The Binomial Distribution

Gary Schurman MBE, CFA

March 2012

The Binomial Distribution is a discrete probability distribution of the random number of successes drawn from a finite population of [yes][no] experiments. A binomially-distributed random variable can take the integer values $k = 0, 1, 2, 3, .., n$ where k is the number of successes and n is the known population size. In each experiment the probability of success is p and the probability of failure is $1 - p$. Each experiment is independent in that the results of previous experiments do not influence the result of the current experiment. An example application of the binomial distribution would be to determine the probability of finding exactly two defective widgets in a box of ten widgets given that the probability of each widget being defective is 0.05. In this example $p = 0.05$, $k = 2$ and $n = 10$.

We will use the binomial distribution to solve a bond portfolio problem where the number of bond defaults is binomially-distributed. Let's get started...

Our Hypothetical Problem

Imagine that we have a bond portfolio that contains three bonds each with a principal balance of $1,000,000 and each with a maturity date one year hence. If at maturity the bond does not default we will receive the principal balance plus 15.00% simple interest. If at maturity the bond does default we will receive the recovery on that bond which is expected to be 40% of the principal balance. Each bond has an annual default probability of 0.10 and defaults are independent.

Question 1: What is the probability that our bond portfolio will experience zero defaults, one default, two defaults and three defaults?

Question 2: How much principal and interest can we expect to collect in one years time and what is our expected total return?

The Mean and Variance of the Binomial Distribution

The binomial distribution is a distribution of discrete random variables. If the random variable k, which is the number of successes out of n trials, is discrete then the moment generating function of the random variable k where $P(k)$ is the probability mass function can be defined as...

$$M_k(t) = \sum_{k=0}^{n} e^{tk} P(k) \tag{1}$$

The equation for the probability mass function of the binomial distribution with the probability of success equal to p and the probability of failure equal to $1 - p$ is...

$$P(k) = \frac{n!}{k!(n-k)!} \, p^k \, (1-p)^{n-k} \tag{2}$$

If we combine Equations (1) and (2) from above we can write the equation for the moment generating function of the binomial distribution as...

$$\begin{aligned} M_k(t) &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \, p^k \, (1-p)^{n-k} e^{tk} \\ &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \, (pe^t)^k \, (1-p)^{n-k} \end{aligned} \tag{3}$$

Pascal's rule says that...

$$\sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\, a^k\, b^{n-k} = (a+b)^n \tag{4}$$

If we define $a = pe^t$ and $b = 1 - p$ then Equation (3) becomes...

$$M_k(t) = (pe^t + 1 - p)^n \tag{5}$$

The first derivative of Equation (5) with respect to the variable t is...

$$\frac{\delta M_k(t)}{\delta t} = n(pe^t + 1 - p)^{n-1} pe^t \tag{6}$$

The first moment of the distribution will be the limit of Equation (6) as t goes to zero

$$\begin{aligned}\mathbb{E}\Bigg[k\Bigg] &= \lim_{t\to 0} \frac{\delta M_k(t)}{\delta t} \\ &= n(pe^0 + 1 - p)^{n-1} pe^0 \\ &= np \end{aligned} \tag{7}$$

The second derivative of Equation (6) with respect to the variable t is...

$$\frac{\delta^2 M_k(t)}{\delta t^2} = n(n-1)(pe^t + 1 - p)^{n-2} pe^t pe^t + n(pe^t + 1 - p)^{n-1} pe^t \tag{8}$$

The second moment of the distribution will be the limit of Equation (6) as t goes to zero

$$\begin{aligned}\mathbb{E}\Bigg[k^2\Bigg] &= \lim_{t\to 0} \frac{\delta^2 M_k(t)}{\delta t^2} \\ &= n(n-1)(pe^0 + 1 - p)^{n-2} pe^0 pe^0 + n(pe^0 + 1 - p)^{n-1} pe^0 \\ &= n(n-1)p^2 + np \\ &= np - np^2 + n^2p^2 \\ &= np(1-p) + n^2p^2 \end{aligned} \tag{9}$$

The mean of the distribution of the binomially-distributed random variable k is...

$$\begin{aligned} mean &= \mathbb{E}\Bigg[k\Bigg] \\ &= np \end{aligned} \tag{10}$$

The variance of the distribution of the binomially-distributed random variable k is...

$$\begin{aligned} variance &= \mathbb{E}\Bigg[k^2\Bigg] - \Bigg[\mathbb{E}\Bigg[k\Bigg]\Bigg]^2 \\ &= np(1-p) + n^2p^2 - (np)^2 \\ &= np(1-p) \end{aligned} \tag{11}$$

Our Hypothetical Problem Solution

The symbols that we will use to represent the three bonds in our bond portfolio are B_1, B_2 and B_3, respectively. The symbols that we will use to represent the three bonds in our bond portfolio given that there is no default at maturity are B_1^N, B_2^N and B_3^N, respectively. The symbols that we will use to represent the three bonds in our bond portfolio given that there is a default at maturity are B_1^D, B_2^D and B_3^D, respectively. The table below presents the payoffs at maturity on our three bonds given default or no default...

Bond Symbol	No Default Symbol	No Default Payoff	Default Symbol	Default Payoff
B_1	B_1^N	\$1,150,000	B_1^D	\$400,000
B_2	B_2^N	\$1,150,000	B_2^D	\$400,000
B_3	B_3^N	\$1,150,000	B_3^D	\$400,000

The first thing that we want to do is to chart out all of the [default][no default] combinations for our bond portfolio. The table below presents all possible combination of defaults and no defaults...

