Basic Terms in Statistics

Nguyen Thi Tu Van - Nguyen Quang Vinh

August 16, 2016

Statistics

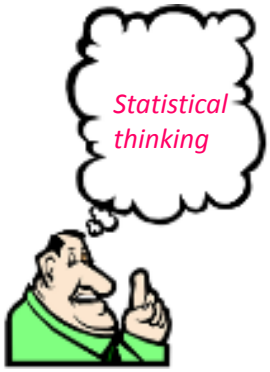
Descriptive *statistics*

- Organize data
- Summarize data

Inferential *statistics*

(drawing of inferences from sample → population)

- Estimation
- Hypothesis testing - reaching a decision:
  - *Parametric tests*
  - *Non-parametric tests (distribution-free)*
- Modeling, Predicting



- Understand medical articles & Use information of published medical evidence.

  → → → Have better choice / given situation.

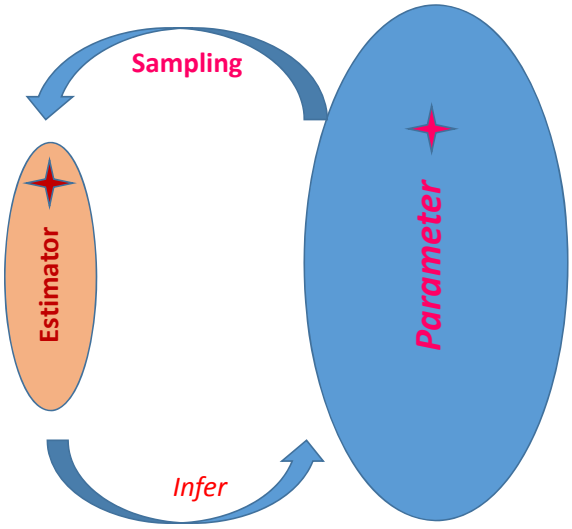
- Be able to communicate with a statistical consultant.

Estimation

Estimation – A point estimate

- Population - Sample
- A parameter may be estimated by more than one estimator:

  Example:
  – *Sample mean → estimate population mean*
  – *Sample median → estimate population mean*



Confidence interval for a population mean

In general, an interval estimate is obtained by the formula

estimator ± (reliability coefficient) x (standard error)

In particular: $\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$ or $\bar{x} \pm t_{\alpha/2}S.E.$

How to interpret the interval given by this expression

- In repeated sampling 100(1- α)% of all intervals of the form will in the long run include the population mean, μ.
- The quantity 1 - α, is called the *confidence coefficient*, &

  The interval $\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$, is called the *confidence interval* for μ.

The practical interpretation

- *We are 100(1 - α)% confident that the single computed interval*

$$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$$

  contains the population mean, μ

- *E = margin error = maximum error = practical / clinical acceptable error:*

$$E = z_{\alpha/2}\sigma_{\bar{x}} = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Hypothesis Testing

Reaching a decision concerning a population

by examining a sample from that population.

Hypothesis

Two types of hypotheses:

(1) *Research Hypotheses:*

- The conjecture or supposition.
- The results of years of observation.
- Research H. leads directly to Statistical H.

(2) *Statistical Hypotheses:*

Hypotheses are stated in such a way that they may be evaluated by appropriate statistical techniques.

Statistical Hypotheses

- Hypothesis to be tested = Null H = H_o = H of no difference.
- *If* H_o *is not rejected*, we will say that the data on which the test is based do not provide sufficient evidence to cause rejection.
- *If the testing process leads to rejection*, we will say that the data at hand are not compatible with the H_o, but are supportive of some other hypothesis & may be designated by H_A.
- H_A: a contradiction statement of H_o (complementary).

Test Statistic

Decision maker: reject or not to reject the H_o depends on the *magnitude* of the test statistic

Test Statistic → p value

Conditions under which type I & type II errors may be committed (the four possibilities)		Actual Situation (Truth in the population)	
		H_o false	**H_o true**
The results in the study sample → Conclusion:	**Reject H_o**	Correct decision	*Type I error*
	Fail to reject H_o	*Type II error*	Correct decision

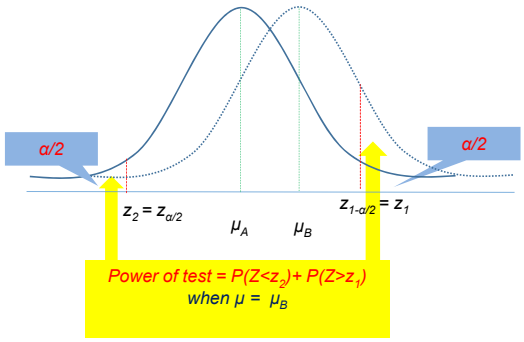
One-sided vs. Two-sided Hypothesis Test

Ho: $\mu 1 - \mu 2 = 0$, HA: $\mu 1 - \mu 2 \neq 0$

Ho: $\mu 1 - \mu 2 \geq 0$, HA: $\mu 1 - \mu 2 < 0$

Ho: $\mu 1 - \mu 2 \leq 0$, HA: $\mu 1 - \mu 2 > 0$

The Power of a Statistical Test



- We can not know the power of a test until the study is complete.
- The power of test should be considered when an opportunity to reject the H_o correctly was lost.