جامعة الأنبار

1987 1408

UNIVERSITY OF ANBAR

- الكلية : كلية طب العام
- الفرع : طب االسرة والمجتمع
- المرحلة : الثالثة
- أستاذ المادة : د بديعه ثامر يحيى
- اسم المادة باللغة العربية : االحصاء الحياتي
- اسم المادة باللفة الانكليزية : Biostatistics

اسم المحاضرة الخامسة باللغة العربيه:اخذ الغينات

اسم المحاضرة الخامسة باللغة الانكليزية : Sampling

Sampling

Badeaa Thamir Yahyaa
L5

- There are many kinds of samples that may be drawn from a population. Not every kind of sample, however, can be used as a basis for making valid inferences about a population.

- In general, in order to make a valid inference about a population, we need a scientific sample from the population. There are also many kinds of scientific samples that may be drawn from a population.

- The simplest of these is the simple random sample. In this section we define a simple random sample and show you how to draw one from a population

- **Statistical inference**: the procedure by which we reach a conclusion about a population on the basis of the information contained in a sample that has been drawn from that population.

If we use the letter N to designate the size of a finite population and the letter n to designate the size of a sample, we may define a simple random sample as follows:

- **If a sample of size n is drawn from a population of size N in such a way that every possible sample of size n has the same chance of being selected.**

Sample: is a selected part of a population which represents the characteristics of the population . We use samples if the population is too large to be studied.

- Sampling is the process of selecting units (people, organizations) from a population of interest so that by studying the sample we may generalize our results back to the population from which they were chosen. The sampling process comprises several stages: defining the population of concern.

- Steps :

- Definition of target population
- Selection of a sampling frame (list)
- Probability or Nonprobability sampling
- Sampling Unit

Population: is the totality of all members or cases. A population is thus composed of individual units / members / cases / elements. A population does not necessarily consists of human beings. If we want to know characteristics of a population •

- Specifying a **sampling frame**, a set of items or events possible to measure specifying **sampling method** for selecting items or events from the frame determining the sample size & implementing the sampling plan and data collecting.

PROBABILITY sampling :

The process of probability sampling has certain steps:

1. Identify a “sampling frame”: this is a complete list of all cases in the population. For example: if you want to study the students of THIRD STATE OF ANBAR MEDICAL COLLEGE , sampling frame is a complete list of ALL students registered at THE 3RD STATE .

. 2. Decide on a “sample size”: the larger your sample, the less likely error thus, selecting a sample size can be a matter of judgment. However, to have a “representative” sample (i.e. a sample that can represent the whole population)

Characteristics of types of sampling:

Probability sampling :

a. quantitative data

b. selection on random way & give chance to every person to participate in the study .

c. results can be generalized to total population.

Non probability sampling :

a. qualitative data

b. selection on non random way & give chance to every person to participate in the study .

c. results can not be generalized to total population.

Types of Sampling :

Probability	Non Probability
1. Simple . R . S	1. accidental(convenience)
2. Systematic = =	2.Purposive S.
3. Multi stages = =	3. Quota S.
4. Stratified = =	4. snow ball S.
5. Cluster = =	5. Volunteer S.

Probability sampling

1- Simple random sample :

- lottery method.
- random numbers of Table.
- computer method.

Total number of third state are 267

- 001 002 003 004 005 006 007 008 009 010
- 011 012 013 014 015 016 017 018 019 020
- 021 022 023 024 025 026 027 028 029 030
- 031 032 033 034 035 036 037 038 039 040
- 041 042 043 044 045 046 047 048 049 050
- 051 052 053 054 055 056 057 058 059 060
- 061 062 063 064 065 066 067 068 069 070
- 071 072 073 074 075 076 077 078 079 080
- 081 082 083 084 085 086 087 088 089 090
- 091 092 093 094 095 096 097 098 099 100
- 101 102---------- 267

TABLE OF RANDOM SAMPLING NUMBERS

10480	15011	01536	02011	81647
22368	46573	25595	85393	30995
24130	48360	22527	97265	76393
42176	93093	06343	61680	07856
37570	37975	81837	16656	06121
77921	06907	11008	42751	27756

