# PROBABILITY

**Why Use Probability.**

- We use it to describe behavior of uncertain things.

  Unlike $2^3$ always gives 8, we can't say whether a throw of a dice will give 2 or not. (We know it won't give 7 or –2.)

- However, one thing is certain: If we collect data from a very large number of throws of a perfect dice (a perfect cube, with its center of gravity at the cube's center), each of the numbers 1, 2, ⋯, 6 will appear approx. 1/6 of the times.

- Probability Theory finds the things that are certain even in presence of uncertainties.

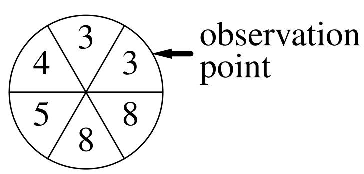
## Sample Space of an Experiment.

- Sample space $S$: set of all possible outcomes of the experiment.

  (a) Experiment: A toss of a coin.
  $S$: $\{H, T\}$; here, $H$ = head and $T$ = tail.

  (b) Experiment: Two tosses of a coin.
  $S$: $\{HH, HT, TH, TT\}$; here $HT$ means 1st toss is $H$ and 2nd is $T$.

  (c) Experiment: The number in the sector pointed to (or on its right if it falls on the line between two sectors) by the fixed arrow when we turn the spinning wheel shown below, with 6 equal size sectors labeled 3, 3, 8, 8, 4, and 5 (in some order).
  $S$: $\{3, 4, 5, 8\}$.




**Basic Probabilities of A Sample Point $x \in S$.**

- It is the idealized ratio of frequency of $x$ and the number repetitions of the experiment, when the later is very very large (tends to infinity) and the experiments are "independent" in that one experiment's result doesn't affect the other experiment's result.

  (a) If the coin is ideal ("unbiased"), then Prob($H$) = ½ = Prob($T$).
  (b) If addition, if the two throws in the experiment are "independent", then Prob($HH$) = Prob($HT$) = Prob($TH$) = Prob($TT$) = 1/4.
  (c) No matter how the numbers in the wheel are arranged, Prob(3) = 2/6 = 1/3 = Prob(8) and Prob(4) = 1/6 = Prob(5).

**Practice Questions.**

1. Create a spinning wheel with its equal size sectors marked $H$ and $T$ so that Prob($H$) = 1/5 and Prob($T$) = 4/5.
2. Suppose our experiment consists of three turns of a spinning wheel as in Problem 1, and we write the outcome of an experiment as a triplet ($s_1$, $s_2$, $s_3$) or simply as $s_1 s_2 s_3$ (in short), where $s_i$ is the outcome ($H$ or $T$) of $i$th turn of the wheel. Show the sample space $S$ and the basic probabilities Prob($x$) for each $x \in S$. (Hint: the three turns can produce $5\times5\times5 = 125$ possible combinations of sectors and the outcome ($H$, $T$, $T$) corresponds to $1\times4\times4 = 16$ of them, giving its probability 16/125 = (1/5)(4/5)(4/5).)
3. Verify the following two properties of Prob($x$)'s for $x \in S$ in Problem 2. (They hold for the sample space of any experiment; for example, they hold for each of the three sample spaces in the previous page.)
   (a) For each $x \in S$, $0 \leq \text{Prob}(x) \leq 1$.
   (b) $\sum_{x \in S} \text{Prob(x)} = 1$.

# COMPLEX EVENTS AND THEIR PROBABILITIES

**Events.**

- An event $E$ is simply a subset of the sample space $S$; thus, there are $2^{|S|}$ many events, including $E = \varnothing$ and $E = S$.
- $\text{Prob}(E) = \sum_{x \in E} \text{Prob}(x)$; in particular, we always have $\text{Prob}(\varnothing) = 0$ and $\text{Prob}(S) = 1$.

**Examples.**

Consider the experiment "turning of the wheel", whose sample space $S = \{3, 4, 5, 8\}$, with Prob(3) = 1/3 = Prob(8) and Prob(4) = 1/6 = Prob(5).

($a$) For $E_1 = \{4, 5, 8\} \subseteq S$, i.e., the outcome is at least 4 (which is the same as "outcome is not 3"), $\text{Prob}(E_1)$ = Prob(4) + Prob(5) + Prob(8) = 1/6 + 1/6 + 1/3 = 4/6 = 2/3.

