CHAPTER 9

Neyman–Pearson Theory of Testing of Hypotheses

9.1 INTRODUCTION

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population distribution $F_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta$, where the functional form of $F_{\boldsymbol{\theta}}$ is known except perhaps for the parameter $\boldsymbol{\theta}$. For example, the X_i's may be a random sample from $\mathcal{N}(\theta, 1)$, where $\theta \in \mathcal{R}$ is not known. In many practical problems the experimenter is interested in testing the validity of an assertion about the unknown parameter θ. For example, in a coin-tossing experiment it is of interest to test, in some sense, whether the (unknown) probability of heads p equals a given number p_0, $0 < p_0 < 1$. Similarly, it is of interest to check the claim of a car manufacturer about the average mileage per gallon of gasoline achieved by a particular model. A problem of this type is usually referred to as a problem of *testing of hypotheses* and is the subject of discussion in this chapter. We develop the fundamentals of Neyman–Pearson theory. In Section 9.2 we introduce the various concepts involved. In Section 9.3 the fundamental Neyman–Pearson lemma is proved, and Sections 9.4 and 9.5 deal with some basic results in the testing of composite hypotheses. Section 9.6 deals with locally optimal tests.

9.2 SOME FUNDAMENTAL NOTIONS OF HYPOTHESES TESTING

In Chapter 8 we discussed the problem of point estimation in sampling from a population whose distribution is known except for a finite number of unknown parameters. Here we consider another important problem in statistical inference, the testing of statistical hypotheses. We begin by considering the following examples.

Example 1. In coin-tossing experiments one frequently assumes that the coin is fair, that is, the probability of getting heads or tails is the same: $\frac{1}{2}$. How does one test whether the coin is fair (unbiased) or loaded (biased)? If one is guided by intuition, a reasonable procedure would be to toss the coin n times say, and count the number of heads. If the proportion of heads observed does not deviate "too much" from $p = \frac{1}{2}$, one would tend to conclude that the coin is fair.

Example 2. It is usual for manufacturers to make quantitative assertions about their products. For example, a manufacturer of 12-volt batteries may claim that a certain brand of their batteries lasts for N hours. How does one go about checking the truth of this assertion? A reasonable procedure suggests itself: Take a random sample of n batteries of the brand in question and note their length of life under more or less identical conditions. If the average length of life is "much smaller" than N, one would tend to doubt the manufacturer's claim.

To fix ideas, let us define formally the concepts involved. As usual, $\mathbf{X} = (X_1, X_2, \dots, X_n)$ and let $\mathbf{X} \sim F_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta \subseteq \mathcal{R}_k$. It will be assumed that the functional form of $F_{\boldsymbol{\theta}}$ is known except for the parameter $\boldsymbol{\theta}$. Also, we assume that Θ contains at least two points.

Definition 1. A parametric hypothesis is an assertion about the unknown parameter $\boldsymbol{\theta}$. It is usually referred to as the *null hypothesis*, $H_0\colon \boldsymbol{\theta} \in \Theta_0 \subset \Theta$. The statement $H_1\colon \boldsymbol{\theta} \in \Theta_1 = \Theta - \Theta_0$ is usually referred to as the *alternative hypothesis*.

Usually, the null hypothesis is chosen to correspond to the smaller or simpler subset Θ_0 of Θ and is a statement of "no difference," whereas the alternative represents change.

Definition 2. If $\Theta_0(\Theta_1)$ contains only one point, we say that $\Theta_0(\Theta_1)$ is *simple*; otherwise, *composite*. Thus, if a hypothesis is simple, the probability distribution of $\mathbf{X}$ is specified completely under that hypothesis.

Example 3. Let $X \sim \mathcal{N}(\mu, \sigma^2)$. If both μ and σ^2 are unknown, $\Theta = \{(\mu, \sigma^2)\colon -\infty < \mu < \infty, \sigma^2 > 0\}$. The hypothesis $H_0\colon \mu \le \mu_0, \sigma^2 > 0$, where μ_0 is a known constant, is a composite null hypothesis. The alternative hypothesis is $H_1\colon \mu > \mu_0$, $\sigma^2 > 0$, which is also composite. Similarly, the null hypothesis $\mu = \mu_0, \sigma^2 > 0$ is composite.

If $\sigma^2 = \sigma_0^2$ is known, the hypothesis $H_0\colon \mu = \mu_0$ is a simple hypothesis.

Example 4. Let $X_1, X_2, \dots, X_n$ be iid $b(1, p)$ RVs. Some hypotheses of interest are $p = \frac{1}{2}$, $p \le \frac{1}{2}$, $p \ge \frac{1}{2}$ or, quite generally, $p = p_0$, $p \le p_0$, $p \ge p_0$, where p_0 is a known number, $0 < p_0 < 1$.

The problem of testing of hypotheses may be described as follows: Given the sample point $\mathbf{x} = (x_1, x_2, \dots, x_n)$, find a decision rule (function) that will lead to a decision to reject or fail to reject the null hypothesis. In other words, partition the sample space into two disjoint sets C and C^c such that if $\mathbf{x} \in C$, we reject H_0, and if $\mathbf{x} \in C^c$, we fail to reject H_0. In the following we write "accept H_0" when we fail to reject H_0. We emphasize that when the sample point $\mathbf{x} \in C^c$ and we fail to reject H_0, it does not mean that H_0 gets our stamp of approval. It simply means that the sample does not have enough evidence against H_0.

Definition 3. Let $\mathbf{X} \sim F_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta$. A subset C of $\mathcal{R}_n$ such that if $\mathbf{x} \in C$, then H_0 is rejected (with probability 1) is called the *critical region* (*set*):

$$C = \{\mathbf{x} \in \mathcal{R}_n : H_0 \text{ is rejected if } \mathbf{x} \in C\}.$$

There are two types of errors that can be made if one uses such a procedure. One may reject H_0 when in fact it is true, called a *type I error*, or accept H_0 when it is false, called a *type II error*:

		True	
		H_0	H_1
Accept	H_0	Correct	Type II error
	H_1	Type I error	Correct

If C is the critical region of a rule, $P_{\boldsymbol{\theta}}C$, $\boldsymbol{\theta} \in \Theta_0$, is a *probability of type* I *error*, and $P_{\boldsymbol{\theta}}C^c$, $\boldsymbol{\theta} \in \Theta_1$, is a *probability of type* II *error*. Ideally, one would like to find a critical region for which both these probabilities are 0. This will be the case if we can find a subset $S \subseteq \mathcal{R}_n$ such that $P_{\boldsymbol{\theta}}S = 1$ for every $\boldsymbol{\theta} \in \Theta_0$ and $P_{\boldsymbol{\theta}}S = 0$ for every $\boldsymbol{\theta} \in \Theta_1$. Unfortunately, situations such as this do not arise in practice, although they are conceivable. For example, let $X \sim \mathcal{C}(1, \theta)$ under H_0 and $X \sim P(\theta)$ under H_1. Usually, if a critical region is such that the probability of type I error is 0, it will be of the form "do not reject H_0" and the probability of type II error will then be 1.

The procedure used in practice is to limit the probability of type I error to a preassigned level α (usually, 0.01 or 0.05) that is small and to minimize the probability of type II error. To restate our problem in terms of this requirement, let us formulate these notions.

Definition 4. Every Borel-measurable mapping φ of $\mathcal{R}_n \to [0, 1]$ is known as a *test function*.

Some simple examples of test functions are $\varphi(\mathbf{x}) = 1$ for all $\mathbf{x} \in \mathcal{R}_n$, $\varphi(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathcal{R}_n$, or $\varphi(\mathbf{x}) = \alpha$, $0 \le \alpha \le 1$, for all $\mathbf{x} \in \mathcal{R}_n$. In fact, Definition 4 includes Definition 3 in the sense that whenever φ is the indicator function of some Borel subset A of $\mathcal{R}_n$, A is called the *critical region* (of the test φ).

Definition 5. The mapping φ is said to be a *test* of hypothesis $H_0\colon \boldsymbol{\theta} \in \Theta_0$ against the alternatives $H_1\colon \boldsymbol{\theta} \in \Theta_1$, with *error probability* α (also called *level of significance* or, simply, *level*) if

$$(1) \qquad E_{\boldsymbol{\theta}}\varphi(\mathbf{X}) \le \alpha \qquad \text{for all } \boldsymbol{\theta} \in \Theta_0.$$

We shall say, in short, that φ is a test for the problem $(\alpha, \Theta_0, \Theta_1)$.

Let us write $\beta_\varphi(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}\varphi(\mathbf{X})$. Our objective, in practice, will be to seek a test φ for a given α, $0 \le \alpha \le 1$, such that

$$\sup_{\boldsymbol{\theta}\in\Theta_0} \beta_\varphi(\boldsymbol{\theta}) \le \alpha. \tag{2}$$

The left-hand side of (2) is usually known as the *size* of the test φ. Condition (1) therefore restricts attention to tests whose size does not exceed a given level of significance α.

The following interpretation may be given to all tests φ satisfying $\beta_\varphi(\boldsymbol{\theta}) \le \alpha$ for all $\boldsymbol{\theta} \in \Theta_0$. To every $\mathbf{x} \in \mathcal{R}_n$ we assign a number $\varphi(\mathbf{x})$, $0 \le \varphi(\mathbf{x}) \le 1$, which is the probability of rejecting H_0 that $\mathbf{X} \sim f_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta_0$, if $\mathbf{x}$ is observed. The restriction $\beta_\varphi(\boldsymbol{\theta}) \le \alpha$ for $\boldsymbol{\theta} \in \Theta_0$ then says that if H_0 were true, φ rejects it with a probability $\le \alpha$. We will call such a test a *randomized test* function. If $\varphi(\mathbf{x}) = I_A(\mathbf{x})$, φ will be called a *nonrandomized test*. If $\mathbf{x} \in A$, we reject H_0 with probability 1; and if $\mathbf{x} \notin A$, this probability is 0. Needless to say, $A \in \mathfrak{B}_n$.

We next turn our attention to the type II error.

Definition 6. Let φ be a test function for the problem $(\alpha, \Theta_0, \Theta_1)$. For every $\boldsymbol{\theta} \in \Theta$, define

$$\beta_\varphi(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}\varphi(\mathbf{X}) = P_{\boldsymbol{\theta}}\{\text{reject } H_0\}. \tag{3}$$

As a function of $\boldsymbol{\theta}$, $\beta_\varphi(\boldsymbol{\theta})$ is called the *power function* of the test φ. For any $\boldsymbol{\theta} \in \Theta_1$, $\beta_\varphi(\boldsymbol{\theta})$ is called the power of φ against the alternative $\boldsymbol{\theta}$.

In view of Definitions 5 and 6, the problem of testing of hypotheses may now be reformulated. Let $\mathbf{X} \sim f_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta \subseteq \mathcal{R}_k$, $\Theta = \Theta_0 + \Theta_1$. Also, let $0 \le \alpha \le 1$ be given. Given a sample point $\mathbf{x}$, find a test $\varphi(\mathbf{x})$ such that $\beta_\varphi(\boldsymbol{\theta}) \le \alpha$ for $\boldsymbol{\theta} \in \Theta_0$, and $\beta_\varphi(\boldsymbol{\theta})$ is a maximum for $\boldsymbol{\theta} \in \Theta_1$.

Definition 7. Let Φ_α be the class of all tests for the problem $(\alpha, \Theta_0, \Theta_1)$. A test $\varphi_0 \in \Phi_\alpha$ is said to be a *most powerful* (MP) *test* against an alternative $\boldsymbol{\theta} \in \Theta_1$ if

$$\beta_{\varphi_0}(\boldsymbol{\theta}) \ge \beta_\varphi(\boldsymbol{\theta}) \qquad \text{for all } \varphi \in \Phi_\alpha. \tag{4}$$

If Θ_1 contains only one point, this definition suffices. If, on the other hand, Θ_1 contains at least two points, as will usually be the case, we will have an MP test corresponding to each $\boldsymbol{\theta} \in \Theta_1$.

Definition 8. A test $\varphi_0 \in \Phi_\alpha$ for the problem $(\alpha, \Theta_0, \Theta_1)$ is said to be a *uniformly most powerful* (UMP) *test* if

$$\beta_{\varphi_0}(\boldsymbol{\theta}) \ge \beta_\varphi(\boldsymbol{\theta}) \qquad \text{for all } \varphi \in \Phi_\alpha, \quad \text{uniformly in } \boldsymbol{\theta} \in \Theta_1. \tag{5}$$

Thus, if Θ_0 and Θ_1 are both composite, the problem is to find a UMP test φ for the problem $(\alpha, \Theta_0, \Theta_1)$. We will see that UMP tests very frequently do not exist, and we will have to place further restrictions on the class of all tests, Φ_α.

Note that, if φ_1, φ_2 are two tests and λ is a real number, $0 < \lambda < 1$, then $\lambda\varphi_1 + (1-\lambda)\varphi_2$ is also a test function, and it follows that the class of all test functions Φ_α is convex.

