Frequently Asked Questions

1. What are random numbers?

Answer:

Random numbers also called as random sampling numbers refer to some well known sequence of digits in which the successive figures appear in a perfectly random order. This means that if a digit is blindly selected from the random number table any of the 10 digits 0,1,2,...,9 is likely to occur with the same probability 1/10. Similarly if two consecutive digits are taken the numbers formed by them in order they appear may be any of the hundred numbers 00, 01,02,....,99 with the same probability 1/100 and son on. The series of random numbers are usually available in groups of 4 digits to facilitate easy reading as follows:

4504	9523	3282	3756	0653
1014	7894	9307	0458	9983
1575	2458	9200	8566	7302

Of course any haphazard arrangement of digits 0 to 9 will not give random numbers. The random numbers have been prepared by special devices and tested for randomness. Some of the famous series of random numbers are those of Trippett (41,600 digits) Kendall and Babibngto Smith (100,000 digits), Rand corporation (1,000,000 digits). Random numbers are used for drawing random samples.

2. Describe in detail how you will select without replacement a random sample of 3 units from a population of 121 units using a procedure which does not involve rejection of a large number of random numbers.

Answer:

The method is described in the following steps:

- All the 121 units of the population are serially numbed
- Let us now take a page of random numbers and with closed eyes select a digit from a page. Starting from this and proceeding horizontally consecutive digits are taken down in groups of three (because largest serial number 121 is three digit) giving several 3 digit numbers
- The random numbers are divided by 135 and the remainders are obtained (Note that 1000 three digit numbers are possible viz. 001,002,...,999 and 000 ; 1000 is exactly divisible by 125.) If a number is less than 125 that number itself is taken as a remainder
- Any remainders in between 122 and 124 and are also 000, if any occurring are rejected if any remainder appears more than once all subsequent repetitions are rejected

- From the numbers now remaining only the first three numbers are taken and the population units corresponding to these as serial numbers give the random sample
- 3. Explain how do you select random samples from a population related to attributes represented in the form of a contingency table.

Answer:

- 1) Cumulate the frequencies of the first, second,, last Rows.
- 2) Inclusive class Intervals corresponding to the cumulative frequencies are formed.
- 3) For each random number selected from a random numbers table tally marks are put corresponding to the intervals containing the value of the number
- 4) This procedure is continued until we get the required sample
- 5) Total of the tally marks in the corresponding intervals represents a sample frequency of the corresponding cell.
- 4. Draw a random sample of size 10 with replacement from the following frequency distribution and compare the sample mean with the population mean

Annual sales	30-39	40-49	50-59	60-69	70-79	80- 89	90- 99	100- 109
(Rs.'000)								
No. of firms	14	36	47	66	41	24	10	2

Answer:

All the 240 firms considered in the frequency distribution are serially numbered -14 observations in the Class interval 30-39 are numbered 1 to 14, 36 in the next interval numbered 15 to 50 and so on as shown in the table given below

Class interval	Frequency	Mid value	Cum frequency	Serial numbers alloted
30-39	14	34.5	14	001-014
40-49	36	44.5	50	015-050
50-59	47	54.5	97	051-097
60-69	66	64.5	163	098-163
70-79	41	74.5	204	164-204
80-89	24	84.5	228	205-228
90-99	10	94.5	238	229-238
100-109	2	104.5	240	239-240

Using Fisher & Yate's random numbers table 3 digit random numbers are selected preceding horizontally. The numbers 250 and above are replaced by their remainders on division by 250

Random numbers	489	572	665	890	501	154	786	475	888	750	147
Remainder s	239	72	165	140	1	154	36	225	138	000	147

The random sample is given by the mid values of the classes in which the selected firms are placed. Random number 750 is not considered here because the remainder is zero

Random sample with replacement

Serial No's	239	72	165	140	1	154	36	225	138	147
Sample values	104.5	54.5	74.5	64.5	34.5	64.5	44.5	84.5	64.5	64.5
(Rs.000's)										

The sample mean is 655/10 = 65.5 (Rs. '000). The mean calculated from the given frequency distribution is 63.08 (Rs.'000)

Sample mean = 65.5 Population mean = 63.08

5. Draw SRSWR samples of size 25 from the following frequency distribution and find the mean and variance of the sampling distribution.

