The Binomial Distribution

Welcome to the Binomial Distribution. It's quite a gentle introduction, because this page is basically about counting. If you're thinking some of this looks familiar, you're dead right — you met the binomial expansion back in Section 4.

n different objects can be arranged in *n!* different ways...

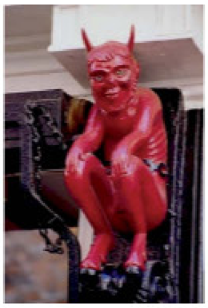
There are $n!$ ("**n factorial**") ways of arranging n **different** objects, where $n! = n \times (n-1) \times (n-2) \times ... \times 3 \times 2 \times 1$.

Example: a) In how many ways can 4 different ornaments be arranged on a shelf?
b) In how many ways can 8 different objects be arranged?

a) You have **4 choices** for the first ornament, **3 choices** for the second ornament, **2 choices** for the third ornament, and **1 choice** for the last ornament.
So there are $4 \times 3 \times 2 \times 1 = 4! =$ **24** arrangements.

b) There are $8! =$ **40 320** arrangements.

Calculators have a factorial button so you don't need to type all the numbers out.



Of course, not all ornaments deserve to go on the shelf.

...but *Divide by r!* if *r* of these objects are the *Same*

If r of your n objects are **identical**, then the total number of possible arrangements is **($n! \div r!$)**.

Example: a) In how many different ways can 5 objects be arranged if 2 of those objects are identical?
b) In how many different ways can 7 objects be arranged if 4 of those objects are identical?

a) Imagine those 2 identical objects were **different**. Then there would be $5! = 120$ possible arrangements. But because those 2 objects are actually **identical**, you can always **swap them round** without making a different arrangement. So there are really only $120 \div 2 =$ **60** different ways to arrange the objects.

b) There are $\frac{n!}{r!} = \frac{7!}{4!} = \frac{5040}{24} =$ **210** different ways to arrange the objects.

Use *Binomial Coefficients* if there are *Only Two Types* of object

See p.50 for more about binomial coefficients.

Binomial Coefficients

$$\binom{n}{r} = {}^nC_r = \frac{n!}{r!(n-r)!}$$

${}^nC_r$ and $\binom{n}{r}$ both mean $\frac{n!}{r!(n-r)!}$

Example: a) In how many different ways can n objects of two types be arranged if r are of the first type?
b) How many ways are there to select 11 players from a squad of 16?
c) How many ways are there to pick 6 lottery numbers from 59?

a) If the objects were all **different**, there would be $n!$ ways to arrange them.
But r of the objects are of the same type and could be **swapped around**, so divide by $r!$.
Since there are only **two types**, the other $(n - r)$ could also be **swapped around**, so divide by $(n - r)!$. This means there are $\frac{n!}{r!(n-r)!}$ arrangements.

b) This is basically a 'number of different **arrangements**' problem. Imagine the 16 players are lined up — then you could '**pick**' or '**not pick**' players by giving each of them a sign marked with a tick or a cross. So just find the number of ways to arrange 11 ticks and 5 crosses — this is $\binom{16}{11} = \frac{16!}{11!5!} =$ **4368**.

c) Again, numbers are either '**picked**' or '**unpicked**', so there are $\binom{59}{6} = \frac{59!}{6!53!} =$ **45 057 474** possibilities.

The Binomial Distribution

Use *Binomial Coefficients* to count arrangements of 'successes' and 'failures'

For this bit, you need to use the fact that if p = **P(something happens)**, then $1 - p$ = **P(that thing doesn't happen)**.

Example: I toss a fair coin 5 times. Find the probability of: a) 0 heads, b) 1 head, c) 2 heads.

First, note that each coin toss is **independent** of the others.
That means you can **multiply** individual probabilities together.