Example 5. Let $X_1, X_2, \dots, X_n$ be iid $\mathcal{N}(\mu, 1)$ RVs, where μ is unknown but it is known that $\mu \in \Theta = \{\mu_0, \mu_1\}$, $\mu_0 < \mu_1$. Let H_0: $X_i \sim \mathcal{N}(\mu_0, 1)$, H_1: $X_i \sim \mathcal{N}(\mu_1, 1)$. Both H_0 and H_1 are simple hypotheses. Intuitively, one would accept H_0 if the sample mean $\overline{X}$ is "closer" to μ_0 than to μ_1; that is, one would reject H_0 if $\overline{X} > k$, and accept H_0 otherwise. The constant k is determined from the level requirements. Note that under H_0, $\overline{X} \sim \mathcal{N}(\mu_0, 1/n)$, and under H_1, $\overline{X} \sim \mathcal{N}(\mu_1, 1/n)$. Given $0 < \alpha < 1$, we have

$$\begin{aligned} P_{\mu_0}\{\overline{X} > k\} &= P\left\{\frac{\overline{X}-\mu_0}{1/\sqrt{n}} > \frac{k-\mu_0}{1/\sqrt{n}}\right\} \\ &= P\{\text{type I error}\} = \alpha, \end{aligned}$$

so that $k = \mu + z_\alpha/\sqrt{n}$. The test, therefore, is (Fig. 1)

$$\varphi(\mathbf{x}) = \begin{cases} 1, & \text{if } \overline{x} > \mu_0 + \dfrac{z_\alpha}{\sqrt{n}}, \\ 0, & \text{otherwise.} \end{cases}$$

Here $\overline{X}$ is known as a *test statistic*, and the test φ is nonrandomized with critical region $C = \{\mathbf{x}: \overline{x} > \mu_0 + z_\alpha/\sqrt{n}\}$. Note that in this case the continuity of $\mathbf{X}$ (that is, the absolute continuity of the DF of $\mathbf{X}$) allows us to achieve any size α, $0 < \alpha < 1$.

The power of the test at μ_1 is given by

$$\begin{aligned} E_{\mu_1}\varphi(\mathbf{X}) &= P_{\mu_1}\left\{\overline{X} > \mu_0 + \frac{z_\alpha}{\sqrt{n}}\right\} \\ &= P\left\{\frac{\overline{X}-\mu_1}{1/\sqrt{n}} > (\mu_0-\mu_1)\sqrt{n} + z_\alpha\right\} \\ &= P\{Z > z_\alpha - \sqrt{n}\,(\mu_1-\mu_0)\}, \end{aligned}$$

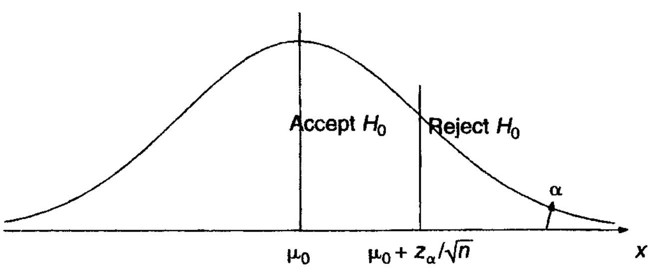

Fig. 1. Rejection region of H_0 in Example 5.

where $Z \sim \mathcal{N}(0, 1)$. In particular, $E_{\mu_1}\varphi(\mathbf{X}) > \alpha$ since $\mu_1 > \mu_0$. The probability of type II error is given by

$$\begin{aligned} P\{\text{type II error}\} &= 1 - E_{\mu_1}\varphi(X) \\ &= P\{Z \le z_\alpha - \sqrt{n}\,(\mu_1 - \mu_0)\}. \end{aligned}$$

Figure 2 gives a graph of the power function $\beta_\varphi(\mu)$ of φ for $\mu > 0$ when $\mu_0 = 0$, and $H_1: \mu > 0$.

Example 6. Let X_1, X_2, X_3, X_4, X_5, be a sample from $b(1, p)$, where p is unknown and $0 \le p \le 1$. Consider the simple null hypothesis H_0: $X_i \sim b(1, \frac{1}{2})$, that is, under H_0, $p = \frac{1}{2}$. Then $H_1: X_i \sim b(1, p)$, $p \ne \frac{1}{2}$. A reasonable procedure would be to compute the average number of 1's, namely, $\overline{X} = \sum_1^5 X_i/5$, and to accept H_0 if $|\overline{X} - \frac{1}{2}| \le c$, where c is to be determined. Let $\alpha = 0.10$. Then we would like to choose c such that the size of our test is α, that is,

$$0.10 = P_{p=1/2}\left\{\left|\overline{X} - \frac{1}{2}\right| > c\right\},$$

or

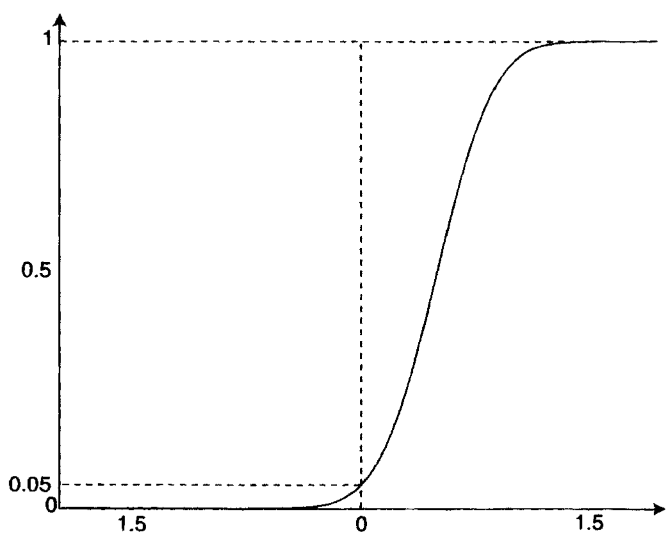

Fig. 2. Power function of φ in Example 5.

$$(6) \qquad 0.90 = P_{p=1/2}\left\{-5c \le \sum_1^5 X_i - \frac{5}{2} \le 5c\right\}$$
$$= P_{p=1/2}\left\{-k \le \sum_1^5 X_i - \frac{5}{2} \le k\right\},$$

where $k = 5c$. Now $\sum_1^5 X_i \sim b(5, \frac{1}{2})$ under H_0, so that the PMF of $\sum_1^5 X_i - \frac{5}{2}$ is given in the following table:

$\sum_1^5 x_i$	$\sum_1^5 x_i - \frac{5}{2}$	$P_{p=1/2}\left\{\sum_1^5 X_i = \sum_1^5 x_i\right\}$
0	−2.5	0.03125
1	−1.5	0.15625
2	−0.5	0.31250
3	0.5	0.31250
4	1.5	0.15625
5	2.5	0.03125

Note that we cannot choose any k to satisfy (6) exactly. It is clear that we have to reject H_0 when $k = \pm 2.5$, that is, when we observe $\sum X_i = 0$ or 5. The resulting size if we use this test is $\alpha = 0.03125 + 0.03125 = 0.0625 < 0.10$. A second procedure would be to reject H_0 if $k = \pm 1.5$ or ± 2.5 ($\sum X_i = 0, 1, 4, 5$), in which case the resulting size is $\alpha = 0.0625 + 2(0.15625) = 0.375$, which is considerably larger than 0.10. If we insist on achieving $\alpha = 0.10$, a third alternative is to randomize on the boundary. Instead of accepting or rejecting H_0 with probability 1 when $\sum X_i = 1$ or 4, we reject H_0 with probability γ where

$$0.10 = P_{p=1/2}\left\{\sum_1^5 X_i = 0 \text{ or } 5\right\} + \gamma P_{p=1/2}\left\{\sum_1^5 X_i = 1 \text{ or } 4\right\}.$$

Thus

$$\gamma = \frac{0.0375}{0.3125} = 0.114.$$

A randomized test of size $\alpha = 0.10$ is therefore given by

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } \sum_1^5 x_i = 0 \text{ or } 5, \\ 0.114 & \text{if } \sum_1^5 x_i = 1 \text{ or } 4, \\ 0 & \text{otherwise.} \end{cases}$$

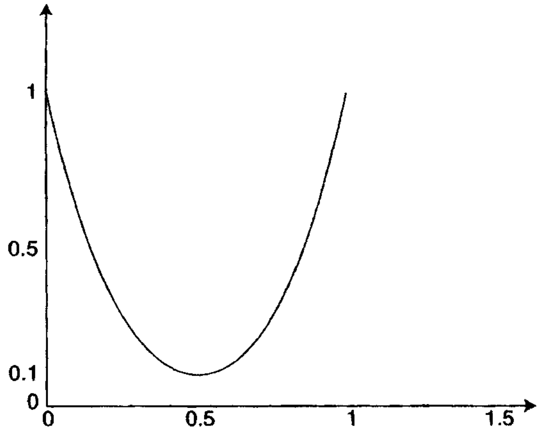

Fig. 3. Power function of φ in Example 6.

The power of this test is

$$E_p\varphi(\mathbf{X}) = P_p\left\{\sum_1^5 X_i = 0 \text{ or } 5\right\} + 0.114P_p\left\{\sum_1^5 X_i = 1 \text{ or } 4\right\},$$

where $p \neq \frac{1}{2}$ and can be computed for any value of p. Figure 3 gives a graph of $\beta_\varphi(p)$.

We conclude this section with the following remarks.

Remark 1. The problem of testing of hypotheses may be considered as a special case of the general decision problem described in Section 8.8. Let $\mathcal{A} = \{a_0, a_1\}$, where a_0 represents the decision to accept H_0: $\boldsymbol{\theta} \in \Theta_0$, and a_1 represents the decision to reject H_0. A decision function δ is a mapping of $\mathcal{R}_n$ into $\mathcal{A}$. Let us introduce the following loss functions:

$$L_1(\boldsymbol{\theta}, a_1) = \begin{cases} 1 & \text{if } \boldsymbol{\theta} \in \Theta_0 \\ 0 & \text{if } \boldsymbol{\theta} \in \Theta_1 \end{cases} \quad \text{and} \quad L_1(\boldsymbol{\theta}, a_0) = 0 \text{ for all } \boldsymbol{\theta},$$

and

$$L_2(\boldsymbol{\theta}, a_0) = \begin{cases} 0 & \text{if } \boldsymbol{\theta} \in \Theta_0 \\ 1 & \text{if } \boldsymbol{\theta} \in \Theta_1 \end{cases} \quad \text{and} \quad L_2(\boldsymbol{\theta}, a_1) = 0 \text{ for all } \boldsymbol{\theta}.$$

Then the minimization of $E_{\boldsymbol{\theta}}L_2(\boldsymbol{\theta}, \delta(\mathbf{X}))$ subject to $E_{\boldsymbol{\theta}}L_1(\boldsymbol{\theta}, \delta(\mathbf{X})) \leq \alpha$ is the hypothesis-testing problem discussed above. We have

$$\begin{aligned} E_{\boldsymbol{\theta}} L_2(\boldsymbol{\theta}, \delta(\mathbf{X})) &= P_{\boldsymbol{\theta}}\{\delta(\mathbf{X}) = a_0\}, \qquad \boldsymbol{\theta} \in \Theta_1, \\ &= P_{\boldsymbol{\theta}}\{\text{accept } H_0 \mid H_1 \text{ true}\}, \end{aligned}$$

and

$$\begin{aligned} E_{\boldsymbol{\theta}} L_1(\boldsymbol{\theta}, \delta(\mathbf{X})) &= P_{\boldsymbol{\theta}}\{\delta(\mathbf{X}) = a_1\}, \qquad \boldsymbol{\theta} \in \Theta_0, \\ &= P_{\boldsymbol{\theta}}\{\text{reject } H_0 \mid \boldsymbol{\theta} \in \Theta_0 \text{ true}\}. \end{aligned}$$

Remark 2. In Example 6 we saw that the size α chosen is often unattainable. The choice of a specific value of α is completely arbitrary and is determined by nonstatistical considerations such as the possible consequences of rejecting H_0 falsely and the economic and practical implications of the decision to reject H_0. An alternative and somewhat subjective approach wherever possible is to report the *P-value* of the test statistic observed. This is the smallest level α at which the sample statistic observed is significant. In Example 6, let $S = \sum_{i=1}^{5} X_i$. If $S = 0$ is observed, then $P_{H_0}(S = 0) = P_0(S = 0) = 0.03125$. By symmetry, if we reject H_0 for $S = 0$, we should also do so for $S = 5$, so the probability of interest is $P_0(S = 0 \text{ or } 5) = .0625$, which is the P-value. If $S = 1$ is observed and we decide to reject H_0, we would also do so for $S = 0$ because $S = 0$ is *more extreme* than $S = 1$. By symmetry considerations,

$$P\text{-value} = P_0(S \le 1 \text{ or } S \ge 4) = 2(0.03125 + 0.15625) = 0.375.$$

This discussion motivates Definition 9 below. Suppose that the appropriate critical region for testing H_0 against H_1 is one-sided. That is, suppose that C is either of the form $\{T \ge c_1\}$ or $\{T \le c_2\}$, where T is the test statistic.

Definition 9. The probability of observing under H_0 a sample outcome at least as extreme as the one observed is called the *P-value*. The smaller the P-value, the more extreme the outcome and the stronger the evidence against H_0.

If α is given, we reject H_0 if $P \le \alpha$ and do not reject H_0 if $P > \alpha$. In the two-sided case when the critical region is of the form $C = \{|T(\mathbf{X})| > k\}$, the one-sided P-value is doubled to obtain the P-value. If the distribution of T is not symmetric, the P-value is not well defined in the two-sided case, although many authors recommend doubling the one-sided P-value.