Grade	90-100	80-89	70-79	60-69	50-59	40-49	30-39
No. of students	9	32	43	21	11	3	1

Answer:

All the 120 students considered in the frequency distribution are serially numbered -09 observations in the Class interval 90-100 are numbered 000 to 009, 32 in the next interval are numbered 010 to 041 and so on as shown in the table given below

				With Replac	ement
Grade (Class interval)	No. of students(Frequ	L.T.C.F	Inclusive C.I	Tally	Frequency
	ency)			marks	
90-100	9	9	000-009	//	2

80-89	32	41	010-041	///	3
70-79	43	84	042-084		12
60-69	21	105	085-105		4
50-59	11	116	106-116		3
40-49	3	119	117-119	1	1
30-39	1	120	120		0
	N=120				

3 digited Random numbers are obtained from page No. 134 of Fisher and Yate's Random numbers table starting from column 1 proceeding vertically. The numbers above 120 are discarded here

034,003,112,052,079,097,112,086,082,084,050,057,047,059,094,060,099,114,026,0 33,003,068,078,119,083

Sampling Distribution with Replacement scheme is

Grade	No. of Sample values(f _i)	Mid points	f _i yi	f _i y _i ²
		())		
90-100	2	95	190	18050
80-89	3	84.5	253.5	21420.75
70-79	12	74.5	894	66603
60-69	4	64 5	258	16641
00 00	т 	04.0	200	100+1
50-59	3	54.5	163.5	8910.75
40-49	1	44.5	44.5	1980.25
30-39	0	34.5	0	0
	n=25		1803.5	133605.75

Sample mean
$$\frac{1}{y} = \frac{\sum_{i=1}^{n} f_i y_i}{n} = \frac{1803.5}{25} = 72.14$$

Variance
$$\sigma^2 = \frac{\sum_{i=1}^{n} f_i y_i^2}{n} - y^2 = \frac{133605.75}{25} - 72.14^2 = 140.0504$$

6. Draw a Random sample of size 10 using SRSWOR from the following contingency Table

Sex	Male	Female
Class		
А	123	153
В	145	303

Answer:

Cumulative Frequencies, Inclusive Class Intervals and tally marks are obtained as follows

Sex	Male	Female
Class		
	123	276
A	000 – 123	124 – 276
	/	////
В	421	724
	277 – 421	422 – 724
	/	///

N=724 (has 3 digits)

Therefore we shall select 3 digited random numbers leaving the numbers 000 and numbers greater than 724 under WOR scheme from Fisher & Yates table from page 134, Column 1

The random numbers are 034, 167, 125, 555, 162, 630, 332, 576, 181, 266

Therefore, the required sample is:

Se	Male	Female
Class		
A	1	5
В	1	3

7. Draw SRSWOR samples of size 25 from the following frequency distribution and find the mean and variance of the sampling distribution.

Grade	90-100	80-89	70-79	60-69	50-59	40-49	30-39
No. of students	9	32	43	21	11	3	1

Answer:

				Without Rep	olacement
Grade (Class interval)	No. of students(Frequ ency)	L.T.C.F	Inclusive C.I	Tally marks	Frequency
90-100	9	9	000-009	/	1
80-89	32	41	010-041	////	4
70-79	43	84	042-084		12
60-69	21	105	085-105		4
50-59	11	116	106-116	///	3
40-49	3	119	117-119	/	1
30-39	1	120	120		0
	N=120				

3 digit Random numbers are obtained from page No. 134 of Fisher and Yate's Random numbers table starting from column 1 proceeding vertically. The numbers above 120 are discarded and since we are in WOR scheme the numbers which are already copied are rejected.

034,003,112,052,079,097,112(rejected), 086, 082, 084, 050, 057, 047,

059, 094, 060, 099, 114, 026, 033, 003 (rejected), 068, 078, 119, 083, 013, 111 Sampling Distribution without Replacement scheme is:

Grade	No. of Sample values(f _i)	Mid points	fi y i	f _i yi ²
		() ¹ /		
90-100	1	95	95	9025
80-89	4	84.5	338	28561
		00		
70-79	12	74.5	894	66603
00.00		045	050	10011
60-69	4	64.5	258	16641
50-59	3	54.5	163.5	8910.25
40-49	1	44.5	44.5	1980.25
30-39	0	34.5	0	0
	n=25		1793	131721

Sample mean
$$\frac{1}{y} = \frac{\sum_{i=1}^{n} f_i y_i}{n} = \frac{1793}{25} = 71.72$$

Variance
$$\sigma^2 = \frac{\sum_{i=1}^{n} f_i y_i^2}{n} - \overline{y}^2 = \frac{131721}{25} - 71.72^2 = 125.0816$$

8. How do you select random samples from a grouped data or a population represented in the form of a frequency distribution?

