Statistical Inference and Multivariate Analysis (MA324)

Lecture Slides
Lecture 17

Null and Alternative Hypotheses, Errors, Power Function

Indian Institute of Technology Guwahati

Jan-May 2023

Statistical Hypothesis

Def: [**Statistical Hypothesis**] A statistical hypothesis or simply hypothesis **is a statement** about the unknown parameters.

Def: [**Null and Alternative Hypotheses**] Suppose that one wants to choose between two reasonable hypotheses $H_0 : \boldsymbol{\theta} \in \Theta_0$ against $H_1 : \boldsymbol{\theta} \in \Theta_1$, where $\Theta_0 \subset \Theta$, $\Theta_1 \subset \Theta$ and $\Theta_1 \cap \Theta_2 = \emptyset$. We call $H_0 : \boldsymbol{\theta} \in \Theta_0$ and $H_1 : \boldsymbol{\theta} \in \Theta_1$ as **null hypothesis** and **alternative hypothesis**, respectively.

Def: [**Simple and Composite Hypothesis**] A hypothesis is called simple if it specifies the underlying distribution. Otherwise, it is called composite hypothesis.

- Simple hypothesis:

$$(1) \qquad H_0 : \mu = \mu_0$$

- Composite hypothesis:

$$\begin{aligned} &(1) \qquad H_1 : \mu > \mu_0 \\ &(2) \qquad H_1 : \mu < \mu_0 \\ &(3) \qquad H_1 : \mu \neq \mu_0 \end{aligned}$$

Our aim is to choose one hypothesis among null and alternative hypotheses. As we will see that the roles of these two hypotheses are asymmetric, we need to be careful while labeling these two hypotheses.

Errors and Errors Probabilities

- As illustrated in Examples, the **decision to accept or reject** null hypothesis will be taken **based on a RS**.
- We will consider a **reasonable statistic** and make the choice based on the statistic. If the observed value of the statistic belongs to an appropriate set, we reject null hypothesis and if the value of the statistic does not belong to the set, we accept null hypothesis.
- Now, the statistic belongs to a set can be **alternatively written as the sample observation belongs** to a subset of $\mathbb{R}^n$.
- In earlier coin-tossing examples, suppose that we reject the null hypothesis if and only if $|T| > 1.96$. The set $\{|T| > 1.96\}$ is the set of all sample points $\boldsymbol{x}$ such that $T(\boldsymbol{x}) > 1.96$ or $T(\boldsymbol{x}) < -1.96$. Thus, the condition $|T| > 1.96$ actually divides $\mathbb{R}^n$ into two parts, *viz.*, $\{\boldsymbol{x} \in \mathbb{R}^n : |T(\boldsymbol{x})| > 1.96\}$ and $\{\boldsymbol{x} \in \mathbb{R}^n : |T(\boldsymbol{x})| \leq 1.96\}$.

- If the sample observation x belongs to the first partition, we reject the null hypothesis. If the sample observation x belongs the second partition, we accept the null hypothesis. Of course, we need to **find a "meaningful" partition** so that we can make correct decision.

Def: [**Critical Region and Acceptance Region**] Suppose that we want to test $H_0 : \boldsymbol{\theta} \in \Theta_0$ against $H_1 : \boldsymbol{\theta} \in \Theta_1$ based on a RS of size n drawn from a population having PMF/PDF $f(\cdot, \boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \Theta_0 \cup \Theta_1$. Let R be a subset of $\mathbb{R}^n$ such that we **reject** H_0 **if and only if** $\boldsymbol{x} \in R$, where $\boldsymbol{x}$ denotes a realization of the RS. Then, R is called **rejection region or critical region**, and R^c is called **acceptance region**.

Type-I and Type-II Errors

- In this process, there are **four possibility as described in the following table**.
- There are two cases, where we do **not commit any error**. These cases are accepting null hypothesis when it is actually true and rejecting null hypothesis when it is actually false.
- The **green ticks** in the following table signify that there are **no errors**.
- We commit errors in other two cases, *viz.*, **accepting null hypothesis when it is actually false** or **rejecting null hypothesis** when it is **actually true**.

