1. Introduction

Welcome to the series of e-learning modules on Practicals. In this module we shall simulate the random samples from the given distribution. We can also compare the mean and variance of theoretical and sample distributions.

By the end of this session, you will be able to:

- Simulate observations from a given distribution
- Compare the mean and variance of an original distribution with the simulated distribution

In the first exercise, we are required to simulate a random sample of size 10 from a discrete uniform distribution with parameter n is equal to 5 using the random numbers given, which are 614, 122, 811, 658, 147, 389, 948, 825, 994, 862. Also we are required to compare the sample mean and variance with the population measures.

Given X follows discrete uniform distribution with parameter n is equal to 5.
Therefore the probability mass function is given by, p of x is equal to 1 by n, where x take values zero, one two three and four.
Is equal to 1 by 5.

To simulate the random sample, first we complete the table as given.
The first column in the table denotes the values taken by x; that is zero, 1, 2, 3 and 4.
The second column denotes the probability of X which takes the value x, which is calculated using probability mass function.
Since p of x is independent of x, all the p of x is equal to 1 by n is equal to 1 by 5 is equal to zero point 2.
The third column gives the probability that X is less than or equal to x, that is the cumulative probabilities. First we write the value of p of x as it is. In the next row we add this probability to the probability of p of x.
That is, zero point 2 plus zero point 2 is equal to zero point 4, the third row will be zero point 4 plus zero point 2 is equal to zero point 6 and so on. Usually the last entry in this column will be 1 since the total probability of any probability mass function is 1.
Now, using the cumulative probabilities, we write range for the random numbers. Since we have given 3-digit random numbers, we write 3 digits in the range.
Usually while writing the range, lower limit of probability is included in the same class and the upper limit in the next class. That is first cumulative frequency we have zero point 2. If we consider 3 digits after point, we get zero point 2 zero zero. Hence first range we write as zero zero zero to 199. 200 is then included in the next class. That is 200 to 399 etc., till the last class which will be 800 to 999.
Once we have filled in the range column, we put tally bars using the given random numbers. The first random number is 614, which belongs to 4th row in the table. Hence we put tally bar in 4th row that is x is equal to 3. Second random number is 122 which belong to the 1st row that is x is equal to zero. Hence we put tally bar in the first row. Likewise, we put tally bars using all the random numbers.
Hence the random observations from discrete uniform distribution is given by,
3, 0, 4, 3, 0, 1, 4, 4, 4, 4

The 6th column gives the frequencies of the values taken by x. We count the number of tally bars in the 5th column and write the frequencies.
Since we need to calculate the sample mean and variance, we calculate 2 more columns f into x and f into x square and find the totals of f, f into x and f into x square.

2. Population Mean and Variance

First let us find the population mean and variance. For a discrete uniform distribution, mean is equal to n minus 1 by 2 is equal to 5 minus 1 divided by 2 is equal to 2.
And variance is given by, n square minus 1 divided by 12 is equal to 5 square minus 1 divided by 12 is equal to 2.
From the sample observations, the sample mean is given by, x bar is equal to summation of f into x by N is equal 27 by 10 is equal to 2 point 7, and
Variance is given by, sigma square is equal to summation of f into x square divided by N minus summation of f into x divided by N the whole square, which is equal to 99 by 10 minus 27 by 10 the whole square, is equal to 2 point 6 one.
Here, observe that the sample mean and variance are greater than that of population mean and variance.

In this exercise, we need to simulate a random sample of size 10 from Bernoulli distribution with parameter p is equal to zero point 4 using the random numbers   845, 449, 339, 797, 345, 937, 573, 194, 476, 968. Also we need to compare the sample mean and variance with population measures.

Given that x follows Bernoulli distribution with parameter p is equal to zero point 4.
Therefore the probability mass function p of x is equal to p power x into q power 1 minus x, where x takes the values zero and one.
Is equal to zero point 4 power x into zero point 6 power 1 minus x.
To simulate the random numbers, first we create the table.