PROBLEMS 9.2

1. A sample of size 1 is taken from a population distribution $P(\lambda)$. To test $H_0: \lambda = 1$ against $H_1: \lambda = 2$, consider the nonrandomized test $\varphi(x) = 1$ if $x > 3$, and $= 0$ if $x \le 3$. Find the probabilities of type I and type II errors and the power

of the test against $\lambda = 2$. If it is required to achieve a size equal to 0.05, how should one modify the test φ?

2. Let $X_1, X_2, \ldots, X_n$ be a sample from a population with finite mean μ and finite variance σ^2. Suppose that μ is not known but σ is known, and it is required to test $\mu = \mu_0$ against $\mu = \mu_1$ ($\mu_1 > \mu_0$). Let n be sufficiently large so that the central limit theorem holds, and consider the test

$$\varphi(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } \overline{x} > k, \\ 0 & \text{if } \overline{x} \le k, \end{cases}$$

where $\overline{x} = n^{-1}\sum_{i=1}^{n} x_i$. Find k such that the test has (approximately) size α. What is the power of this test at $\mu = \mu_1$? If the probabilities of type I and type II errors are fixed at α and β, respectively, find the smallest sample size needed.

3. In Problem 2, if σ is not known, find k such that the test φ has size α.

4. Let $X_1, X_2, \ldots, X_n$ be a sample from $\mathcal{N}(\mu, 1)$. For testing $\mu \le \mu_0$ against $\mu > \mu_0$, consider the test function

$$\varphi(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } \overline{x} > \mu_0 + \dfrac{z_\alpha}{\sqrt{n}}, \\ 0 & \text{if } \overline{x} \le \mu_0 + \dfrac{z_\alpha}{\sqrt{n}}. \end{cases}$$

Show that the power function of φ is a nondecreasing function of μ. What is the size of the test?

5. A sample of size 1 is taken from an exponential PDF with parameter θ, that is, $X \sim G(1, \theta)$. To test $H_0\colon \theta = 1$ against $H_1\colon \theta > 1$, the test to be used is the nonrandomized test

$$\varphi(x) = \begin{cases} 1 & \text{if } x > 2, \\ 0 & \text{if } x \le 2. \end{cases}$$

Find the size of the test. What is the power function?

6. Let $X_1, X_2, \ldots, X_n$ be a sample from $\mathcal{N}(0, \sigma^2)$. To test $H_0\colon \sigma = \sigma_0$ against $H_1 = \sigma \ne \sigma_0$, it is suggested that the test

$$\varphi(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } \sum x_i^2 > c_1 \text{ or } \sum x_i^2 < c_2, \\ 0 & \text{if } c_2 \le \sum x_i^2 \le c_1, \end{cases}$$

be used. How will you find c_1 and c_2 such that the size of φ is a preassigned number α, $0 < \alpha < 1$? What is the power function of this test?

7. An urn contains 10 marbles, of which M are white and $10 - M$ are black. To test that $M = 5$ against the alternative hypothesis that $M = 6$, one draws 3 marbles

from the urn without replacement. The null hypothesis is rejected if the sample contains 2 or 3 white marbles; otherwise, it is accepted. Find the size of the test and its power.

9.3 NEYMAN–PEARSON LEMMA

In this section we prove the fundamental lemma due to Neyman and Pearson [74], which gives a general method for finding a best (most powerful) test of a simple hypothesis against a simple alternative. Let $\{f_\theta, \theta \in \Theta\}$, where $\Theta = \{\theta_0, \theta_1\}$, be a family of possible distributions of $\mathbf{X}$. Also, f_θ represents the PDF of $\mathbf{X}$ if $\mathbf{X}$ is a continuous RV, and the PMF of $\mathbf{X}$ if $\mathbf{X}$ is of the discrete type. Let us write $f_0(\mathbf{x}) = f_{\theta_0}(\mathbf{x})$ and $f_1(\mathbf{x}) = f_{\theta_1}(\mathbf{x})$ for convenience.

Theorem 1 (Neyman–Pearson Fundamental Lemma)

(a) Any test φ of the form

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } f_1(\mathbf{x}) > k\, f_0(\mathbf{x}), \\ \gamma(\mathbf{x}) & \text{if } f_1(\mathbf{x}) = k\, f_0(\mathbf{x}), \\ 0 & \text{if } f_1(\mathbf{x}) < k\, f_0(\mathbf{x}), \end{cases} \tag{1}$$

for some $k \geq 0$ and $0 \leq \gamma(\mathbf{x}) \leq 1$, is most powerful of its size for testing $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$. If $k = \infty$, the test

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } f_0(\mathbf{x}) = 0, \\ 0 & \text{if } f_0(\mathbf{x}) > 0, \end{cases} \tag{2}$$

is most powerful of size 0 for testing H_0 against H_1.

(b) Given α, $0 \leq \alpha \leq 1$, there exists a test of form (1) or (2) with $\gamma(\mathbf{x}) = \gamma$ (a constant) for which $E_{\theta_0}\varphi(\mathbf{X}) = \alpha$.

Proof. Let φ be a test satisfying (1) and φ^* be any test with $E_{\theta_0}\varphi^*(\mathbf{X}) \leq E_{\theta_0}\varphi(\mathbf{X})$. In the continuous case

$$\int [\varphi(\mathbf{x}) - \varphi^*(\mathbf{x})][f_1(\mathbf{x}) - k\, f_0(\mathbf{x})]\, d\mathbf{x}$$

$$= \left(\int_{f_1 > kf_0} + \int_{f_1 < kf_0}\right) [\varphi(\mathbf{x}) - \varphi^*(\mathbf{x})][f_1(\mathbf{x}) - k\, f_0(\mathbf{x})]\, d\mathbf{x}.$$

For any $\mathbf{x} \in \{f_1(\mathbf{x}) > kf_0(\mathbf{x})\}$, $\varphi(\mathbf{x}) - \varphi^*(\mathbf{x}) = 1 - \varphi^*(\mathbf{x}) \geq 0$, so that the integrand is ≥ 0. For $\mathbf{x} \in \{f_1(\mathbf{x}) < kf_0(\mathbf{x})\}$, $\varphi(\mathbf{x}) - \varphi^*(\mathbf{x}) = -\varphi^*(\mathbf{x}) \leq 0$, so that the integrand is again ≥ 0. It follows that

$$\int [\varphi(\mathbf{x}) - \varphi^*(\mathbf{x})][f_1(\mathbf{x}) - k\ f_0(\mathbf{x})]\,d\mathbf{x}$$
$$= E_{\theta_1}\varphi(\mathbf{X}) - E_{\theta_1}\varphi^*(\mathbf{X}) - k(E_{\theta_0}\varphi(\mathbf{X}) - E_{\theta_0}\varphi^*(\mathbf{X})) \geq 0,$$

which implies that

$$E_{\theta_1}\varphi(\mathbf{X}) - E_{\theta_1}\varphi^*(\mathbf{X}) \geq k(E_{\theta_0}\varphi(\mathbf{X}) - E_{\theta_0}\varphi^*(\mathbf{X})) \geq 0$$

since $E_{\theta_0}\varphi^*(\mathbf{X}) \leq E_{\theta_0}\varphi(\mathbf{X})$.

If $k = \infty$, any test φ^* of size 0 must vanish on the set $\{f_0(\mathbf{x}) > 0\}$. We have

$$E_{\theta_1}\varphi(\mathbf{X}) - E_{\theta_1}\varphi^*(\mathbf{X}) = \int_{\{f_0(\mathbf{x})=0\}} [1 - \varphi^*(\mathbf{x})]f_1(\mathbf{x})\,d\mathbf{x} \geq 0.$$

The proof for the discrete case requires the usual change of integral by a sum throughout.

To prove (b) we need to restrict ourselves to the case where $0 < \alpha \leq 1$, since the MP size 0 test is given by (2). Let $\gamma(\mathbf{x}) = \gamma$, and let us compute the size of a test of form (1). We have

$$E_{\theta_0}\varphi(\mathbf{X}) = P_{\theta_0}\{f_1(\mathbf{X}) > kf_0(\mathbf{X})\} + \gamma P_{\theta_0}\{f_1(\mathbf{X}) = kf_0(\mathbf{X})\}$$
$$= 1 - P_{\theta_0}\{f_1(\mathbf{X}) \leq kf_0(\mathbf{X})\} + \gamma\ P_{\theta_0}\{f_1(\mathbf{X}) = kf_0(\mathbf{X})\}.$$

Since $P_{\theta_0}\{f_0(\mathbf{X}) = 0\} = 0$, we may rewrite $E_{\theta_0}\varphi(\mathbf{X})$ as

$$(3)\qquad E_{\theta_0}\varphi(\mathbf{X}) = 1 - P_{\theta_0}\left\{\frac{f_1(\mathbf{X})}{f_0(\mathbf{X})} \leq k\right\} + \gamma P_{\theta_0}\left\{\frac{f_1(\mathbf{X})}{f_0(\mathbf{X})} = k\right\}.$$

Given $0 < \alpha \leq 1$, we wish to find k and γ such that $E_{\theta_0}\varphi(\mathbf{X}) = \alpha$, that is,

$$(4)\qquad P_{\theta_0}\left\{\frac{f_1(\mathbf{X})}{f_0(\mathbf{X})} \leq k\right\} - \gamma P_{\theta_0}\left\{\frac{f_1(\mathbf{X})}{f_0(\mathbf{X})} \leq k\right\} = 1 - \alpha.$$

Note that

$$P_{\theta_0}\left\{\frac{f_1(\mathbf{X})}{f_0(\mathbf{X})} \leq k\right\}$$

is a DF so that it is a nondecreasing and right continuous function of k. If there exists a k_0 such that

$$P_{\theta_0}\left\{\frac{f_1(\mathbf{X})}{f_0(\mathbf{X})} \leq k_0\right\} = 1 - \alpha,$$

we choose $\gamma = 0$ and $k = k_0$. Otherwise, there exists a k_0 such that

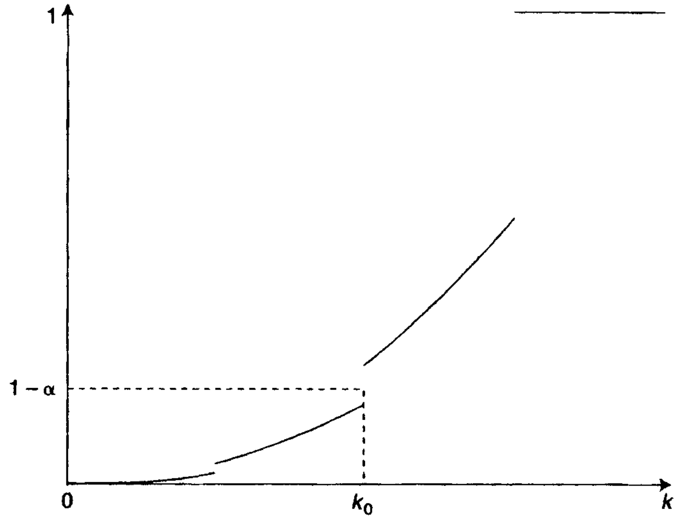

Fig. 1.

(5)
$$P_{\theta_0}\left\{\frac{f_1(\mathbf{X})}{f_0(\mathbf{X})} < k_0\right\} \le 1-\alpha < P_{\theta_0}\left\{\frac{f_1(\mathbf{X})}{f_0(\mathbf{X})} \le k_0\right\};$$

that is, there is a jump at k_0 (see Fig. 1). In this case we choose $k = k_0$ and

(6)
$$\gamma = \frac{P_{\theta_0}\{f_1(\mathbf{X})/f_0(\mathbf{X}) \le k_0\} - (1-\alpha)}{P_{\theta_0}\{f_1(\mathbf{X})/f_0(\mathbf{X}) = k_0\}}.$$

Since γ given by (6) satisfies (4), and $0 \le \gamma \le 1$, the proof is complete.

Remark 1. It is possible to show (see Problem 6) that the test given by (1) or (2) is unique (except on a null set), that is, if φ is an MP test of size α of H_0 against H_1, it must have form (1) or (2), except perhaps for a set A with $P_{\theta_0}(A) = P_{\theta_1}(A) = 0$.

Remark 2. An analysis of proof of part (a) of Theorem 1 shows that test (1) is MP even if f_1 and f_0 are not necessarily densities.

Theorem 2. If a sufficient statistic T exists for the family $\{f_\theta\colon \theta \in \Theta\}$, $\Theta = \{\theta_0, \theta_1\}$, the Neyman–Pearson MP test is a function of T.

The proof of this result is left as an exercise.

Remark 3. If the family $\{f_\theta : \theta \in \Theta\}$ admits a sufficient statistic, one can restrict attention to tests based on the sufficient statistic, that is, to tests that are functions of the sufficient statistic. If φ is a test function and T is a sufficient statistic, $E\{\varphi(\mathbf{X}) \mid$

$T\}$ is itself a test function, $0 \le E\{\varphi(\mathbf{X}) \mid T\} \le 1$, and

$$E_\theta\{E\{\varphi(\mathbf{X}) \mid T\}\} = E_\theta\varphi(\mathbf{X}),$$

so that φ and $E\{\varphi \mid T\}$ have the same power function.

