Chapter 10 – Hypothesis Testing

0.1 – Introduction

In Chapter 9, we studied the method of inference called *point estimation.*

In this chapter, we will study another method of inference called *hypothesis testing.*

Definition 1 *A* ***hypothesis*** *is a statement or a claim about a population parameter.*

Definition 2 *The two complementary hypotheses in a hypothesis testing problem are called the* null hypothesis *and the* alternative hypothesis, *denoted as H_0 and H_1, respectively.*

Setting: Let θ be a parameter of interest with $\theta \in \Theta$:

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta \in \Theta_1 = \Theta_0^c$$

where $\Theta_0 \subset \Theta$ and $\Theta_0^c = \Theta_1$.

Definition 3 *A* hypothesis testing procedure *or* hypothesis test *is a rule that specifies:*

1. For which sample values, the decision is made to accept H_0 as true.

2. For which sample values, H_0 is rejected and H_1 is accepted as true.

Definition 4 ***Rejection Region or Critical Region****: The set of (sample) points, $R = R(\mathbf{x})$ in the sample space for which H_0 will be rejected. Clearly, R is a subset of the sample space.*

Definition 5 ***Acceptance Region****: Complement of the rejection region, R^c.*

Definition 6 ***Decision Rule****: The rule specifying when to reject H_o. Typically, it is "reject H_0 if $\mathbf{x} \in R(\mathbf{x})$" (i.e., when the observed sample $\mathbf{x}$ falls in $R(\mathbf{x})$.*

Philosophical Issues:

"Reject H_0" and "accepting H_1"

"Accepting H_0" and "not rejecting H_0"

Our Main Concern is the assertion/claim in H_0 or H_1, and we will say "reject H_0" and "accept H_0" (with the above caveats in effect).

Definition 7 ***Test Statistic*** $W(\mathbf{X})$***:*** *A function of the sample* $\mathbf{X}$*; used together with the rejection region to make decisions about the hypotheses being tested.*

Question: How?

Types of Hypotheses:

1. Simple Hypothesis: $H : \theta = \theta_0, \quad H : \theta = \theta_1$ (can be null or alternative hypotheses)

2. Composite Hypothesis: more than one possible parameter (or distribution), (can be null or alternative hypotheses)

a. One-sided Hypothesis:

null hypotheses $H_0 : \theta \leq \theta_0, \quad H_0 : \theta \geq \theta_0,$

alternative hypotheses $H_1 : \theta < \theta_0, \quad H_1 : \theta > \theta_0$

b. Two-sided Hypothesis: $H : \theta \neq \theta_0$ (can only be alternative hypothesis)

0.2 Likelihood Ratio Tests

Recall that, if $X_1, X_2, \ldots, X_n$ is a random sample from a pdf or pmf $f(x|\theta)$, the likelihood function is

$$L(\theta|x_1, \ldots, x_n) = L(\theta|\mathbf{x}) = f(\mathbf{x}|\theta) = \prod_{i=1}^{n} f(x_i|\theta).$$

Definition 8 *The* ***likelihood ratio test statistic*** *for testing* $H_0 : \theta \in \Theta_0$ *versus* $H_1 : \theta \in \Theta_1 = \Theta_0^c$ *(based on sample* $\mathbf{x}$*) is*

$$\lambda(\mathbf{x}) = 2 \log \left(\frac{\sup_{\theta \in \Theta} L(\theta|\mathbf{x})}{\sup_{\theta \in \Theta_0} L(\theta|\mathbf{x})} \right).$$

A likelihood ratio test *(*LRT*) is any test that has a rejection region of the form* $\{\mathbf{x} : \lambda(\mathbf{x}) > c\}$, *where* c *is any number satisfying* $0 \leq c \leq 1$. *So, the decision rule is "reject* H_0 *if* $\lambda(\mathbf{x}) > c$*".*

Qu: Why is such a region reasonable to reject H_0 from the likelihood perspective?

Notes:

1. **Connection between LRT and MLE**: Let $\hat{\theta}$ be the MLE of θ under the unrestricted (i.e., entire) parameter space Θ and $\hat{\theta}_0$ be the MLE of θ under the restricted parameter space Θ_0. Then

$$\lambda(\mathbf{x}) = 2 \log \left(\frac{L\left(\hat{\theta}|\mathbf{x}\right)}{L\left(\hat{\theta}_0|\mathbf{x}\right)} \right).$$

Here, $\hat{\theta}$ is also called the *global maximizer* and $\hat{\theta}_0$ is also called the *restricted maximizer* or *null maximizer* of $L(\theta|\mathbf{x})$.

2. Proper choice of c depends on the dimensions of the null and alternative hypotheses.

Suppose Θ is of dimension r and Θ_0 is of dimension q with $q < r$. Then, by Theorem 10.22 in AoS, we have $\lambda(\mathbf{x}) \xrightarrow{d} \chi^2_{r-q}$. Thus, the to have a size α LRT, the critical value (or the cut-off value) in Definition 8 should be $c = \chi^2_{r-q,\alpha}$ where $\chi^2_{r-q,\alpha}$ is the $1 - \alpha$ quantile of χ^2_{r-q} distribution.

Example 1 ***(Poisson LRT)*** *Let* $X_1, X_2, \ldots, X_n$ *be iid from Poisson*(λ) *distribution. We want to test* $H_0 : \lambda = \lambda_0$ *versus* $H_1 : \lambda \neq \lambda_0$ *where* λ_0 *is a number specified by the experimenter. Construct the LRT for this setting.*

Solution: