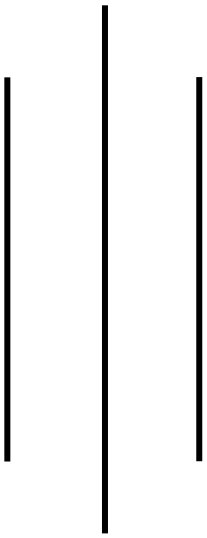
NCCS
College of IT & Management
[Affiliated to TU]

National College of Computer Studies

Paknajol, Kathmandu



Lab Report of Statistics II

Submitted By:

Name: Aakash Dhakal

Program: BSc. CSIT (3rd Semester)

Section: B

Roll No: 1

Submitted To:

Santosh Chhathkuli Sir

Lab - 1

The following data represent the number of days absent per year in a population of the employees of a small community hospital: 8, 3, 1, 11, 4 and 7.

a) Find the population mean and population standard deviation.
b) Consider all possible samples of size two i.e., n = 2 which can be drawn with simple random sampling with replacement (SRSWR) and sampling without replacement (SRSWOR)
c) Compute mean for each possible samples for both sampling plans.
d) Draw histogram of distribution of population values and sampling distribution of means. Comment on the shape of the distribution of population values and shape of sampling distribution of means.
e) Find mean of these samples means and verify that population mean is equals to mean of the sample means i.e., $E(\overline{x}) = \mu$ for both sampling plans.
f) Find the standard error of means for each sample and verify following results.
    a) S.E. $(\overline{X}) = \frac{\sigma^2}{\sqrt{n}}$ for SRSWR
    b) S.E. $(\overline{X}) = \frac{\sigma^2}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ for SRSWOR

SOLUTION

Random variable **X** = No. of days an employee of a small community hospital remains absent in a year.

a) To find the population mean and population standard deviation.

Step 1: The data is typed into the worksheet into cells B5 to B10.

Step 2: The population mean (μ) and population standard deviation (σ) are calculated using the following formulas:

Population Mean (μ) = **AVERAGE (B5:B10)**

= 5.6667

Population Standard Deviation (σ) = **STDEV.P(B5:B10)**

= 3.3500

b) To find all possible samples that can be drawn using SRSWOR and SRSWR

Population Size (N) = 6

Sample Size (n) = 2

No. of possible sample using SRSWOR = **COMBIN (6,2)**

= 15

No. of possible sample using SRSWOR = **6^2**

= 36

c) To compute mean for each possible samples for both sampling plans.

All the possible samples drawn using SRSWOR and SRSWR are kept in the table and their mean and standard deviation is calculated as follows:

(X1 and X2 are the samples drawn)

Sample Mean ($\bar{X}$) = **AVERAGE (X1, X2)**

Sample Standard Deviation (S) = **STDEV.S(X1, X2)**

SRSWOR

Sample no	X1	X2	Sample mean	Sample SD
1	8	3	5.5	3.5355
2	8	1	4.5	4.9497
3	8	11	9.5	2.1213
4	8	4	6	2.8284
5	8	7	7.5	0.7071
6	3	1	2	1.4142
7	3	11	7	5.6569
8	3	4	3.5	0.7071
9	3	7	5	2.8284
10	1	11	6	7.0711
11	1	4	2.5	2.1213
12	1	7	4	4.2426
13	11	4	7.5	4.9497
14	11	7	9	2.8284
15	4	7	5.5	2.1213

SRSWR

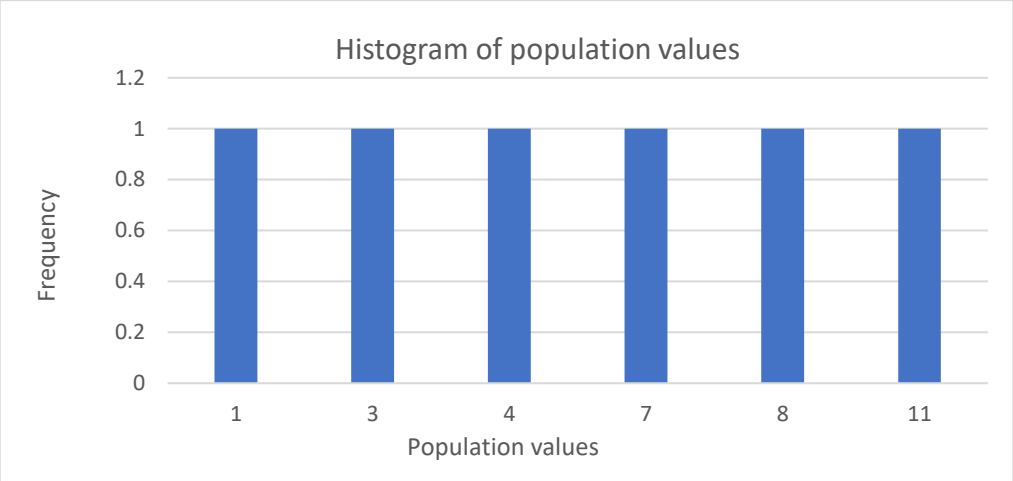
Sample No	X1	X2	Sample Mean	Sample SD
1	8	8	8	0.0000
2	8	3	5.5	3.5355
3	8	1	4.5	4.9497
4	8	11	9.5	2.1213
5	8	4	6	2.8284
6	8	7	7.5	0.7071
7	3	8	5.5	3.5355

