Chapter 18

TEST OF STATISTICAL HYPOTHESES FOR PARAMETERS

18.1. Introduction

Inferential statistics consists of estimation and hypothesis testing. We have already discussed various methods of finding point and interval estimators of parameters. We have also examined the goodness of an estimator.

Suppose $X_1, X_2, ..., X_n$ is a random sample from a population with probability density function given by

$$f(x;\theta) = \begin{cases} (1+\theta)\, x^{\theta} & \text{for } 0 < x < 1 \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is an unknown parameter. Further, let $n = 4$ and suppose $x_1 = 0.92$, $x_2 = 0.75$, $x_3 = 0.85$, $x_4 = 0.8$ is a set of random sample data from the above distribution. If we apply the maximum likelihood method, then we will find that the estimator $\widehat{\theta}$ of θ is

$$\widehat{\theta} = -1 - \frac{4}{\ln(X_1) + \ln(X_2) + \ln(X_3) + \ln(X_2)}.$$

Hence, the maximum likelihood estimate of θ is

$$\begin{aligned} \widehat{\theta} &= -1 - \frac{4}{\ln(0.92) + \ln(0.75) + \ln(0.85) + \ln(0.80)} \\ &= -1 + \frac{4}{0.7567} = 4.2861 \end{aligned}$$

Therefore, the corresponding probability density function of the population is given by

$$f(x) = \begin{cases} 5.2861\, x^{4.2861} & \text{for } 0 < x < 1 \\ \\ 0 & \text{otherwise.} \end{cases}$$

Since, the point estimate will rarely equal to the true value of θ, we would like to report a range of values with some degree of confidence. If we want to report an interval of values for θ with a confidence level of 90%, then we need a 90% confidence interval for θ. If we use the pivotal quantity method, then we will find that the confidence interval for θ is

$$\left[-1 - \frac{\chi^2_{\frac{\alpha}{2}}(8)}{2\sum_{i=1}^4 \ln X_i}, \quad -1 - \frac{\chi^2_{1-\frac{\alpha}{2}}(8)}{2\sum_{i=1}^4 \ln X_i} \right].$$

Since $\chi^2_{0.05}(8) = 2.73$, $\chi^2_{0.95}(8) = 15.51$, and $\sum_{i=1}^4 \ln(x_i) = -0.7567$, we obtain

$$\left[-1 + \frac{2.73}{2(0.7567)}, \quad -1 + \frac{15.51}{2(0.7567)} \right]$$

which is

$$[\,0.803, \quad 9.249\,].$$

Thus we may draw inference, at a 90% confidence level, that the population X has the distribution

$$f(x;\theta) = \begin{cases} (1+\theta)\, x^{\theta} & \text{for } 0 < x < 1 \\ \\ 0 & \text{otherwise,} \end{cases} \tag{$\star$}$$

where $\theta \in [0.803, 9.249]$. If we think carefully, we will notice that we have made one assumption. The assumption is that the observable quantity X can be modeled by a density function as shown in ($\star$). Since, we are concerned with the parametric statistics, our assumption is in fact about θ.

Based on the sample data, we found that an interval estimate of θ at a 90% confidence level is $[0.803, 9.249]$. But, we assumed that $\theta \in [0.803, 9.249]$. However, we can not be sure that our assumption regarding the parameter is real and is not due to the chance in the random sampling process. The validation of this assumption can be done by the hypothesis test. In this chapter, we discuss testing of statistical hypotheses. Most of the ideas regarding the hypothesis test came from Jerry Neyman and Karl Pearson during 1928-1938.

Definition 18.1. A *statistical hypothesis* H is a conjecture about the distribution $f(x;\theta)$ of a population X. This conjecture is usually about the parameter θ if one is dealing with a parametric statistics; otherwise it is about the form of the distribution of X.

Definition 18.2. A hypothesis H is said to be a *simple hypothesis* if H completely specifies the density $f(x; \theta)$ of the population; otherwise it is called a *composite hypothesis*.

Definition 18.3. The hypothesis to be tested is called the null hypothesis. The negation of the null hypothesis is called the alternative hypothesis. The null and alternative hypotheses are denoted by H_o and H_a, respectively.

