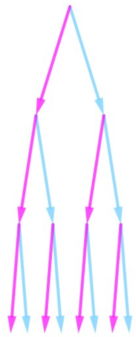
Binomial Distribution

Binomial Distribution

A discrete random variable based on repeated Bernoulli trials (independent trials where the two outcomes are success or failure). It is binomial in the sense that each possible probability is a term in the binomial expansion of $\left(p+(1-p)\right)^n$.



Consider a tree diagram with complementary outcomes at each trial. How many of each combination of outcomes do you get after n trials? What probabilities get multiplied along the way? How does this relate to binomial expansion?

Notation

A binomial distribution, X, with n trials and a probability of success of p is written as $X \sim \mathrm{Bi}(n,p)$

Probability Mass Function for the Binomial Distribution

The probability of x successes out of n trials is $\Pr(X=x) = {}^n\mathrm{C}_x p^x (1-p)^{n-x}$, the product of

- the number of ways of getting x successes out of n trials, ${}^n\mathrm{C}_x$
- the probability of success to the power of x (the number of successful trials), p^x for $0<p<1$
- the probability of failure to the power of $n-x$ (the remaining trials), $(1-p)^{n-x}$

Example

A random variable X is binomially distributed such that $X \sim \mathrm{Bi}(4,0.2)$
The probability distribution of X is given by the table below.

x	0	1	2	3	4
$\Pr(X=x)$	${}^4\mathrm{C}_0(0.2)^0(0.8)^4$ $=(0.8)^4$ $=0.4096$	${}^4\mathrm{C}_1(0.2)^1(0.8)^3$ $=4(0.2)(0.8)^3$ $=0.4096$	${}^4\mathrm{C}_2(0.2)^2(0.8)^2$ $=6(0.2)^2(0.8)^2$ $=0.1536$	${}^4\mathrm{C}_3(0.2)^3(0.8)^1$ $=4(0.2)^3(0.8)$ $=0.0256$	${}^4\mathrm{C}_4(0.2)^4(0.8)^0$ $=(0.2)^4$ $=0.0016$

Example Modified VCAA 2017 Exam 2 Question 3e

The probability that Jennifer spends more than 50 minutes on her homework on any given day is $\frac{8}{25}$. Assume that the amount of time spent on her homework on any day is independent of the time spent on her homework on any other day. The probability that Jennifer spends more than 50 minutes on her homework on three of seven randomly chosen days is

$X \sim \mathrm{Bi}\left(7,\frac{8}{25}\right)$, $\quad \Pr(X=3) = {}^7\mathrm{C}_3\left(\frac{8}{25}\right)^3\left(\frac{17}{25}\right)^4 = 0.2452$

or on CAS: Binomial Pdf function
$\Pr(X=3): n=7, p=\frac{8}{25}, x=3$

The probability that Jennifer spends more than 50 minutes on her homework on at least two of seven randomly chosen days, given that she spends more than 50 minutes on her homework on at least one of those days is

$$\Pr(X\ge 2 | X \ge 1) = \frac{\Pr(X\ge 2 \cap X \ge 1)}{\Pr(X\ge 1)}$$

$$= \frac{\Pr(\mathrm{X}\ge 2)}{\Pr(X\ge 1)} = \frac{0.7113}{0.9328} = 0.7626$$

CAS: Binomial Cdf function

$\Pr(X\ge 2): n=7, p=\frac{8}{25}, \text{lower}=2, \text{upper}=7$

$\Pr(X\ge 1): n=7, p=\frac{8}{25}, \text{lower}=1, \text{upper}=7$

Proof for the Probability Mass Function for the Binomial Distribution

Proof that Probabilities are Non-Negative

${}^{n}\mathrm{C}_x > 0 \qquad p^x \geq 0,$ for all $p, 0 \leq x \leq n \qquad (1-p)^{n-x} \geq 0,$ for all $p \leq 1, 0 \leq x \leq n$

Since p is a probability, it must be between 0 and 1. $\therefore 0 \leq p \leq 1$
The values of x that are used for a binomial distribution are the integers from 0 to n. $\therefore 0 \leq x \leq n$
$\therefore {}^{n}\mathrm{C}_x p^x (1-p)^{n-x} \geq 0$

Proof that Total Probability Equals One

Each probability is a term in the binomial expansion of $\big(p + (1-p)\big)^n$, which is also equal to 1.
$\sum_{x=0}^{n} \left({}^{n}\mathrm{C}_x p^x (1-p)^{n-x}\right) = \big(p + (1-p)\big)^n = 1^n = 1$

Determining the Number of Successes and Inverse Binomial

When the cumulative probability $\Pr(X \leq a)$ is known but the number of successes is not known for a binomial distribution, we can use trial and error or inverse binomial on the CAS to determine the value of a that would give that cumulative probability.

Example VCAA 2019 Exam 2 Question 4f

The Lorenz birdwing is the largest butterfly in Town A. Each year, a detailed study is conducted on a random sample of 36 Lorenz birdwing butterflies in Town A. The probability that a Lorenz birdwing butterfly in Town A is considered to be **very large** is 0.0527, correct to four decimal places.

The probability that n or more butterflies, in a random sample of 36 Lorenz birdwing butterflies from Town A, are **very large** is less than 1%. Find the smallest value of n, where n is an integer.