Figure 1

x	P(X=x)	P(X≤x)	Range	Tally bars	f	fx	fx2
0	0.6	0.6	000-599	llll l	6	0	0
1	0.4	1.0	600-999	llll	4	4	4
			Total		10	4	4

Here x takes only 2 values zero and 1.
Using p of x we find the corresponding probabilities and in the next column we find the cumulative probabilities.
Using the cumulative probabilities we find range to put tally bars to the given random numbers

so that we can simulate the random observations.
The first random number corresponds to x is equal to 1, the 2nd random number corresponds to x is equal to 0, and so on we can simulate 10 random observations from Bernoulli distribution as:
1, 0, 0, 1, 0, 1, 0, 0, 0, 1.
To compare the population mean and variance with sample mean and variance, we calculate 3 more columns f, f into x and f into x square and find the corresponding totals.

Now let us find the population mean and variance.
Population mean is given by, p is equal to zero point 4 and variance is equal to p into q is equal to zero point 4 into zero point 6 is equal to zero point 2 four.
For the sample, mean is given by, x bar is equal to summation of f into x divided by N is equal to 4 by 10 is equal to zero point 4.
Variance is given by, sigma square is equal to summation of f into x square by N minus summation of f into x divided N the whole square is equal to 4 by 10 minus 4 by 10 whole square is equal to zero point 2 four.

Here we need to note that in this distribution for given random numbers, the sample mean and variance exactly coincide with population mean and variance.

3. Simulate Random Sample – Part 1

In this exercise, we need to simulate a random sample of size 15 from binomial distribution with parameters n is equal to 5 and p is equal to zero point 5 using the random numbers given: 443, 207, 761, 503, 575, 227, 222, 108, 489, 255, 877, 182, 190, 380, 758. Also we need to compare the sample mean and variance with population measures.

Given that X follows binomial distribution with parameters n is equal to 5 and p is equal to zero point 5. Therefore probability mass function is given by,
p of x is equal to n c x into p power x into q power n minus x where x takes the values zero, one, two and so on to n.
That is, p of x is equal to 5 c x into zero point 5 to the power 5.

To simulate random observations from binomial distribution, first we find the following table.

Figure 2

x	P(X=x)	P(X≤x)	Range	Tally bars	f	fx	fx2
0	0.031	0.031	000-030		0	0	0
1	0.156	0.187	031-186	ll	2	2	2
2	0.313	0.5	187-499	llll lll	8	16	32
3	0.313	0.813	500-812	llll	4	12	36
4	0.156	0.969	813-968	l	1	4	16
5	0.031	1	969-999		0	0	0
				Total	15	34	86

Here x take values zero, one, two, three, four and five.
Using the probability mass function of binomial distribution, we find p of x.
Then we find the cumulative frequencies as we have done in the earlier problems and fill in the ranges as zero zero zero to zero 3 zero, zero 3 one to 186 etc and lastly 969 to 999.
Now using the random numbers given in the problem, we can generate random observations from the binomial distribution as:
2, 2, 3, 3, 3, 2, 2, 1, 2, 2, 4, 1, 2, 2, and 3.
To compare the population mean and variance with sample mean and variance, we calculate 3 more columns f, f into x and f into x square. And find the corresponding totals.

Now let us find the population mean and variance.
Mean of the binomial distribution is given by, n into p is equal to 5 into zero point 5 is equal to 2 point 5
And variance is n into p into q is equal to 5 into zero point 5 into zero point 5 is equal to 1 point 2 five.

Next, the sample mean and variance are found as follows.
Sample mean is given by, x bar is equal to summation of f into x divided by N is equal to 34 by 15 is equal to 2 point 2 six 6 seven and
Variance sigma square is equal to summation of f into x square divided by N minus summation of f into x by N the whole square is equal to 86 by 15 minus 43 by 15 whole square is equal to 2 point zero 2 eight 9.

Observe that in this problem, mean of the population is greater than that of sample mean and variance of sample is greater than that of variance of population.

Simulate a random sample of size 15 from binomial distribution with parameters n is equal to 6 and p is equal to zero point 3 using following random numbers. 4462, 6791, 6102, 0215, 7398, 2671, 6175, 0513, 1413, 8771, 1596, 1747, 2005, 2493, 9285.

Given x follows binomial distribution with parameters n is equal to 6 and p is equal to zero point 3.
Therefore probability mass function is given by,
P of x is equal to n c x into p to the power x into q to the power n minus x where x takes the values zero, one etc., to n.
Substituting the values in the equation, we have P of x is equal to 6 c x into zero point 3 to the power x into zero point 7 to the power 6 minus x.

Before we simulate the random observations, first we find the following table.