Answer:

We follow the following procedure for the selection of the sample from grouped data or frequency distribution.

1) We first cumulate the given frequencies

- 2) Corresponding to the cumulative frequencies the inclusive class intervals are formed
- 3) For each random number selected from a random numbers table tally marks are put corresponding to the intervals containing the value of the number
- 4) This procedure is continued until we get the required sample
- 5) Total of the tally marks in the corresponding intervals represents a sample frequency of the corresponding value of the variables.
- 9. How do you select simple random samples from an ungrouped data?

Answer:

All the units of the population are numbered from 1 to N or from 0 to N-1. We consult the random number table to take a simple random sample. Suppose the size of the population is 80 and we have to select a random sample of 8 units.

The units of the population are numbered from 01 to 80. We read two-digit numbers from the table of random numbers. We can take a start from any columns or rows of the table. Let us consult random number table given in this content.

Two-digit numbers are taken from the table. Any number above 80 will be ignored and if any number is repeated, we shall not record it if sampling is done without replacement. Let us read the first two columns of the table. The random number from the table is 10, 37, 08, 12, 66, 31, 63 and 73.

The two numbers 99 and 85 have not been recorded because the population does not contain these numbers. The units of the population whose numbers have been selected constitute the simple random sample. If we have numbers which are already copied then the scheme is called with Replacement (WR) scheme.

If we do not copy the numbers which are already copied then the scheme is called Without Replacement (WOR) scheme

10. Draw a Random sample of size 15 from the following frequency distribution by explaining the procedure.

Class Interval:	0-5	5-10	10-15	15-20	20-25	25-30	30-35
Frequency:	10	12	13	8	14	16	7

Answer:

Draw a table consisting of following columns

Column1: Class intervals or the variable values given in the population

Column 2: Frequencies or number of observations taking particular values of the variables in the population

Column 3: Obtain the cumulative frequencies.

Column 4: Obtain the Inclusive Class intervals for the cumulative frequencies obtained in the third column. All limits are two digit because total frequency N = 80 is a two digited number.

Column 5: Therefore, we select 2 digit 15 random numbers leaving the numbers 00 and the numbers greater than 80 under WOR scheme from the Fisher & Yates table page No. 134, Col. 1. The random number are

03, 49, 43, 36, 56, 70, 68, 59, 13, 40, 71, 78, 09, 58 and 10

Select first random number 03 put a tally mark against the interval 00 to 10. Like this continue the process until you reach a fifteenth random number 10, which falls in the interval 00 to 10.

Column 6: Count the total number of tally marks of each intervals. Observe the total of the column 6 must be equal to 15, a required sample size.

C.I	Freq.	Cum. Freq	Incl.C.I	Tally Marks	Total
0 - 5	10	10	00 - 10	///	03
5 - 10	12	22	11 - 22	/	01
10 - 15	13	35	23 - 35		00
15 - 20	8	43	36 - 43	///	03
20 - 25	14	57	44 - 57	//	02
25 - 30	16	73	58 - 73	////	05
30 - 35	7	80	74 - 80	/	01

Here N = 80 has 2 digits.

Hence, the required sample is:

C.I	0-5	5-10	10-15	15-20	20-25	25-30	30-35
frequency	3	1	0	3	2	5	1

11. Draw a random sample of size 15 from the following contingency table

good	Very	Good	Medium	Poor	
------	------	------	--------	------	--

Method A	80	70	50	10
Method B	26	34	76	25

Use the random numbers 164, 041, 351, 348, 296, 341, 225, 236, 371, 115, 241, 281, 202, 254, 151

Answer:

	Very good	Good	Medium	Poor
Method A	80	70	50	10
	80 000-080 / (1)	150 081-150 / (1)	200 151-200 // (2)	210 201-210 / (1)
Method B	26 236 211-236 // (2)	34 270 237-270 // (2)	76 346 271-346 /// (3)	25 371 347-371 /// (3)

The required sample is:

	Very good	Good	Medium	Poor
Method	1	1	2	1
A				
Method	2	2	3	3
В				

12. The following frequency distribution relates to the number of children per family for 210 families

 No. of children:
 1
 2
 3
 4
 5
 6
 7

 Frequency:
 4
 35
 78
 52
 30
 9
 2

Given the following random numbers draw an SRSWOR samples of 10 families from the given population:

190, 044, 193, 181, 121, 179, 121, 194, 209, 009, 156, 113, 038, 199, 144, 172, 124, 179, 200, 088.

