Testing of Hypotheses

Let $X_1, \ldots, X_n$ be a random sample from a population distribution F_θ, $\theta \in \Theta$, where the functional form of F_θ is known, except, for the parameter θ. For example, we may have a random sample from $N(\mu, 1)$ population where the value μ is unknown. One may be interested in examining the validity of assertion that the value of μ lies in a certain known range, say (μ_1, μ_2) on the basis of the sample drawn from the population. A problem of this type is usually referred to as a problem of testing of hypotheses.

Definition 1. *A parametric hypothesis is a statement about the unknown parameter θ. It is usually referred to as the null hypothesis $H_0 : \theta \in \Theta_0$ where $\Theta_0 \subset \Theta$. The statement $H_1 : \theta \in \Theta_1$ is usually referred to as the alternative hypothesis.*

Our task is to test H_0 against H_1. Here θ can be vector valued also.

Definition 2. *If Θ_0 contains only one point, say θ_0, the hypothesis H_0 is said to be a simple hypothesis. Otherwise, i.e., if Θ_0 contains more than one point, the hypothesis H_0 is said to be composite hypothesis. Similar definition hold for alternative hypothesis.*

Under simple hypothesis the probability density function(pdf) or probability mass function(pmf) of a random variable X is completely specified.

Example 3. *Suppose X follows $N(\mu, \sigma^2)$, where σ^2 is known. The hypothesis $H_0 : \mu = \mu_0$ is a simple hypothesis. The hypotheses $H : \mu > \mu_0$, $H : \mu \le \mu_0$ are composite hypothesis. If σ^2 is also unknown, $H_0 : \mu = \mu_0$ is composite hypothesis, because, here, $\Theta = \{(\mu, \sigma^2), -\infty < \mu < \infty, \sigma^2 > 0\}$ and $\Theta_0 = \{\mu = \mu_0, \sigma^2 > 0\}$ contains infinitely many points.*

Often we are interested in testing a simple hypothesis $H_0(\theta = \theta_0)$ against alternative composite hypothesis $H_1(\theta \ne \theta_0)$ called two/both sided alternative or one sided composite alternatives $H_1(\theta < \theta_0)$, $H_2(\theta > \theta_0)$.
The problem of testing of hypothesis H_0 against H_1 may be described as follows. Given the sample observations, $\underline{x} = (x_1, \ldots, x_n)'$, we make a decision which will either lead to either acceptance or rejection of H_0. The sample space χ_n of the random vector $\underline{X} = (X_1, \ldots, X_n)'$, is divided into two disjoint subsets w and $w^c = \chi_n - w$ such that H_0 is rejected if $x \in w$ and is accepted if $x \in w^c$. The region w is called the critical region and w^c the region of acceptance. Such a test is called a non-randomized test of H_0 against H_1.
Let $\delta(x)$ is be a function denoting the probability of rejecting the null hypothesis H_0 when $\underline{x}$ is the sample observation. Then, for a non-randomized test

$$\delta(x) = \begin{cases} 1, & if x \in w \\ 0, & if x \in w^c. \end{cases}$$

Definition 4. *Every Borel-measurable mapping $\phi : \mathbb{R}^n \to [0, 1]$ is called a test function.*

Definition 5. *A non-randomized test for testing H_0 against H_1 is a test function $\delta(x)$ defined for all $\underline{x} \in \chi_n$ such that*

$$\delta(x) = \begin{cases} 1, & if x \in w \\ 0, & if x \in w^c. \end{cases}$$

Here, the critical region w depends on the test function δ. If δ changes, $w(\delta)$ will be different. In a non-randomized test if $\underline{x}$ is observed we either accept H_0 or reject it with probability 1.

Definition 6. *A randomised test for testing H_0 against H_1 is a test function $\delta(x)$ defined for all $\underline{x} \in \chi_n$ such that $0 \le \delta(\underline{x}) \le 1$ $\forall \underline{x}$. If we observe $\underline{x}$ we make a Bernoulli experiment with probability of success $\delta(\underline{x})$. If a success occurs we reject H_0, otherwise we accept it.*

Randomised test will be needed in general only if X is discrete random variable. But we shall only consider non-randomised tests.

In the problem of testing of hypotheses, the true value of θ remains unknown. We are only aiming at testing whether the observation x supports our assertion $\theta \in \Theta_0$ i.e. $\underline{x}$ is a random sample from the pdf $f_\theta(\underline{x})$, $\theta \in \Theta_0$. Hence, the acceptance (rejection) of H_0 on the basis of $\underline{x}$ does not necessarily imply that H_0 is true (false). Therefore, we may reject H_0 when it is, in fact, true; or we may accept it H_0 when it is false. In both the cases we commit some error.