($b$) For $E_2 = \{3\} \subseteq S$, i.e., the outcome is less than 4, $\text{Prob}(E_2)$ = Prob(3) = 1/3.

($c$) For $E_3 = \{4, 5\} \subseteq S$, $\text{Prob}(E_3)$ = Prob(4) + Prob(5) = 1/6 + 1/6 = 1/3.

($d$) For $E_4 = \{3, 4, 5\} \subseteq S$, $\text{Prob}(E_4)$ = Prob(3) + Prob(4) + Prob(5) = 1/3 + 1/6 + 1/6 = 2/3.

**Two Basic Rules of Probability of Events.**

- Sum-rule: $\text{Prob}(E \cup E') = \text{Prob}(E) + \text{Prob}(E') - \text{Prob}(E \cap E')$.

  A special case: for disjoint events $E$ and $E'$, $\text{Prob}(E \cup E') = \text{Prob}(E) + \text{Prob}(E')$.
- Complement-rule: For an event $E$ and its complement $E^c$ = not $E$, $\text{Prob}(E) + \text{Prob}(E^c) = 1$, i.e., $\text{Prob}(E^c) = 1 - \text{Prob}(E)$.

  It is a special case of sum-rule; $\text{Prob}(E) + \text{Prob}(E^c) = \text{Prob}(E \cup E^c) = \text{Prob}(S) = 1$.

**Notes.**

- The sum-rule is a generalization of $|A \cup B| = |A| + |B| - |A \cap B|$ for set-cardinalities.
- The sum-rule is one of those certain things in presence of uncertainties.

**Examples.**

(i) In Examples ($a$), ($c$), and ($d$) above, we have $\text{Prob}(S) = \text{Prob}(E_1 \cup E_4) = 1 = 2/3 + 2/3 - 1/3 = \text{Prob}(E_1) + \text{Prob}(E_4) - \text{Prob}(E_1 \cap E_4)$ because $E_1 \cap E_4 = \{4\}$.

(ii) In Examples (a)-(b) above, we have $E_2 = E_1^c$ and $\text{Prob}(E_1) + \text{Prob}(E_1^c) = 2/3 + \text{Prob}(E_2) = 2/3 + 1/3 = 1$.

**Practice Problems.**

1. Consider three turns of a spinning wheel with 5 equal size sectors labeled $H$, $T$, $T$, $T$, and $T$ in some order, giving Prob($H$) = 1/5 and Prob($T$) = 4/5 in each turn of the wheel. Determine the set of sample points $s_1s_2s_3$, where each $s_i = H$ or $T$ (outcome of $i$th turn of the wheel), for each of the following events and compute probabilities of those events.
   ($a$) $E_1 = \{s_1s_2s_3: \#(H) \le 2\}$.
   ($b$) $E_2 = \{s_1s_2s_3: \#(H) \ge 2\}$.
   ($c$) $E_3 = \{s_1s_2s_3: \#(H) \text{ equals } 2\} = E_1 \cap E_2$.
   (d) $E_4 = \{s_1s_2s_3: \#(H) \le 1\} = E_2^c$.
2. Verify the sum-rule and the complement-rule involving the events in Problem 1.
3. Consider three turns of a spinning wheel with 6 equal size sectors, where two sectors are labeled 3, two sectors are labeled 8, and the other two sectors are labeled 4 and 5, one each. Determine the sample points $(s_1, s_2, s_3)$ corresponding to the following events and determine probabilities of the events.
   (a) There is an equilateral triangle with the sides $s_1$, $s_2$, and $s_3$.
   (b) There is a right-angled triangle with the sides $s_1$, $s_2$, and $s_3$.
   (c) There is an equilateral or right-angled triangle with the sides $s_1$, $s_2$, and $s_3$.
   (d) There is an isosceles triangle with the sides $s_1$, $s_2$, and $s_3$.
   (e) There is no triangle with the sides $s_1$, $s_2$, and $s_3$.
4. Verify the sum-rule involving the events (a)-(c) in Problem 3.

