Binomial Distribution

Definition: Binomial Distribution

To say that a random variable X has a *Binomial* distribution with parameters n and p, denoted:

$$X \sim \mathrm{B}(n, p)$$

means that:

1. The underlying experiment is composed of n repeated Bernoulli trials.
2. The n trials are independent.
3. Each of the n trials has fixed probability p for success.
4. X counts the number of successes resulting from the n trials.

Note that a Binomial distribution for selection from a population implies replacement.

Examples: Binomial Distributions

1. Flip a fair coin 10 times: $X =$ the number of heads.

$$X \sim \mathrm{B}\left(10, \frac{1}{2}\right)$$

2. Answer 10 multiple choice questions, each with 4 possible answers, by random guessing: $Y =$ the number of correct answers.

$$Y \sim \mathrm{B}\left(10, \frac{1}{4}\right)$$

3. Select (with replacement) 10 balls from an urn that has 30 red balls and 20 blue balls: $Z =$ the number of selected red balls.

$$Z \sim \mathrm{B}\left(10, 0.6\right)$$

Theorem

Let X be a random variable with a Binomial distribution with parameters n and p:

- $f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$
- $E(X) = np$
- $V(X) = np(1-p)$

Proof. For $P(X = x)$, any x of the n trials are successful: $\binom{n}{x}$. The x successes have probability p^x. That leaves $n - x$ failures with probability $(1-p)^{n-x}$. Therefore:

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Now, let X_i be an indicator variable for the i^{th} Bernoulli trial:

$$E(X) = E\left(\sum_{i=0}^{n} X_i\right) = \sum_{i=0}^{n} E(X_i) = \sum_{i=0}^{n} p = np$$

Finally, since the trials are independent:

$$V(X) = V\left(\sum_{i=0}^{n} X_i\right) = \sum_{i=0}^{n} V(X_i) = \sum_{i=0}^{n} p(1-p) = np(1-p)$$

∎

Example

Flip a fair coin 10 times ($n = 10, p = \frac{1}{2}$):

$$X \sim \mathrm{B}\left(10, \frac{1}{2}\right)$$

$$P(X = 0) = \binom{10}{0}\left(\frac{1}{2}\right)^0\left(1 - \frac{1}{2}\right)^{10-0} = \left(\frac{1}{2}\right)^{10} = \frac{1}{1024}$$
$$P(X = 1) = \binom{10}{1}\left(\frac{1}{2}\right)^1\left(1 - \frac{1}{2}\right)^{10-1} = 10\left(\frac{1}{2}\right)^{10} = \frac{10}{1024}$$
$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1) = 1 - \frac{1}{1024} - \frac{10}{1024} = \frac{1013}{1024}$$

$$E(X) = np = 10\left(\frac{1}{2}\right) = 5$$

$$V(X) = np(1-p) = 10\left(\frac{1}{2}\right)\left(1 - \frac{1}{2}\right) = 10\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{10}{4} = 2.5$$

$$\sigma = \sqrt{2.5} \approx 1.58$$

Example

Answer 10 multiple choice questions, each with 4 possible answers, by random guessing.

$$X \sim \mathrm{B}\left(10, \frac{1}{4}\right)$$

$$P(X=0) = \binom{10}{0}\left(\frac{1}{4}\right)^{0}\left(1-\frac{1}{4}\right)^{10-0} = 0.0563$$
$$P(X=2) = \binom{10}{2}\left(\frac{1}{4}\right)^{1}\left(1-\frac{1}{4}\right)^{10-2} = 0.2816$$
$$P(X=8) = \binom{10}{8}\left(\frac{1}{4}\right)^{1}\left(1-\frac{1}{4}\right)^{10-8} = 0.0004$$

$$E(X) = np = 10\left(\frac{1}{4}\right) = 2.5$$

$$V(X) = np(1-p) = 10\left(\frac{1}{4}\right)\left(1-\frac{1}{4}\right) = 10\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = \frac{30}{16} = 1.875$$

$$\sigma = \sqrt{1.875} \approx 1.37$$

The expected value and variance can also be calculated directly:

$$\begin{aligned}
E(X) &= \sum_{x=0}^{n} xP(X=x) \\
&= \sum_{x=0}^{n} x\binom{n}{x}p^{x}(1-p)^{n-x} \\
&= \sum_{x=1}^{n} x\frac{n!}{x!(n-x)!}p^{x}(1-p)^{n-x} \\
&= np\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)![(n-1)-(x-1)]!}p^{x-1}(1-p)^{[(n-1)-(x-1)]} \\
&= np\sum_{x=0}^{n-1} \frac{(n-1)!}{x![(n-1)-x]!}p^{x}(1-p)^{[(n-1)-x]} \\
&= np[p+(1-p)]^{n-1} \\
&= np(1)^{n-1} \\
E(X) &= np
\end{aligned}$$

For the variance, start with:

$$
\begin{aligned}
E(X^2 - X) &= E(X(X-1)) \\
&= \sum_{x=0}^{n} x(x-1)P(X=x) \\
&= \sum_{x=0}^{n} x(x-1)\binom{n}{x}p^x(1-p)^{n-x} \\
&= \sum_{x=2}^{n} x(x-1)\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} \\
&= n(n-1)p^2\sum_{x=2}^{n}\frac{(n-2)!}{(x-2)![(n-2)-(x-2)]!}p^{x-2}(1-p)^{[(n-2)-(x-2)]} \\
&= n(n-1)p^2\sum_{x=0}^{n-2}\frac{(n-2)!}{x!((n-2)-x)!}p^{x}(1-p)^{(n-2)-x} \\
&= n(n-1)p^2[p+(1-p)]^{n-2} \\
&= n(n-1)p^2(1)^{n-2} \\
E(X^2 - X) &= n(n-1)p^2
\end{aligned}
$$

And now:

$$
\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= E(X^2 - X + X) - E(X)^2 \\
&= E(X^2 - X) + E(X) - E(X)^2 \\
&= n(n-1)p^2 + np - n^2p^2 \\
&= n^2p^2 - np^2 + np - n^2p^2 \\
&= np - np^2 \\
V(X) &= np(1-p)
\end{aligned}
$$