Figure 3

x	P(X=x)	P(X≤x)	Range
0	0.1176	0.1176	0000-1175
1	0.3025	0.4201	1176-4200
2	0.3241	0.7442	4201-7441
3	0.1852	0.9294	7442-9293
4	0.0595	0.9889	9294-9888
5	0.0103	0.9992	9889-9991
6	0.0008	1	9992-9999

Here x take values zero one two etc., to 6.
By substituting the values of x in p of x, we can find the second column p of x is equal to x.
Third column gives the cumulative probabilities and using these probabilities we write the 4th column which is nothing but the range for random numbers.
Since we have been given the 4 digit random numbers in this problem, we write 4 digits for the range of the random numbers. That is zero zero zero zero to 1175, 1176 to 4200 and so on.
Now using the random numbers given in the problem, we can generate random observations from the binomial distribution as:
2, 2, 3, 3, 3, 2, 2, 1, 2, 2, 4, 1, 2, 2, and 3.

4. Simulate Random Sample – Part 2

In this exercise, we are required to simulate a random sample of size 15 from Poisson distribution with parameter lambda is equal to zero point 2 five using the following random numbers: 230, 298, 864, 891, 498, 462, 635, 618, 214, 939, 601, 832, 158, 635, 984.

Given that X follows Poisson distribution with parameter lambda is equal to zero point 2 five. Therefore the probability mass function is given by,
P of x is equal to e power minus lambda, into lambda power x by x factorial which is equal to e power minus zero point 2 five into zero point 2 five to the power x divided by x factorial, where x takes the values zero one two etc to infinity.

To simulate the random observations, first we find the following table.

Figure 4

x	P(X=x)	P(X≤x)	Range
0	0.779	0.779	000-778
1	0.195	0.974	779-973
2	0.024	0.998	974-997
≥3	0.002	1	997-999

Since we do not know the last value of x, we find p of x is equal to x and the cumulative probabilities simultaneously.
Observe that, since lambda is small we get very large probabilities in the beginning.
We reach cumulative probability zero point 9 nine 8 when x is equal to 2 and hence the remaining probability zero point zero zero two by putting x greater than or equal to 3 and we get the last entry in 3^{rd} column as one.
Since we have given only 3 digit random numbers, we write 3 digits in range column.
Once we have the range column ready, we can generate the random observations from Poisson distribution as:
0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0 and 2

Simulate a random sample of size 25 from Poisson distribution with parameter λ is equal to 5 using following random numbers. 633, 570, 128, 966, 219, 045, 054, 260, 001, 042, 986, 685, 763, 564, 982, 593, 652, 003, 080, 861, 812, 679, 627, 430, 896.

Given that x follows Poisson distribution with parameter lambda is equal to 5.

Therefore the probability mass function is given by,
P of x is equal to e power minus lambda, into lambda power x by x factorial is equal to e power minus five, into five power x divided by x factorial, where x takes the values zero one two etc to infinity.
To simulate the random observations, first we find the following table.

Figure 5

x	P(X=x)	P(X≤x)	Range
0	0.007	0.007	000-006
1	0.034	0.041	007-040
2	0.084	0.125	041-124
3	0.140	0.265	125-264
4	0.176	0.441	265-440
5	0.176	0.617	441-616
6	0.146	0.763	617-762
7	0.104	0.867	763-866
8	0.065	0.932	867-931
9	0.036	0.968	932-967
10	0.018	0.986	968-985
11	0.008	0.994	986-993
≥12	0.006	1	994-999

Observe that p of x values are very small, hence we find probabilities of x and the cumulative probabilities simultaneously.
When x is equal to 11, the cumulative probability becomes zero point 9 nine 4. Hence the remaining probability, that is 1 minus zero point 9 nine 4 is equal to zero point zero zero six and we assign to x value as greater than or equal to 12.

Once we obtain the cumulative probabilities, we can find the range for the random numbers and simulate random observations from the Poisson distribution. Hence we get the random sample as:
6, 5, 3, 9, 3, 2, 2, 3, 0, 2, 11, 6, 7, 5, 10, 5, 6, 1, 2, 7, 7, 6, 6, 4 and 8.

5. Simulate Random Sample & Summary – Part 3

Simulate a random sample of size 15 from geometric distribution with parameter p is equal to zero point 6 using following random numbers. 271, 478, 866, 828, 436, 250, 005, 939, 048, 031, 759, 124, 601, 105, 715.

