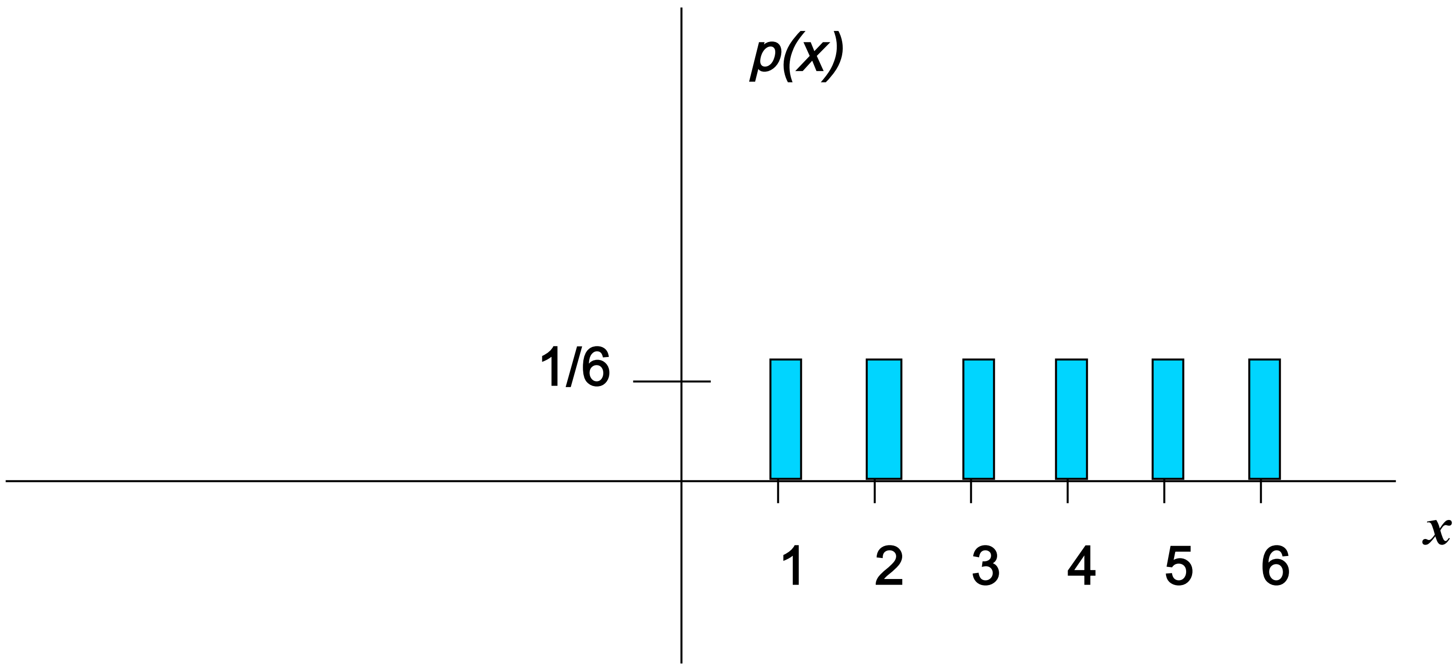
Random Variable

- A random variable *x* takes on a defined set of values with different probabilities.
    - For example, if you roll a die, the outcome is random (not fixed) and there are 6 possible outcomes, each of which occur with probability one-sixth.
    - For example, if you poll people about their voting preferences, the percentage of the sample that responds "Yes on Proposition 100" is a also a random variable (the percentage will be slightly differently every time you poll).
- Roughly, <u>probability</u> is how frequently we expect different outcomes to occur if we repeat the experiment over and over ("frequentist" view)

Probability functions

- A probability function maps the possible values of *x* against their respective probabilities of occurrence, $p(x)$
- $p(x)$ is a number from 0 to 1.0.
- The area under a probability function is always 1.