Example 1. Let X be an RV with PMF under H_0 and H_1 given by

x	1	2	3	4	5	6
$f_0(x)$	0.01	0.01	0.01	0.01	0.01	0.95
$f_1(x)$	0.05	0.04	0.03	0.02	0.01	0.85

Then $\lambda(x) = f_1(x)/f_0(x)$ is given by

x	1	2	3	4	5	6
$\lambda(x)$	5	4	3	2	1	0.89

If $\alpha = 0.03$, for example, then Neyman–Pearson MP size 0.03 test rejects H_0 if $\lambda(X) \ge 3$, that is, if $X \le 3$ and has power

$$P_1(X \le 3) = 0.05 + 0.04 + 0.03 = 0.12$$

with $P(\text{type II error}) = 1 - 0.12 = 0.88$.

Example 2. Let $X \sim \mathcal{N}(0, 1)$ under H_0 and $X \sim C(1, 0)$ under H_1. To find an MP size α test of H_0 against H_1,

$$\begin{aligned}\lambda(x) = \frac{f_1(x)}{f_0(x)} &= \frac{(1/\pi)[1/(1+x^2)]}{(1/\sqrt{2\pi})e^{-x^2/2}}\\ &= \sqrt{\frac{2}{\pi}}\frac{e^{x^2/2}}{1+x^2}.\end{aligned}$$

Figure 2 gives a graph of $\lambda(x)$ and we note that λ has a maximum at $x = 0$ and two minima at $x = \pm 1$. Note that $\lambda(0) = 0.7979$ and $\lambda(\pm 1) = 0.6578$, so for $k \in (0.6578, 0.7989)$, $\lambda(x) = k$ intersects the graph at four points and the critical region is of the form $|X| \le k_1$ or $|X| \ge k_2$, where k_1 and k_2 are solutions of $\lambda(x) = k$. For $k = 0.7979$, the critical region is of the form $|X| \ge k_0$, where k_0 is the positive solution of $e^{-k_0^2/2} = 1 + k_0^2$, so that $k_0 \approx 1.59$ with $\alpha = 0.1118$. For $k < 0.6578$, $\alpha = 1$, and for $k = 0.6578$, the critical region is $|X| \ge 1$ with $\alpha = 0.3413$. For the traditional level $\alpha = 0.05$, the critical region is of the form $|X| \ge 1.96$.

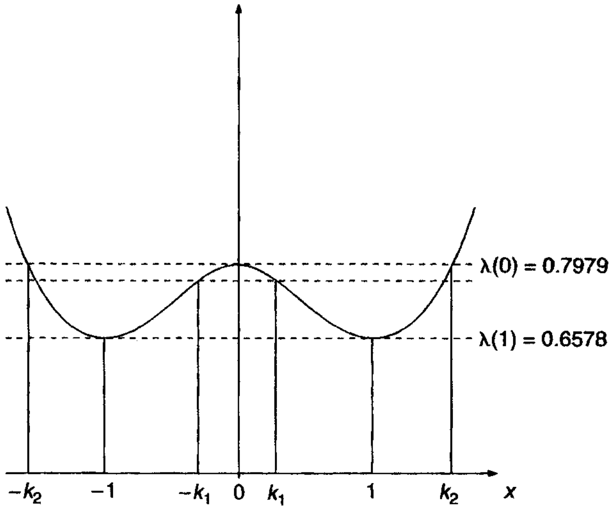

Fig. 2. Graph of $\lambda(x) = (2/\pi)^{1/2}[\exp(x^2/2)/(1+x^2)]$.

Example 3. Let $X_1, X_2, \dots, X_n$ be iid $b(1, p)$ RVs, and let $H_0\colon p = p_0$, $H_1\colon p = p_1, p_1 > p_0$. The MP size α test of H_0 against H_1 is of the form

$$\varphi(x_1, x_2, \dots, x_n) = \begin{cases} 1, & \lambda(\mathbf{x}) = \dfrac{p_1^{\sum x_i}(1-p_1)^{n-\sum x_i}}{p_0^{\sum x_i}(1-p_0)^{n-\sum x_i}} > k, \\ \gamma, & \lambda(\mathbf{x}) = k, \\ 0, & \lambda(\mathbf{x}) < k, \end{cases}$$

where k and γ are determined from

$$E_{p_0}\varphi(\mathbf{X}) = \alpha.$$

Now

$$\lambda(\mathbf{x}) = \left(\frac{p_1}{p_0}\right)^{\sum x_i}\left(\frac{1-p_1}{1-p_0}\right)^{n-\sum x_i},$$

and since $p_1 > p_0$, $\lambda(\mathbf{x})$ is an increasing function of $\sum x_i$. It follows that $\lambda(\mathbf{x}) > k$ if and only if $\sum x_i > k_1$, where k_1 is a constant. Thus the MP size α test is of the form

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } \sum x_i > k_1, \\ \gamma & \text{if } \sum x_i = k_1, \\ 0 & \text{otherwise.} \end{cases}$$

Also, k_1 and γ are determined from

$$\alpha = E_{p_0}\varphi(\mathbf{X}) = P_{p_0}\left\{\sum_1^n X_i > k_1\right\} + \gamma P_{p_0}\left\{\sum_1^n X_i = k_1\right\}$$

$$= \sum_{r=k_1+1}^{n} \binom{n}{r} p_0^r (1-p_0)^{n-r} + \gamma \binom{n}{k_1} p_0^{k_1}(1-p_0)^{n-k_1}.$$

Note that the MP size α test is independent of p_1 as long as $p_1 > p_0$; that is, it remains an MP size α test against any $p > p_0$ and is therefore a UMP test of $p = p_0$ against $p > p_0$.

In particular, let $n = 5$, $p_0 = \frac{1}{2}$, $p_1 = \frac{3}{4}$, and $\alpha = 0.05$. Then the MP test is given by

$$\varphi(\mathbf{x}) = \begin{cases} 1, & \sum x_i > k, \\ \gamma, & \sum x_i = k, \\ 0, & \sum x_i < k, \end{cases}$$

where k and γ are determined from

$$0.05 = \alpha = \sum_{k+1}^{5} \binom{5}{r}\left(\frac{1}{2}\right)^5 + \gamma \binom{5}{k}\left(\frac{1}{2}\right)^5.$$

It follows that $k = 4$ and $\gamma = 0.122$. Thus the MP size $\alpha = 0.05$ test is to reject $p = \frac{1}{2}$ in favor of $p = \frac{3}{4}$ if $\sum_1^n X_i = 5$ and reject $p = \frac{1}{2}$ with probability 0.122 if $\sum_1^n X_i = 4$.

It is simply a matter of reversing inequalities to see that the MP size α test of $H_0\colon p = p_0$ against $H_1\colon p = p_1$ $(p_1 < p_0)$ is given by

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } \sum x_i < k, \\ \gamma & \text{if } \sum x_i = k, \\ 0 & \text{if } \sum x_i > k, \end{cases}$$

where γ and k are determined from $E_{p_0}\varphi(\mathbf{X}) = \alpha$.

We note that $T(X) = \sum X_i$ is minimal sufficient for p, so that in view of Remark 3, we could have considered tests based only on T. Since $T \sim b(n, p)$,

$$\lambda(t) = \frac{f_1(t)}{f_0(t)} = \frac{\binom{n}{t} p_1^t (1-p_1)^{n-t}}{\binom{n}{t} p_0^t (1-p_0)^{n-t}} = \left(\frac{p_1}{p_0}\right)^t \left(\frac{1-p_1}{1-p_0}\right)^{n-t}$$

so that an MP test is of the same form as above but the computation is somewhat simpler.

We remark that in both cases ($p_1 > p_0$, $p_1 < p_0$) the MP test is quite intuitive. We would tend to accept the larger probability if a larger number of "successes" showed up, and the smaller probability if a smaller number of "successes" were observed. See, however, Example 2.

Example 4. Let $X_1, X_2, \dots, X_n$ be iid $\mathcal{N}(\mu, \sigma^2)$ RVs where both μ and σ^2 are unknown. We wish to test the null hypothesis H_0: $\mu = \mu_0$, $\sigma^2 = \sigma_0^2$ against the alternative H_1: $\mu = \mu_1$, $\sigma^2 = \sigma_0^2$. The fundamental lemma leads to the following MP test:

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } \lambda(\mathbf{x}) > k, \\ 0 & \text{if } \lambda(\mathbf{x}) < k, \end{cases}$$

where

$$\lambda(\mathbf{x}) = \frac{(1/\sigma_0\sqrt{2\pi})^n \exp\{-[\sum(x_i - \mu_1)^2/2\sigma_0^2]\}}{(1/\sigma_0\sqrt{2\pi})^n \exp\{-[\sum(x_i - \mu_0)^2/2\sigma_0^2]\}},$$

and k is determined from $E_{\mu_0,\sigma_0}\varphi(\mathbf{X}) = \alpha$. We have

$$\lambda(\mathbf{x}) = \exp\left[\sum x_i \left(\frac{\mu_1}{\sigma_0^2} - \frac{\mu_0}{\sigma_0^2}\right) + n\left(\frac{\mu_0^2}{2\sigma_0^2} - \frac{\mu_1^2}{2\sigma_0^2}\right)\right].$$

If $\mu_1 > \mu_0$, then

$$\lambda(\mathbf{x}) > k \quad \text{if and only if} \quad \sum_{i=1}^{n} x_i > k',$$

where k' is determined from

$$\alpha = P_{\mu_0,\sigma_0}\left\{\sum_{i=1}^{n} X_i > k'\right\} = P\left\{\frac{\sum X_i - n\mu_0}{\sqrt{n}\,\sigma_0} > \frac{k' - n\mu_0}{\sqrt{n}\,\sigma_0}\right\},$$

giving $k' = z_\alpha\sqrt{n}\,\sigma_0 + n\mu_0$. The case $\mu_1 < \mu_0$ is treated similarly. If σ_0 is known, the test determined above is independent of μ_1 as long as $\mu_1 > \mu_0$, and it follows that the test is UMP against H_1': $\mu > \mu_1, \sigma^2 = \sigma_0^2$. If, however, σ_0 is not known, that is, the null hypothesis is a composite hypothesis H_0'': $\mu = \mu_0$, $\sigma^2 > 0$ to be tested against the alternatives H_1'': $\mu = \mu_1$, $\sigma^2 > 0$ ($\mu_1 > \mu_0$), the MP test determined above depends on σ^2. In other words, an MP test against the alternative μ_1, σ_0^2 will not be MP against μ_1, σ_1^2, where $\sigma_1^2 \neq \sigma_0^2$.

PROBLEMS 9.3

1. A sample of size 1 is taken from PDF

$$f_\theta(x) = \begin{cases} \dfrac{2}{\theta^2}(\theta - x) & \text{if } 0 < x < \theta, \\ 0 & \text{otherwise.} \end{cases}$$

Find an MP test of $H_0: \theta = \theta_0$ against $H_1: \theta_1$ ($\theta_1 < \theta_0$).

2. Find the Neyman–Pearson size α test of $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ ($\theta_1 < \theta_0$), based on a sample of size 1 from the PDF

$$f_\theta(x) = 2\theta x + 2(1-\theta)(1-x), \qquad 0 < x < 1, \quad \theta \in [0, 1].$$

3. Find the Neyman–Pearson size α test of $H_0: \beta = 1$ against $H_1: \beta = \beta_1$ (> 1), based on a sample of size 1 from

$$f(x; \beta) = \begin{cases} \beta x^{\beta-1}, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

4. Find an MP size α test of $H_0: X \sim f_0(x)$, where $f_0(x) = (2\pi)^{-1/2}e^{-x^2/2}$, $-\infty < x < \infty$, against $H_1: X \sim f_1(x)$, where $f_1(x) = 2^{-1}e^{-|x|}$, $-\infty < x < \infty$, based on a sample of size 1.

5. For the PDF $f_\theta(x) = e^{-(x-\theta)}$, $x \geq \theta$, find an MP size α test of $\theta = \theta_0$ against $\theta = \theta_1$ ($> \theta_0$), based on a sample of size n.

6. If φ^* is an MP size α test of $H_0: \mathbf{X} \sim f_0(\mathbf{x})$ against $H_1: \mathbf{X} \sim f_1(\mathbf{x})$, show that it has to be either of form (1) or form (2) (except for a set of $\mathbf{x}$ that has probability 0 under H_0 and H_1).

7. Let φ^* be an MP size α ($0 < \alpha \leq 1$) test of H_0 against H_1, and let $k(\alpha)$ denote the value of k in (1). Show that if $\alpha_1 < \alpha_2$, then $k(\alpha_2) \leq k(\alpha_1)$.

8. For the family of Neyman–Pearson tests, show that the larger the α, the smaller the β (= P{type II error}).

9. Let $1 - \beta$ be the power of an MP size α test, where $0 < \alpha < 1$. Show that $\alpha < 1 - \beta$ unless $P_{\theta_0} = P_{\theta_1}$.

10. Let α be a real number, $0 < \alpha < 1$, and φ^* be an MP size α test of H_0 against H_1. Also, let $\beta = E_{H_1}\varphi^*(\mathbf{X}) < 1$. Show that $1 - \varphi^*$ is an MP test for testing H_1 against H_0 at level $1 - \beta$.

11. Let $X_1, X_2, \ldots, X_n$ be a random sample from the PDF

$$f_\theta(x) = \frac{\theta}{x^2} \qquad \text{if } 0 < \theta \leq x < \infty.$$

Find an MP test of $\theta = \theta_0$ against $\theta = \theta_1 (\neq \theta_0)$.

12. Let X be an observation in $(0, 1)$. Find an MP size α test of $H_0\colon X \sim f(x) = 4x$ if $0 < x < \frac{1}{2}$, and $= 4 - 4x$ if $\frac{1}{2} \le x < 1$, against $H_1\colon X \sim f(x) = 1$ if $0 < x < 1$. Find the power of your test.

