Introduction &overview of the course

Professor Dr Abubakir M. Saleh

Biostatistics NUR 304

1st week

Outline

- Definitions: Statistics & Biostatistics
- Aim and objectives of the course
- Course Syllabus
- Lecture and Practical Schedules
- Course outline
- Examinations

Objectives

- To describe the content of the course to the students.

- To familiarize students with assessment methods of the students in the course.

Statistics

Statistics is the art and science of data. It deals with:
- Collecting data
- Organizing data
- Describing data
- Summarizing and presenting data
- Analyzing data
- Interpreting results
- Reaching decisions or discovering new knowledge

Biostatistics

- Biostatistics is application of statistical methods to a wide range of topics in biology including medicine and health sciences

Aim of the course

- To introduce basic descriptive and inferential statistical concepts and provide an understanding of the relevance and importance of statistics to health research.

Biostatistics

- The course will particularly emphasize the practical applications of basic statistical concepts

- The mathematical formulae will be kept to a minimum.

- The theory and the methods of analysis will be extensively illustrated with examples relating to a wide variety of relevant fields.

Course objectives

By the end of this course the students should be able to:

- Describe the role of statistical methods in health research;

- Present results effectively by making appropriate displays, summaries, and tables of data;

- Appreciate the problem of sampling variation and the role of statistical methods in quantifying it

- Select an appropriate statistical method for the analysis of simple data sets;

Objectives-Cont.

- Correctly interpret and critically appraise the results of statistical analyses reported in the health literature;

- Present and discuss the findings from statistical analyses in a clear, concise and logical manner

Syllabus

- Course length: 15 weeks (including mid and end course examination)

- Course language: English

- Assessment (grading):

    - Mid-course exam : 30%
    - Homework: 10%
    - Participation : 10%
    - Quiz: 10%
    - End of course exam: 40%

- **Course Reading List:**

    Comprehensive notes are provided

    Extra reading

- **Essential Medical Statistics**, by Betty Kirkwood & Jonathan Sterne
  (Published by Blackwell)

    Statistics Without Tears, a Primer for Non-mathematicians, by Derek Rowntree
    (Published by Penguin)

- All the statistical tables you need for this course are provided.

Lecture Schedule

Week No.	Topic
1	Introduction & overview of the course.
2	Variable& data presentation
3	Population and sample
4	Descriptive measures
5	Confidence intervals
6	Mid-course examination
7	Hypothesis testing
8	Comparison of two means (t-test)
9	Comparison of two proportions(Chi-square)
10	Sample size calculation

Lecture Schedule

Week No.	Topic
11	Using computer software in data analysis (SPSS).
12	Using computer software in data analysis(SPSS).
13	Using computer software in data analysis(SPSS).
14	Review of the course.
15	End-course examination.

Course outline:

| Subject |
| --- |
| **Introduction**
- Course description
- Course objectives
- Course syllabus
-Students evaluation (examination) |
| **Population and sample**
- Study population
- Sample
- Reasons for sampling
- Probability (random) sampling
- Probability sampling methods
•   Simple random sampling
•   Systematic sampling
•   Stratified sampling
•   Cluster sampling
- Non-probability sampling |

Types of variables
- Qualitative (categorical) variables
- Nominal
- Ordinal
- Quantitative variables
- Continuous
- Discrete
- Dependent (outcome) variable
- Independent (exposure/risk factor) variable

Data presentation
- Types of data
- Frequency distributions
- Graphical methods
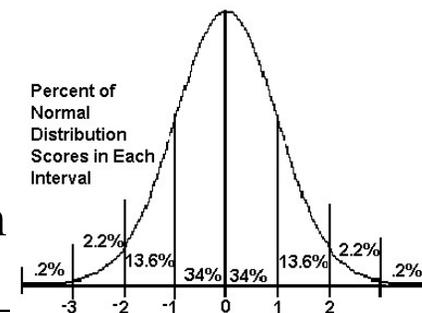- Bar graph
- Histogram
- Pie chart
- Line graph