If θ denotes a population parameter, then the general format of the null hypothesis and alternative hypothesis is

$$H_o : \theta \in \Omega_o \qquad \text{and} \qquad H_a : \theta \in \Omega_a \tag{$\star$}$$

where Ω_o and Ω_a are subsets of the parameter space Ω with

$$\Omega_o \cap \Omega_a = \emptyset \qquad \text{and} \qquad \Omega_o \cup \Omega_a \subseteq \Omega.$$

Remark 18.1. If $\Omega_o \cup \Omega_a = \Omega$, then $(\star)$ becomes

$$H_o : \theta \in \Omega_o \qquad \text{and} \qquad H_a : \theta \notin \Omega_o.$$

If Ω_o is a singleton set, then H_o reduces to a simple hypothesis. For example, $\Omega_o = \{4.2861\}$, the null hypothesis becomes $H_o : \theta = 4.2861$ and the alternative hypothesis becomes $H_a : \theta \neq 4.2861$. Hence, the null hypothesis $H_o : \theta = 4.2861$ is a simple hypothesis and the alternative $H_a : \theta \neq 4.2861$ is a composite hypothesis.

Definition 18.4. A *hypothesis test* is an ordered sequence

$$(X_1, X_2, ..., X_n; H_o, H_a; C)$$

where $X_1, X_2, ..., X_n$ is a random sample from a population X with the probability density function $f(x; \theta)$, H_o and H_a are hypotheses concerning the parameter θ in $f(x; \theta)$, and C is a Borel set in $\mathbb{R}^n$.

Remark 18.2. Borel sets are defined using the notion of σ-algebra. A collection of subsets $\mathcal{A}$ of a set S is called a σ-algebra if (i) $S \in \mathcal{A}$, (ii) $A^c \in \mathcal{A}$, whenever $A \in \mathcal{A}$, and (iii) $\bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$, whenever $A_1, A_2, ..., A_n, ... \in \mathcal{A}$. The Borel sets are the member of the smallest σ-algebra containing all open sets of $\mathbb{R}^n$. Two examples of Borel sets in $\mathbb{R}^n$ are the sets that arise by countable union of closed intervals in $\mathbb{R}^n$, and countable intersection of open sets in $\mathbb{R}^n$.

The set C is called the critical region in the hypothesis test. The critical region is obtained using a *test statistics* $W(X_1, X_2, ..., X_n)$. If the outcome of $(X_1, X_2, ..., X_n)$ turns out to be an element of C, then we decide to accept H_a; otherwise we accept H_o.

Broadly speaking, a hypothesis test is a rule that tells us for which sample values we should decide to accept H_o as true and for which sample values we should decide to reject H_o and accept H_a as true. Typically, a hypothesis test is specified in terms of a test statistics W. For example, a test might specify that H_o is to be rejected if the sample total $\sum_{k=1}^n X_k$ is less than 8. In this case the critical region C is the set $\{(x_1, x_2, ..., x_n) \,|\, x_1 + x_2 + \cdots + x_n < 8\}$.

18.2. A Method of Finding Tests

There are several methods to find test procedures and they are: (1) Likelihood Ratio Tests, (2) Invariant Tests, (3) Bayesian Tests, and (4) Union-Intersection and Intersection-Union Tests. In this section, we only examine likelihood ratio tests.

Definition 18.5. The *likelihood ratio test statistic* for testing the simple null hypothesis $H_o : \theta \in \Omega_o$ against the composite alternative hypothesis $H_a : \theta \notin \Omega_o$ based on a set of random sample data $x_1, x_2, ..., x_n$ is defined as

$$W(x_1, x_2, ..., x_n) = \frac{\max\limits_{\theta \in \Omega_o} L(\theta, x_1, x_2, ..., x_n)}{\max\limits_{\theta \in \Omega} L(\theta, x_1, x_2, ..., x_n)},$$

where Ω denotes the parameter space, and $L(\theta, x_1, x_2, ..., x_n)$ denotes the likelihood function of the random sample, that is

$$L(\theta, x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f(x_i; \theta).$$

A *likelihood ratio test* (LRT) is any test that has a critical region C (that is, rejection region) of the form

$$C = \{(x_1, x_2, ..., x_n) \mid W(x_1, x_2, ..., x_n) \leq k\},$$

where k is a number in the unit interval $[0, 1]$.