Answer:

No.of children	Frequency	Cum. Freq	Incl.C.I	Tally Marks	Total
1	4	4	000-004		0

2	35	39	005-039	/	1
3	78	117	040-117	/	1
4	52	169	118-169	//	2
5	30	199	170-199	##	5
6	9	208	200-208		0
7	2	210	209-210	/	1

Hence a random sample of 10 families are obtained as follows

No.of children	1	2	3	4	5	6	7
No.of families	0	1	1	2	5	0	1

13. The following contingency table relates to the smoking ability per day and the economic status of people located in a region.

Draw a random sample of size 10 using WOR scheme from the above table using the following random numbers.

234, 228, 190, 44, 93, 282, 321, 179, 221, 294, 310, 009, 256, 113, 124, 179, 26, 88. Economic Status

No. of		Royal	Rich	Middle
Cigarates				class
	0-5	18	51	40
	5-10	40	14	12
	10-15	21	22	10

Answer:

ſ

Economic status

No. of		Royal	Rich	Middle class
Cigarates	0-5	18 000-018 /	69 019-069 /	109 070-109 /
	5-10	149 110-149 //	163 150-163	175 164-175

10-15	196 176-196 ///	218 197-218	228 219-228 //

The required sample is:

Economic status						
No. of Cigarates		Royal	Rich	Middle class		
	0-5	1	1	1		
	5-10	2	0	0		
	10-15	3	0	2		

14. Nine villagers in a certain administrative area contain 793, 170, 970, 657, 1721, 1603, 864, 383 and 826 fields respectively. Make a random selection of 6 fields using the random numbers.

7358, 922, 4112, 3596, 633 and 3999

Answer:

Villages	Fields	Cum. Freq	Incl.C.I	Tally Marks	Total
1	793	793	0000-0793	/	1
2	170	963	0794-0963	/	1
3	970	1933	0964-1933		0
4	657	2590	1934-2590		0
5	1721	4311	2591-4311		3
6	1603	5914	4312-5914		0
7	864	6778	5915-6778		0
8	383	7161	6779-7161		0
9	826	7987	7162-7987	/	1

Hence a random sample of 6 fields are obtained as follows

Villages	1	2	3	4	5	6	7	8	9
No.of fields	1	1	0	0	3	0	0	0	1

15. Show how to select 10 random samples of 4 students each with replacement from the table below using random numbers and also find the mean and S.D of the sampling distribution of means

Height	Frequency
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8

Answer:

Use two digit numbers each of the 100 students 00,01,02,...,99. Thus the five students with heights 60-62 in are numbered 00-04, 18 students with height 63-65 in are numbered 05-22, etc.

We now draw sampling numbers from the random umber table.

From the first line we find the sequence 51, 77, 27, 46, 40, etc. which we take as a random sampling numbers each which yields the height of a particular student. Thus 51 correspond to a student having height 66-68 in which we take as 67 in the class mark. Similarly 77,27 and 46 yields the height of 70,67 and 67 in respectively. By this process we obtain below table which shows the sample numbers drawn the corresponding heights and the mean height of each of 10 samples. It should be mentioned that although we have entered the random number table on the first line, we could have started any where and chosen any specified pattern.

SI.	Sample drawn	Corresponding heights	Mean height
no			
1	51,77,27,46	67,70,67,67	67.75
2	40,42,33,12	67,67,67,64	66.25
3	90,44,46,62	70,67,67,67	67.75
4	16,28,98,93	64.67.73.73	69.25
5	58.20.41.86	67.64.67.70	67
6	19.64.08.70	64.67.64.70	66.25
7	56,24,03,32	67,67,61,67	65.5

8	34,91,83,58	67,70,70,67	68.5
9	70,65,68,21,	70,70,70,64	68.5
10	96,02,13,87	73,61,64,70	67

Sample mean (y _{i)}	Tally	f	fyi	fy _i ²
65.50	/	1	65.50	4290.25
66.25	//	2	132.5	8778.125
67	//	2	134	8978
67.75	//	2	135.5	9180.125
68.5	//	2	137	9384.5
69.25	/	1	69.25	4795.5625
	Total	10	673.75	47406.5625

The above table gives the frequency distribution of sample mean of heights. This is a sampling distribution of the means. The mean and S.D are obtained as usual

$$\overline{y} = \frac{\prod_{i=1}^{n} fy_i}{n} = \frac{673.75}{10} = 67.375$$

S.D = $\sqrt{\frac{\prod_{i=1}^{n} fy_i^2}{n} - \frac{1}{y^2}} = \sqrt{\frac{47406.5625}{10}} - (67.375)^2 = 14.187$