# BINOMIAL PROBABILITY DISTRIBUTION

**Probability of $X$ = #($H$) in $n$ = 2 Tosses of A Coin.**

- Assume $p = \text{Prob}(H)$ and $q = 1 - p = \text{Prob}(T)$.
- Shown below are values of $X$, the associated sample points, and the probabilities.

| $X$ | Related sample points in $S$ | Probability |
|---|---|---|
| 0 | $TT$ | $\text{Prob}(X = 0) = q^2$ |
| 1 | $HT$, $TH$ | $\text{Prob}(X = 1) = pq + qp = 2pq$ |
| 2 | $HH$ | $\text{Prob}(X = 2) = p^2$ |

**Binomial Probability Distributions.**

- The case $n = 1$:

  Sample space of values of $X$ is $\{0, 1\}$;
  Probabilities $\text{Prob}(0) = q$, $\text{Prob}(1) = p$.
  The sum of these probabilities = $p + q = 1$.

- The case $n = 2$:

  Sample space of values of $X$ is $\{0, 1, 2\}$;
  Probabilities $\text{Prob}(0) = q^2$, $\text{Prob}(1) = 2pq$, and $\text{Prob}(2) = p^2$.
  The sum of these probabilities = $p^2 + 2pq + q^2 = (p + q)^2 = 1$.

- The case $n = 3$:

  Sample space of values of $X$ is $\{0, 1, 2, 3\}$;
  Probabilities $\text{Prob}(0) = q^3$, $\text{Prob}(1) = 3pq^2$, $\text{Prob}(2) = 3p^2q$, and $\text{Prob}(3) = p^3$.
  The sum of these probabilities = $p^3 + 3p^2q + 3pq^2 + q^3 = (p + q)^3 = 1$.

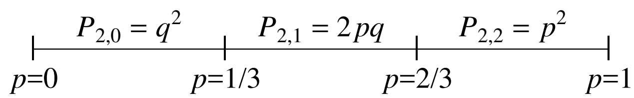
- The general case $n \geq 1$:

  Sample space of values of $X$ is $\{0, 1, \cdots, n\}$;
  Probabilities $\text{Prob}(i) = C(n, i)p^i q^{n-i}$ for $0 \leq i \leq n$.
  The sum of these probabilities = $\sum_{0 \leq i \leq n} C(n, i)p^i q^{n-i} = (p + q)^n = 1$.

**Practice Questions.**

Below we use the short notation $P_{n,i}$ = Prob($X$ = $i$) = $C(n, i) p^i q^{n-i}$; we use a shorter notation $P_i$ when $n$ is understood from the context.

1. Carry out the analysis shown below for $n = 2$ to the case of $n = 3$ and the proabilities $P_{3,i}$.
   (a) First, $P_{2,0} = q^2 < 2pq = P_{2,1}$ when $q < 2p$, i.e., $1 = p + q < 3p$, i.e., $p > 1/3$.
   (b) Next, $P_{2,1} = 2pq < p^2 = P_{2,2}$ when $2q < p$, i.e., $2 = 2(p + q) < 3p$, i.e., $p > 2/3$.
   (c) The diagram below shows the 3-way partition of the interval [0, 1] into 3 subintervals of length 1/3 each and the successive probabilities $P_{2,0}$, $P_{2,1}$, and $P_{2,2}$ that are maximum in the successive subintervals.




Subintervals in which each $P_{2,i}$ is maximum among $P_{2,0}$, $P_{2,1}$, and $P_{2,2}$.

One use of the above analysis is to answer the question "Given a value of $p$, $0 \le p \le 1$, which value of $X = \#(H)$ in $n = 2$ tosses has the highest probability?". For $0 \le p < 1/3$, $X = 0$ has the highest probability, for $1/3 < p < 2/3$, $X = 1$ has the highest probability, and for $2/3 < p \le 1$, $X = 2$ has the highest probability. At the crossing point $p = 1/3$, both $X = 0$ and $X = 1$ have the highest probability and at the crossing-point $p = 2/3$, both $X = 1$ and $X = 2$ have the highest probability.