Definition 7. ***Type I Error:*** *Rejecting H_0, when it is true is known as type I error.*

Definition 8. ***Type II Error:*** *Accepting H_0, when it is false is known as type II error.*

Let $H_0 : \theta = \theta_0$ so that $\Theta_0 = \{\theta_0\}$. In this case probability of type I error is given by

$$P_{\theta_0}(w) = P\{\underline{x} \in w | H_0\} = \int_w f_{\theta_0}(\underline{x}) d\underline{x} = \int \delta(\underline{x}) f_{\theta_0}(\underline{x}) d\underline{x} = E_{\theta_0}(\delta(\underline{x})) \tag{1}$$

and probability of type II error is given by

$$P_\theta(w^c) = P\{\underline{x} \in w^c | H_1\} = \int_w^c f_\theta(\underline{x}) d\underline{x}, \quad \theta \in \Theta_1. \tag{2}$$

The probability of rejecting a true H_0 depends on the test function δ and the value θ_0 and is called the level of significance of test or the size of the critical region w.

The probability of rejecting H_0 when it is false i.e. when $\theta \in \bar{\Theta}_0 = \Theta - \{\theta_0\}$, is

$$P_\theta(x \in w) = 1 - P_\theta(x \in w^c) = 1 - P_\theta(w^c), \quad \theta \in \bar{\Theta_0} = \gamma_\theta(w), \quad \theta \in \bar{\Theta_0}. \tag{3}$$

$\gamma_\theta(w)$ is called the power of the test w. It depends on the test function δ and the value of the parameter $\theta \in \bar{\Theta}_0$. Note that $\gamma_\theta(w) = E_\theta(\delta(\underline{x}))$ where $\theta \in \bar{\Theta}_0$.

For $\theta \in \Theta_1(\subseteq (\bar{\Theta}_0)$,

$$\gamma_\theta(w) = 1 - P_\theta(w^c) = 1 - \text{Probability of type II error.} \tag{4}$$

In an ideal test procedure both types of errors should be minimum. However, simultaneous minimization of both both the errors is not possible. Therefore, we try to fix an upper bound on one error and then find a test procedure for which the second probability is minimum. In practice, we pre-assign a small value α (usually 0.05 or 0.01) to probability of type I error and minimize the probability of type II error (β_θ)subject to this constraint.

Example 9. *Let $X_1, \cdots, X_n$ be a random sample from $N(\mu, 1)$. We want to test $H_0 : \mu = -1/2$ against $H_1 : \mu = 1/2$.*
Here, the acceptance region is $w^c = (-\infty, 0]$, i.e., accept H_0 if $\bar{X} \leq 0$. The rejection region is $w = (0, \infty)$, i.e., reject H_0 if $\bar{X} > 0$. Now, we calculate both the errors.

$$\begin{aligned} \alpha = & \quad \textit{Prob(Type I error)= Prob(Rejecting } H_0\textit{, when it is true)} \\ = & \quad P_{\mu=\frac{-1}{2}}(\bar{X} > 0) = P_{\mu=\frac{-1}{2}}(\sqrt{n}(\bar{X} + \tfrac{1}{2}) > \tfrac{\sqrt{n}}{2}) = P(Z > \tfrac{\sqrt{n}}{2}) \\ = & \quad P(Z > 2) = 0.0228, \qquad \textit{for } n = 16. \end{aligned}$$

β=Prob(Type II error)= Prob(Accepting H_0, when it is false) $\Rightarrow$

$$\beta = P_{\mu=\frac{1}{2}}(\bar{X} \leq 0) = P_{\mu=\frac{-1}{2}}(\sqrt{n}(\bar{X} - \frac{1}{2}) \leq \frac{-\sqrt{n}}{2}) = P(Z \leq \frac{\sqrt{n}}{2}) = P(Z \leq 2) = 0.0228$$

, for $n = 16$.
Here, α and β are same.

Now let us modify the test procedure. Let the acceptance and rejection region be $w_1^c = \{\bar{X} < \frac{-1}{4}\}$ and $w_1 = \{\bar{X} \geq \frac{-1}{4}\}$, respectively. Therefore, the probability of type I and type II errors are

$$\alpha^* = P_{\mu=\frac{-1}{2}}(\bar{X} \geq \frac{-1}{4}) = P_{\mu=\frac{-1}{2}}(\sqrt{n}(\bar{X} + \frac{1}{2}) > \frac{\sqrt{n}}{4}) = P(Z \geq \frac{\sqrt{n}}{4}) = 0.1587,$$

for $n = 16$.

$$\beta^* = P_{\mu=\frac{1}{2}}(\bar{X} < \frac{-1}{4}) = P(Z < -3) = 0.0013,$$

for $n = 16$.
Here, we observe that $\beta^* < \beta$ but $\alpha^* > \alpha$. Hence, it is clear that the simultaneous minimization of both the errors α and β is not possible.

Exercise 10. *Let $X_1, \ldots, X_{20}$ be a random sample from the exponential distribution with pdf $f(x;\theta) = \theta e^{-\theta x}$, $0 < x < \infty$, $\theta > 0$. Calculate type I error and type II error for testing $H_0 : \theta = 1$ against $H_1 : \theta = 2$.*