# Probability, and Mathematical Statistics

Miskolc, 2025.

Dr. Tamás Glavosits

Lecture 12

# Statistics III., Hypothesis Testing

# 1. Hypothesis Testing

**Definition**

Let $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ be a sample with joint distribution function $\mathbb{F}(\cdot, \vartheta)$, where $\vartheta \in \Theta$ is the unknown parameter. We partition the parameter space $\Theta$ into two parts $\Theta_0$ and $\Theta_1$ such that
$\Theta_0 \cup \Theta_1 = \Theta$.
**Null hypothesis** $H_0$**:** $\vartheta \in \Theta_0$;
**Alternative hypothesis** $H_1$**:** $\vartheta \in \Theta_1$.
If $\vartheta \in \Theta_0 \cap \Theta_1$, then the decision is made randomly.

**Definition**

Keeping the previous notation
**Simple null hypothesis** $H_0$, if $\Theta_0 = \{\vartheta_0\}$, i.e. $\vartheta = \vartheta_0$;
**Alternative hypothesis** $H_1$, if $\Theta_1 = \Theta \setminus \{\vartheta_0\}$, i.e. $\vartheta \neq \vartheta_0$.

# 2. Decision, Significance Level, Type I and Type II Error

# Critical Region

The sample space
$\chi = \{(x_1, x_2, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \text{ are sample realizations}\} \subseteq \mathbb{R}^N$
is divided into two disjoint sets $\chi = C \cup \overline{C}$, where the set $C$ is said to be the **critical region**.

- If $(x_1, x_2, \ldots, x_n) \in C$, then we reject $H_0$,
- If $(x_1, x_2, \ldots, x_n) \in \overline{C}$, then we accept $H_0$.

### Definition

The significance level of the test is $\alpha$, if

$$\mathbb{P}(\xi) \in C \mid H_0) \leqslant \alpha.$$

If we make a decision we can make errors.

- **Type I error** is the probability that we reject $H_0$ although it is true, that is, the Type I error is the probability

$$\mathbb{P}(\xi) \in C \mid H_0).$$

- **Type II error** is the probability that we accept $H_0$ although it is false, that is, the Type II error is the probability

$$\mathbb{P}(\xi) \in \overline{C} \mid H_1).$$

# Power Function, Consistent Power Function

## Definition

The function $W : \Theta \to [0,1]$ ($W = W(C, \vartheta, n)$) defined by

$$W(\vartheta) = \mathbb{P}(\{\xi \in C \mid \vartheta\}) \qquad (\vartheta \in \Theta)$$

is called the **power function corresponding to the critical region** $C$, where $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ is the independent sample.

## Definition

The power function $W$ corresponding to the critical region $C$ is called **consistent** if

$$\lim_{n \to \infty} W(\vartheta) = 1 \qquad (\vartheta \in \Theta_1).$$

# Unbiased Test at Level $\alpha$, More Powerful Tests

## Definition

A test is called **unbiased at level** $\alpha$ if

$$\begin{aligned} \mathbb{P}(\xi \in C \mid \vartheta) &\leqslant \alpha \qquad (\vartheta \in \Theta_0), \\ \mathbb{P}(\xi \in C \mid \vartheta) &\geqslant \alpha \qquad (\vartheta \in \Theta_1), \end{aligned}$$

## Definition

Among the level $\alpha$ tests defined by the critical regions $C_1$ and $C_2$, we say that $C_1$ **is more powerful than** $C_2$ if

$$\mathbb{P}(\xi \in C_1 \mid \vartheta) \geqslant \mathbb{P}(\xi \in C_2 \mid \vartheta) \qquad (\vartheta \in \Theta),$$

that is, for every $\vartheta \in \Theta$, the power function corresponding to $C_1$ is greater than or equal to the power function corresponding to $C_2$.

# 2. Classical Tests for the Expected Value

# Classical Tests for the Expected Value

- one-sample u-test;
- one-sample t-test;
- two-sample u-test;
- two-sample t-test (For this, the two unknown theoretical variances must be equal, which is checked using the F-test.)

# 2.1. One-sample u-test

# One-sample u-test

Let $\xi_1, \dots, \xi_n \sim \mathcal{N}(\mu, \sigma^2)$ be an independent sample, where $\sigma^2$ is known, and $\mu_0 \in \mathbb{R}$ is a given number:
**The simple null and alternative hypothesis:**
$H_0 : \mu = \mu_0$;
$H_1 : \mu \neq \mu_0$.
**The decision:**

1. We assume that $\mu = \mu_0$, we have that

$$\frac{\overline{\xi} - \mu_0}{\sigma}\sqrt{n} \sim \mathcal{N}(0, 1)$$

2. As in the construction of confidence intervals, we need to determine the value of $u$ such that

$$\mathbb{P}\left(-u < \frac{\overline{\xi} - \mu}{\sigma}\sqrt{n} < u, \mid \mu = \mu_0\right) = 1 - \alpha$$

that is, the appropriate value of $u$ can be calculated from

$$\Phi(u) = 1 - \frac{\alpha}{2}.$$

3. Based on the sample, we calculate

$$u_{\text{obs}} := \frac{\overline{\xi} - \mu_0}{\sigma}\sqrt{n}.$$

4. Decision: If $|u_{\text{obs}}| < u$ (obs = observed), then we accept the null hypothesis $H_0$; otherwise, we reject it $H_0$.

# 2.2. One-sample t-test

# One-sample t-test

Let $\xi_1, \ldots, \xi_n \sim \mathcal{N}(\mu, \sigma^2)$ be an independent sample, where $\sigma^2$ is unknown, and $\mu_0 \in \mathbb{R}$ is a given number.
The simple null and alternative hypotheses:
$H_0 : \mu = \mu_0$;
$H_1 : \mu \neq \mu_0$.
**Decision:**

1. We assume that $\mu = \mu_0$ and have

$$\frac{\overline{\xi} - \mu_0}{s_n^*}\sqrt{n} \sim t_{n-1}$$

# One-sample t-test

2. In the same way as constructing confidence intervals, we need to determine the $t$ value for which
$$\mathbb{P}\left(-t\frac{\overline{\xi}-\mu_0}{s_n^*}\sqrt{n} < t, \mid \mu = \mu_0\right) = 1-\alpha,$$
that is, if $\mathbb{F}$ denotes the cumulative distribution function of a $t_{n-1}$ random variable, then we look for the $t$ value satisfying
$$\mathbb{F}(t) = 1 - \frac{\alpha}{2} \qquad \left(\text{or } G(t) = \frac{\alpha}{2}\right).$$
3. Based on the sample, we calculate
$$t_{\text{obs}} := \frac{\overline{\xi}-\mu_0}{s_n^*}\sqrt{n}$$
(obs = observed).
4. The decision is made as follows: if $|t_{\text{obs}}| < t$, we accept the null hypothesis $H_0$; otherwise, we reject it.

# 2.3 Two-sample u-test

# Two-sample u-test

Given $\xi_1, \xi_2, \ldots, \xi_m \sim \mathcal{N}(\mu_1, \sigma^2)$ and $\eta_1, \ldots, \eta_n \sim \mathcal{N}(\mu_2, \sigma^2)$ independent samples. (The $\xi_i$ and $\eta_j$ are also independent from each other.)

The simple null and alternative hypotheses:

$H_0 : \mu_1 = \mu_2$;

$H_1 : \mu_1 \neq \mu_2$.

**Decision:**

1. Assuming $\mu = \mu_0$, we have

$$\frac{\overline{\xi} - \overline{\eta}}{\sigma}\sqrt{\frac{mn}{m+n}} \sim \mathcal{N}(0, 1).$$

# Two-sample u-test

2. We find the $u$ value for which

$$\mathbb{P}\left(-u < \frac{\overline{\xi} - \overline{\eta}}{\sigma}\sqrt{\frac{mn}{m+n}} < u \,\middle|\, \mu = \mu_0\right) = 1 - \alpha,$$

that is,

$$\Phi(u) = 1 - \frac{\alpha}{2}.$$

3. Based on the sample, we calculate

$$u_{\text{obs}} := \frac{\overline{\xi} - \overline{\eta}}{\sigma}\sqrt{\frac{mn}{m+n}}$$

4. The decision is made as follows: if $|u_{\text{obs}}| < u$, we accept the null hypothesis $H_0$, otherwise we reject it.

# 2.4 Two-sample t-test

# Two-sample t-test

Given $\xi_1, \ldots, \xi_m \sim \mathcal{N}(\mu_1, \sigma^2)$ and $\eta_1, \ldots, \eta_n \sim \mathcal{N}(\mu_2, \sigma^2)$ independent samples (where $\xi_i$ and $\eta_j$ are also independent of each other). (It is important that the samples come from distributions with a common variance.)