2. Repeat Problem 1 for the general case of $n$.

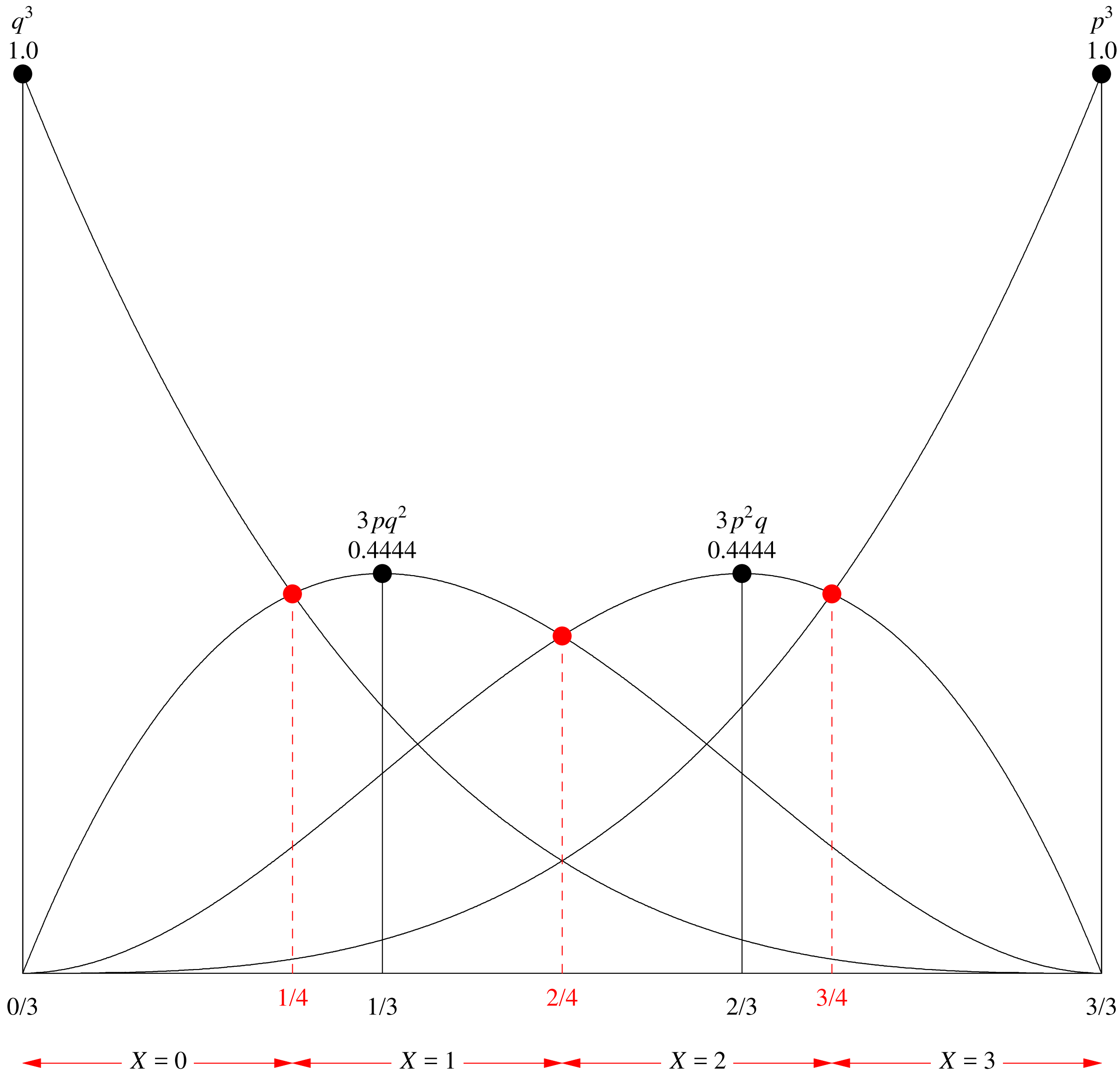
3. To find the maximum value $M_{n,i}$ of $P_{n,i} = C(n, i) p^i q^{n-i}$ for $0 < i < n$, we take its derivative with respepect to $p$ and set it equal to 0. This gives $i.\, p^{i-1} q^{n-i} - (n - i) p^i q^{n-i-1} = 0$, i.e., $iq - (n - i)p = 0$, i.e., $p = i/n$. Thus, $M_{n,i} = C(n, i)(i/n)^i((n - i)/n)^{n-i}$. The maximum of $P_{n,0}$ is at $p = 0$ and that of $P_{n,n}$ is at $p = 1$; both these maximum values are 1. The figures in the next 2 pages indicate the maximum values $M_{n,i}$ for $P_{n,i}$'s for $n = 3$ and 4.