13. In each of the following cases of simple versus simple hypotheses $H_0\colon X \sim f_0$, $H_1\colon X \sim f_1$, draw a graph of the ratio $\lambda(x) = f_1(x)/f_0(x)$ and find the form of the Neyman–Pearson test:

(a) $f_0(x) = \frac{1}{2}\exp(-|x+1|)$; $f_1(x) = \frac{1}{2}\exp(-|x-1|)$.

(b) $f_0(x) = \frac{1}{2}\exp(-|x|)$; $f_1(x) = 1/[\pi(1+x^2)]$.

(c) $f_0(x) = (1/\pi)[1 + (1+x)^2]^{-1}$; $f_1(x) = (1/\pi)[1 + (1-x)^2]^{-1}$.

14. Let $X_1, X_2, \ldots, X_n$ be a random sample with common PDF

$$f_\theta(x) = \frac{1}{2\theta}\exp\left(-\frac{|x|}{\theta}\right), \qquad x \in \mathcal{R}, \quad \theta > 0.$$

Find a size α MP test for testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$ ($> \theta_0$).

15. Let $X \sim f_j$, $j = 0, 1$, where

	x = 1	2	3	4	5
$f_0(x)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
$f_1(x)$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$

(a) Find the form of the MP test of its size.

(b) Find the size and the power of your test for various values of the cutoff point.

(c) Consider now a random sample of size n from f_0 under H_0 or f_1 under H_1. Find the form of the MP test of its size.

9.4 FAMILIES WITH MONOTONE LIKELIHOOD RATIO

In this section we consider the problem of testing one-sided hypotheses on a single real-valued parameter. Let $\{f_\theta, \theta \in \Theta\}$ be a family of PDFs (PMFs), $\Theta \subseteq \mathcal{R}$, and suppose that we wish to test $H_0\colon \theta \le \theta_0$ against the alternatives $H_1\colon \theta > \theta_0$ or its dual, $H_0'\colon \theta \ge \theta_0$, against $H_1'\colon \theta < \theta_0$. In general, it is not possible to find a UMP test for this problem. The MP test of $H_0\colon \theta \le \theta_0$, say, against the alternative $\theta = \theta_1$ ($> \theta_0$) depends on θ_1 and cannot be UMP. Here we consider a special class of distributions that is large enough to include the one-parameter exponential family, for which a UMP test of a one-sided hypothesis exists.

Definition 1. Let $\{f_\theta, \theta \in \Theta\}$ be a family of PDFs (PMFs), $\theta \subseteq \mathcal{R}$. We say that $\{f_\theta\}$ has a *monotone likelihood ratio* (MLR) in statistic $T(\mathbf{x})$ if for $\theta_1 < \theta_2$, whenever

f_{θ_1}, f_{θ_2} are distinct, the ratio $f_{\theta_2}(\mathbf{x})/f_{\theta_1}(\mathbf{x})$ is a nondecreasing function of $T(\mathbf{x})$ for the set of values $\mathbf{x}$ for which at least one of f_{θ_1} and f_{θ_2} is > 0.

It is also possible to define families of densities with nonincreasing MLR in $T(\mathbf{x})$, but such families can be treated by symmetry.

Example 1. Let $X_1, X_2, \ldots, X_n \sim U[0, \theta], \theta > 0$. The joint PDF of $X_1, \ldots, X_n$ is

$$f_\theta(\mathbf{x}) = \begin{cases} \dfrac{1}{\theta^n}, & 0 \le \max x_i \le \theta, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\theta_2 > \theta_1$ and consider the ratio

$$\begin{aligned} \frac{f_{\theta_2}(\mathbf{x})}{f_{\theta_1}(\mathbf{x})} &= \frac{(1/\theta_2^n) I_{[\max x_i \le \theta_2]}}{(1/\theta_1^n) I_{[\max x_i \le \theta_1]}} \\ &= \left(\frac{\theta_1}{\theta_2}\right)^n \frac{I_{[\max x_i \le \theta_2]}}{I_{[\max x_i \le \theta_1]}}. \end{aligned}$$

Let

$$R(\mathbf{x}) = I_{[\max x_i \le \theta_2]} / I_{[\max x_i \le \theta_1]} = \begin{cases} 1, & \max x_i \in [0, \theta_1], \\ \infty, & \max x_i \in [\theta_1, \theta_2]. \end{cases}$$

Define $R(\mathbf{x}) = \infty$ if $\max x_i > \theta_2$. It follows that $f_{\theta_2}/f_{\theta_1}$ is a nondecreasing function of $\max_{1 \le i \le n} x_i$, and the family of uniform densities on $[0, \theta]$ has an MLR in $\max_{1 \le i \le n} x_i$.

Theorem 1. The one-parameter exponential family

$$f_\theta(\mathbf{x}) = \exp[Q(\theta)T(\mathbf{x}) + S(\mathbf{x}) + D(\theta)], \tag{1}$$

where $Q(\theta)$ is nondecreasing, has an MLR in $T(\mathbf{x})$.

The proof is left as an exercise.

Remark 1. The nondecreasingness of $Q(\theta)$ can be obtained by a reparametrization, putting $\vartheta = Q(\theta)$, if necessary.

Theorem 1 includes normal, binomial, Poisson, gamma (one parameter fixed), beta (one parameter fixed), and so on. In Example 1 we have already seen that $U[0, \theta]$, which is not an exponential family, has an MLR.

Example 2. Let $X \sim \mathcal{C}(1, \theta)$. Then

$$\frac{f_{\theta_2}(x)}{f_{\theta_1}(x)} = \frac{1 + (x - \theta_1)^2}{1 + (x - \theta_2)^2} \to 1 \qquad \text{as } x \to \pm\infty,$$

and we see that $\mathcal{C}(1, \theta)$ does not have an MLR.

Theorem 2. Let $\mathbf{X} \sim f_\theta$, $\theta \in \Theta$, where $\{f_\theta\}$ has an MLR in $T(\mathbf{x})$. For testing $H_0: \theta \le \theta_0$ against $H_1: \theta > \theta_0$, $\theta_0 \in \Theta$, any test of the form

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } T(\mathbf{x}) > t_0, \\ \gamma & \text{if } T(\mathbf{x}) = t_0, \\ 1 & \text{if } T(\mathbf{x}) < t_0, \end{cases} \tag{2}$$

has a nondecreasing power function and is UMP of its size $E_{\theta_0}\varphi(\mathbf{X}) = \alpha$ (provided that the size is not 0).

Moreover, for every $0 \le \alpha \le 1$ and every $\theta_0 \in \Theta$, there exists a t_0, $-\infty \le t_0 \le \infty$, and $0 \le \gamma \le 1$ such that the test described in (2) is the UMP size α test of H_0 against H_1.

Proof. Let $\theta_1, \theta_2 \in \Theta$, $\theta_1 < \theta_2$. By the fundamental lemma, any test of the form

$$\varphi(\mathbf{x}) = \begin{cases} 1, & \lambda(\mathbf{x}) > k, \\ \gamma(\mathbf{x}), & \lambda(\mathbf{x}) = k, \\ 0, & \lambda(\mathbf{x}) < k, \end{cases} \tag{3}$$

where $\lambda(\mathbf{x}) = f_{\theta_2}(\mathbf{x})/f_{\theta_1}(\mathbf{x})$, is MP of its size for testing $\theta = \theta_1$ against $\theta = \theta_2$, provided that $0 \le k < \infty$; and if $k = \infty$, the test

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } f_{\theta_1}(\mathbf{x}) = 0, \\ 0 & \text{if } f_{\theta_1}(\mathbf{x}) > 0, \end{cases} \tag{4}$$

is MP of size 0. Since f_θ has an MLR in T, it follows that any test of form (2) is also of form (3), provided that $E_{\theta_1}\varphi(\mathbf{X}) > 0$, that is, provided that its size is > 0. The trivial test $\varphi'(\mathbf{x}) \equiv \alpha$ has size α and power α, so that the power of any test (2) is at least α, that is,

$$E_{\theta_2}\varphi(\mathbf{X}) \ge E_{\theta_2}\varphi'(\mathbf{X}) = \alpha = E_{\theta_1}\varphi(\mathbf{X}).$$

It follows that if $\theta_1 < \theta_2$ and $E_{\theta_1}\varphi(\mathbf{X}) > 0$, then $E_{\theta_1}\varphi(\mathbf{X}) \le E_{\theta_2}\varphi(\mathbf{X})$, as asserted.

Let $\theta_1 = \theta_0$ and $\theta_2 > \theta_0$, as above. We know that (2) is an MP test of its size $E_{\theta_0}\varphi(\mathbf{X})$ for testing $\theta = \theta_0$ against $\theta = \theta_2$ ($\theta_2 > \theta_0$), provided that $E_{\theta_0}\varphi(\mathbf{X}) > 0$. Since the power function of φ is nondecreasing,

$$E_\theta\varphi(\mathbf{X}) \le E_{\theta_0}\varphi(\mathbf{X}) = \alpha_0 \qquad \text{for all } \theta \le \theta_0. \tag{5}$$

Since, however, φ does not depend on θ_2 (it depends only on constants k and γ), it follows that φ is the UMP size α_0 test for testing $\theta = \theta_0$ against $\theta > \theta_0$. Thus φ is UMP among the class of tests φ'' for which

(6) $$E_{\theta_0}\varphi''(\mathbf{X}) \le E_{\theta_0}\varphi(\mathbf{X}) = \alpha_0.$$

Now the class of tests satisfying (5) is contained in the class of tests satisfying (6) [there are more restrictions in (5)]. It follows that φ, which is UMP in the larger class satisfying (6), must also be UMP in the smaller class satisfying (5). Thus, provided that $\alpha_0 > 0$, φ is the UMP size α_0 test for $\theta \le \theta_0$ against $\theta > \theta_0$.

We ask the reader to complete the proof of the final part of the theorem, using the fundamental lemma.

Remark 2. By interchanging inequalities throughout in Theorem 2, we see that this theorem also provides a solution of the dual problem $H_0'\colon \theta \ge \theta_0$ against $H_1'\colon \theta < \theta_0$.

Example 3. Let X have the hypergeometric PMF

$$P_M\{X = x\} = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \qquad x = 0, 1, 2, \ldots, M.$$

Since

$$\frac{P_{M+1}\{X = x\}}{P_M\{X = x\}} = \frac{M+1}{N-M}\,\frac{N-M-n+x}{M+1-x},$$

we see that $\{P_M\}$ has an MLR in $x(P_{M_2}/P_{M_1}$, where $M_2 > M_1$ is just a product of such ratios). It follows that there exists a UMP test of $H_0\colon M \le M_0$ against $H_1 : M > M_0$, which rejects H_0 when X is too large; that is, the UMP size α test is given by

$$\varphi(x) = \begin{cases} 1, & x > k, \\ \gamma, & x = k, \\ 0, & x < k, \end{cases}$$

where (integer) k and γ are determined from

$$E_{M_0}\varphi(X) = \alpha.$$

For the one-parameter exponential family, UMP tests also exist for some two-sided hypotheses of the form

(7) $$H_0\colon \theta \le \theta_1 \quad \text{or} \quad \theta \ge \theta_2 (\theta_1 < \theta_2).$$

We state the following result without proof.

Theorem 3. For the one-parameter exponential family (1), there exists a UMP test of the hypothesis $H_0: \theta \le \theta_1$ or $\theta \ge \theta_2$ ($\theta_1 < \theta_2$) against $H_1: \theta_1 < \theta < \theta_2$ that is of the form

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } c_1 < T(\mathbf{x}) < c_2, \\ \gamma_i & \text{if } T(\mathbf{x}) = c_i, \qquad i = 1, 2 \quad (c_1 < c_2), \\ 0 & \text{if } T(\mathbf{x}) < c_1 \text{ or } > c_2, \end{cases} \tag{8}$$

where the c's and the γ's are given by

$$E_{\theta_1}\varphi(\mathbf{X}) = E_{\theta_2}\varphi(\mathbf{X}) = \alpha. \tag{9}$$

See Lehmann [63, pp. 101–103] for proof.

Example 4. Let $X_1, X_2, \dots, X_n$ be iid $\mathcal{N}(\mu, 1)$ RVs. To test $H_0: \mu \le \mu_0$ or $\mu \ge \mu_1$ ($\mu_1 > \mu_0$) against $H_1: \mu_0 < \mu < \mu_1$, the UMP test is given by

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } c_1 < \sum_1^n x_i < c_2, \\ \gamma_i & \text{if } \sum x_i = c_1 \text{ or } c_2, \\ 0 & \text{if } \sum x_i < c_1 \text{ or } > c_2, \end{cases}$$

where we determine c_1, c_2 from

$$\alpha = P_{\mu_0}\{c_1 < \sum X_i < c_2\} = P_{\mu_1}\{c_1 < \sum X_i < c_2\}$$

and $\gamma_1 = \gamma_2 = 0$. Thus

$$\begin{aligned} \alpha &= P\left\{\frac{c_1 - n\mu_0}{\sqrt{n}} < \frac{\sum X_i - n\mu_0}{\sqrt{n}} < \frac{c_2 - n\mu_0}{\sqrt{n}}\right\} \\ &= P\left\{\frac{c_1 - n\mu_1}{\sqrt{n}} < \frac{\sum X_i - n\mu_1}{\sqrt{n}} < \frac{c_2 - n\mu_1}{\sqrt{n}}\right\} \\ &= P\left\{\frac{c_1 - n\mu_0}{\sqrt{n}} < Z < \frac{c_2 - n\mu_0}{\sqrt{n}}\right\} \\ &= P\left\{\frac{c_1 - n\mu_1}{\sqrt{n}} < Z < \frac{c_2 - n\mu_1}{\sqrt{n}}\right\}, \end{aligned}$$

where Z is $\mathcal{N}(0, 1)$. Given α, n, μ_0, and μ_1, we can solve for c_1 and c_2 from the simultaneous equations

$$\Phi\left(\frac{c_2 - n\mu_0}{\sqrt{n}}\right) - \Phi\left(\frac{c_1 - n\mu_0}{\sqrt{n}}\right) = \alpha,$$

$$\Phi\left(\frac{c_2 - n\mu_1}{\sqrt{n}}\right) - \Phi\left(\frac{c_1 - n\mu_1}{\sqrt{n}}\right) = \alpha,$$

where Φ is the DF of Z.