The simple null and alternative hypotheses:

$H_0 : \mu_1 = \mu_2;$

$H_1 : \mu_1 \neq \mu_2.$

**Decision:**

1. If $\mu_1 = \mu_2$, then

$$\frac{\bar{\xi} - \bar{\eta}}{\sqrt{\frac{(m-1)s_m^{*} + (n-1)s_n^{*2}}{m+n-2}}} \sqrt{\frac{mn}{m+n}} \sim t_{m+n-2}.$$

# Two-sample t-test

2. Find the $t$ value for which

$$\mathbb{F}(t) = 1 - \frac{\alpha}{2}, \qquad \left(\text{or } G(t) = \frac{\alpha}{2}\right).$$

3. Based on the sample, calculate

$$t_{\text{obs}} := \frac{\overline{\xi} - \overline{\eta}}{\sqrt{\frac{(m-1)s_m^* + (n-1)s_n^{*2}}{m+n-2}}} \sqrt{\frac{mn}{m+n}}$$

4. If $|t_{\text{obs}}| < t$, then we accept the null hypothesis $H_0$; otherwise, we reject it.

# 2.5. F-test

# F-próba

Legyen $\xi_1, \xi_2 \dots, \xi_m \sim \mathcal{N}(\mu_1, \sigma_1^2)$ és $\eta_1, \eta_2 \dots, \eta_n \sim \mathcal{N}(\mu_2, \sigma_2^2)$.
$H_0 : \sigma_1^2 = \sigma_2^2$;
$H_1 : \sigma_1^2 \neq \sigma_2^2$.
**Statistic:**

$$f = \frac{ns_n^2}{ms_m^2} \sim F_{m-1,n-1} \qquad \text{ha} \qquad ns_n^2 > ms_m^2.$$

Let $\mathbb{F}$ denote the cumulative distribution function of the $F_{m-1,n-1}$ distribution. The values $x_a$ and $x_f$ are determined so that

$$\mathbb{F}(x_f) - \mathbb{F}(x_a) = 1 - \alpha,$$

which is difficult to realize exactly. Instead, we can restrict to

$$\mathbb{F}(x_f) = 1 - \frac{\alpha}{2}, \qquad \mathbb{F}(x_a) = \frac{\alpha}{2},$$

or (depending on the table)

$$G(x_f) = \frac{\alpha}{2}, \quad G(x_a) = \frac{\alpha}{2}.$$

If $x_a < f_{\text{obs}} < x_f$, then $H_0$ is accepted; otherwise, it is rejected.

# 3. Test for the Variance

# $\chi^2$ Test

Given $\xi_1, \xi_2 \ldots, \xi_n \sim \mathcal{N}(\mu, \sigma^2)$ independent sample.
$H_0 : \sigma = \sigma_0$;
$H_1 : \sigma \neq \sigma_0$.
**Statistic:**

$$\frac{ns_n^2}{\sigma_0^2} \sim \chi_m^2, \quad \text{provided that } \sigma = \sigma_0.$$

Let $\mathbb{F}$ denote the $\chi_{n-1}^2$ cumulative distribution function. Determine $x_a$, $x_f$ such that

$$\mathbb{F}(x_f) - \mathbb{F}(x_a) = 1 - \alpha,$$

it suffices that

$$\mathbb{F}(x_f) = 1 - \frac{\alpha}{2}, \qquad \mathbb{F}(x_a) = \frac{\alpha}{2}.$$

In the table, $\mathbb{F}(x) + G(x) = 1$, so $G(x_f) = \frac{\alpha}{2}$, $G(x_a) = 1 - \frac{\alpha}{2}$.
**Decision:**
If $x_a < \chi < x_f$, then accept $H_0$; otherwise, reject it.

# 4. Goodness-of-Fit Test

# Goodness-of-Fit Test

Let $\xi_1, \xi_2, \ldots, \xi_n$ be an independent sample. With distribution function $\mathbb{F}$.
$H_0 : \mathbb{F} = \mathbb{F}_0$;
$H_1 : \mathbb{F} \neq \mathbb{F}_0$.
Divide the real line into $k$ parts with breakpoints

$$-\infty = a_0 < a_1 < a_2 < \cdots < a_{k-1} < a_k = +\infty$$

$$\mathbb{P}(a_{i-1} \leqslant \xi < a_i) = \mathbb{F}_0(a_i) - \mathbb{F}_0(a_{i-1}) = p_i \qquad (i = 1, 2, \ldots, k)$$

Based on the sample, we know that $\eta_i$ sample elements fall into the $i$-th interval. Then

$$\chi := \mathbb{F}(1-\alpha) = G(\alpha);$$

where $\mathbb{F}$ denotes the distribution function of a chi-square distribution with $k-1$ degrees of freedom, and $G(x) = 1 - \mathbb{F}(x)$ $(x \in \mathbb{R})$ is defined accordingly.

# Goodness-of-Fit Test

Let

$$\chi_{\text{obs}} := \sum_{i=1}^{k} \frac{(\eta_i - np_i)^2}{np_i}.$$

**Decision:**

We accept $H_0$ if $\chi < \chi_{\text{obs}}$, otherwise we reject it.

End of Lecture 12