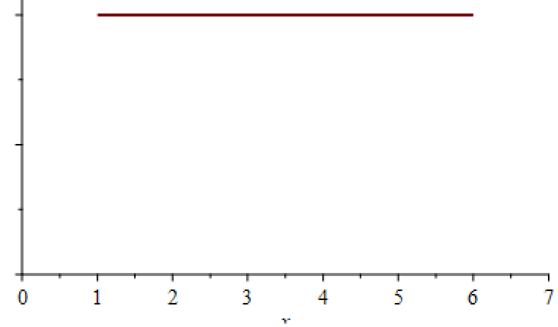
$P(3 \le X < 6)$

$3 \cdot \frac{1}{5} = \boxed{\frac{3}{5}}$



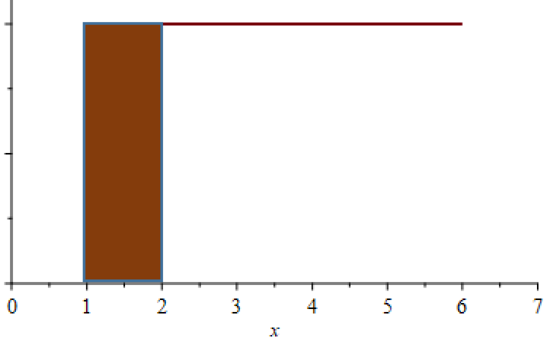
$P(X > 6)$

0



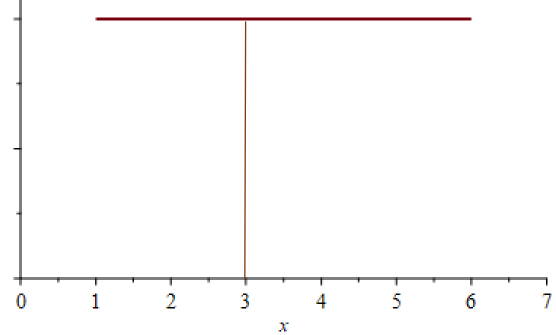
$P(0 \le X < 2)$

$1 \cdot \frac{1}{5} = \boxed{\frac{1}{5}}$



$P(X = 3)$

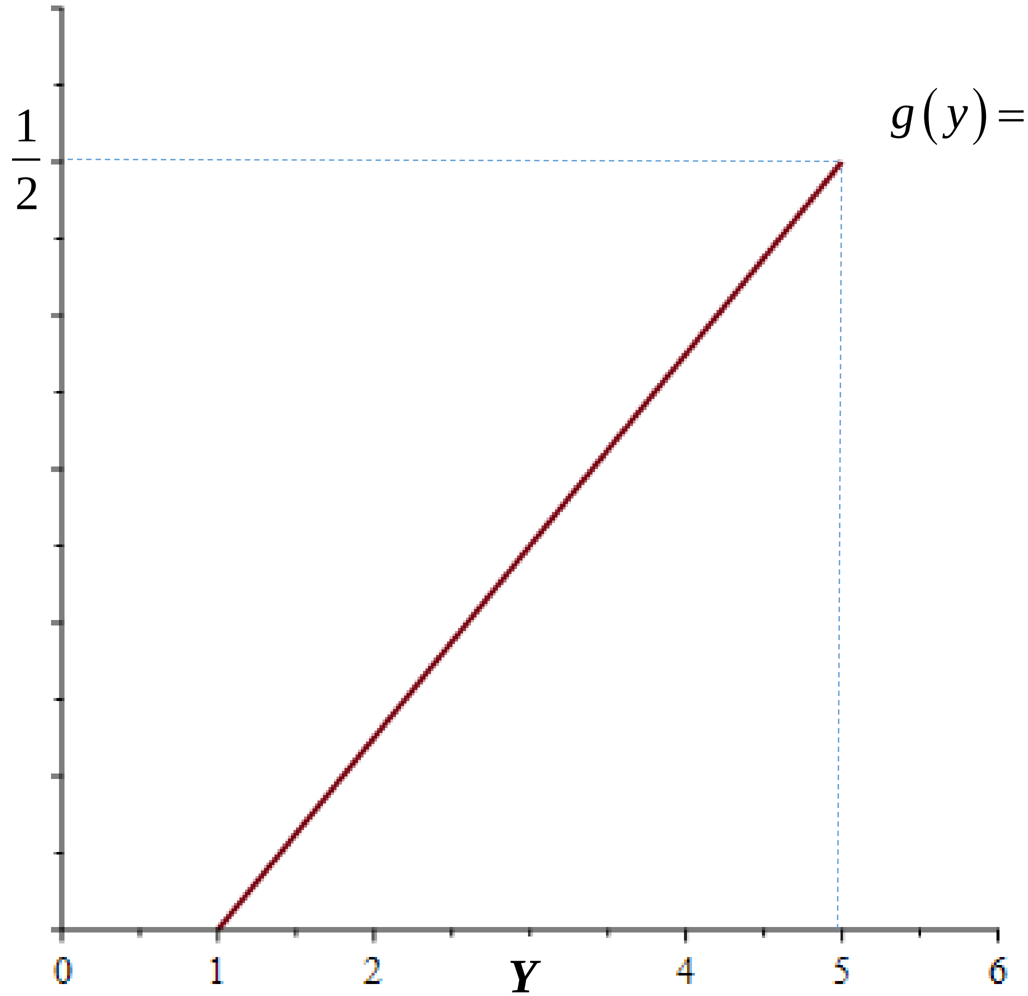
$0 \cdot \frac{1}{5} = \boxed{0}$



If *X* represents measurements from a population, then what proportion of the measurements fall between 1 and 2? $\frac{1}{5} = \boxed{20\%}$