Discrete example: roll of a die

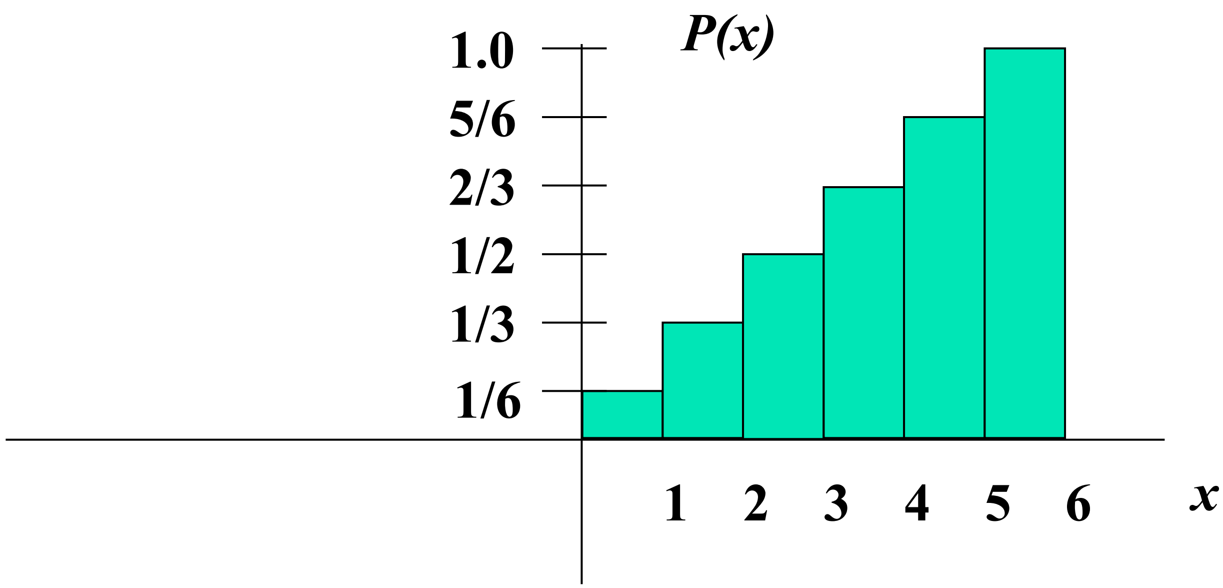


$$\sum_{\text{all } x} P(x) = 1$$

Probability mass function (pmf)

x	*p(x)*
1	$p(x=1)=1/6$
2	$p(x=2)=1/6$
3	$p(x=3)=1/6$
4	$p(x=4)=1/6$
5	$p(x=5)=1/6$
6	$p(x=6)=1/6$

Cumulative distribution function (CDF)



Cumulative distribution function

x	$P(x \leq A)$
1	$P(x \leq 1)$=1/6
2	$P(x \leq 2)$=2/6
3	$P(x \leq 3)$=3/6
4	$P(x \leq 4)$=4/6
5	$P(x \leq 5)$=5/6
6	$P(x \leq 6)$=6/6

Practice Problem:

- The number of patients in any given hour is a random variable represented by *x*. The probability distribution for *x* is:

x	10	11	12	13	14
P(x)	.4	.2	.2	.1	.1

Find the probability that in a given hour:

a. exactly 14 patients arrive $p(x=14) = .1$

b. At least 12 patients arrive $p(x \geq 12) = (.2 + .1 + .1) = .4$

c. At most 11 patients arrive $p(x \leq 11) = (.4 + .2) = .6$

Review Question 1

If you toss a die, what's the probability that you roll a 3 or less?

a. 1/6
b. 1/3
c. 1/2
d. 5/6
e. 1.0

Review Question 1

If you toss a die, what's the probability that you roll a 3 or less?