	H_0 true	H_1 true
Accept H_0	✓	Type-II Error
Reject H_0	Type-I Error	✓

Type-I and Type-II Errors

Def: [**Type-I and Type-II Errors**] The error committed by **rejecting** H_0 when it is **actually true** is called Type-I Error. The error committed by **accepting** H_0 when it is **actually false** is called Type-II Error.

Example 1: Let $X_1, X_2, \ldots, X_9 \overset{i.i.d.}{\sim} N(\theta, 1)$. Suppose that we are want to test $H_0 : \theta = 5.5$ against $H_1 : \theta = 7.5$. For comparison purpose, let us consider four critical regions:

$$R_1 = \emptyset, \quad R_2 = \left\{ \boldsymbol{x} \in \mathbb{R}^9 : \bar{x} > 7 \right\}, \quad R_3 = \left\{ \boldsymbol{x} \in \mathbb{R}^9 : \bar{x} > 6 \right\}, \quad R_4 = \mathbb{R}^9.$$

As we are trying to test hypotheses regarding **population mean**, it is intuitively make sense to use **sample mean**. That is the reason to take R_2 and R_3 in terms of sample mean $\bar{x}$. On the other hand, R_1 and R_4 are two extremes. The critical regions R_1 **accepts** null hypothesis irrespective of the realization of the RS. Similarly, we **always reject** the null hypothesis if the critical region is R_4.

Note that Type-I or Type-II errors are events. Thus, we may talk about the probabilities of these errors. For critical region R_3,

$$P\,(\text{Type-I Error}) = P_{\theta=5.5}\left(\bar{X} > 6\right) = 1 - \Phi\left(3(6 - 5.5)\right) = 0.06681$$

$$P\,(\text{Type-II Error}) = P_{\theta=7.5}\left(\bar{X} \leq 6\right) = \Phi\left(3(6 - 7.5)\right) \sim 0.$$

Similarly, the probabilities of errors for other critical regions can be calculated and given in following table.

	R_1	R_2	R_3	R_4
P(Type-I)	0	~ 0	0.06681	1
P(Type-II)	1	0.06681	~ 0	0

In this example, $R_1 \subset R_2 \subset R_3 \subset R_4$. Notice that as we **increase the size of the critical region**, probability of Type-I error **increases** and that of Type-II error **decreases**. In other words, if we **try to reduce probability of one error**, probability of the other one increases.

Power Function

Def: [**Power Function**] The power function of a critical region, denoted by $\beta : \Theta_1 \cup \Theta_0 \to [0,\, 1]$, is the **probability of rejecting the null hypothesis** H_0 when $\boldsymbol{\theta}$ is the true value of the parameter, *i.e.*,

$$\beta(\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(\boldsymbol{X} \in R).$$

It is clear that $\beta(\cdot)$ is the probability of Type-I error if $\boldsymbol{\theta} \in \Theta_0$. For $\boldsymbol{\theta} \in \Theta_1$, $\beta(\cdot)$ is one minus probability of Type-II error.

Example 2: Let $X_1,\, X_2,\, \ldots,\, X_9 \overset{i.i.d.}{\sim} N(\theta,\, 1)$. Suppose that we are want to test $H_0 : \theta = 5.5$ against $H_1 : \theta = 7.5$. The power function of the critical region $R_2 = \left\{\boldsymbol{x} \in \mathbb{R}^9 : \bar{x} > 7\right\}$ is

$$\beta(\theta) = P_\theta\left(\boldsymbol{X} \in R\right) = P_\theta\left(\bar{X} > 7\right) = P_\theta\left(3\left(\bar{X} - \theta\right) > 21 - 3\theta\right) = 1 - \Phi(21 - 3\theta).$$

for $\theta = 5.5$ and 7.5.