Example: Suppose that 0.1% of individuals in a population have a particular disease. There is a diagnostic test for the disease, however, the test is not completely accurate: 95% of those who have the disease will test positive and 90% of those who don't have the disease will test negative. Suppose a person is randomly selected from the population and given the test.

$P(\text{ has disease }) =$ 0.001 $P(\text{ doesn't have disease }) =$ $1 - 0.001 = 0.999$

$P(\text{ positive} \mid \text{has disease })$

$=$ 0.95

$P(\text{ negative} \mid \text{doesn't have disease })$

$=$ 0.9

false negative

$P(\text{ negative} \mid \text{has disease })$

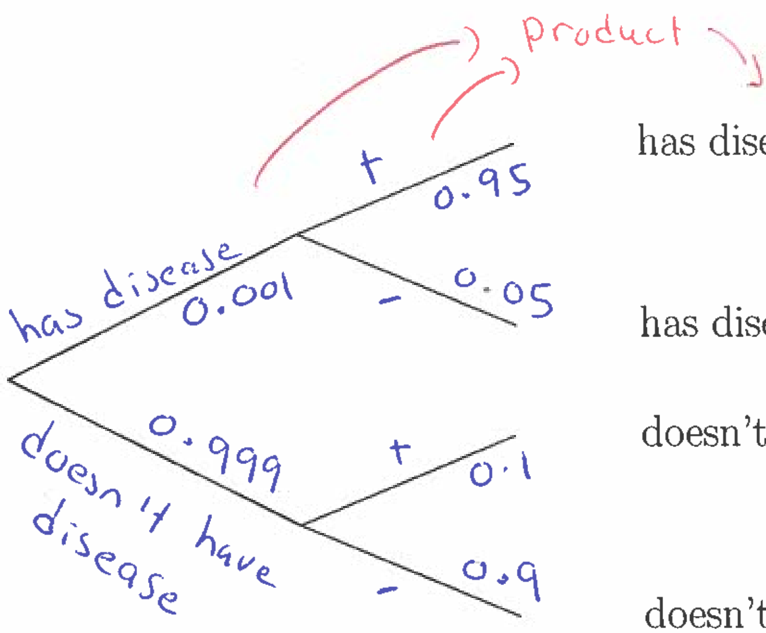
$=$ $1 - 0.95$

$=$ 0.05

false positive

$P(\text{ positive} \mid \text{doesn't have disease })$

$=$ $1 - 0.9 = 0.1$



has disease and positive

has disease and negative

doesn't have disease and positive

doesn't have disease and negative

conditional probabilities

$P(\text{ has disease } \cap \text{ positive })$
$= P(\text{ has disease })P(\text{ positive} \mid \text{has disease }) = (0.001)(0.95)$
$= 0.00095$

$P(\text{ has disease } \cap \text{ negative })$
$= P(\text{ has disease })P(\text{ negative} \mid \text{has disease }) = (0.001)(0.05)$
$= 0.00005$

$P(\text{ doesn't have disease } \cap \text{positive })$
$= P(\text{ doesn't have disease })P(\text{ positive} \mid \text{doesn't have disease })$
$= (0.999)(0.1) = 0.0999$

$P(\text{ doesn't have disease } \cap \text{negative })$
$= P(\text{ doesn't have disease })P(\text{ negative} \mid \text{doesn't have disease })$
$= (0.999)(0.9) = 0.8991$

	Test positive	Test negative	Total
Has disease	0.00095	0.00005	0.001
Doesn't have disease	0.0999	0.8991	0.999
Total	0.10085	0.89915	1

a) What is the probability that this person will test positive?

$P(\text{positive})$ disjoint
$= P(\text{ has disease } \cap \text{ positive }) + P(\text{ doesn't have disease } \cap \text{positive })$
$= 0.00095 + 0.0999 = 0.10085$

b) What is the probability that this person has the disease, given that they have tested positive?

$$P(\text{ has disease} \mid \text{positive }) = \frac{P(\text{ has disease } \cap \text{ positive })}{P(\text{positive})} = \frac{0.00095}{0.10085} \approx 0.0094$$

∴ a person who tests positive for the disease only has a 0.94% chance of having the disease.

Probability Rules:

Let S be a sample space of a random phenomenon and let A and B be two events consisting of outcomes of S.