Number Defaults	Number Combinations	Combinations
0	1	B_1^N, B_2^N, B_3^N
1	3	B_1^D, B_2^N, B_3^N and B_1^N, B_2^D, B_3^N and B_1^N, B_2^N, B_3^D
2	3	B_1^D, B_2^D, B_3^N and B_1^D, B_2^N, B_3^D and B_1^N, B_2^D, B_3^D
3	1	B_1^D, B_2^D, B_3^D

The equation for the total number of [default][no default] combinations is...

$$\text{Total number of combinations} = 2^n = 2^3 = 8 \tag{12}$$

Because the problem assumes that defaults are independent (*) we can use the mathematics from the Compound Probability of Independent Events, which states that the probability of the compound event is simply the product of the single event probabilities. As stated above the single event probability of default is...

$$\text{Probability of default} = P[D] = 0.10 \tag{13}$$

* Note that correlation (i.e. defaults are not independent) is not only a problem in credit risk management it is **the** problem in credit risk management and therefore the independent defaults assumption is not a valid real-world assumption.

The single event probability of no-default is therefore...

$$\text{Probability of no-default} = P[N] = 1 - P[D] = 0.90 \tag{14}$$

Using the compound probability of independent events the probability of zero defaults, one default, two defaults and three defaults is simply product of the single event probabilities times the number of combinations. The probability table that answers question one of our hypothetical problem is...

Number Defaults	Probability Calculation Equation		Probability Calculation		Event Probability
0	$P[N] \times P[N] \times P[N] \times 1$	=	$0.90 \times 0.90 \times 0.10 \times 1$	=	0.7290
1	$P[D] \times P[N] \times P[N] \times 3$	=	$0.10 \times 0.90 \times 0.90 \times 3$	=	0.2430
2	$P[D] \times P[D] \times P[N] \times 3$	=	$0.10 \times 0.10 \times 0.90 \times 3$	=	0.0270
3	$P[D] \times P[D] \times P[D] \times 1$	=	$0.10 \times 0.10 \times 0.10 \times 1$	=	0.0010
Total				=	1.0000

To count combinations as we did above will prove problematic for values of large values of n. We can also further simplify the probability calculation in the table above by using exponents. Using Equation (2) above we can rewrite the probability table above as...

Number Defaults	Probability Equation		Probability Calculation		Event Probability
0	$\binom{n}{k} p^k (1-p)^{n-k}$	=	$\frac{3!}{0!(3-0)!} \times 0.10^0 \times 0.90^3$	=	0.7290
1	$\binom{n}{k} p^k (1-p)^{n-k}$	=	$\frac{3!}{1!(3-1)!} \times 0.10^1 \times 0.90^2$	=	0.2430
2	$\binom{n}{k} p^k (1-p)^{n-k}$	=	$\frac{3!}{2!(3-2)!} \times 0.10^2 \times 0.90^1$	=	0.0270
3	$\binom{n}{k} p^k (1-p)^{n-k}$	=	$\frac{3!}{3!(3-3)!} \times 0.10^3 \times 0.90^0$	=	0.0010
Total		=		=	1.0000

The table below presents the bond portfolio payoff table given the number of defaults and the expected payoff...

Number Defaults	Payoff Calculation	Total Payoff	Event Probability	Expected Payoff
0	$1,150,000 \times 3$	3,450,000	0.7290	2,515,050
1	$400,000 \times 1 + 1,150,000 \times 2$	2,700,000	0.2430	656,100
2	$400,000 \times 2 + 1,150,000 \times 1$	1,950,000	0.0270	52,650
3	$400,000 \times 3$	1,200,000	0.0010	1,200
Total			1.0000	3,225,000

Per the payoff table above we can expect to collect \$3,225,000 in principal and interest with the expected total return being...

$$\text{Expected return} = \frac{3,225,000}{3,000,000} - 1 = 7.50\% \tag{15}$$

This answers question two of our hypothetical problem.

Tip - When Population Size Is Very Large

When population size n is very large the factorials of n, k and $n-k$ may be too large for your computer or hand-held calculator to handle thus making the probabilities impossible to calculate. The solution to this problem is to calculate the natural log of the probability and then take the exponential. Let's use the following example...

Given that our population size is 1,000 what is the probability of exactly 600 successes given an individual probability of success of 0.60? For this example we have...

$$n = 1000 \text{ ...and... } k = 600 \text{ ...and... } p = 0.60 \tag{16}$$

Using Equation (2) above the probability of 600 successes is...

$$P\Bigg[k = 600\Bigg] = \frac{1000!}{600!(1000-600)!} \times 0.60^{600} \times (1-0.60)^{1000-600} \tag{17}$$

The factorial of n (and maybe even k and $n-k$) in the equation above is too large for mosts computers and hand-helds to handle. We can work around this problem by rewriting Equation (2) as...

$$P(k) = exp\Bigg[\sum_{i=1}^{n} \ln i - \sum_{i=1}^{k} \ln i - \sum_{i=1}^{n-k} \ln i + k \ln p + (n-k)\ln(1-p)\Bigg] \tag{18}$$

The solution to the problem is...

$$\begin{aligned} P(600) &= exp\Bigg[\sum_{i=1}^{1000} \ln i - \sum_{i=1}^{600} \ln i - \sum_{i=1}^{1000-600} \ln i + 600 \ln 0.60 + (1000-600)\ln(1-0.60)\Bigg] \\ &= exp\Bigg[5,912.13 - 3,242.28 - 2,000.50 + (-306.50) + (-366.52)\Bigg] \\ &= exp\Bigg[-3.66\Bigg] \\ &= 0.02574 \end{aligned} \tag{19}$$