Descriptive measures
- Measure of central tendency
- Mean/Median/Mode
- Measure of dispersion
- Standard deviation/Range
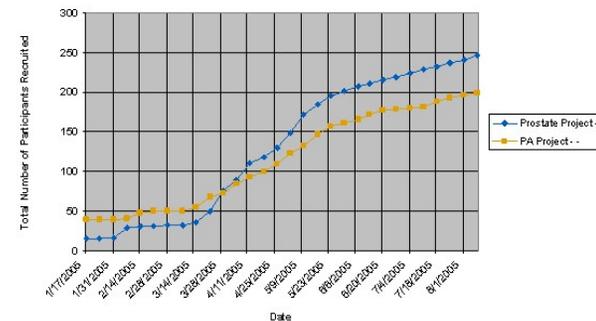
Confidence intervals

- Statistical inference
- Estimation
- Confidence interval for a population mean
- Confidence interval for a population proportion
- Standard error and standard deviation

Percent of Normal Distribution Scores in Each Interval

.2%   2.2%   13.6%   34%   34%   13.6%   2.2%   .2%
-3   -2   -1   0   1   2

Hypothesis testing

- Null hypothesis
- Level of significance
- Test of significance

Comparison of two means

- The t-test for difference between two means (independent samples)
- Paired t-test: difference between paired samples

Comparison of two proportions
- Contingency tables
- Chi-square test: Relationship between two categorical variables

Sample size calculation
- One sample
  - Mean
  - Proportion
- Two samples
  - Difference between two means
  - Difference between two proportions

 - Application of computer software for data analysis (SPSS).
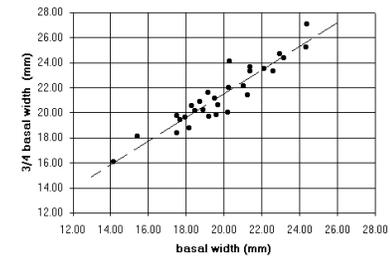
Examination

The examinations that will be held during the course include:

- Quizzes - during the practical sessions
- Mid-course exam – Theoretical including basic statistical exercises
- End of course exam – Theoretical including basic statistical exercises

Examinations(Cont.)

- Theoretical exams will include different types of question.

- Samples of the expected questions and their answers:

-

Q1- Mark each of the following by (T) if true or (F) if false, in front of each of them.

- 1. In simple random sampling, each unit has an equal and known probability of being selected.

- 2. If we study the association between eating salty food and having high blood pressure, then high blood pressure is the independent variable.

- 3. A histogram can be used to display frequency distribution but not relative frequency distribution.

- 4. The best way to visualize correlation is by using scatter-plot.

- 5. When the confidence intervals of two values overlap, this means there is evidence of having a real difference between the two values.

Q2- Encircle the most appropriate answer in the following statements:

Examples of multiple choice questions, with notes and solutions

1. A study was carried out to investigate low birth weight in a large city. Pregnant women were recruited early in pregnancy and followed up until the birth. Each baby's birth weight was classified as low if it was 2.5 kg or less and normal otherwise. This variable is:

- a   Continuous
- b   Ordered categorical
- **c**   Binary
- d   Skewed

Q3. Statistical questions

The weights of 120 successive babies born alive to residents of a Health District were recorded to the nearest 100g. The results were as follows:

Wt (cgm)	Number		Wt (cgm)	Number		Wt (cgm)	Number
19	1		29	2		39	6
20	0		30	7		40	1
21	0		31	4		41	5
22	1		32	12		42	3
23	0		33	11		43	1
24	1		34	15		44	0
25	2		35	5		45	0
26	6		36	5		46	1
27	2		37	13		47	1
28	11		38	4			

Q3(Cont.)

(a) Draw up the cumulative percentage frequency table, and find the quartiles and the median.

(b) Illustrate the distribution using (i) a histogram, and (ii) a box plot. Were these birth weights symmetrically distributed?

(c) Compute the mean and standard deviation of the birth weights, and compare the mean and the median. Also compute the mean ± SD, and the mean ± 2SD, and see how these relate to the percentiles of the distribution.

References

- Essential Medical Statistics, by Betty Kirkwood & Jonathan Sterne
  (Published by Blackwell)

  Statistics Without Tears, a Primer for Non-mathematicians, by Derek Rowntree
  (Published by Penguin)