- **Total Probability Rule:**

$$P(S) = 1$$

- **Complement Rule:**

$$P(A) = 1 - P(A^C)$$

- **General Addition Rule:**

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- **Disjoint Events:**

$$P(A \cap B) = 0$$

- **Addition Rule for Disjoint Events:**

$$P(A \cup B) = P(A) + P(B)$$

- **Conditional Probability:**

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- **General Multiplication Rule:**

$$\begin{aligned} P(A \cap B) &= P(B)P(A|B) \\ &= P(A)P(B|A) \end{aligned}$$

- **Independent Events:**

$$P(A|B) = P(A)$$

- **Multiplication Rule for Independent Events:**

$$P(A \cap B) = P(A)P(B)$$

Chapter 13: Random Variables

A **random variable** is a variable that associates a numerical value with each outcome of a random event. ↳ Capital letters X

→ specific value x

A random variable is said to be:

- **discrete** if its set of possible values is finite or countably infinite (isolated points on a number line).

↑ gap

randomly Selected

  - number of defective tires on a car
  - number of voters in a sample who support impeaching the president
  - seating capacity of an airplane.
  - number of bacteria in a cubic centimeter of water.

- **continuous** if its set of possible values is uncountably infinite (an entire interval on a number line).

andomly Selected

  - body temperature of a hospital patient
  - fuel efficiency of an automobile
  - length of time an employee is late for work
  - distance traveled by a student from home to the university

Example: Flip a coin.

X = number of heads observed.

Outcome	Value of X
T	0
H	1

The **probability distribution** (or probability model) of a random variable is a function that associates a probability to:

- each value of a discrete random variable X.
- any interval of values of a continuous random variable X (using a density curve). → Chapter 5

Notation: $P(X = x)$ or $P(x)$

The probability distribution $P(x)$ of a discrete random variable satisfies:

- $0 \leq P(x) \leq 1$, for each value x of X
- $\sum_{x} P(x) = 1$

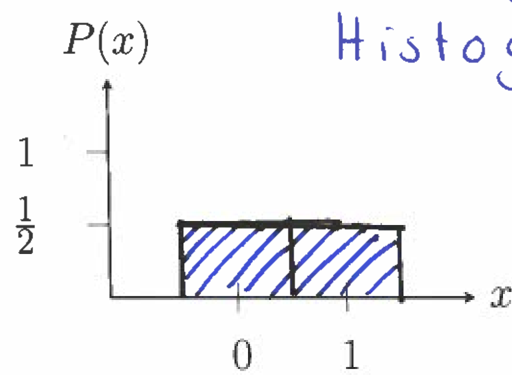
↑ sum using <u>all</u> values x of X

Example: Flip a fair coin.

X = number of heads observed.

x	$P(x)$
0	$\frac{1}{2} = 0.5$
1	$\frac{1}{2}$

Probability Histogram



$6 \times 6 = 36$

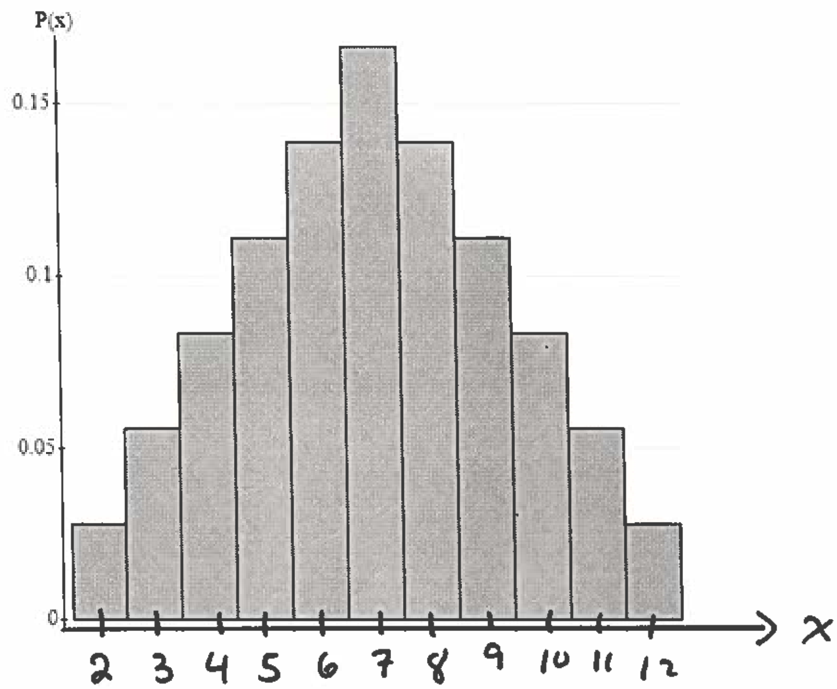
Example: Roll two fair dice.

