Hypothesis testing

Introduction

- H_0 null hypothesis
- H_A alternative hypothesis
- AH_0 acceptation of the null hypothesis
- RH_0 rejection of the null hypothesis
- α $P(RH_0|H_0)$ probability to reject the null hypothesis when it is true
- β $P(AH_0|H_A)$ probability to accept the null hypothesis when it is false
- 1-β $P(RH_0|H_A)$ Power of the test (also called rejection power)

	H0	HA
AH0	Correct acception	Type II error β error
RH0	Type I error α error	Correct rejection

Examples of tests

- **Test of significance**: assess the significance of the difference between expectation and observation.
  - Example : a sequence contains 26 occurrences of the word GATAAG, whereas 3 occurrences would be expected by chance. Is this word over-represented ? How significant is the over-representation ?
- **Test of conformity** : compare a sample with a theoretical distribution, on the basis of some parameters (mean, variance).
  - Two-tailed test on the mean $H_0: m = 0;\ H_A: m \neq 0$
  - One-tailed test on the mean $H_0: m \geq 5;\ H_A: m < 5$
  - Two-tailed test on the variance $H_0: \sigma^2 = 1;\ H_A: \sigma^2 \neq 0$
- **Test of homogeneity (or of equality)**: compare two populations, on the basis of some statistics calculated on their samples.
  - Two-tailed test on the means $H_0: m_1 = m_2;\ H_A: m_1 \neq m_2$
  - One-tailed test on the means $H_0: m_1 \geq m_2;\ H_A: m_1 < m_2$
  - Two-tailed test on the variance $H_0: \sigma_1^2 = \sigma_2^2;\ H_A: \sigma_1^2 \neq \sigma_2^2$
- **Goodness of fit**: test if a given sample can be considered as extracted from a population, on the basis of its whole distribution.
  - Example: assess whether the log-ratios from a control DNA chip (no treatment) follow a normal distribution.