Remark 3. We caution the reader that UMP tests for testing $H_0\colon \theta_1 \le \theta \le \theta_2$ and $H_0'\colon \theta = \theta_0$ for the one-parameter exponential family do not exist. An example will suffice.

Example 5. Let $X_1, X_2, \ldots, X_n$ be a sample from $\mathcal{N}(0, \sigma^2)$. Since the family of joint PDFs of $\mathbf{X} = (X_1, \ldots, X_n)$ has an MLR in $T(\mathbf{X}) = \sum_1^n X_i^2$, it follows that UMP tests exist for one-sided hypotheses $\sigma \ge \sigma_0$ and $\sigma \le \sigma_0$.

Consider now the null hypotheses $H_0\colon \sigma = \sigma_0$ against the alternative $H_1\colon \sigma \ne \sigma_0$. We will show that a UMP test of H_0 does not exist. For testing $\sigma = \sigma_0$ against $\sigma > \sigma_0$, a test of the form

$$\varphi_1(\mathbf{x}) = \begin{cases} 1, & \sum x_i^2 > c_1, \\ 0, & \text{otherwise}, \end{cases}$$

is UMP, and for testing $\sigma = \sigma_0$ against $\sigma < \sigma_0$, a test of the form

$$\varphi_2(\mathbf{x}) = \begin{cases} 1, & \sum x_i^2 < c_2, \\ 0, & \text{otherwise}, \end{cases}$$

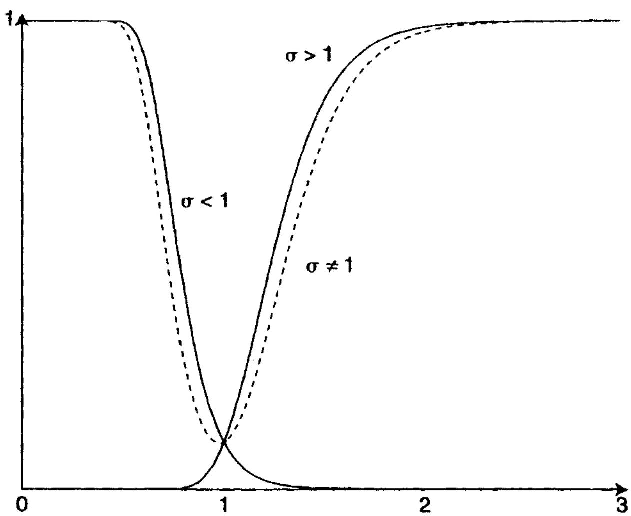

Fig. 1. Power functions of chi-square tests of $H_0\colon \sigma = \sigma_0$ against H_1.

is UMP. If the size is chosen as α, then $c_1 = \sigma_0^2 \chi^2_{n,\alpha}$ and $c_2 = \sigma_0^2 \chi^2_{n,1-\alpha}$. Clearly, neither φ_1 nor φ_2 is UMP for H_0 against $H_1: \sigma \neq \sigma_0$. The power of any test of H_0 for values $\sigma > \sigma_0$ cannot exceed that of φ_1, and for values of $\sigma < \sigma_0$ it cannot exceed the power of test φ_2. Hence no test of H_0 can be UMP (see Fig. 1).

PROBLEMS 9.4

1. For the following families of PMFs (PDFs) $f_\theta(x), \theta \in \Theta \subseteq \mathcal{R}$, find a UMP size α test of $H_0: \theta \le \theta_0$ against $H_1: \theta > \theta_0$, based on a sample of n observations:
   (a) $f_\theta(x) = \theta^x(1-\theta)^{1-x}, x = 0, 1; 0 < \theta < 1$.
   (b) $f_\theta(x) = (1/\sqrt{2\pi})\exp[-(x-\theta)^2/2], -\infty < x < \infty, -\infty < \theta < \infty$.
   (c) $f_\theta(x) = e^{-\theta}(\theta^x/x!), x = 0, 1, 2, \ldots; \theta > 0$.
   (d) $f_\theta(x) = (1/\theta)e^{-x/\theta}, x > 0, \theta > 0$.
   (e) $f_\theta(x) = [1/\Gamma(\theta)]x^{\theta-1}e^{-x}, x > 0, \theta > 0$.
   (f) $f_\theta(x) = \theta x^{\theta-1}, 0 < x < 1, \theta > 0$.

2. Let $X_1, X_2, \ldots, X_n$ be a sample of size n from the PMF

$$P_N(x) = \frac{1}{N}, \qquad x = 1, 2, \ldots, N; N \in \{1, 2, \ldots\}.$$

   (a) Show that the test

$$\varphi(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } \max(x_1, x_2, \ldots, x_n) > N_0, \\ \alpha & \text{if } \max(x_1, x_2, \ldots, x_n) \le N_0, \end{cases}$$

   is UMP size α for testing $H_0: N \le N_0$ against $H_1: N > N_0$.

   (b) Show that

$$\varphi(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } \max(x_1, x_2, \ldots, x_n) > N_0 \text{ or} \\ & \quad \max(x_1, x_2, \ldots, x_n) \le \alpha^{1/n} N_0, \\ 0 & \text{otherwise,} \end{cases}$$

   is a UMP size α test of $H_0': N = N_0$ against $H_1': N \neq N_0$.

3. Let $X_1, X_2, \ldots, X_n$ be a sample of size n from $U(0, \theta), \theta > 0$. Show that the test

$$\varphi_1(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } \max(x_1, \ldots, x_n) > \theta_0, \\ \alpha & \text{if } \max(x_1, x_2, \ldots, x_n) \le \theta_0, \end{cases}$$

   is UMP size α for testing $H_0: \theta \le \theta_0$ against $H_1: \theta > \theta_0$ and that the test

$$\varphi_2(x_1, x_2, \dots, x_n) = \begin{cases} 1 & \text{if } \max(x_1, \dots, x_n) > \theta_0 \text{ or} \\ & \quad \max(x_1, x_2, \dots, x_n) \le \theta_0 \alpha^{1/n}, \\ 0 & \text{otherwise,} \end{cases}$$

is UMP size α for $H_0'\colon \theta = \theta_0$ against $H_1'\colon \theta \neq \theta_0$.

4. Does the Laplace family of PDFs

$$f_\theta(x) = \tfrac{1}{2}\exp(-|x - \theta|), \qquad -\infty < x < \infty, \qquad \theta \in \mathcal{R},$$

possess an MLR?

5. Let X have logistic distribution with the PDF

$$f_\theta(x) = e^{-x-\theta}(1 + e^{-x-\theta})^{-2}, \qquad x \in \mathcal{R}.$$

Does $\{f_\theta\}$ belong to the exponential family? Does $\{f_\theta\}$ have MLR?

6. (a) Let f_θ be the PDF of a $\mathcal{N}(\theta, \theta)$ RV. Does $\{f_\theta\}$ have MLR?
   (b) Do the same as in part (a) if $X \sim \mathcal{N}(\theta, \theta^2)$.

9.5 UNBIASED AND INVARIANT TESTS

We have seen that if we restrict ourselves to the class Φ_α of all size α tests, there do not exist UMP tests for many important hypotheses. This suggests that we reduce the class of tests under consideration by imposing certain restrictions.

Definition 1. A size α test φ of $H_0\colon \theta \in \Theta_0$ against the alternatives $H_1\colon \theta \in \Theta_1$ is said to be *unbiased* if

(1) $$E_\theta \varphi(\mathbf{X}) \ge \alpha \qquad \text{for all } \theta \in \Theta_1.$$

It follows that a test φ is unbiased if and only if its power function $\beta_\varphi(\theta)$ satisfies

(2) $$\beta_\varphi(\theta) \le \alpha \qquad \text{for } \theta \in \Theta_0$$

and

(3) $$\beta_\varphi(\theta) \ge \alpha \qquad \text{for } \theta \in \Theta_1.$$

This seems to be a reasonable requirement to place on a test. An unbiased test rejects a false H_0 more often than a true H_0.

Definition 2. Let U_α be the class of all unbiased size α tests of H_0. If there exists a test $\varphi \in U_\alpha$ that has maximum power at each $\theta \in \Theta_1$, we call φ a *UMP unbiased* size α test.

Clearly, $U_\alpha \subset \Phi_\alpha$. If a UMP test exists in Φ_α, it is UMP in U_α. This follows by comparing the power of the UMP test with that of the trivial test $\varphi(\mathbf{x}) = \alpha$. It is convenient to introduce another class of tests.

Definition 3. A test φ is said to be *α-similar* on a subset Θ^* of Θ if

$$\beta_\varphi(\theta) = E_\theta\varphi(\mathbf{X}) = \alpha \qquad \text{for } \theta \in \Theta^*. \tag{4}$$

A test is said to be *similar* on a set $\Theta^* \subseteq \Theta$ if it is α-similar on Θ^* for some α, $0 \le \alpha \le 1$.

It is clear that there exists at least one similar test on every Θ^*, namely, $\varphi(x) \equiv \alpha$, $0 \le \alpha \le 1$.

Theorem 1. Let $\beta_\varphi(\theta)$ be continuous in θ for any φ. If φ is an unbiased size α test of $H_0: \theta \in \Theta_0$ against $H_1: \theta \in \Theta_1$, it is α-similar on the boundary $\Lambda = \overline{\Theta}_0 \cap \overline{\Theta}_1$. (Here $\overline{A}$ is the closure of set A.)

Proof. Let $\theta \in \Lambda$. Then there exists a sequence $\{\theta_n\}$, $\theta_n \in \Theta_0$, such that $\theta_n \to \theta$. Since $\beta_\varphi(\theta)$ is continuous, $\beta_\varphi(\theta_n) \to \beta_\varphi(\theta)$; and since $\beta_\varphi(\theta_n) \le \alpha$ for $\theta_n \in \Theta_0$, $\beta_\varphi(\theta) \le \alpha$. Similarly, there exists a sequence $\{\theta_n'\}$, $\theta_n' \in \Theta_1$, such that $\beta_\varphi(\theta_n') \ge \alpha$ (φ is unbiased) and $\theta_n' \to \theta$. Thus $\beta_\varphi(\theta_n') \to \beta_\varphi(\theta)$, and it follows that $\beta_\varphi(\theta) \ge \alpha$. Hence $\beta_\varphi(\theta) = \alpha$ for $\theta \in \Lambda$, and φ is α-similar on Λ.

Remark 1. Thus if $\beta_\varphi(\theta)$ is continuous in θ for any φ, an unbiased size α test of H_0 against H_1 is also α-similar for the PDFs (PMFs) of Λ, that is, for $\{f_\theta, \theta \in \Lambda\}$. If we can find an MP similar test of $H_0: \theta \in \Lambda$ against H_1, and if this test is unbiased size α, then necessarily it is MP in the smaller class.

Definition 4. A test φ that is UMP among all α-similar tests on the boundary $\Lambda = \overline{\Theta}_0 \cap \overline{\Theta}_1$ is said to be a *UMP α-similar* test.

It is frequently easier to find a UMP α-similar test. Moreover, tests that are UMP similar on the boundary are often UMP unbiased.

Theorem 2. Let the power function of every test φ of $H_0: \theta \in \Theta_0$ against $H_1: \theta \in \Theta_1$ be continuous in θ. Then a UMP α-similar test is UMP unbiased, provided that its size is α for testing H_0 against H_1.

Proof. Let φ_0 be UMP α-similar. Then $E_\theta\varphi_0(\mathbf{X}) \le \alpha$ for $\theta \in \Theta_0$. Comparing its power with that of the trivial similar test $\varphi(\mathbf{x}) \equiv \alpha$, we see that φ_0 is unbiased also. By the continuity of $\beta_\varphi(\theta)$, we see that the class of all unbiased size α tests is a subclass of the class of all α-similar tests. It follows that φ_0 is a UMP unbiased size α test.

Remark 2. The continuity of power function $\beta_\varphi(\theta)$ is not always easy to check, but sufficient conditions may be found in most advanced calculus texts (see, for example, Widder [116, p. 356]). If the family of the PDF (PMF) f_θ is an exponential family, a proof is given in Lehmann [63, p. 59].

Example 1. Let $X_1, X_2, \dots, X_n$ be a sample from $\mathcal{N}(\mu, 1)$. We wish to test $H_0\colon \mu \le 0$ against $H_1 : \mu > 0$. Since the family of densities has an MLR in $\sum_1^n X_i$, we can use Theorem 9.4.2 to conclude that a UMP test rejects H_0 if $\sum_1^n X_i > c$. This test is also UMP unbiased. Nevertheless, we use this example to illustrate the concepts introduced above.