Outcomes:

$$\begin{array}{cccccc} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{array}$$

a) X = sum of the two dice

x	$P(x)$ $\rightarrow P(X = x)$
2	$\frac{1}{36}$
3	$\frac{2}{36}$
4	$\frac{3}{36}$
5	$\frac{4}{36}$
6	$\frac{5}{36}$
7	$\frac{6}{36}$
8	$\frac{5}{36}$
9	$\frac{4}{36}$
10	$\frac{3}{36}$
11	$\frac{2}{36}$
12	$\frac{1}{36}$
	total = 1



Compute the following:

$P(X=2)$

i) $P(X \leq 4)$ $= p(2) + p(3) + p(4)$

include 4

$= \frac{1}{36} + \frac{2}{36} + \frac{3}{36}$

$= \frac{6}{36} \simeq 0.167$

ii) $P(6 \leq X < 11)$ $= P(6) + P(7) + P(8) + P(9) + P(10)$

include 6 don't include 11

$= \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} + \frac{3}{36}$

$= \frac{23}{36} \simeq 0.639$

iii) $P(X \geq 10)$ $= p(10) + p(11) + P(12)$

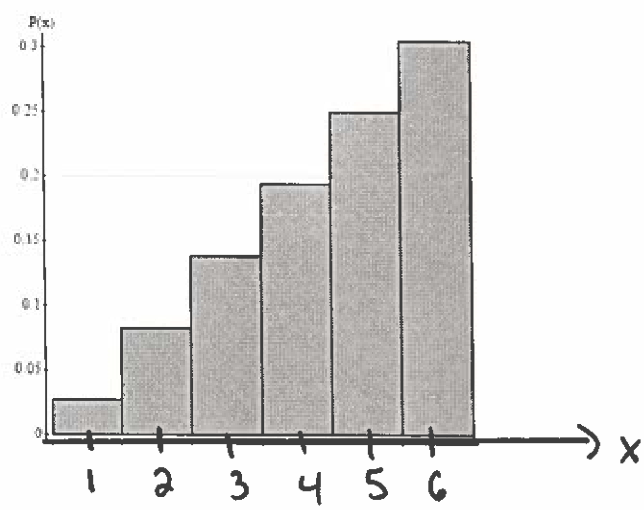
$= \frac{3}{36} + \frac{2}{36} + \frac{1}{36}$

$= \frac{6}{36} \simeq 0.167$

b) X = maximum of the numbers showing on the two dice

x	$P(x)$
1	$\frac{1}{36}$
2	$\frac{3}{36}$
3	$\frac{5}{36}$
4	$\frac{7}{36}$
5	$\frac{9}{36}$
6	$\frac{11}{36}$

total = 1



Expected Value or Mean

To describe the center and spread of a probability distribution of a random variable, we often use the mean and the standard deviation, respectively.

The probability distribution is a model, so the mean and the standard deviation are parameters of this model.

Notation: $\mu =$ mean $\sigma =$ standard deviation

Let X be a discrete random variable with probability distribution $P(x)$. The mean or **expected value** of X is given by

$$\mu = E(X) = \sum_x xP(x)$$

Example: Roll two fair dice.

a) $X =$ sum of the two dice.

$$E(X) = 2\left(\tfrac{1}{36}\right) + 3\left(\tfrac{2}{36}\right) + 4\left(\tfrac{3}{36}\right) + 5\left(\tfrac{4}{36}\right) + 6\left(\tfrac{5}{36}\right) + 7\left(\tfrac{6}{36}\right)$$
$$+ 8\left(\tfrac{5}{36}\right) + 9\left(\tfrac{4}{36}\right) + 10\left(\tfrac{3}{36}\right) + 11\left(\tfrac{2}{36}\right) + 12\left(\tfrac{1}{36}\right)$$
$$= \frac{252}{36} = 7$$

b) $X =$ maximum of the numbers showing on the two dice

$$E(X) = 1\left(\tfrac{1}{36}\right) + 2\left(\tfrac{3}{36}\right) + 3\left(\tfrac{5}{36}\right) + 4\left(\tfrac{7}{36}\right) + 5\left(\tfrac{9}{36}\right) + 6\left(\tfrac{11}{36}\right)$$
$$= \frac{161}{36} \simeq 4.47$$

Example: You draw a card from a standard deck of 52 cards. If you draw a black card, you win nothing. If you draw a diamond, you win \$5. If you draw a heart, you win \$10, unless it is the Queen of Hearts. You win \$30 for the Queen of Hearts.

a) Create a probability distribution for the amount of money you win at this game. X = amount of winnings