P(tails) = P(heads) = 0.5

a) P(0 heads) = P(tails) × P(tails) × P(tails) × P(tails) × P(tails) = 0.5^5 = **0.03125**

These are the $\binom{5}{1} = 5$ ways to arrange 1 head and 4 tails.

b) P(1 head) = [P(heads) × P(tails) × P(tails) × P(tails) × P(tails)] + [P(tails) × P(heads) × P(tails) × P(tails) × P(tails)]
+ [P(tails) × P(tails) × P(heads) × P(tails) × P(tails)] + [P(tails) × P(tails) × P(tails) × P(heads) × P(tails)]
+ [P(tails) × P(tails) × P(tails) × P(tails) × P(heads)]

So $P(1 \text{ head}) = 0.5 \times (0.5)^4 \times \binom{5}{1} = 0.03125 \times \frac{5!}{1!4!} = \mathbf{0.15625}$

= P(heads) × $[P(tails)]^4$ × ways to arrange 1 head and 4 tails

c) $P(2 \text{ heads}) = [P(heads)]^2 \times [P(tails)]^3 \times$ ways to arrange 2 heads and 3 tails $= (0.5)^2 \times (0.5)^3 \times \binom{5}{2} = \mathbf{0.3125}$

The *Binomial Probability Function* gives P(r successes out of n trials)

The previous example really just shows why this thing-in-a-box must be true.

Binomial Probability Function

$$P(r \text{ successes in } n \text{ trials}) = \binom{n}{r} \times [P(\text{success})]^r \times [P(\text{failure})]^{n-r}$$

This is the probability function for a binomial distribution — see below for more info.

Example: I roll a fair six-sided dice 5 times. Find the probability of rolling:
a) 2 sixes, b) 3 sixes, c) 4 numbers less than 3.

Again, note that each roll of a dice is **independent** of the other rolls.

a) For this part, call "roll a 6" a success, and "roll anything other than a 6" a failure.

Then $P(\text{roll 2 sixes}) = \binom{5}{2} \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^3 = \frac{5!}{2!3!} \times \frac{1}{36} \times \frac{125}{216} = \mathbf{0.161}$ (3 d.p.)

b) Again, call "roll a 6" a success, and "roll anything other than a 6" a failure.

Then $P(\text{roll 3 sixes}) = \binom{5}{3} \times \left(\frac{1}{6}\right)^3 \times \left(\frac{5}{6}\right)^2 = \frac{5!}{3!2!} \times \frac{1}{216} \times \frac{25}{36} = \mathbf{0.032}$ (3 d.p.)

Notice how $\binom{5}{2} = \binom{5}{3}$. In fact, $\binom{n}{r} = \binom{n}{n-r}$.

c) This time, success means "roll a 1 or a 2", while failure is now "roll a 3, 4, 5 or 6".

Then $P(\text{roll 4 numbers less than 3}) = \binom{5}{4} \times \left(\frac{1}{3}\right)^4 \times \frac{2}{3} = \frac{5!}{4!1!} \times \frac{1}{81} \times \frac{2}{3} = \mathbf{0.041}$ (3 d.p.)

There are *5 Conditions* for a *Binomial Distribution*

Binomial Distribution: B(n, p)

A random variable X follows a binomial distribution if these **5 conditions** are satisfied:

1) There is a **fixed number** (n) of trials.
2) Each trial results in either "**success**" or "**failure**".
3) All the trials are **independent**.
4) The probability of "success" (p) is the **same** in each trial.
5) The variable is the **total number of successes** in the n trials.

Binomial variables are discrete — they only take values 0, 1, 2... n.

n and p are the **parameters** of the binomial distribution.

Then, $P(X = x) = \binom{n}{x} \times p^x \times (1-p)^{n-x}$ for $x = 0, 1, 2, \ldots, n$, and you can write $X \sim B(n, p)$.

If you're asked to **comment on the appropriateness** of a binomial model, you should check whether the variable satisfies **all** of these conditions.

The **expected** number of successes is given by (n × p).

The Binomial Distribution

Use your **Calculator** to find **Binomial Probabilities**

Example: I have an unfair coin. When I toss this coin, the probability of getting heads is 0.35. Find the probability that it will land on heads fewer than 3 times when I toss it 12 times in total.