If $H_o : \theta = \theta_0$ and $H_a : \theta = \theta_a$ are both simple hypotheses, then the likelihood ratio test statistic is defined as

$$W(x_1, x_2, ..., x_n) = \frac{L(\theta_o, x_1, x_2, ..., x_n)}{L(\theta_a, x_1, x_2, ..., x_n)}.$$

Now we give some examples to illustrate this definition.

Example 18.1. Let X_1, X_2, X_3 denote three independent observations from a distribution with density

$$f(x;\theta) = \begin{cases} (1+\theta)\, x^{\theta} & \text{if } 0 \leq x \leq 1 \\ \\ 0 & \text{otherwise.} \end{cases}$$

What is the form of the LRT critical region for testing $H_o : \theta = 1$ versus $H_a : \theta = 2$?

Answer: In this example, $\theta_o = 1$ and $\theta_a = 2$. By the above definition, the form of the critical region is given by

$$\begin{aligned} C &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \;\middle|\; \frac{L\,(\theta_o, x_1, x_2, x_3)}{L\,(\theta_a, x_1, x_2, x_3)} \leq k \right\} \\ &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \;\middle|\; \frac{(1+\theta_o)^3 \, \prod_{i=1}^3 x_i^{\theta_o}}{(1+\theta_a)^3 \, \prod_{i=1}^3 x_i^{\theta_a}} \leq k \right\} \\ &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \;\middle|\; \frac{8 x_1 x_2 x_3}{27 x_1^2 x_2^2 x_3^2} \leq k \right\} \\ &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \;\middle|\; \frac{1}{x_1 x_2 x_3} \leq \frac{27}{8} k \right\} \\ &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \;\mid\; x_1 x_2 x_3 \geq a, \right\} \end{aligned}$$

where a is some constant. Hence the likelihood ratio test is of the form: "Reject H_o if $\displaystyle\prod_{i=1}^{3} X_i \geq a$."

Example 18.2. Let $X_1, X_2, ..., X_{12}$ be a random sample from a normal population with mean zero and variance σ^2. What is the form of the LRT critical region for testing the null hypothesis $H_o : \sigma^2 = 10$ versus $H_a : \sigma^2 = 5$?

Answer: Here $\sigma_o^2 = 10$ and $\sigma_a^2 = 5$. By the above definition, the form of the critical region is given by (with ${\sigma_o}^2 = 10$ and ${\sigma_a}^2 = 5$)

$$
\begin{aligned}
C &= \left\{ (x_1, x_2, ..., x_{12}) \in \mathbb{R}^{12} \;\middle|\; \frac{L\left({\sigma_o}^2, x_1, x_2, ..., x_{12}\right)}{L\left({\sigma_a}^2, x_1, x_2, ..., x_{12}\right)} \leq k \right\} \\
&= \left\{ (x_1, x_2, ..., x_{12}) \in \mathbb{R}^{12} \;\middle|\; \prod_{i=1}^{12} \frac{\frac{1}{\sqrt{2\pi\sigma_o^2}}\, e^{-\frac{1}{2}\left(\frac{x_i}{\sigma_o}\right)^2}}{\frac{1}{\sqrt{2\pi\sigma_a^2}}\, e^{-\frac{1}{2}\left(\frac{x_i}{\sigma_a}\right)^2}} \leq k \right\} \\
&= \left\{ (x_1, x_2, ..., x_{12}) \in \mathbb{R}^{12} \;\middle|\; \left(\frac{1}{2}\right)^6 e^{\frac{1}{20}\sum_{i=1}^{12} x_i^2} \leq k \right\} \\
&= \left\{ (x_1, x_2, ..., x_{12}) \in \mathbb{R}^{12} \;\middle|\; \sum_{i=1}^{12} x_i^2 \leq a \right\},
\end{aligned}
$$

where a is some constant. Hence the likelihood ratio test is of the form: "Reject H_o if $\displaystyle\sum_{i=1}^{12} X_i^2 \leq a$."