10480↓	15011←	01536←	02011←	81647x	91646x	69179x	14194x	62590x	36207x ←	20969←	99570
22368↓	46573	25595	85393	30995x →	89198x →	27982x →	53402x →	93965x↓	34095	52666x	19174
24130↓	48360	22527	97265	76393x	64809	15179	24830	49340x	32081	30680x	19655
42167x	93093	06243	61680	07856↑	16376	39440	53537	71341x	57004	00849↑	74917
37570x	39975	81837	16656	06121↑	91782	60468	81305	49684x	60672	14110↑	06927
77921x	06907	11008	42751	27756x	53498	18602	70659	90655x	15053	21916↑	81825
99562x	72905	56420	69994	98872x	31016	71194	18738	44013x	48840	63213x	21069
96301x	91977	05463	07972	18876↑	20922	94595	56869	69014x	60045	18425↑	84903
89579x	14342	63661	10281	74553x	18103→	57740x	84378x	25331X	12566	58678X	44947x↓
85475x	36857	53342	53988	53060x	59533	38867	62300	08158↓	17983	16439↑	11458↓
28918x	69578	88231	33276	70997x	79936	56865	05859	90106x	31595	01547↑	85590x
63553x	40961	48235	03427	49626x	69445	18663	72695	52180x	20847	12234↑	90511x
09429↓	93969	52636	92737	88974x	33488	36320	17617	30015x	08272	84115x	27156x
10365↓	61129	87529	85689	48237x	52267	67689	93394	01511↓	26358	85104x	20285↓
07119↓	97336	71048	08178	77233x	13916	47564	81056	97735x	85977	29372x	74461x
51085x	12765	51821	51259	77452x	16308	60756	92144	49442x	53900	70960x	63990x
02368↓	21382	52404	60268	89368x	19885↑	55322	44819	01188↓	65255	64835x	44919x
01011↓	54092	33362	94904	31273x	04146↑	18594	29852	71585x	85030	51132x	01915↓
52162x	53916	46369	58586	23216↑	14513↑	83149	98736	23495x	64350	94738x	17752↓
07056↓	97628	33787	09998	42698x	06691↑	76988	13602	51851x	46104	88916x	19509↓
48663x	91245	85828	14346	09172↑	30168x↑	90229	04734	59193x	22178	30421x	61611x
54164x	58492	22421	74103	47070x	25306x↑	76468	26384	58151x	06646	21524↑	15227↓
32639x	32363	05597	24200	13363↑	38005x↑	94342	28728	35806x	06912	17012↑	64161x
29334x	27001	87637	87308	58731x	00256↑	05834	15398	46557x ←	41135	10367↑	07684↓
02488→	33062x	28834x	08751→	19731↑	92420	60952	61280	50001	67658	32586x↑	86679x

Sample :

181, 253, 125, 114, 202, 019, 177, 195, 152, 076 , 103, 170, 215,122, 015, 184, 219, 141, 008, 209, 020, 015, 150, 104 , 223, 241, 094, 103, 071, 023, 070, 024, 087, 197,133,091, 232, 188, 061,078, 081, 015, 011,153, 058, 066 , 145, 041,198 , 163 ,

2- Systemic random sample :

select numbers of your sampling frame at regular intervals

- Number each case/member of your sampling frame
- Select the first number randomly (close your eyes and point!)
- Calculate the sampling fraction
- Select the cases/members using the sampling fraction

$$\text{SAMPLE FRACTION} = \frac{\text{TOTAL POPULATION}}{\text{SAMPLE SIZE}} = \frac{267}{25} = 11$$

Total number of third state are 267

- 001 002 003 004 005 006 007 008 009 010
- 011 012 013 014 015 016 017 018 019 020
- 021 022 023 024 025 026 027 028 029 030
- 031 032 033 034 035 036 037 038 039 040
- 041 042 043 044 045 046 047 048 049 050
- 051 052 053 054 055 056 057 058 059 060
- 061 062 063 064 065 066 067 068 069 070
- 071 072 073 074 075 076 077 078 079 080
- 081 082 083 084 085 086 087 088 089 090
- 091 092 093 094 095 096 097 098 099 100
- 101 102---------- 267