2. Verify that the following curve can be a probability density curve for the random variable *Y* that takes on values between 1 and 5.

$$g(y) = \begin{cases} \frac{(y-1)}{8} & ;1 \le y \le 5 \\ 0 & ;\text{ otherwise} \end{cases}$$

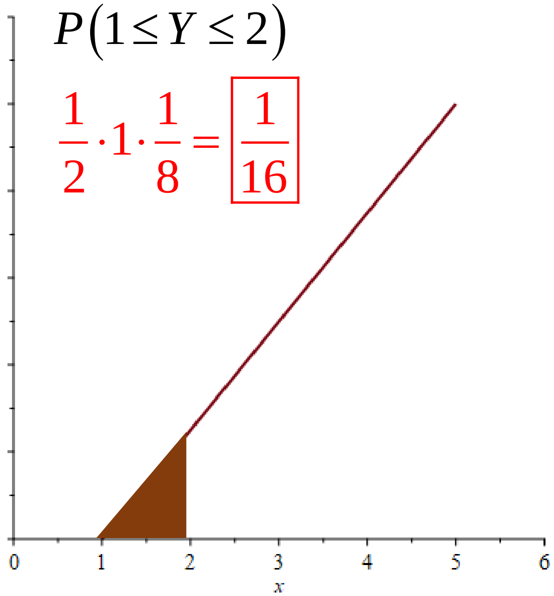


The region between the curve and the horizontal axis is a triangle. The area of a triangle is half the product of its base and height. For this triangle we get $\frac{1}{2} \cdot 4 \cdot \frac{1}{2} = 1$. So this curve can be a probability density curve.

Use the probability density curve to find the following probabilities:

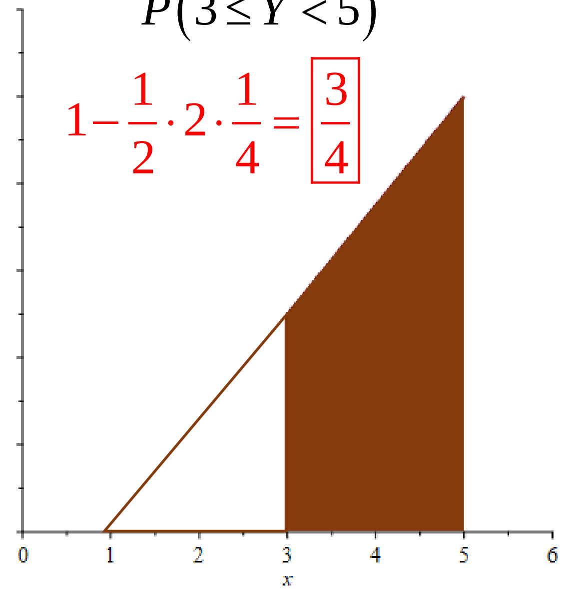
$P(1 \leq Y \leq 2)$

$\frac{1}{2} \cdot 1 \cdot \frac{1}{8} = \boxed{\frac{1}{16}}$



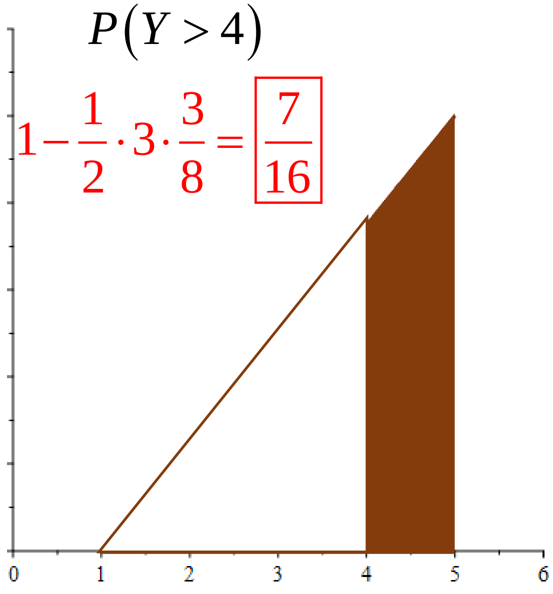
$P(3 \leq Y < 5)$

$1 - \frac{1}{2} \cdot 2 \cdot \frac{1}{4} = \boxed{\frac{3}{4}}$



$P(Y > 4)$

$1 - \frac{1}{2} \cdot 3 \cdot \frac{3}{8} = \boxed{\frac{7}{16}}$



$P(Y = 2)$

0

If *Y* represents measurements from a population, then what proportion of the measurements fall between 3 and 5?

$$\frac{3}{4} = \boxed{75\%}$$