   Do you notice any "trends" in these maximum values $M_{n,i}$'s?

   Is there a value of $p$ such that $P_{2,1} > 1/2$? How about a value of $p$ such that $P_{3,1} > (2/3)^2$?

4. If $p$ is slightly above 1/2, what should be the your lucky choice for the number of heads in $n = 3$ tosses so that you have the best chance of winning? What are the answers for $n = 4$ and 5?

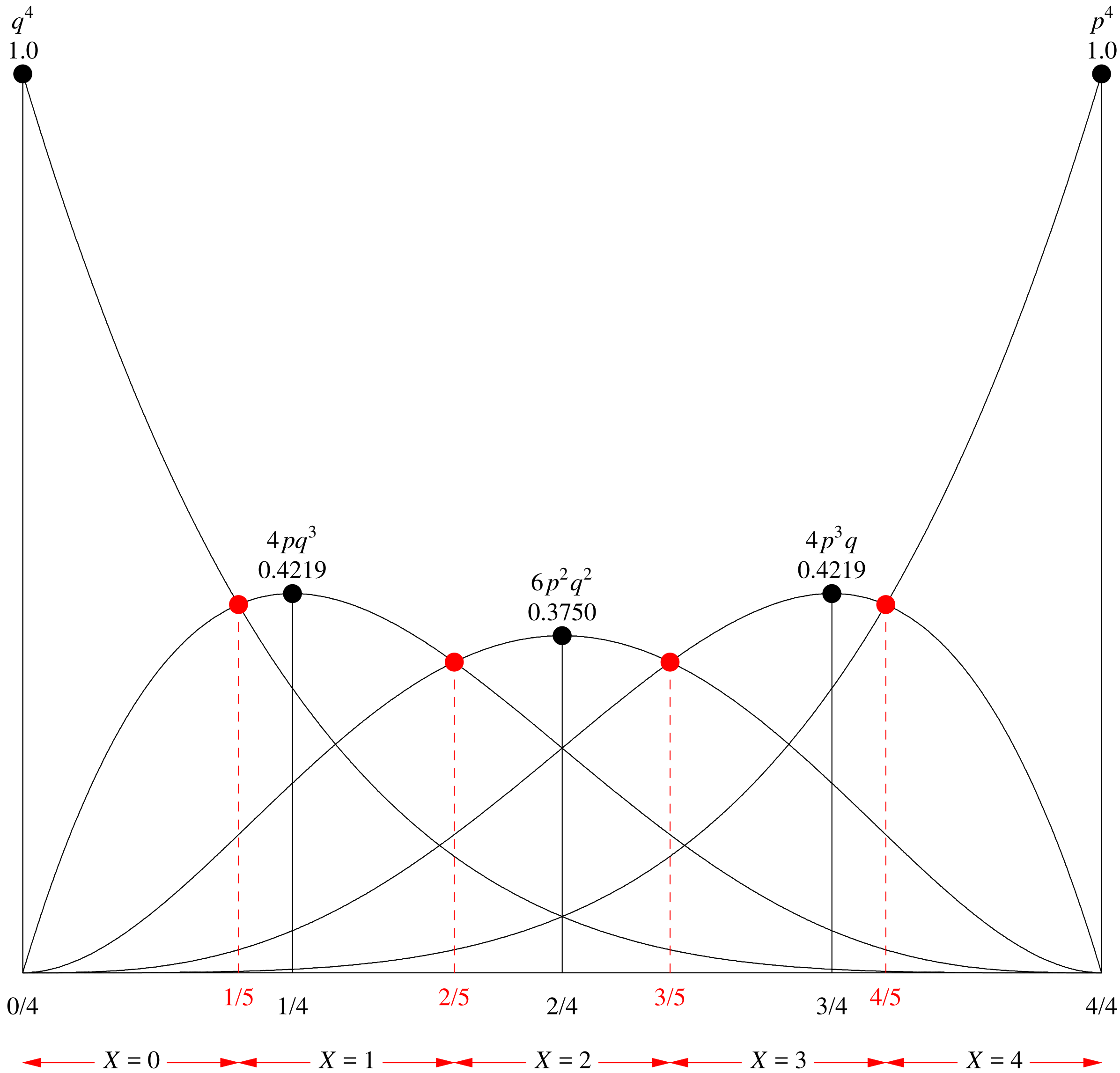
# THE CURVES $B_{3,m}(p) = C(3, m)p^m q^{3-m}$, THEIR CROSSING-POINTS, AND THEIR MAXIMUM-POINTS




Max. probability choice for $X = \#(H)$ in $n = 3$ tosses for various intervals of values of $p = \text{Prob}(H)$.