Outcome	x	$P(x)$	
black	0	$\frac{1}{2}$	$\left(\frac{26}{52}\right)$
diamond	5	$\frac{1}{4}$	$\left(\frac{13}{52}\right)$
heart (not queen)	10	$\frac{12}{52}$	
Queen of Hearts	30	$\frac{1}{52}$	

b) Find your expected winnings.

$$E(x) = 0\left(\tfrac{1}{2}\right) + 5\left(\tfrac{1}{4}\right) + 10\left(\tfrac{12}{52}\right) + 30\left(\tfrac{1}{52}\right)$$
$$= 0 + \frac{5}{4} + \frac{120}{52} + \frac{30}{52}$$
$$= \frac{215}{52} \simeq \$4.13$$

c) Would you pay \$6 to play this game? How about \$2?

\$6: No, \$6 > E(x) = \$4.13

\$2: Yes, \$2 < E(x) = \$4.13

Example: A company has five applicants for two positions. Three of the applicants are male and two of the applicants are female. Suppose that all applicants are equally qualified and that no preference is given for choosing either gender. → randomly Selected

a) Create a probability distribution for the number of females chosen to fill the two positions.

X = number of females hired

x	$P(x)$
0	0.3
1	0.6
2	0.1
	1

0: mm

$$\frac{3}{5} \cdot \frac{2}{4} = \frac{6}{20} = 0.3$$

1: mf or fm

$$\frac{3}{5} \cdot \frac{2}{4} + \frac{2}{5} \cdot \frac{3}{4} = \frac{12}{20} = 0.6$$

2: ff

$$\frac{2}{5} \cdot \frac{1}{4} = \frac{2}{20} = 0.1$$

b) Find the expected number of females hired.

$$E(x) = 0(0.3) + 1(0.6) + 2(0.1)$$
$$= 0.8$$

Variance and Standard Deviation

The variance of a random variable is the expected value of the squared deviations from the mean.

Let X be a discrete random variable with probability distribution $P(x)$ and mean μ. The **variance** of X is:

$$\sigma^2 = \mathrm{Var}(X) = \sum_x (x-\mu)^2 P(x)$$

and the **standard deviation** is

$$\sigma = SD(X) = \sqrt{\mathrm{Var}(X)}$$

Alternative formula for variance:

$$\sigma^2 = \mathrm{Var}(X) = \sum_x x^2 P(x) - \mu^2$$

Example: Card Game $\mu = E(X) = \frac{215}{52} \simeq 4.13$

Outcome	x	$P(x)$	$(x-\mu)^2 P(x)$
Black	0	$\frac{1}{2}$	$\left(0 - \frac{215}{52}\right)^2 \cdot \frac{1}{2}$
Diamond	5	$\frac{1}{4}$	$\left(5 - \frac{215}{52}\right)^2 \cdot \frac{1}{4}$
Heart (not queen)	10	$\frac{3}{13}$	$\left(10 - \frac{215}{52}\right)^2 \cdot \frac{3}{13}$
Queen of Hearts	30	$\frac{1}{52}$	$\left(30 - \frac{215}{52}\right)^2 \cdot \frac{1}{52}$

$\mathrm{Var}(X) \simeq 29.54$

$SD(X) = \sqrt{\mathrm{Var}(X)} = \sqrt{29.54} \simeq 5.44$

Example: Roll two fair dice. $\mu = E(x) = \frac{161}{36}$

X = maximum of the numbers showing on the two dice

x	$P(x)$	$x^2P(x)$	
1	$\frac{1}{36}$	$1^2\left(\frac{1}{36}\right)$	$\frac{1}{36}$
2	$\frac{1}{12}$	$2^2\left(\frac{1}{12}\right)$	$\frac{12}{36}$
3	$\frac{5}{36}$	$3^2\left(\frac{5}{36}\right)$	$\frac{45}{36}$
4	$\frac{7}{36}$	$4^2\left(\frac{7}{36}\right)$	$\frac{112}{36}$
5	$\frac{1}{4}$	$5^2\left(\frac{1}{4}\right)$	$\frac{225}{36}$
6	$\frac{11}{36}$	$6^2\left(\frac{11}{36}\right)$	$\frac{396}{36}$
		$\frac{791}{36}$	

$$Var(x) = \sum_x x^2 P(x) - \mu^2$$

$$= \frac{791}{36} - \left(\frac{161}{36}\right)^2$$

$$\simeq 1.97$$

$$SD(x) = \sqrt{Var(x)}$$

$$\simeq 1.4$$