Here $\Theta_0 = \{\mu \le 0\}$, $\Theta_1 = \{\mu > 0\}$, and $\Lambda = \overline{\Theta}_0 \cap \overline{\Theta}_1 = \{\mu = 0\}$. Since $T(\mathbf{X}) = \sum_{i=1}^n X_i$ is sufficient, we focus attention on tests based on T alone. Note that $T \sim \mathcal{N}(n\mu, n)$, which is one-parameter exponential. Thus the power function of any test φ based on T is continuous in μ. It follows that any unbiased size α test of H_0 has the property $\beta_\varphi(0) = \alpha$ of similarity over Λ. In order to use Theorem 2, we find a UMP test of $H_0' : \mu \in \Lambda$ against H_1. Let $\mu_1 > 0$. By the fundamental lemma, an MP test of $\mu = 0$ against $\mu = \mu_1 > 0$ is given by

$$\varphi(t) = \begin{cases} 1 & \text{if } \exp\left[\dfrac{t^2}{2n} - \dfrac{(t-n\mu)^2}{2n}\right] > k', \\ 0 & \text{otherwise}, \end{cases}$$

$$= \begin{cases} 1 & \text{if } t > k, \\ 0 & \text{if } t \le k, \end{cases}$$

where k is determined from

$$\alpha = P_0\{T > k\} = P\left\{Z > \frac{k}{\sqrt{n}}\right\}.$$

Thus $k = \sqrt{n}\, z_\alpha$. Since φ is independent of μ_1 as long as $\mu_1 > 0$, we see that the test

$$\varphi(t) = \begin{cases} 1, & t > \sqrt{n}\, z_\alpha, \\ 0, & \text{otherwise}, \end{cases}$$

is UMP α-similar. We need only check that φ is of the right size for testing H_0 against H_1. We have for $\mu \le 0$,

$$\begin{aligned} E_\mu \varphi(T) &= P_\mu\{T > \sqrt{n}\, z_\alpha\} \\ &= P\left\{\frac{T - n\mu}{\sqrt{n}} > z_\alpha - \sqrt{n}\,\mu\right\} \\ &\le P\{Z > z_\alpha\}, \end{aligned}$$

since $-\sqrt{n}\,\mu \geq 0$. Here Z is $\mathcal{N}(0, 1)$. It follows that

$$E_\mu \varphi(T) \leq \alpha \qquad \text{for } \mu \leq 0,$$

hence φ is UMP unbiased.

Theorem 2 can be used only if it is possible to find a UMP α-similar test. Unfortunately, this requires heavy use of conditional expectation, and we will not pursue the subject any further. We refer to Lehmann [63, Chaps. 4 and 5], and Ferguson [25, pp. 224–233], for further details.

Yet another reduction is obtained if we apply the principle of invariance to hypothesis-testing problems. We recall that a class of distributions is invariant under a group of transformations $\mathcal{G}$ if for every $g \in \mathcal{G}$ and every $\boldsymbol{\theta} \in \Theta$ there exists a unique $\boldsymbol{\theta}' \in \Theta$ such that $g(\mathbf{X})$ has distribution $P_{\boldsymbol{\theta}'}$, whenever $\mathbf{X} \sim P_{\boldsymbol{\theta}}$. We rewrite $\boldsymbol{\theta}' = \bar{g}\boldsymbol{\theta}$.

In a hypothesis-testing problem we need to reformulate the principle of invariance. First, we need to ensure that under transformations $\mathcal{G}$, not only does $\mathcal{P} = \{P_{\boldsymbol{\theta}} \colon \boldsymbol{\theta} \in \Theta\}$ remain invariant but also the problem of testing $H_0 \colon \boldsymbol{\theta} \in \Theta_0$ against $H_1 \colon \boldsymbol{\theta} \in \Theta_1$ remains invariant. Second, since the problem has not changed by application of $\mathcal{G}$, the decision also must not change.

Definition 5. A group $\mathcal{G}$ of transformations on the space of values of X leaves a hypothesis-testing problem *invariant* if $\mathcal{G}$ leaves both $\{P_{\boldsymbol{\theta}} \colon \boldsymbol{\theta} \in \Theta_0\}$ and $\{P_{\boldsymbol{\theta}} \colon \boldsymbol{\theta} \in \Theta_1\}$ invariant.

Definition 6. We say that φ is *invariant* under $\mathcal{G}$ if

$$\varphi(g(\mathbf{x})) = \varphi(\mathbf{x}) \qquad \text{for all } \mathbf{x} \text{ and all } g \in \mathcal{G}.$$

Definition 7. Let $\mathcal{G}$ be a group of transformations on the space of values of the RV $\mathbf{X}$. We say that a statistic $T(\mathbf{x})$ is *maximal invariant* under $\mathcal{G}$ if (a) T is invariant; (b) T is maximal, that is, $T(\mathbf{x}_1) = T(\mathbf{x}_2) \Rightarrow \mathbf{x}_1 = g(\mathbf{x}_2)$ for some $g \in \mathcal{G}$.

Example 2. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, and $\mathcal{G}$ be the group of translations

$$g_c(\mathbf{x}) = (x_1 + c, \ldots, x_n + c), \qquad -\infty < c < \infty.$$

Here the space of values of $\mathbf{X}$ is $\mathcal{R}_n$. Consider the statistic

$$T(\mathbf{x}) = (x_n - x_1, \ldots, x_n - x_{n-1}).$$

Clearly,

$$T(g_c(\mathbf{x})) = (x_n - x_1, \ldots, x_n - x_{n-1}) = T(\mathbf{x}).$$

If $T(\mathbf{x}) = T(\mathbf{x}')$, then $x_n - x_i = x_n' - x_i'$, $i = 1, 2, \dots, n-1$, and we have $x_i - x_i' = x_n - x_n' = c$ $(i = 1, 2, \dots, n-1)$; that is, $g_c(\mathbf{x}') = (x_1'+c, \dots, x_n'+c) = \mathbf{x}$ and T is maximal invariant.

Next consider the group of scale changes

$$g_c(\mathbf{x}) = (cx_1, \dots, cx_n), \qquad c > 0.$$

Then

$$T(\mathbf{x}) = \begin{cases} 0 & \text{if all } x_i = 0, \\ \left(\dfrac{x_1}{z}, \dots, \dfrac{x_n}{z}\right) & \text{if at least one } x_i \neq 0, \quad z = \left(\displaystyle\sum_1^n x_i^2\right)^{1/2}, \end{cases}$$

is maximal invariant; for

$$T(g_c(\mathbf{x})) = T(cx_1, \dots, cx_n) = T(\mathbf{x}),$$

and if $T(\mathbf{x}) = T(\mathbf{x}')$, then either $T(\mathbf{x}) = T(\mathbf{x}') = 0$, in which case $x_i = x_i' = 0$, or $T(\mathbf{x}) = T(\mathbf{x}') \neq 0$, in which case $x_i/z = x_i'/z'$, implying that $x_i' = (z'/z)x_i = cx_i$, and T is maximal.

Finally, if we consider the group of translation and scale changes,

$$g(\mathbf{x}) = (ax_1 + b, \dots, ax_n + b), \qquad a > 0, \quad -\infty < b < \infty,$$

a maximal invariant is

$$T(\mathbf{x}) = \begin{cases} 0 & \text{if } \beta = 0, \\ \left(\dfrac{x_1 - \overline{x}}{\beta}, \dfrac{x_2 - \overline{x}}{\beta}, \dots, \dfrac{x_n - \overline{x}}{\beta}\right) & \text{if } \beta \neq 0, \end{cases}$$

where $\overline{x} = n^{-1} \sum_1^n x_i$ and $\beta = n^{-1} \sum_1^n (x_i - \overline{x})^2$.

Definition 8. Let I_α denote the class of all invariant size α tests of $H_0: \boldsymbol{\theta} \in \Theta_0$ against $H_1: \boldsymbol{\theta} \in \Theta_1$. If there exists a UMP member in I_α, we call the test a *UMP invariant test* of H_0 against H_1.

The search for UMP invariant tests is greatly facilitated by use of the following result.

Theorem 3. Let $T(\mathbf{x})$ be maximal invariant with respect to $\mathcal{G}$. Then φ is invariant under $\mathcal{G}$ if and only if φ is a function of T.

Proof. Let φ be invariant. We have to show that $T(\mathbf{x}_1) = T(\mathbf{x}_2) \Rightarrow \varphi(\mathbf{x}_1) = \varphi(\mathbf{x}_2)$. If $T(\mathbf{x}_1) = T(\mathbf{x}_2)$, there is a $g \in \mathcal{G}$ such that $\mathbf{x}_1 = g(\mathbf{x}_2)$, so that $\varphi(\mathbf{x}_1) = \varphi(g(\mathbf{x}_2)) = \varphi(\mathbf{x}_2)$.

Conversely, if φ is a function of T, $\varphi(\mathbf{x}) = h[T(\mathbf{x})]$, then

$$\varphi(g(\mathbf{x})) = h[T(g(\mathbf{x}))] = h[T(\mathbf{x})] = \varphi(\mathbf{x}),$$

and φ is invariant.

Remark 3. The use of Theorem 3 is obvious. If a hypothesis-testing problem is invariant under a group $\mathcal{G}$, the *principle of invariance* restricts attention to invariant tests. According to Theorem 3, it suffices to restrict attention to test functions that are functions of maximal invariant T.

Example 3. Let $X_1, X_2, \ldots, X_n$ be a sample from $\mathcal{N}(\mu, \sigma^2)$, where both μ and σ^2 are unknown. We wish to test $H_0\colon \sigma \geq \sigma_0$, $-\infty < \mu < \infty$, against $H_1\colon \sigma < \sigma_0$, $-\infty < \mu < \infty$. The family $\{\mathcal{N}(\mu, \sigma^2)\}$ remains invariant under translations $x_i' = x_i + c$, $-\infty < c < \infty$. Moreover, since $\operatorname{var}(X + c) = \operatorname{var}(X)$, the hypothesis-testing problem remains invariant under the group of translations; that is, both $\{\mathcal{N}(\mu, \sigma^2)\colon \sigma^2 \geq \sigma_0^2\}$ and $\{\mathcal{N}(\mu, \sigma^2)\colon \sigma^2 < \sigma_0^2\}$ remain invariant. The joint sufficient statistic is $(\overline{X}, \sum(X_i - \overline{X})^2)$, which is transformed to $(\overline{X} + c, \sum(X_i - \overline{X})^2)$ under translations. A maximal invariant is $\sum(X_i - \overline{X})^2$. It follows that the class of invariant tests consists of tests that are functions of $\sum(X_i - \overline{X})^2$.

Now $\sum(X_i - \overline{X})^2/\sigma^2 \sim \chi^2(n-1)$, so that the PDF of $Z = \sum(X_i - \overline{X})^2$ is given by

$$f_{\sigma^2}(z) = \frac{\sigma^{-(n-1)}}{\Gamma[(n-1)/2]2^{(n-1)/2}} z^{(n-3)/2} e^{-z/2\sigma^2}, \qquad z > 0.$$

The family of densities $\{f_{\sigma^2}\colon \sigma^2 > 0\}$ has an MLR in z, and it follows that a UMP test is to reject $H_0\colon \sigma^2 \geq \sigma_0^2$ if $z \leq k$, that is, a UMP invariant test is given by

$$\varphi(\mathbf{x}) = \begin{cases} 1 & \text{if } \sum(x_i - \overline{x})^2 \leq k, \\ 0 & \text{if } \sum(x_i - \overline{x})^2 > k, \end{cases}$$

where k is determined from the size restriction

$$\alpha = P_{\sigma_0}\left\{\sum(X_i - \overline{X})^2 \leq k\right\} = P\left\{\frac{\sum(X_i - \overline{X})^2}{\sigma_0^2} \leq \frac{k}{\sigma_0^2}\right\},$$

that is,

$$k = \sigma_0^2 \chi^2_{n-1,1-\alpha}.$$

Example 4. Let $\mathbf{X}$ have PDF $f_i(x_1 - \theta, \ldots, x_n - \theta)$ under H_i $(i = 0, 1)$, $-\infty < \theta < \infty$. Let $\mathcal{G}$ be the group of translations

$$\mathbf{g}_c(\mathbf{x}) = (x_1 + c, \ldots, x_n + c), \qquad -\infty < c < \infty, \quad n \geq 2.$$

Clearly, g induces $\bar{g}$ on Θ, where $\bar{g}\theta = \theta + c$. The hypothesis-testing problem remains invariant under $\mathcal{G}$. A maximal invariant under $\mathcal{G}$ is $T(\mathbf{X}) = (X_1 - X_n, \ldots, X_{n-1} - X_n) = (T_1, T_2, \ldots, T_{n-1})$. The class of invariant tests coincides with the class of tests that are functions of T. The PDF of T under H_i is independent of θ and is given by $\int_{-\infty}^{\infty} f_i(t_1+z, \ldots, t_{n-1}+z, z)\,dz$. The problem is thus reduced to testing a simple hypothesis against a simple alternative. By the fundamental lemma the MP test

$$\varphi(t_1, t_2, \ldots, t_{n-1}) = \begin{cases} 1 & \text{if } \lambda(\mathbf{t}) > c, \\ 0 & \text{if } \lambda(\mathbf{t}) < c, \end{cases}$$

where $\mathbf{t} = (t_1, t_2, \ldots, t_{n-1})$ and

$$\lambda(\mathbf{t}) = \frac{\displaystyle\int_{-\infty}^{\infty} f_1(t_1 + z, \ldots, t_{n-1} + z, z)\,dz}{\displaystyle\int_{-\infty}^{\infty} f_0(t_1 + z, \ldots, t_{n-1} + z, z)\,dz}$$

is UMP invariant.