The maximum-points are shown in black-circles and the crossing-points between $B_{3,m}(p)$ and $B_{3,m+1}(p)$ are shown in red-circles.

# THE CURVES $B_{4,m}(p) = C(4, m)p^m q^{4-m}$, THEIR CROSSING-POINTS, AND THEIR MAXIMUM-POINTS




Max. probability choice for $X = \#(H)$ in $n = 4$ tosses for various intervals of values of $p = \text{Prob}(H)$.

The maximum-points are shown in black-dots and the crossing-points between $B_{4,m}(p)$ and $B_{4,m+1}(p)$ are shown in black-dots.

# EXPECTED VALUE *E(X)*

## A Finite Discrete Random Variable $X$.

- A random variable $X$ has a probability associated with each of its possible value and those values form a finite sample space $S_X$ of numbers.

  This is obtained by defining a function $X$ on a sample space $S$ with given probabilities Prob($s$) for $s \in S$. The set $S_X$ = values of $X$ on sample points in $S$ and we define

  $$\text{Prob}(X = x) = \sum_{s \in S \text{ and } X(s)=x} \text{Prob}(s).$$

  For example, $X = \#(H$ in $n$ tosses of a coin), $S_X = \{0, 1, 2, \cdots, n\}$, and Prob$(X = i) = C(n, i)p^i q^{n-i}$. (There should be no confusion between $S_X$ and the sample space $S$ of the underlying experiment.)

- We allow both values 0 and 1 for Prob$(X = x)$, $x \in S_X$. (If $S_X = \{x\}$, Prob$(X = x) = 1$; allowing Prob$(X = x) = 0$ for some $x \in S_X$ makes comparison of different random variables easier.)

Note: An ordinary variable $x$ has no notion of an associated probability for it.

## Expected Value $E(X)$ of A Random Variable $X$.

- $E(X) = \sum_{x \in S_X} x.\text{Prob}(X = x)$.
- For $X = \#(H$ in $n$ tosses), we can write
  $E(\#(H \text{ in } n \text{ tosses})) = \sum_{0 \le i \le n} i.\text{Prob}(\#(H \text{ in } n \text{ tosses}) = i)$.

## Example.

Consider Binomial probability distributions.

- For $n = 1$, $E(X) = 0.q + 1.p = p$.
- For $n = 2$, $E(X) = 0.q^2 + 1.2pq + 2.p^2 = 2p(q + p) = 2p$.
- For $n = 3$, $E(X) = 0.q^3 + 1.3pq^2 + 2.3p^2q + 3.p^3 = 3p(q^2 + 2pq + q^2) = 3p$.
- For the general case $n \ge 1$,

$$\begin{aligned}
E(X) &= \sum_{0 \le i \le n} i.\,C(n, i)p^i q^{n-i} \\
&= \sum_{1 \le i \le n} i.\,C(n, i)p^i q^{n-i} \\
&= np.\sum_{1 \le i \le n} C(n-1, i-1)p^{i-1}q^{n-i} \text{ because } i.\,C(n, i) = n.\,C(n-1, i-1) \\
&= np.\sum_{0 \le j \le n-1} C(n-1, j)p^j q^{(n-1)-j}, \text{ putting } j = i-1 \\
&= np.\,(p+q)^{n-1} = np.
\end{aligned}$$

