Chapter 4 - Part 3

Binomial Random Variable

Binomial Random Variable

Toss a coin (not necessarily fair) with

$$p(H) = p, \qquad p(T) = 1 - p$$

n times. Let

$$X: \text{ the number of heads}$$

A random variable equivalent to this is called a binomial random variable.

Binomial Random Variable

Formal definition of a binomial random variable X:

- The experiment consists of n identical trials.
- There are only two possible outcomes on each trial:

$$S \text{ (success)}, \qquad F \text{ (failure)}.$$

- $P(S)$ and $P(F)$ remain the same from trial to trial.
- $$P(S) + P(F) = 1.$$
- The trials are independent.
- The random variable x is the number of successes in n trials.

These trials are called Bernoulli trials.

Notation

If X is a binomial random variable where number of trials is n and probabilty of success at each trial is p, then we denote

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

Example

We have a coin with

$$P(H) = 0.65, \qquad P(T) = 0.35.$$

It is tossed $n = 4$ times.

Let X be the random variable representing the number of heads in 4 tosses.

We construct its probability distribution by grouping outcomes according to the value of X.

Example (cont'd)

We have, $|S| = 2^4 = 16$.

$$x = 0: \quad \{TTTT\}$$

$$x = 1: \quad \{TTTH,\ TTHT,\ THTT,\ HTTT\}$$

$$x = 2: \quad \{TTHH,\ THTH,\ THHT,\ HTTH,\ HTHT,\ HHTT\}$$

$$x = 3: \quad \{THHH,\ HTHH,\ HHTH,\ HHHT\}$$

$$x = 4: \quad \{HHHH\}$$

The number of outcomes giving each value of x is:

$$\binom{4}{0} = 1, \qquad \binom{4}{1} = 4, \qquad \binom{4}{2} = 6, \qquad \binom{4}{3} = 4, \qquad \binom{4}{4} = 1.$$

Example (cont'd)

So the probabilities are:

$$P(X=0) = (0.35)^4 = \binom{4}{0}(0.65)^0(0.35)^4,$$

$$P(X=1) = 4(0.65)(0.35)^3 = \binom{4}{1}(0.65)^1(0.35)^3,$$

$$P(X=2) = 6(0.65)^2(0.35)^2 = \binom{4}{2}(0.65)^2(0.35)^2,$$

$$P(X=3) = 4(0.65)^3(0.35) = \binom{4}{3}(0.65)^3(0.35)^1,$$

$$P(X=4) = (0.65)^4 = \binom{4}{4}(0.65)^4(0.35)^0.$$

Binomial Formula

Let X be the binomial random variable with

p is the probability of success (at each trial),

n is the number of trials.

If $0 \le m \le n$, then

$$P(X = m) = \binom{n}{m} p^m (1 - p)^{n-m}.$$

Binomial Formula

In other words,

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

is the probabilty of

exactly m successes

in n trials

where p is the probability of success (at each trial).

Therefore,

$1 - p$ is the probability of failure (at each trial),

and $n - m$ is the number of failure.

In the previous example, number of elements in the sample space was

$$\binom{4}{0} + \binom{4}{1} + \binom{4}{2} + \binom{4}{3} + \binom{4}{4} = 16 = 2^4.$$

In general, number of elements in the sample space of n Bernoulli trials is

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n.$$

We can also check that

$$
\begin{aligned}
1 = 1^n &= (p + (1-p))^n \\
&= \underbrace{\binom{n}{0} p^0 (1-p)^n}_{P(X=0)} + \underbrace{\binom{n}{1} p^1 (1-p)^{n-1}}_{P(X=1)} + \underbrace{\binom{n}{2} p^2 (1-p)^{n-2}}_{P(X=2)} + \cdots + \underbrace{\binom{n}{n} p^n (1-p)^0}_{P(X=n)}
\end{aligned}
$$

Expected Value and Variance of a Binomial Random Variable

Let X_i be the indicator random variable of success on the ith Bernoulli trial:

$$X_i = \begin{cases} 1, & \text{if trial } i \text{ is a success,} \\ 0, & \text{if trial } i \text{ is a failure.} \end{cases}$$

x_i	0	1
$p(x_i)$	1-p	p

Then

$$E(X_i) = 0 \cdot (1 - p) + 1 \cdot p = p$$

and

$$\mathrm{Var}(X_i) = E(X_i^2) - E(X_i)^2 = p - p^2 = p(1 - p).$$

Expected Value and Variance of a Binomial Random Variable

Since

$$X = X_1 + X_2 + \cdots + X_n,$$

we get

$$E(X) = \underbrace{E(X_1)}_{p} + \underbrace{E(X_2)}_{p} + \cdots + \underbrace{E(X_n)}_{p} = np$$

and, because the trials are independent,

$$\mathrm{Var}(X) = \underbrace{\mathrm{Var}(X_1)}_{p(1-p)} + \underbrace{\mathrm{Var}(X_2)}_{p(1-p)} + \cdots + \underbrace{\mathrm{Var}(X_n)}_{p(1-p)} = np(1-p).$$

Example

What is the expected number and variance of the number of tails in 17 coin tosses with a fair coin?