Given, x follows geometric distribution with parameter p is equal to zero point 6.
Therefore the probability mass function is given by p of x is equal to q power x into p where x takes the values zero, one, two etc and p lies between zero and one
This is equal to zero point 4 power x into zero point 6.

Next, to simulate the random observations, first we find the table as given.

Figure 6

x	P(X=x)	P(X≤x)	Range
0	0.600	0.600	0-599
1	0.240	0.840	600-839
2	0.096	0.936	840-935
3	0.038	0.974	936-973
4	0.015	0.989	974-988
≥5	0.011	1	989-999

Here, we do not have the last value of x.
Since x takes the values from zero to infinity, we find p of x and cumulative probabilities p of x less than or equal to x simultaneously.
Observe that when x is equal to 4 we get cumulative probability as zero point 9 eight 9.
Hence we combine other values of x and write greater than or equal to 5 and the remaining probability we assign to this class that is, zero point zero 1 one.
Using these cumulative probabilities, we find the range for the random numbers and simulate the random observations from this distribution as:
0, 0, 2, 1, 0, 0, 0, 3, 0, 0, 1, 0, 1, 0 and 1

In this problem, we have to simulate a random sample of size 15 from geometric distribution with parameter p is equal to zero point 3 using the given random numbers: 521, 000, 009, 209, 818, 833, 180, 161, 714, 426, 804, 510, 734, 891 and 912.

Given that x follows geometric distribution with parameter p is equal to zero point 3.
Therefore the probability mass function is given by p of x is equal to q to the power x into p

where x takes the values zero, one, two etc and p lies between zero and one.
Is equal to zero point 7 to the power x into zero point 3.

To simulate the random observations, first we find the following table.

Figure 7

x	P(X=x)	P(X ≤ x)	Range
0	0.300	0.300	000-299
1	0.210	0.510	300-509
2	0.147	0.657	510-656
3	0.103	0.760	657-759
4	0.072	0.832	760-831
5	0.050	0.882	832-881
6	0.035	0.917	882-916
7	0.025	0.912	917-941
≥8	0.058	1	942-999

Like in the previous problem, here too we do not have the last value of x.
Since x takes values from zero to infinity, we find p of x and cumulative probabilities p of x less than or equal to x simultaneously.
Observe that our highest random number here is 912 which corresponds to x is equal to 7.
Hence we can write the table either up to x is equal to 7 or 8. Thus, we stop the table by writing greater than or equal to 8.
The next column gives the range for random numbers using which we generate the random observations for the distribution as:
2, 0, 0, 0, 4, 5, 1, 1, 3, 1, 4, 2, 3, 6 and 6.

Simulate a random sample of size 20 from negative binomial distribution with parameters r is equal to and p is equal to zero point 7 five based on the following random numbers. 989, 761, 511, 755, 386, 894, 565, 999, 979, 525, 083, 912, 994, 865, 915, 052, 208, 785, 073, 690.

Given that x follows a negative binomial distribution with parameters r is equal to 5 and p is equal to zero point 7 five.
Hence, the probability mass function is given by, p of x is equal to x plus r minus 1 c r minus 1 into p to the power r into q to the power x
Is equal to x plus 4 c 4 into zero point seven 5 to the power 5 into zero point 2 five to the power x, where x takes the values zero, one, two etc and p lies between zero and one.

Now, to simulate the random observations, first we find the following table.

Figure 8

x	P(X=x)	P(X≤x)	Range
0	0.237	0.237	000-236
1	0.297	0.534	237-533
2	0.222	0.756	534-755
3	0.130	0.886	756-885
4	0.065	0.951	886-950
5	0.029	0.980	951-979
6	0.012	0.992	980-991
7	0.005	0.997	992-996
8	0.002	0.999	997-998
9	0.001	1.000	999-999

As we have done in earlier problems, here also we give find p of x for different values of x and write it in the 2^{nd} column and find cumulative probabilities in the third column.
Next, using cumulative probabilities we write the range for random numbers to generate random observations from the distribution.
Hence the random observations are given by:
6, 3, 1, 2, 2, 4, 2, 9, 5, 1, 0, 4, 7, 3, 4, 0, 0, 3, 0 and 2.

Here's a summary of our learning from this session:

- Simulate observations from any given distribution
- Compare the mean and variance of the original distribution with the simulated distribution.