Example 18.3. Suppose that X is a random variable about which the hypothesis $H_o : X \sim UNIF(0,1)$ against $H_a : X \sim N(0,1)$ is to be tested. What is the form of the LRT critical region based on one observation of X?

Answer: In this example, $L_o(x) = 1$ and $L_a(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$. By the above definition, the form of the critical region is given by

$$
\begin{aligned}
C &= \left\{ x \in \mathbb{R} \;\middle|\; \frac{L_o(x)}{L_a(x)} \leq k \right\}, \qquad \text{where } k \in [0, \infty) \\
&= \left\{ x \in \mathbb{R} \;\middle|\; \sqrt{2\pi}\, e^{\frac{1}{2}x^2} \leq k \right\} \\
&= \left\{ x \in \mathbb{R} \;\middle|\; x^2 \leq 2 \ln\left(\frac{k}{\sqrt{2\pi}}\right) \right\} \\
&= \{ x \in \mathbb{R} \mid x \leq a, \}
\end{aligned}
$$

where a is some constant. Hence the likelihood ratio test is of the form: "Reject H_o if $X \leq a$."

In the above three examples, we have dealt with the case when null as well as alternative were simple. If the null hypothesis is simple (for example, $H_o : \theta = \theta_o$) and the alternative is a composite hypothesis (for example, $H_a : \theta \neq \theta_o$), then the following algorithm can be used to construct the likelihood ratio critical region:

(1) Find the likelihood function $L(\theta, x_1, x_2, ..., x_n)$ for the given sample.

(2) Find $L(\theta_o, x_1, x_2, ..., x_n)$.

(3) Find $\max_{\theta \in \Omega} L(\theta, x_1, x_2, ..., x_n)$.

(4) Rewrite $\frac{L(\theta_o, x_1, x_2, ..., x_n)}{\max_{\theta \in \Omega} L(\theta, x_1, x_2, ..., x_n)}$ in a "suitable form".

(5) Use step (4) to construct the critical region.

Now we give an example to illustrate these steps.

Example 18.4. Let X be a single observation from a population with probability density

$$f(x;\theta) = \begin{cases} \frac{\theta^x \, e^{-\theta}}{x!} & \text{for } x = 0, 1, 2, ..., \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta \geq 0$. Find the likelihood ratio critical region for testing the null hypothesis $H_o : \theta = 2$ against the composite alternative $H_a : \theta \neq 2$.

Answer: The likelihood function based on one observation x is

$$L(\theta, x) = \frac{\theta^x \, e^{-\theta}}{x!}.$$

Next, we find $L(\theta_o, x)$ which is given by

$$L(2, x) = \frac{2^x \, e^{-2}}{x!}.$$

Our next step is to evaluate $\max_{\theta \geq 0} L(\theta, x)$. For this we differentiate $L(\theta, x)$ with respect to θ, and then set the derivative to 0 and solve for θ. Hence

$$\frac{dL(\theta, x)}{d\theta} = \frac{1}{x!} \left[e^{-\theta} \, x \, \theta^{x-1} - \theta^x \, e^{-\theta} \right]$$

and $\frac{dL(\theta,x)}{d\theta} = 0$ gives $\theta = x$. Hence

$$\max_{\theta \geq 0} L(\theta, x) = \frac{x^x \, e^{-x}}{x!}.$$

To do the step (4), we consider

$$\frac{L(2, x)}{\max_{\theta \in \Omega} L(\theta, x)} = \frac{\frac{2^x \, e^{-2}}{x!}}{\frac{x^x \, e^{-x}}{x!}}$$

which simplifies to

$$\frac{L(2,x)}{\underset{\theta\in\Omega}{\max} L(\theta,x)} = \left(\frac{2e}{x}\right)^x e^{-2}.$$

Thus, the likelihood ratio critical region is given by

$$C = \left\{ x \in \mathbb{R} \;\middle|\; \left(\frac{2e}{x}\right)^x e^{-2} \leq k \right\} = \left\{ x \in \mathbb{R} \;\middle|\; \left(\frac{2e}{x}\right)^x \leq a \right\}$$

where a is some constant. The likelihood ratio test is of the form: "Reject H_o if $\left(\frac{2e}{X}\right)^X \leq a$."