a. 1/6
b. 1/3
c. 1/2
d. 5/6
e. 1.0

Review Question 2

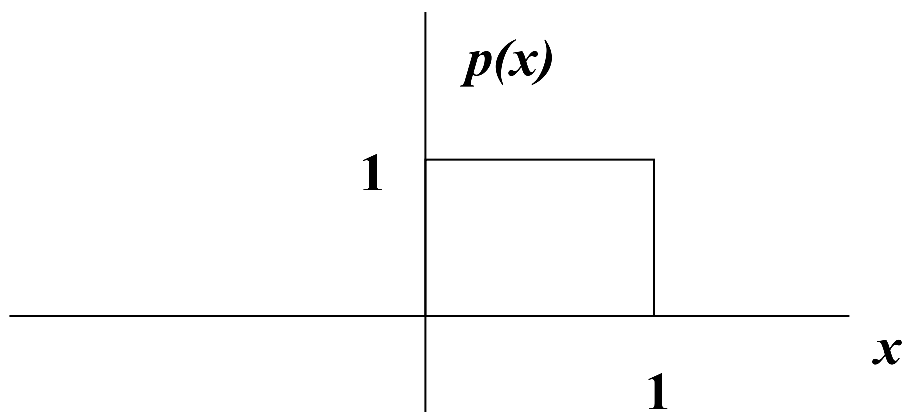
Two dice are rolled and the sum of the face values is six? What is the probability that at least one of the dice came up a 3?

a. 1/5
b. 2/3
c. 1/2
d. 5/6
e. 1.0

Example 2: Uniform distribution

The uniform distribution: all values are equally likely.

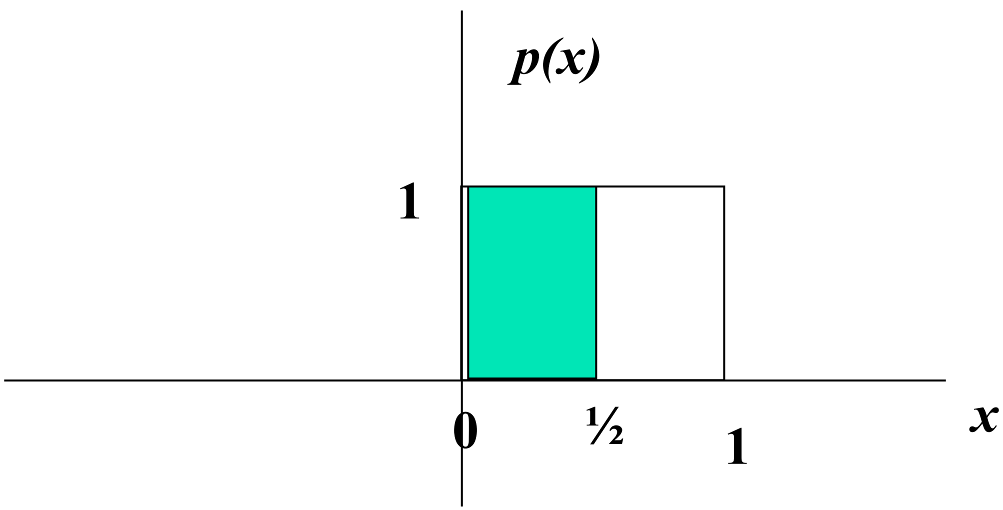
$f(x) = 1$, for $1 \geq x \geq 0$



We can see it's a probability distribution because the area of the rectangle is 1.

Example: Uniform distribution

What's the probability that x is between 0 and ½?



Clinical Research Example: When randomizing patients in an RCT, we often use a random number generator on the computer. These programs work by randomly generating a number between 0 and 1 (with equal probability of every number in between). Then a subject who gets X<.5 is control and a subject who gets X>.5 is treatment.

$$P(\tfrac{1}{2} \geq x \geq 0) = \tfrac{1}{2}$$

Expected Value and Variance

- All probability distributions are characterized by an expected value (mean) and a variance (standard deviation squared).

Expected value of a random variable

- Expected value is just the average or mean (μ) of random variable *x*.
- It’s sometimes called a “weighted average” because more frequent values of X are weighted more highly in the average.
- It’s also how we expect X to behave on-average over the long run (“frequentist” view again).

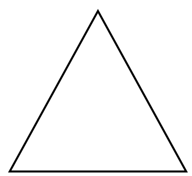
Expected value, formally

$$E(X) = \sum_{\text{all } x} x_i \, p(x_i)$$

Example: expected value

- Recall the following probability distribution of ER arrivals:

x	10	11	12	13	14
P(x)	.4	.2	.2	.1	.1



$$\sum_{i=1}^{5} x_i p(x) = 10(.4) + 11(.2) + 12(.2) + 13(.1) + 14(.1) = 11.3$$