Using Inverse Binomial

$X \sim \mathrm{Bi}(36, 0.0527)$

$\Pr(X \geq n) < 0.01$

$\Pr(X \leq n-1) = \Pr(X < n) > 0.99$

Use Inverse Binomial: $n - 1 = 6$, so $n = 7$

Check to confirm:
$\Pr(X \leq 5) = \Pr(X < 6) = 0.98934 \ldots$
$\Pr(X \leq 6) = \Pr(X < 7) = 0.99756 \ldots$

Using Trial and Error

$X \sim \mathrm{Bi}(36, 0.0527)$

$\Pr(X \geq n) < 0.01$

$\Pr(X \geq 6) = 0.01065 \ldots$
$\Pr(X \geq 7) = 0.00243 \ldots$

$\therefore n = 7$

Determining the Probability of Success

When the probability of success is unknown, we create a polynomial equation that can be solved.

Example VCAA 2004 Exam 1 Question 6ab

A barrel contains 100 balls, some of which are rainbow-coloured. Four balls are randomly selected from the barrel, **with replacement** (that is, a ball is selected, its colour noted, and the ball replaced before the next ball is selected). Let p be the proportion of rainbow-coloured balls in the barrel.

An expression for the probability that exactly one of the four balls selected is rainbow coloured.
$X \sim \mathrm{Bi}(4, p), \qquad \Pr(X = 1) = {}^{4}\mathrm{C}_1 p^1 (1-p)^3 = 4p(1-p)^3$

Use calculus to find the exact value of p for which this probability will be a maximum.
$$\frac{d}{dp}\left(4p(1-p)^3\right) = 4(1-p)^3 - 12p(1-p)^2 = 4(p-1)^2(1-p-3p) = 4(p-1)^2(1-4p)$$

Maximums occur when the derivative is equal to zero.
$4(p-1)^2(1-4p) = 0 \;\Rightarrow\; p - 1 = 0, 1 - 4p = 0 \;\Rightarrow\; p = \frac{1}{4},$ reject $p = 1$ since $0 < p < 1$

Determining the Number of Trials

We often want to know how many trials we would need to have to ensure at least a certain probability to have a certain number of successes. To solve this, we can solve the resulting inequality using logarithms or by CAS, or Inverse Binomial N on TI-nSpire CAS.

Example VCAA 2006 Sample Exam 2 Question 20

The probability of winning a single game of chance is 0.15, and whether or not the game is won is independent of any other game. Suppose Jodie plays a sequence of n games. If the probability of Jodie winning at least one game is more than 0.95, then the smallest value n can take is closest to

$$\Pr(X \geq 1) = {}^{n}\mathrm{C}_1(0.15)^1(0.85)^{n-1} + {}^{n}\mathrm{C}_2(0.15)^2(0.85)^{n-2} + \cdots + {}^{n}\mathrm{C}_n(0.15)^n(0.85)^0 > 0.95$$

$$\Pr(X = 0) = 1 - {}^{n}\mathrm{C}_0(0.15)^0(0.85)^n = 1 - (0.85)^n > 0.95$$

Method 1

$1 - (0.85)^n > 0.95$

$(0.85)^n < 0.05$

$\log_e\left((0.85)^n\right) < \log_e(0.05)$

$n\log_e(0.85) < \log_e(0.05)$

$n > \dfrac{\log_e(0.05)}{\log_e(0.85)}$

(sign change since $\log_e(0.85) < 0$)

$n > 18.4331$ ..., so $n = 19$

Method 2

$1 - (0.85)^n > 0.95$

$(0.85)^n < 0.05$

$\log_{0.85}\left((0.85)^n\right) > \log_{0.85}(0.05)$

(sign change since $\log_{0.85}(x)$ is a decreasing function)

$n > \log_{0.85}(0.05)$

$n > 18.4331$ ..., so $n = 19$

Example Modified VCAA 2015 Exam 2 Question 3dii

Mani also grows lemons, which are sold to a food factory. When a truckload of lemons arrives at the food factory, the manager randomly selects and weighs four lemons from the load. If one or more of these lemons is underweight, the load is rejected. Otherwise it is accepted. It is known that 3% of Mani's lemons are underweight.

Suppose that instead of selecting only four lemons, n lemons are selected at random from a particular load. Find the smallest integer value of n such that the probability of more than one lemon being underweight exceeds 0.5.

$\Pr(X > 1) = 1 - \Pr(X = 0) - \Pr(X = 1)$

$= 1 - {}^{n}\mathrm{C}_0(0.03)^0(0.97)^n - {}^{n}\mathrm{C}_1(0.03)^1(0.97)^{n-1} > 0.5$

$1 - (0.97)^n - n(0.03)(0.97)^{n-1} > 0.5$

Solve $1 - (0.97)^n - n(0.03)(0.97)^{n-1} = 0.5$ on CAS

$n = 55.605$ ... for $n > 0$

Since $1 - {}^{n}\mathrm{C}_0(0.03)^0(0.97)^n - {}^{n}\mathrm{C}_1(0.03)^1(0.97)^{n-1}$ is increasing, the smallest integer value of n is 56.

$y = 1 - (0.97)^x - x(0.03)(0.97)^{x-1}$