If the random variable X represents the number of heads I get in 12 tosses, then $X \sim B(12, 0.35)$.
You need to find $P(X \leq 2)$. You **could** work this out 'manually'...

$$P(0 \text{ heads}) + P(1 \text{ head}) + P(2 \text{ heads}) = \left[\binom{12}{0}\times 0.35^0 \times 0.65^{12}\right] + \left[\binom{12}{1}\times 0.35^1 \times 0.65^{11}\right] + \left[\binom{12}{2}\times 0.35^2 \times 0.65^{10}\right]$$

$$= 0.00568... + 0.03675... + 0.10884... = 0.15128... = \mathbf{0.151} \text{ (3 s.f.)}$$

However, it's much quicker to use the **binomial cumulative distribution function** (cdf) on your calculator.
This calculates $P(X \leq x)$, for $X \sim B(n, p)$ — just enter the **correct values of n, p and x**.
For example, here, $n = 12$ and $p = 0.35$, and you need $P(X \leq 2)$ (i.e. $x = 2$).
The calculator tells you that this is **0.15128...**, which is what you worked out above.

Be careful though:

- Some calculators have both a binomial **probability distribution function** (pdf) and a binomial **cumulative distribution function** (cdf). You use the **pdf** to find e.g. $\mathbf{P(X = 2)}$ (as on the previous page) and the **cdf** to find e.g. $\mathbf{P(X \leq 2)}$ (as above).
- The cdf gives you $P(X \leq x)$ — if you want $P(X \geq x)$ (or $P(X < x)$ etc.) you'll have to do some fancy probability-wrangling. E.g. $P(X < 7) = P(X \leq 6)$, or $P(X > 4) = 1 - P(X \leq 4)$.

Countless secrets contained within...

Example: I have a different unfair coin. When I toss this coin, the probability of getting tails is 0.6. The random variable X represents the number of tails in 12 tosses, so $X \sim B(12, 0.6)$.
If I toss this coin 12 times, find the probability that:
a) it will land on tails more than 8 times, b) it will land on heads exactly 9 times,
c) it will land on tails more than 3 but fewer than 6 times.

a) You're looking for $P(X > 8)$, which is $1 - P(X \leq 8)$ (since '$X > 8$' and '$X \leq 8$' are complementary events — see p.153). So, from your calculator: $P(X > 8) = 1 - 0.77466... = \mathbf{0.225}$ (3 s.f.)

b) If the coin lands on **heads** 9 times, then it lands on **tails** 3 times. You could use the binomial pdf to go straight to the answer, or if your calculator doesn't have one, you can use the cdf instead:
$P(\text{heads 9 times}) = P(X = 3) = P(X \leq 3) - P(X \leq 2) = 0.01526... - 0.00281... = \mathbf{0.0125}$ (3 s.f.)

c) $P(3 < X < 6) = P(X < 6) - P(X \leq 3) = P(X \leq 5) - P(X \leq 3) = 0.15821... - 0.01526... = \mathbf{0.143}$ (3 s.f.)

Practice Questions

Q1 Find the probability of: a) getting exactly 9 heads when you toss a fair coin 10 times,
b) getting at least 9 heads when you toss a fair coin 10 times.

Q2 Find, to 4 decimal places: a) $P(X = 4)$ if $X \sim B(14, 0.27)$ b) $P(Y \leq 15)$ if $Y \sim B(20, 0.4)$

Exam Questions

Q1 The random variable X follows the binomial distribution $X \sim B(12, 0.6)$. Find:
a) $P(X < 8)$ [2 marks] b) $P(X = 5)$ [1 mark] c) $P(3 < X \leq 7)$ [2 marks]

Q2 Apples are stored in crates of 40. The probability of any apple containing a maggot is 0.15, and is independent of any other apple containing a maggot. In a random sample of 40 apples, find the probability that:
a) fewer than 6 apples contain maggots, [2 marks] b) more than 2 apples contain maggots. [2 marks]
c) Jin has 3 crates. Find the probability that more than 1 crate contains more than 2 apples with maggots. [3 marks]
d) Give one criticism of the assumption that apples contain maggots independently of each other. [1 mark]

I used up all my binomial jokes in Section 4...

Here's a handy trick that might save some time on certain questions: if the number of successes is $X \sim B(n, p)$, then the number of failures is $Y \sim B(n, 1 - p)$. For example, the number of heads in the blue example is $Y \sim B(12, 0.4)$.