Probability and Statistics

6 – Hypothesis Testing

Stefan Heiss

Technische Hochschule Ostwestfalen-Lippe
Dep. of Electrical Engineering and Computer Science

January 9, 2024

Null Hypothesis, Tests

The objective of statistical hypothesis tests, is to check whether a hypothesis is consistent with data from a random sample.

Notation (6.1)

A hypothesis about a parameter θ of a random distribution is called *null hypothesis* H_0.

If H_0 completely determines the distribution (e.g. $H_0 : \theta = 0$), H_0 is called a *simple hypothesis*, otherwise (e.g. $H_0 : \theta > 0$), H_0 is called a *composite hypothesis*.

A *test of a null hypothesis* is always based on a random sample $X_1, \ldots, X_n$ and the (a priori) definition of a so-called *critical region* $C \subseteq \mathbb{R}^n$. A particular test consists of the observation of n values $x_1, \ldots, x_n$ from the random variables $X_1, \ldots, X_n$ and:

- H_0 is *accepted* if $(x_1, \ldots, x_n) \notin C$.
- H_0 is *rejected* if $(x_1, \ldots, x_n) \in C$.

Example (Ross, Chapter 8, Exc. 4)

In a certain chemical process, it is very important that a particular solution that is to be used as a reactant have a pH of exactly 8.20. A method for determining pH that is available for solutions of this type is known to give measurements that are normally distributed with a mean equal to the actual pH and with a standard deviation of 0.02. Suppose 10 independent measurements yielded the following pH values:

$$8.18, \quad 8.17, \quad 8.16, \quad 8.15, \quad 8.17, \quad 8.21, \quad 8.22, \quad 8.16, \quad 8.19, \quad 8.18$$

$X_i \sim \mathcal{N}(\mu, \sigma)$, $\sigma = 0.02$

$\bar{x} = 8.179$

$H_0: \mu = 8.20$

$\alpha = 0.05$

$\Phi^{-1}(1 - \frac{\alpha}{2}) \approx 1.96$

$\mu \pm \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}(1 - \frac{\alpha}{2})$

$\mu \pm \frac{0.02}{\sqrt{10}} \cdot 1.96 \approx 0.012$ with confidence 95%

Error Types, Significance Level

Notation (6.1 contd.)

In a test, two different types of errors may occur:

- Errors of *type I*: H_0 is rejected (i. e. $(x_1, \ldots, x_n) \in C$), although H_0 is correct.
- Errors of *type II*: H_0 is accepted (i. e. $(x_1, \ldots, x_n) \notin C$), although H_0 is false.

A test has *significance level* α if the probability of false rejection of H_0 (i.e. of a type I error) is at most α.

Tests concerning the mean of a normal distribution

Let $X_1, \dots, X_n$ be a random sample with:

$$X_i \sim \mathcal{N}(\mu, \sigma)$$

(1) σ known $\rightarrow$ $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma}{\sqrt{n}})$ cdf: Φ

(2) σ unknown $\rightarrow$ $\bar{X}, S$ cdf: $F_{t_{n-1}}$

Two-sided tests concerning μ

$$H_0 : \mu = \mu_0$$

As the sample mean $\overline{X}$ is a suitable estimator for μ, the hypothesis should be rejected if the mean value $\overline{x}$ of a set of sampled data $x_1, \ldots, x_n$ differs significantly from μ_0. A critical region is therefore given by

$$C := \{(x_1, \ldots, x_n) \mid |\overline{x} - \mu_0| > c\}$$

for a suitable constant c.

Given a significance level α, c has to be determined, such that:

$$\Pr_{H_0}(|\overline{X} - \mu_0| > c) := \Pr(|\overline{X} - \mu_0| > c \mid \mu = \mu_0) = \alpha$$

conditional probability

$X_i \sim \mathcal{N}(\mu, \sigma)$: Two-sided tests concerning μ

(1)

$$\Pr\left(|\overline{X} - \mu_0| \underset{<}{\cancel{>}} c \mid \mu = \mu_0\right) = \underset{1-\alpha}{\alpha}$$

$$\Pr\left(-c < \overline{X} - \mu < c\right) = 1 - \alpha$$

$$\Pr\left(-\frac{c}{\sigma/\sqrt{n}} < \underset{\sim \mathcal{N}(0,1)}{\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}} < \frac{c}{\sigma/\sqrt{n}}\right) = 1 - \alpha \iff \frac{c}{\sigma/\sqrt{n}} = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

$$\iff c = \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

---

(2)

$$\Pr\left(-\frac{c}{S/\sqrt{n}} < \overset{\sim t_{n-1}}{\frac{\overline{X} - \mu}{S/\sqrt{n}}} < \frac{c}{S/\sqrt{n}}\right) = 1 - \alpha \iff c = \frac{S}{\sqrt{n}} \cdot F^{-1}_{t_{n-1}}\left(1 - \frac{\alpha}{2}\right)$$

$X_i \sim \mathcal{N}(\mu, \sigma)$: Two-sided tests concerning μ

$\Phi(v) = 1 - \frac{\alpha}{2} \Leftrightarrow \alpha = 2\cdot(1 - \Phi(v))$

- H_0 is rejected if:

$$|\overline{x} - \mu_0| > c \iff \overbrace{\frac{\sqrt{n}}{\sigma} \cdot |\overline{x} - \mu_0|}^{v} \overset{\Updownarrow}{\overset{=}{>}} \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

$$\frac{\sqrt{n}}{S} \cdot |\overline{x} - \mu_0| > F^{-1}_{t_{n-1}}\left(1 - \frac{\alpha}{2}\right)$$

- H_0 is accepted if:

$$|\overline{x} - \mu_0| < c \iff \frac{\sqrt{n}}{\sigma} \cdot |\overline{x} - \mu_0| < \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

$$\frac{\sqrt{n}}{S} \cdot |\overline{x} - \mu_0| < F^{-1}_{t_{n-1}}\left(1 - \frac{\alpha}{2}\right)$$

$X_i \sim \mathcal{N}(\mu, \sigma)$: Two-sided tests concerning μ

Starting with a set of sampled data,

$$v \quad := \quad \frac{\sqrt{n}}{\sigma} \cdot |\overline{x} - \mu_0|$$

can be calculated and the maximal value for α can be determined, such that H_0 will be accepted on the basis of the data sampled. This value

$$\alpha_{\overline{x}} \quad := \quad 2(1 - \Phi(v)) \qquad \checkmark$$

is called the *p-value* of the sample.