A particular case of Example 4 will be, for instance, to test $H_0\colon X \sim \mathcal{N}(\theta, 1)$ against $H_1\colon X \sim \mathcal{C}(1, \theta)$, $\theta \in \mathcal{R}$ (see Problem 1).

Example 5. Suppose that (X, Y) has joint PDF

$$f_{\boldsymbol{\theta}}(x, y) = \lambda\mu \exp(-\lambda x - \mu y), \qquad x > 0, \quad y > 0,$$

and $= 0$ elsewhere, where $\boldsymbol{\theta} = (\lambda, \mu)$, $\lambda > 0$, $\mu > 0$. Consider scale group $\mathcal{G} = \{\{0, c\}, c > 0\}$ which leaves $\{f_{\boldsymbol{\theta}}\}$ invariant. Suppose that we wish to test $H_0\colon \mu \geq \lambda$ against $H_1\colon \mu < \lambda$. It is easy to see that $\bar{\mathcal{G}}\Theta_0 = \Theta_0$, so that $\mathcal{G}$ leaves $(\alpha, \Theta_0, \Theta_1)$ invariant and $T = Y/X$ is maximal invariant. The PDF of T is given by

$$f_{\boldsymbol{\theta}}^T(t) = \frac{\lambda\mu}{(\lambda + \mu t)^2}, \qquad t > 0, \quad = 0 \text{ for } t < 0.$$

The family $\{f_\theta^T\}$ has MLR in T, and hence a UMP invariant test of H_0 is of the form

$$\varphi(t) = \begin{cases} 1, & t > c(\alpha), \\ \gamma, & t = c(\alpha), \\ 0, & t < c(\alpha), \end{cases}$$

where

$$\alpha = \int_{c(\alpha)}^{\infty} \frac{1}{(1 + t)^2}\,dt \Rightarrow c(\alpha) = \frac{1 - \alpha}{\alpha}.$$

PROBLEMS 9.5

1. To test $H_0\colon X \sim \mathcal{N}(\theta, 1)$, against $H_1\colon X \sim \mathcal{C}(1, \theta)$, a sample of size 2 is available on X. Find a UMP invariant test of H_0 against H_1.

2. Let $X_1, X_2, \ldots, X_n$ be a sample from $P(\lambda)$. Find a UMP unbiased size α test for the null hypothesis $H_0\colon \lambda \leq \lambda_0$ against alternatives $\lambda > \lambda_0$ by the methods of this section.

3. Let $X \sim NB(1; \theta)$. By the methods of this section, find a UMP unbiased size α test of $H_0\colon \theta \geq \theta_0$ against $H_1\colon \theta < \theta_0$.

4. Let $X_1, X_2, \ldots, X_n$ iid $\mathcal{N}(\mu, \sigma^2)$ RVs. Consider the problem of testing $H_0\colon \mu \leq 0$ against $H_1\colon \mu > 0$.
   (a) It suffices to restrict attention to sufficient statistic (U, V), where $U = \overline{X}$ and $V = S^2$. Show that the problem of testing H_0 is invariant under $\mathcal{G} = \{\{a, 1\}, a \in \mathcal{R}\}$ and a maximal invariant is $T = U/\sqrt{V}$.
   (b) Show that the distribution of T has MLR, and a UMP invariant test rejects H_0 when $T > c$.

5. Let $X_1, X_2, \ldots, X_n$ be iid RVs and let H_0 be that $X_i \sim \mathcal{N}(\theta, 1)$ and H_1 be that the common PDF is $f_\theta(x) = \frac{1}{2}\exp(-|x - \theta|)$. Find the form of the UMP invariant test of H_0 against H_1.

6. Let $X_1, X_2, \ldots, X_n$ be iid RVs and suppose that $H_0\colon X_i \sim \mathcal{N}(0, 1)$ and $H_1\colon X_i \sim f_1(x) = \exp(-|x|)/2$.
   (a) Show that the problem of testing H_0 against H_1 is invariant under scale changes $g_c(\mathbf{x}) = c\mathbf{x}, c > 0$ and a maximal invariant is $T(\mathbf{X}) = (X_1/X_n, \ldots, X_{n-1}/X_n)$.
   (b) Show that the MP invariant test reject H_0 when

$$\frac{\sqrt{1 + \sum_{i=1}^{n-1} Y_i^2}}{1 + \sum_{i=1}^{n-1} |Y_i|} < k,$$

where $Y_j = X_j/X_n$, $j = 1, 2, \ldots, n-1$, or equivalently, when

$$\frac{\left(\sum_{j=1}^{n} X_j^2\right)^{1/2}}{\sum_{j=1}^{n} |X_j|} < k.$$

9.6 LOCALLY MOST POWERFUL TESTS

In the preceding section we argued that whenever a UMP test does not exist, we restrict the class of tests under consideration and then find a UMP test in the subclass. Yet another approach when no UMP test exists is to restrict the parameter set to a subset of Θ_1. In most problems, the parameter values that are close to the null hypothesis are the hardest to detect. Tests that have good power properties for "local alternatives" may also retain good power properties for "nonlocal" alternatives.

Definition 1. Let $\Theta \subseteq \mathcal{R}$. Then a test φ_0 with power function $\beta_{\varphi_0}(\theta) = E_\theta\varphi_0(\mathbf{X})$ is said to be a *locally most powerful* (LMP) *test* of $H_0: \theta \le \theta_0$ against $H_1: \theta > \theta_0$ if there exists a $\Delta > 0$ such that for any other test φ with

$$\text{(1)} \qquad \beta_\varphi(\theta_0) = \beta_{\varphi_0}(\theta_0) = \int \varphi(\mathbf{x}) f_{\theta_0}(\mathbf{x})\, d\mathbf{x},$$

$$\text{(2)} \qquad \beta_{\varphi_0}(\theta) \ge \beta_\varphi(\theta) \qquad \text{for every } \theta \in (\theta_0, \theta_0 + \Delta].$$

We assume that the tests under consideration have continuously differentiable power function at $\theta = \theta_0$ and the derivative may be taken under the integral sign. In that case, an LMP test maximizes

$$\text{(3)} \qquad \frac{\partial}{\partial\theta}\beta_\varphi(\theta)\Big|_{\theta=\theta_0} = \beta'_\varphi(\theta)\Big|_{\theta=\theta_0} = \int \varphi(\mathbf{x}) \frac{\partial}{\partial\theta} f_\theta(\mathbf{x})\Big|_{\theta=\theta_0} d\mathbf{x}$$

subject to the size constraint (1). A slight extension of the Neyman–Pearson lemma (Remark 9.3.2) implies that a test satisfying (1) and given by

$$\text{(4)} \qquad \varphi_0(\mathbf{x}) = \begin{cases} 1 & \text{if } \dfrac{\partial}{\partial\theta} f_\theta(\mathbf{x})\Big|_{\theta_0} > k f_{\theta_0}(\mathbf{x}), \\ \gamma & \text{if} \dfrac{\partial}{\partial\theta} f_\theta(\mathbf{x})\Big|_{\theta_0} = k f_{\theta_0}(\mathbf{x}), \\ 0 & \text{if} \dfrac{\partial}{\partial\theta} f_\theta(\mathbf{x})\Big|_{\theta_0} < k f_{\theta_0}(\mathbf{x}) \end{cases}$$

will maximize $\beta'_\varphi(\theta_0)$. It is possible that a test that maximizes $\beta'_\varphi(\theta_0)$ is not LMP, but if the test maximizes $\beta'(\theta_0)$ and is unique, it must be an LMP test (see Kallenberg et al. [47, p. 290] and Lehmann [63, p. 528]).

Note that for $\mathbf{x}$ for which $f_{\theta_0}(\mathbf{x}) \ne 0$, we can write

$$\frac{\dfrac{\partial}{\partial\theta} f_\theta(\mathbf{x})\Big|_{\theta_0}}{f_{\theta_0}(\mathbf{x})} = \frac{\partial}{\partial\theta} \log f_\theta(\mathbf{x})\Big|_{\theta_0},$$

and we can rewrite

$$\varphi_0(\mathbf{x}) = \begin{cases} 1 & \text{if } \left.\dfrac{\partial}{\partial\theta}\log f_\theta(\mathbf{x})\right|_{\theta_0} > k, \\ \gamma & \text{if } \left.\dfrac{\partial}{\partial\theta}\log f_\theta(\mathbf{x})\right|_{\theta_0} = k, \\ 0 & \text{if } \left.\dfrac{\partial}{\partial\theta}\log f_\theta(\mathbf{x})\right|_{\theta_0} < k. \end{cases} \tag{5}$$

Example 1. Let $X_1, X_2, \ldots, X_n$ be iid with common normal PDF with mean μ and variance σ^2. If one of these parameters is unknown while the other is known, the family of PDFs has MLR, and UMP tests exist for one-sided hypotheses for the unknown parameter. Let us derive the LMP test in each case.

First consider the case when σ^2 is known, say $\sigma^2 = 1$ and $H_0: \mu \le 0$, $H_1: \mu > 0$. An easy computation shows that an LMP test is of the form

$$\varphi_0(\mathbf{x}) = \begin{cases} 1 & \text{if } \overline{x} > k, \\ 0 & \text{if } \overline{x} \le k, \end{cases}$$

which, of course, is the form of the UMP test obtained in Problem 9.4.1 by an application of Theorem 9.4.2.

Next consider the case when μ is known, say $\mu = 0$ and $H_0: \sigma \le \sigma_0$, $H_1: \sigma > \sigma_0$. Using (5), we see that an LMP test is of the form

$$\varphi_1(\mathbf{x}) = \begin{cases} 1 & \text{if } \sum_{i=1}^n x_i^2 > k, \\ 0 & \text{if } \sum_{i=1}^n x_i^2 \le k, \end{cases}$$

which coincides with the UMP test.

In each case the power function is differentiable and the derivatives may be taken inside the integral sign because the PDF is a one-parameter exponential type PDF.

Example 2. Let $X_1, X_2, \ldots, X_n$ be iid RVs with common PDF

$$f_\theta(x) = \frac{1}{\pi}\frac{1}{1+(x-\theta)^2}, \qquad x \in \mathcal{R},$$

and consider the problem of testing $H_0: \theta \le 0$ against $H_1: \theta > 0$.

In this case $\{f_\theta\}$ does not have MLR. A direct computation using the Neyman–Pearson lemma shows that an MP test of $\theta = 0$ against $\theta = \theta_1$, $\theta_1 > 0$, depends on θ_1 and hence cannot be MP for testing $\theta = 0$ against $\theta = \theta_2$, $\theta_2 \ne \theta_1$. Hence a UMP test of H_0 against H_1 does not exist. An LMP test of H_0 against H_1 is of the form

$$\varphi_0(\mathbf{x}) = \begin{cases} 1 & \text{if } \displaystyle\sum_{i=1}^n \frac{2x_i}{1+x_i^2} > k, \\ 0 & \text{otherwise,} \end{cases}$$

where k is chosen so that the size of φ_0 is α. For small n it is hard to compute k but for large n it is easy to compute k using the central limit theorem. Indeed, $X_i/(1+X_i^2)$ are iid RVs with mean 0 and finite variance $(=\frac{3}{8})$, so that $k = z_\alpha\sqrt{n/2}$ will give an (approximate) level α test for large n.

The test φ_0 is good at detecting small departures from $\theta \le 0$, but it is quite unsatisfactory in detecting values of θ away from 0. In fact, for $\alpha < \frac{1}{2}$, $\beta_{\varphi_0}(\theta) \to 0$ as $\theta \to \infty$.

This procedure for finding locally best tests has applications in nonparametric statistics. We refer the reader to Randles and Wolfe [83, Sec. 9.1] for details.

PROBLEMS 9.6

1. Let $X_1, X_2, \dots, X_n$ be iid $C(1,\theta)$ RVs. Show that $E_0(1+X_1^2)^{-k} = (1/\pi)B(k+\frac{1}{2}, \frac{1}{2})$. Hence or otherwise, show that

$$E_0\left[\frac{X_1^2}{(1+X_1^2)^2}\right] = \text{var}\left(\frac{X_1}{1+X_1^2}\right) = \frac{1}{8}.$$

2. Let $X_1, X_2, \dots, X_n$ be a random sample from the logistic PDF

$$f_\theta(x) = \frac{1}{2[1+\cosh(x-\theta)]} = \frac{e^{x-\theta}}{(1+e^{x-\theta})^2}.$$

Show that the LMP test of $H_0\colon \theta = 0$ against $H_1\colon \theta > 0$ rejects H_0 if $\sum_{i=1}^n \tanh(x_i/2) > k$.

3. Let $X_1, X_2, \dots, X_n$ be iid RVs with the common Laplace PDF

$$f_\theta(x) = \tfrac{1}{2}\exp(-|x-\theta|).$$

For $n \ge 2$, show that a UMP size α $(0 < \alpha < 1)$ test of $H_0\colon \theta \le 0$ against $H_1\colon \theta > 0$ does not exist. Find the form of the LMP test.