3- STRATIFIED RANDOM SAMPLING :

- • Divide your sampling frame into “subsets” or “strata” according to different attributes.
- • Choose stratification variable(s). For example: 3RD state of anbar medical college students are classified by the year of enrolment. Or, classified by gender (male or female)
- • Divide your sampling frame into subsets according to the variables
- • Do a simple random or systematic sampling of each subset.
- • Examples :- age , sex , social class, marital status

- Age group
- > 20
- 20 – 24
- 25+
-
- Sex :
- Male =
- Female =
- Residence:

Rural or Urban

4- Cluster random sampling

- select by using simple random S or systematic
- the clusters or groups of elements such as
- classes , schools ,districts, streets, places ,houses and each cluster can select a subsample by using simple random S or systemic

5- Multi Stage random sampling

- This method involves drawing samples .
- draw the sample from target population ,
- continue drawing till reach required sample by using simple R.S or systematic R.S that get the primary sample then continue to draw another sample until get the required sample .
- **Stage 1 = 1000 from 10,000**
- **Stage 2 = 500 from 1000** **Stage 3 = 100 from 500**

Non Probability sampling

1- Accidental (convenance) sampling:

- called haphazard ,
- that the researcher interviews the respondents for the study who comes in contact
- accidentally during the research time .

2- Purposive sampling

- the researcher chooses respondents who in the opinion of the researcher thought to be relevant to the subject under the study

3- Quota sampling

- • Is non-random
- • Used for interview surveys, especially if there is a large population
- • Steps:
- 1. Divide the population into specific groups
- 2. Calculate a quota for each group
- 3. Interview the cases in each quota group
- 4. Combine the results

4- Snow ball sampling

- the researcher begins with few of respondents who are available
- Then ask the respondent to recommend other persons who meets the criteria of the research & who are willing to participate in the study.

5- Volunteer sampling

- base on the acceptance

Confidence Interval & Confidence Level

- The confidence interval is the number of percentage points above or below the proportion that find in the study that the true proportion should be within.
- For example, if confidence interval is 3.5 % and study reveals a proportion of 57 %, the true proportion is likely between 53.5 %and 60.5%

- **The general expression for confidence interval is:**
- Confidence interval = point estimate ± (confidence multiplier x SE)
- For 90% 95%, 99% confidence level the C. multiplier is 1.64 ,1.96, 2.58 respectively.

Confidence limits:

- Are the lower and upper boundaries of a confidence interval
- **Confidence level: Is the probability value attached to a given confidence interval. It can be expressed as a percentage (95%, 99%). Confidence level** .

- **Factors that Affect Confidence Intervals; -**
  There are three factors that determine the size of the confidence interval for a given confidence level:

1- Sample size, population size &Percentage

Sample Size; -The larger sample size, the more sure can be that their answers truly reflect the population. This indicates that for a given confidence level, the larger sample size, the smaller confidence interval .

2- Percentage - Accuracy also depends on the percentage of sample that picks a particular answer. If 99% of sample said "Yes" and 1% said "No," the chances of error are remote, irrespective of sample size. However, if the percentages are 51% and 49% the chances of error are much greater.

-

- 3-Population Size :How many people are there in the group sample represents? This may be the number of people in a city ,Using 10% OF Prevalence of certain disease or problem.

The sample size :

- The appropriate sample size for a population is determined by

(1) the estimated prevalence of the variable of interest , when prevalence is high in community can choose small sample size.
(2) the desired level of confidence and the acceptable margin of error.

(3) design of study ; for cohort & clinical trail need large sample size while in case control can choose small sample size.

$$n = \frac{CF \text{ of } 95\% (1.96) \times P(P-1)}{m \text{ (Standard value of } 0.05)^2}$$

n= required sample size

CF = critical factor at level 95% (standard value of 1.96)

P = estimated prevalence of disease in the project area

m = margin of error at 5% (standard value of 0.05)

Example In a survey has been estimated that roughly 30% (0.3) of the children suffer from chronic malnutrition. This figure has been taken from national statistic s on malnutrition in rural areas. Use of the standard values listed above to find sample size ;

$$n = \frac{(1.96) \times 0.3(1 - 0.3)}{(0.05)^2}$$