Here,

$$n = 17, \qquad p = \frac{1}{2}.$$

Therefore,

$$E(X) = np = 17 \cdot \frac{1}{2} = \frac{17}{2} = 8.5$$

and

$$\mathrm{Var}(X) = np(1-p) = 17 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{17}{4} = 4.25.$$

Example

A fair die is rolled 3600 times.
Let

$$X = \text{ number of 6's.}$$

Then

$$n = 3600, \qquad p = \frac{1}{6}.$$

Hence,

$$E(X) = 3600 \cdot \frac{1}{6} = 600$$

and

$$\mathrm{Var}(X) = 3600 \cdot \frac{1}{6} \cdot \frac{5}{6} = 500.$$

Example

A coin with

$$P(H) = 0.2, \qquad P(T) = 0.8$$

is tossed 200 times.
Let

$$X = \text{ number of heads.}$$

Then

$$E(X) = 200 \cdot 0.2 = 40$$

and

$$\text{Var}(X) = 200 \cdot 0.2 \cdot 0.8 = 32.$$

Also,

$$P(X = 90) = \binom{200}{90}(0.2)^{90}(0.8)^{110} \approx 0.000004443.$$

is the probabilty of getting exactly 90 heads.

Example

What is the probability of getting 7 or 8 heads if a coin with

$$P(H) = 0.65$$

is tossed 11 times?

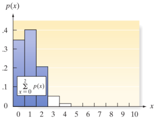
$$\begin{aligned} P(X = 7 \text{ or } X = 8) &= P(X = 7) + P(X = 8) \\ &= \binom{11}{7}(0.65)^7(0.35)^4 + \binom{11}{8}(0.65)^8(0.35)^3 \\ &\approx 0.468. \end{aligned}$$

Binomial Table

Binomial tables give cumulative probabilities.

Example: If $n = 5$, $k = 2$, $p = 0.1$, the table gives 0.991.

Table I **Binomial Probabilities**



Tabulated values are $\sum_{x=0}^{k} p(x)$. *(Computations are rounded at the third decimal place.)*

a. $n = 5$

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.951	.774	.590	.328	.168	.078	.031	.010	.002	.000	.000	.000	.000
1	.999	.977	.919	.737	.528	.337	.188	.087	.031	.007	.000	.000	.000
2	1.000	.999	.991	.942	.837	.683	.500	.317	.163	.058	.009	.001	.000
3	1.000	1.000	1.000	.993	.969	.913	.812	.663	.472	.263	.081	.023	.001
4	1.000	1.000	1.000	1.000	.998	.990	.969	.922	.832	.672	.410	.226	.049

b. $n = 6$

k \ p	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99
0	.941	.735	.531	.262	.118	.047	.016	.004	.001	.000	.000	.000	.000
1	.999	.967	.886	.655	.420	.233	.109	.041	.011	.002	.000	.000	.000
2	1.000	.998	.984	.901	.744	.544	.344	.179	.070	.017	.001	.000	.000
3	1.000	1.000	.999	.983	.930	.821	.656	.456	.256	.099	.016	.002	.000
4	1.000	1.000	1.000	.998	.989	.959	.891	.767	.580	.345	.114	.033	.001
5	1.000	1.000	1.000	1.000	.999	.996	.984	.953	.882	.738	.469	.265	.059

This means

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.991.$$

Example

Check:

$$
\begin{aligned}
P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\
&= \binom{5}{0}(0.1)^0(0.9)^5 + \binom{5}{1}(0.1)^1(0.9)^4 + \binom{5}{2}(0.1)^2(0.9)^3 \\
&= 0.99144.
\end{aligned}
$$

Example

Find the probability of having 9, 10, or 11 successes in 20 Bernoulli trials if the probability of success is 0.6.

From the table:

$$P(X \leq 11) = 0.404, \qquad P(X \leq 8) = 0.057.$$

Hence,

$$\begin{aligned} P(9 \leq X \leq 11) &= P(X \leq 11) - P(X \leq 8) \\ &= 0.404 - 0.057 \\ &= 0.347. \end{aligned}$$

Example

Let

$$n = 10, \qquad p = 0.4.$$

If we want to find $P(X = 4)$ from a cumulative table, then

$$P(X = 4) = P(X \leq 4) - P(X \leq 3).$$

Using the table values,

$$P(X = 4) = 0.633 - 0.382 = 0.251.$$

Directly,

$$P(X = 4) = \binom{10}{4}(0.4)^4(0.6)^6 \approx 0.25082.$$

Binomial in Excel

In Excel, use

```
=BINOM.DIST(number_s, trials, probability_s, cumulative)
```

Here:

- `number_s` = number of successes,
- `trials` = number of trials,
- `probability_s` = probability of success,
- `cumulative = TRUE` gives the CDF,
- `cumulative = FALSE` gives the PDF.

Binomial in Excel - Example

$$\texttt{=BINOM.DIST(4,10,40\%,FALSE)}$$

gives

$$0.25082\ldots$$

$$\texttt{=BINOM.DIST(4,10,40\%,TRUE)}$$

gives

$$0.63310\ldots$$

$$\texttt{=BINOM.DIST(3,10,40\%,TRUE)}$$

gives

$$0.38228\ldots$$