So far, we have learned how to find tests for testing the null hypothesis against the alternative hypothesis. However, we have not considered the goodness of these tests. In the next section, we consider various criteria for evaluating the goodness of an hypothesis test.

18.3. Methods of Evaluating Tests

There are several criteria to evaluate the goodness of a test procedure. Some well known criteria are: (1) Powerfulness, (2) Unbiasedness and Invariancy, and (3) Local Powerfulness. In order to examine some of these criteria, we need some terminologies such as error probabilities, power functions, type I error, and type II error. First, we develop these terminologies.

A statistical hypothesis is a conjecture about the distribution $f(x;\theta)$ of the population X. This conjecture is usually about the parameter θ if one is dealing with a parametric statistics; otherwise it is about the form of the distribution of X. If the hypothesis completely specifies the density $f(x;\theta)$ of the population, then it is said to be a simple hypothesis; otherwise it is called a composite hypothesis. The hypothesis to be tested is called the null hypothesis. We often hope to reject the null hypothesis based on the sample information. The negation of the null hypothesis is called the alternative hypothesis. The null and alternative hypotheses are denoted by H_o and H_a, respectively.

In hypothesis test, the basic problem is to decide, based on the sample information, whether the null hypothesis is true. There are four possible situations that determines our decision is correct or in error. These four situations are summarized below:

	H_o is true	H_o is false
Accept H_o	Correct Decision	Type II Error
Reject H_o	Type I Error	Correct Decision

Definition 18.6. Let $H_o : \theta \in \Omega_o$ and $H_a : \theta \notin \Omega_o$ be the null and alternative hypothesis to be tested based on a random sample $X_1, X_2, ..., X_n$ from a population X with density $f(x; \theta)$, where θ is a parameter. The *significance level* of the hypothesis test

$$H_o : \theta \in \Omega_o \qquad \text{and} \qquad H_a : \theta \notin \Omega_o,$$

denoted by α, is defined as

$$\alpha = P\,(\text{Type I Error})\,.$$

Thus, the significance level of a hypothesis test we mean the probability of rejecting a true null hypothesis, that is

$$\alpha = P\,(\text{Reject H}_\text{o} \;/\; \text{H}_\text{o} \text{ is true})\,.$$

This is also equivalent to

$$\alpha = P\,(\text{Accept H}_\text{a} \;/\; \text{H}_\text{o} \text{ is true})\,.$$

Definition 18.7. Let $H_o : \theta \in \Omega_o$ and $H_a : \theta \notin \Omega_o$ be the null and alternative hypothesis to be tested based on a random sample $X_1, X_2, ..., X_n$ from a population X with density $f(x; \theta)$, where θ is a parameter. The *probability of type II error* of the hypothesis test

$$H_o : \theta \in \Omega_o \qquad \text{and} \qquad H_a : \theta \notin \Omega_o,$$

denoted by β, is defined as

$$\beta = P\,(\text{Accept H}_\text{o} \;/\; \text{H}_\text{o} \text{ is false})\,.$$

Similarly, this is also equivalent to

$$\beta = P\,(\text{Accept H}_\text{o} \;/\; \text{H}_\text{a} \text{ is true})\,.$$

Remark 18.3. Note that α can be numerically evaluated if the null hypothesis is a simple hypothesis and rejection rule is given. Similarly, β can be evaluated if the alternative hypothesis is simple and rejection rule is known. If null and the alternatives are composite hypotheses, then α and β become functions of θ.

Example 18.5. Let $X_1, X_2, ..., X_{20}$ be a random sample from a distribution with probability density function

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

where $0 < p \leq \frac{1}{2}$ is a parameter. The hypothesis $H_o : p = \frac{1}{2}$ to be tested against $H_a : p < \frac{1}{2}$. If H_o is rejected when $\sum_{i=1}^{20} X_i \leq 6$, then what is the probability of type I error?

