

Lecture 3

L3.1. Hypotheses about Mean and Proportion of a Population

- a. Assume, that an average survival time for glioblastoma patients (early state, age<50) is 18 months. You developed a new treatment which should increase the survival time and performed a pilot clinical trial (10 patients) in order to determine the positive effect. As a result, you observed the average survival of 20 months. Standard deviation was equal to 5 months.

pt, mean, sd, sqrt, t.test

L2.2. Hypotheses about Means and Proportions of 2 Populations

- a. Based on complete “*mice*” data set, means of which parameters (except weight) are significantly different for male and female population?

- b. Look for dataset “*cancer*” at <http://edu.sablab.net/data/txt/cancer.txt>

This data contains results of survey aimed at survival and life quality of patients with advanced lung cancer, performed by the North Central Cancer Treatment Group (Loprinzi CL et al, J. of Clinical Oncology. 12(3):601-7, 1994). Look for the survival time (column time, given in days) and compare the survival for male and female populations. Apply both parametric and non-parametric testing.

- c. See “*leukemia*” data at <http://edu.sablab.net/biostat2/leukemia.txt>

This dataset contains information about two groups of patients who died of acute myelogenous leukemia. Patients were classified into the two groups according to the presence or absence of a morphologic characteristic of white cells. Patients termed AG positive were identified by the presence of Auer rods and/or significant granulation of the leukemic cells in the bone marrow at diagnosis. For AG-negative patients, these factors were absent. Leukemia is a cancer characterized by an overproliferation of white blood cells; the higher the white blood count (WBC), the more severe the disease. Investigate the data using parametric and non-parametric tests.

t.test, wilcox.test