8	3	3	3	0.0000
9	3	1	2	1.4142
10	3	11	7	5.6569
11	3	4	3.5	0.7071
12	3	7	5	2.8284
13	1	8	4.5	4.9497
14	1	3	2	1.4142
15	1	1	1	0.0000
16	1	11	6	7.0711
17	1	4	2.5	2.1213
18	1	7	4	4.2426
19	11	8	9.5	2.1213
20	11	3	7	5.6569
21	11	1	6	7.0711
22	11	11	11	0.0000
23	11	4	7.5	4.9497
24	11	7	9	2.8284
25	4	8	6	2.8284
26	4	3	3.5	0.7071
27	4	1	2.5	2.1213
28	4	11	7.5	4.9497
29	4	4	4	0.0000
30	4	7	5.5	2.1213
31	7	8	7.5	0.7071
32	7	3	5	2.8284
33	7	1	4	4.2426
34	7	11	9	2.8284
35	7	4	5.5	2.1213
36	7	7	7	0.0000

d) **To draw histogram of distribution of population values and sampling distribution of means**

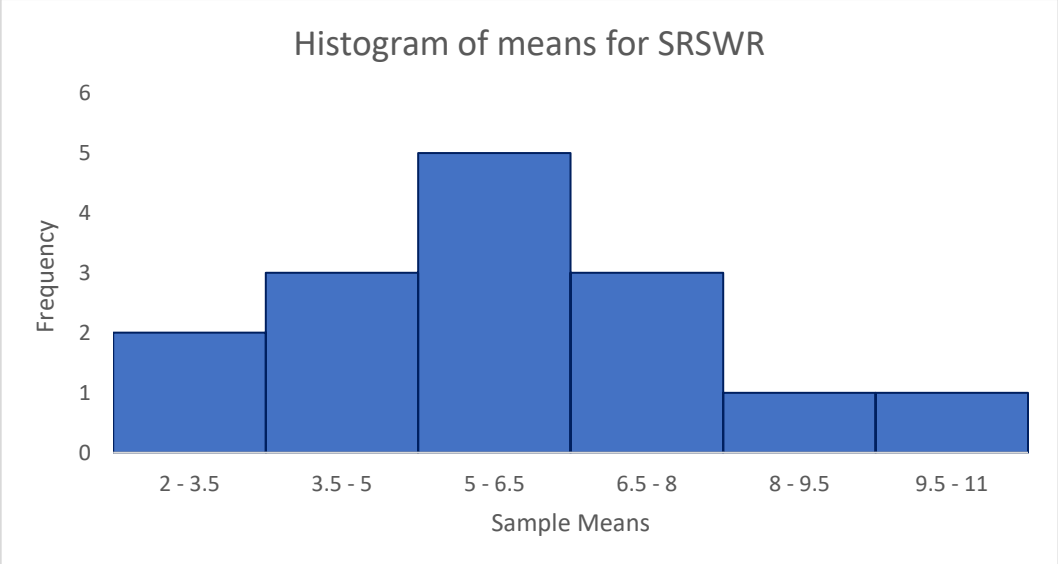
The histogram of the population values can be obtained from the frequency table:

X	F
1	1
3	1
4	1
7	1
8	1
11	1



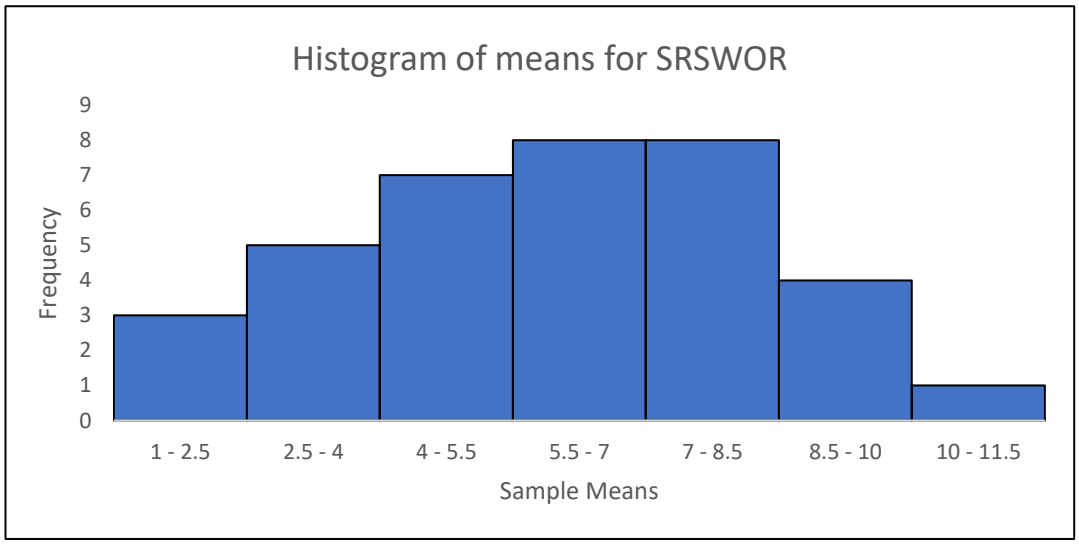
SRSWR

Sample Mean	Bin Limit	Frequency
1 - 2.5	2.4	3
2.5 - 4	3.9	5
4 - 5.5	5.4	7
5.5 - 7	6.9	8
7 - 8.5	8.4	8
8.5 - 10	9.9	4
10 - 11.5	11.4	1



SRSWOR

Sample Mean	Bin Limit	Frequency
1 - 2.5	2.4	3
2.5 - 4	3.9	5
4 - 5.5	5.4	7
5.5 - 7	6.9	8
7 - 8.5	8.4	8
8.5 - 10	9.9	4
10 - 11.5	11.4	1



The shape of values of population has uniform distribution while the shape of the sample means in both SRSWOR and SRSWR sampling plans are approximately normally distributed.

e) To find mean of these samples means.

For SRSWR: E($\bar{X}$) = **AVERAGE (E41:E76) = 5.6667**

For SRSWOR: E($\bar{X}$) = **AVERAGE (E41:E76) = 5.6667**

The expected value of sample mean is equal to the population mean for both sampling plans i.e. SRSWOR and SRSWR. It means that the sample mean is unbiased estimator of population mean.

i.e. E($\bar{X}$) = μ

f) To find the standard error of means

To find the standard error of both sampling plans the total $(\bar{X} - \mu)^2$ of each sample in both plans is calculated.

SRSWOR

Sample No	Sample Mean	$(\bar{X} - \mu)^2$
1	5.5	0.0278
2	4.5	1.3612
3	9.5	14.6942
4	6	0.1111
5	7.5	3.3610
6	2	13.4447
7	7	1.7777
8	3.5	4.6946
9	5	0.4445
10	6	0.1111
11	2.5	10.0280
12	4	2.7779
13	7.5	3.3610
14	9	11.1109
15	5.5	0.0278
Total		67.3333

S.E ($\bar{X}$) = **SQRT(67.333/15) = 2.118699811**

S.E ($\bar{X}$) = **E6/SQRT(2)*SQRT((6-2)/(6-1)) = 2.118699811**

SRSWR

Sample No.	Sample Mean	$(\bar{X} - \mu)^2$
1	8	5.4444
2	5.5	0.0278
3	4.5	1.3611
4	9.5	14.6944
5	6	0.1111
6	7.5	3.3611
7	5.5	0.0278
8	3	7.1111
9	2	13.4444
10	7	1.7778
11	3.5	4.6944
12	5	0.4444
13	4.5	1.3611
14	2	13.4444
15	1	21.7778
16	6	0.1111
17	2.5	10.0278
18	4	2.7778
19	9.5	14.6944
20	7	1.7778
21	6	0.1111
22	11	28.4444
23	7.5	3.3611
24	9	11.1111
25	6	0.1111
26	3.5	4.6944
27	2.5	10.0278
28	7.5	3.3611
29	4	2.7778
30	5.5	0.0278
31	7.5	3.3611
32	5	0.4444
33	4	2.7778
34	9	11.1111
35	5.5	0.0278
36	7	1.7778
Total		202.0000

S.E ($\bar{X}$) = **SQRT(202.0000/36) = 2.368778401**

S.E ($\bar{X}$) = **E6/SQRT(2) = 2.368778401**

Hence for both sampling plans

S.E. $(\overline{X}) = \frac{\sigma^2}{\sqrt{n}}$ for SRSWR

S.E. $(\overline{X}) = \frac{\sigma^2}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ for SRSWOR

is verified.