Answer: Since each observation $X_i \sim BER(p)$, the sum the observations $\displaystyle\sum_{i=1}^{20} X_i \sim BIN(20, p)$. The probability of type I error is given by

$$\begin{aligned}
\alpha &= P\,(\text{Type I Error}) \\
&= P\,(\text{Reject H}_\text{o} \;/\; \text{H}_\text{o} \text{ is true}) \\
&= P\left(\sum_{i=1}^{20} X_i \leq 6 \;\Big/\; \text{H}_\text{o} \text{ is true} \right) \\
&= P\left(\sum_{i=1}^{20} X_i \leq 6 \;\Big/\; H_o : p = \frac{1}{2} \right) \\
&= \sum_{k=0}^{6} \binom{20}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{20-k} \\
&= 0.0577 \qquad\qquad\qquad\qquad \text{(from binomial table).}
\end{aligned}$$

Hence the probability of type I error is 0.0577.

Example 18.6. Let p represent the proportion of defectives in a manufacturing process. To test $H_o : p \leq \frac{1}{4}$ versus $H_a : p > \frac{1}{4}$, a random sample of size 5 is taken from the process. If the number of defectives is 4 or more, the null hypothesis is rejected. What is the probability of rejecting H_o if $p = \frac{1}{5}$?

Answer: Let X denote the number of defectives out of a random sample of size 5. Then X is a binomial random variable with $n = 5$ and $p = \frac{1}{5}$. Hence, the probability of rejecting H_o is given by

$$\begin{aligned}
\alpha &= P\,(\text{Reject H}_\text{o} \;/\; \text{H}_\text{o} \text{ is true}) \\
&= P\,(X \geq 4 \;/\; \text{H}_\text{o} \text{ is true}) \\
&= P\left(X \geq 4 \;\Big/\; p = \frac{1}{5}\right) \\
&= P\left(X = 4 \;\Big/ p = \frac{1}{5}\right) + P\left(X = 5 \;\Big/ p = \frac{1}{5}\right) \\
&= \binom{5}{4} p^4 (1-p)^1 + \binom{5}{5} p^5 (1-p)^0 \\
&= 5\left(\frac{1}{5}\right)^4 \left(\frac{4}{5}\right) + \left(\frac{1}{5}\right)^5 \\
&= \left(\frac{1}{5}\right)^5 \;[20+1] \\
&= \frac{21}{3125}.
\end{aligned}$$

Hence the probability of rejecting the null hypothesis H_o is $\frac{21}{3125}$.

Example 18.7. A random sample of size 4 is taken from a normal distribution with unknown mean μ and variance $\sigma^2 > 0$. To test $H_o : \mu = 0$ against $H_a : \mu < 0$ the following test is used: "Reject H_o if and only if $X_1 + X_2 + X_3 + X_4 < -20$." Find the value of σ so that the significance level of this test will be closed to 0.14.

Answer: Since

$$\begin{aligned}
0.14 &= \alpha \qquad\qquad\qquad\qquad \text{(significance level)} \\
&= P\,(\text{Type I Error}) \\
&= P\,(\text{Reject H}_\text{o} \;/\; \text{H}_\text{o} \text{ is true}) \\
&= P\,(X_1 + X_2 + X_3 + X_4 < -20 \;/ H_o : \mu = 0) \\
&= P\left(\overline{X} < -5 \;/ H_o : \mu = 0\right) \\
&= P\left(\frac{\overline{X} - 0}{\frac{\sigma}{2}} < \frac{-5 - 0}{\frac{\sigma}{2}}\right) \\
&= P\left(Z < -\frac{10}{\sigma}\right),
\end{aligned}$$

we get from the standard normal table

$$1.08 = \frac{10}{\sigma}.$$

Therefore

$$\sigma = \frac{10}{1.08} = 9.26.$$

Hence, the standard deviation has to be 9.26 so that the significance level will be closed to 0.14.

Example 18.8. A normal population has a standard deviation of 16. The critical region for testing $H_o : \mu = 5$ versus the alternative $H_a : \mu = k$ is $\bar{X} > k - 2$. What would be the value of the constant k and the sample size n which would allow the probability of Type I error to be 0.0228 and the probability of Type II error to be 0.1587.