**Practice Questions.**

1. Compute expected value of $X$, where $S_X = \{3, 4, 5, 8\}$ with Prob(3) = 1/3 = Prob(8) and Prob(4) = 1/6 = Prob(5).
2. John has three friends $A$, $B$, and $C$ and he likes to go out for dinner with one of them each night. Assume that the probability of choosing them are, respectively, 1/2, 1/3, and 1/6. How many times he is likely to go to dinner in April with each of them? If John had a fourth friend and all we knew was that probability of going out with $A$ was 1/2, then with whom John went out for dinner most often (justify your answer)?
3. Let $c \neq 0$ be a constant and let $Y = cX$, where $X$ is as in Problem 1. Then, the possible values of $Y$ are $S_Y = \{3c, 4c, 5c, 8c\}$ and Prob($Y = 3c$) = Prob($X = 3$) = Prob(3) = 2/3, etc. Compute $E(Y)$ and show that it equals $cE(X)$.

   Note that if $c = 0$, then $S_Y = \{0\}$ and thus $E(Y) = 0$.Prob($Y = 0$) $= 0 \times 1 = 0 = 0.E(X)$. Thus, $E(cX) = c.\,E(X)$ for all $c$.

   The equality $E(cX) = c.\,E(X)$ holds for all random variable $X$. (This is one of those certain things in presence of uncertainties that Probability Theory finds out.)
4. If we consider $Y = X^2$, where $X$ is as in Problem 1, then what are the possible values of $Y$ and what are the probabilities associated with those values?
5. Compute $E(Y)$ based on your solution of Problem 4.

# THE EQUATION $E(X+Y) = E(X) + E(Y)$

The equation $E(X+Y) = E(X) = E(Y)$ holds for any two random variables $X$ and $Y$ defined on a sample space $S$ with any given probabilities Prob($s$) for $s \in S$. (This is also one of those certain things in presence of uncertainties that Probability Theory finds out.)

**Example.**

Consider the sample space $S = \{HH, HT, TH, TT\}$ of the experiment "2 (independent) tosses of a coin" with Prob($H$) = $p$, $0 \le p \le 1$, for each toss.

- Let $X = 4.\#(H)$ and $Y = 5.\#(T)$. Shown below are the values of $X$ and $Y$, the related sample points in $S$, and the probabilities different values of $X$ and $Y$.

Here, $q = 1 - p$

| $X$ | Related sample points in $S$ | Probability | $Y$ | Related sample points in $S$ | Probability |
|---|---|---|---|---|---|
| 0 | $TT$ | $\text{Prob}(X=0) = q^2$ | 0 | $HH$ | $\text{Prob}(Y=0) = p^2$ |
| 4 | $HT$, $TH$ | $\text{Prob}(X=4) = 2pq$ | 5 | $HT$, $TH$ | $\text{Prob}(Y=5) = 2pq$ |
| 8 | $HH$ | $\text{Prob}(X=8) = p^2$ | 10 | $TT$ | $\text{Prob}(Y=10) = q^2$ |

- We have $Z = X + Y = 4.\#(H) + 5.\#(T)$ for each sample point in $S$. This means $Z(HH) = 8 + 0 = 8$, $Z(HT) = 4 + 5 = 9 = Z(TH)$, and $Z(TT) = 0 + 10 = 10$. This gives the following table for values of $Z = X + Y$, the related sample points in $S$, and the probabilities for different values of $X + Y$.

Here, $q = 1 - p$.

| $Z$ | Related sample points in $S$ | Probability |
|---|---|---|
| 8 | $HH$ | $\text{Prob}(Z=8) = p^2$ |
| 9 | $HT$, $TH$ | $\text{Prob}(Z=9) = 2pq$ |
| 10 | $TT$ | $\text{Prob}(Z=10) = q^2$ |

- This gives $E(X+Y) = E(Z) = 8.p^2 + 9.2pq + 10.q^2 = 9(p^2 + 2pq + q^2) + (q^2 - p^2) = 9(p+q)^2 + (q+p)(q-p) = 9 + q - p = 10 - 2p$.

  We also have $E(X) = E(4.\#(H)) = 4.\,E(\#(H)) = 4(2p) = 8p$ and
  $E(Y) = E(5.\#(T)) = 5.\,E(\#(T)) = 5(2q) = 10q$ and
  $E(X) + E(Y) = 8p + 10q = 10(p+q) - 2p = 10 - 2p = E(X+Y)$.

  This is a verification of the equation $E(X+Y) = E(X) + E(Y)$ for the above $X$ and $Y$.