BINOMIAL DISTRIBUTION

The Binomial distribution is used for a discrete random variable that can only have two results: *pass* or *fail*. Or *success* or *failure*. For example, if we wish to use the binomial distribution to analyze baseball, we must require the batter to either get a hit, or not. We must include sacrifices, bunts, foul balls and so on in one of our two chosen categories: hit or no hit.

Here are the rules:

Each trial or test can only have one of two possible outcomes

The probability of *success*, *p*, does not change from test to test. We either need to know *p*, or we estimate it from a sample

The complete test is for *n* independent trials

The probability of getting exactly *r* successes in *n* trials is given by:

$$P(r) = \binom{n}{r} p^r (1-p)^{n-r}$$

with

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

This function is read as *"n combination r"* and is the number of ways we can pick *r* items from *n* items. For example, how many ways can you select 4 chocolates from a box of 12 assorted chocolates?

$$n = 12; r = 4$$

$$^{12}C_4 = \binom{12}{4} = \frac{12!}{4!(12-4)!} = 495$$

Thus there are 495 different ways of selecting 4 chocolates from the box.

The expected value (mean) of a Binomial distribution is

$$m = np$$

Example: The probability of your car getting you home at Thanksgiving without breaking down is 95%. There are 21 midshipmen in class. What is the expected number of midshipmen who will make it home?

$$p = 0.95; n = 21$$

$$m = np = 21 \times 0.95 = 19.95$$

What does “19.95” mean? Are we really expecting 0.95 of a midshipman to make it home? Should we have rounded the answer off to “20”?

The variance is

$$s^2 = np(1-p)$$

So for this example:

$$s^2 = 21 \times 0.95 \times (1-0.95) = 0.9975$$

But what are the units?

If we want to calculate the probability of getting *r or less* successes, we need to calculate:

$$\begin{aligned} P(r \le k) &= P(0) + P(1) + \ldots\ldots + P(k) \\ &= \sum_{i=0}^{k} P(i) \\ &= \sum_{i=0}^{k} \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \end{aligned}$$

EXAMPLE: The probability of an individual MathCAD disk having a problem during installation is 6%. In a section issue of 21 disks, what is the probability that:

a) None will be faulty
b) Exactly 2 will be faulty
c) 2 or less will be faulty
d) 3 or more will be faulty

a) $P(r=0) = \frac{21!}{0!(21-0)!} 0.06^0 (1-0.06)^{21-0} = 0.2727 = 27.27\%$

b) $P(r=2) = \frac{21!}{2!(21-2)!} 0.06^2 (1-0.06)^{21-2} = 0.23332 = 23.332\%$

c) The probability that there are 2 or less problems is the same as the probability that there will be exactly 2, or exactly 1, or exactly zero problems.

$$
\begin{aligned}
P(r \le 2) &= P(0) + P(1) + P(2) \\
&= \sum_{i=0}^{2} \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \\
&= \frac{n!}{0!(n-0)!} p^0 (1-p)^{n-0} + \frac{n!}{1!(n-1)!} p^1 (1-p)^{n-1} + \frac{n!}{2!(n-2)!} p^2 (1-p)^{n-2} \\
&= 0.2727 + 0.36553 + 0.23332 = 0.872 = 87.2\%
\end{aligned}
$$

d) We can calculate this as the probability of there being exactly 3 problems plus the probability of exactly 4 problems, etc up to exactly 21 problems:

$$
\begin{aligned}
P(r > 2) &= P(3) + P(4) + \ldots.. + P(21) \\
&= \sum_{i=3}^{21} \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} = 0.128 \text{ (using MathCAD)}
\end{aligned}
$$

That's a lot of work. An alternative quicker method is to use the rule of probability that $P(A) + P(\bar{A}) = 1$. We now have:

$$
\begin{aligned}
P(r < 3) + P(r \ge 3) &= 1 \\
P(r \ge 3) &= 1 - P(r < 3) \\
&= 1 - 0.872 = 12.8\%
\end{aligned}
$$

EXAMPLE: During any one hour, there are 5000 aircraft in flight over the USA. What is the probability that all flights will be successful (what does that mean?)?

We solve the binomial distribution for n = 5000 and r = 5000.

P(an individual flight is successful)	P(all OK in any one hour)	P(1 or more fail in any one hour)	8-hr days/year with all successful flights
99.9%	0.67%	99.33%	10^{-15} (zero!)
99.99%	60.6%	39.4%	6.68
99.999%	95.1%	4.9%	245
99.9999%	99.5%	0.5%	351

PROBLEMS

Bin-1:
The basketball coach wants to identify each practice ball separately. He wants to do this by putting colored paint spots on the balls. He has 6 different colored paints: red, green, blue, yellow, pink, and white. Each ball must have a unique combination of paint spots (e.g., red/red/red, or yellow/green/pink). Note that red/green/blue and blue/green/red are NOT DIFFERENT!

a) If he only puts one paint spot on each ball, how many different balls can be marked?
b) If he puts three paint spots on each ball, how many different balls can be marked? Write down all the possible color combinations.
c) Coach has 30 balls he wishes to mark. What is the smallest number of different colored paints that is needed to uniquely identify each ball, and how many paint spots are required on each ball?

Bin-2:
The probability of a plebe ultimately graduating from the academy is 85%. Calculate the following:

a) In a squad of 20 plebes, what is the expected number that will graduate?
b) What is the probability that all 20 will graduate?
c) What is the probability that 18 or more will graduate?
d) What is the probability that less than 18 will graduate?

Bin-3:
When the class of 2009 ran their first PRT, 1050 plebes started. Of these 825 ran fast enough to pass the run.

a) What is the probability that a plebe chosen at random can run fast enough to pass the PRT?
b) In a squad of 16 plebes, what is the probability that all of them will pass the run?
c) And what is the probability that none of them will pass?