Answer: It is given that the population $X \sim N\left(\mu, \, 16^2\right)$. Since

$$\begin{aligned}
0.0228 &= \alpha \\
&= P\,(\text{Type I Error}) \\
&= P\,(\text{Reject H}_\text{o} \ / \ \text{H}_\text{o} \text{ is true}) \\
&= P\left(\overline{X} > k - 2 \ / H_o : \mu = 5\right) \\
&= P\left(\frac{\overline{X} - 5}{\sqrt{\frac{256}{n}}} > \frac{k-7}{\sqrt{\frac{256}{n}}}\right) \\
&= P\left(Z > \frac{k-7}{\sqrt{\frac{256}{n}}}\right) \\
&= 1 - P\left(Z \leq \frac{k-7}{\sqrt{\frac{256}{n}}}\right)
\end{aligned}$$

Hence, from standard normal table, we have

$$\frac{(k-7)\sqrt{n}}{16} = 2$$

which gives

$$(k-7)\sqrt{n} = 32.$$

Similarly

$$\begin{aligned}
0.1587 &= P\,(\text{Type II Error})\\
&= P\,(\text{Accept H}_\text{o} \;/\; \text{H}_\text{a} \text{ is true})\\
&= P\left(\overline{X} \leq k-2 \;/ H_a : \mu = k\right)\\
&= P\left(\frac{\overline{X}-\mu}{\sqrt{\frac{256}{n}}} \leq \frac{k-2-\mu}{\sqrt{\frac{256}{n}}} \;\Big/\; H_a : \mu = k\right)\\
&= P\left(\frac{\overline{X}-k}{\sqrt{\frac{256}{n}}} \leq \frac{k-2-k}{\sqrt{\frac{256}{n}}}\right)\\
&= P\left(Z \leq -\frac{2}{\sqrt{\frac{256}{n}}}\right)\\
&= 1 - P\left(Z \leq \frac{2\sqrt{n}}{16}\right).
\end{aligned}$$

Hence $0.1587 = 1 - P\left(Z \leq \frac{2\sqrt{n}}{16}\right)$ or $P\left(Z \leq \frac{2\sqrt{n}}{16}\right) = 0.8413$. Thus, from the standard normal table, we have

$$\frac{2\sqrt{n}}{16} = 1$$

which yields

$$n = 64.$$

Letting this value of n in

$$(k-7)\sqrt{n} = 32,$$

we see that $k = 11$.

While deciding to accept H_o or H_a, we may make a wrong decision. The probability γ of a wrong decision can be computed as follows:

$$\begin{aligned}
\gamma &= P\,(\text{H}_\text{a} \text{ accepted and } \text{H}_\text{o} \text{ is true}) + P\,(\text{H}_\text{o} \text{ accepted and } \text{H}_\text{a} \text{ is true})\\
&= P\,(\text{H}_\text{a} \text{ accepted / H}_\text{o} \text{ is true})\, P\,(\text{H}_\text{o} \text{ is true})\\
&\qquad + P\,(\text{H}_\text{o} \text{ accepted / H}_\text{a} \text{ is true})\, P\,(\text{H}_\text{a} \text{ is true})\\
&= \alpha\, P\,(\text{H}_\text{o} \text{ is true}) + \beta\, P\,(\text{H}_\text{a} \text{ is true})\,.
\end{aligned}$$

In most cases, the probabilities $P\,(\text{H}_\text{o} \text{ is true})$ and $P\,(\text{H}_\text{a} \text{ is true})$ are not known. Therefore, it is, in general, not possible to determine the exact numerical value of the probability γ of making a wrong decision. However, since γ is a weighted sum of α and β, and $P(\mathrm{H_o} \text{ is true}) + P(\mathrm{H_a} \text{ is true}) = 1$, we have

$$\gamma \leq \max\{\alpha, \beta\}.$$

A good decision rule (or a good test) is the one which yields the smallest γ. In view of the above inequality, one will have a small γ if the probability of type I error as well as probability of type II error are small.

The alternative hypothesis is mostly a composite hypothesis. Thus, it is not possible to find a value for the probability of type II error, β. For composite alternative, β is a function of θ. That is, $\beta : \Omega_o^c :\to [0,1]$. Here Ω_o^c denotes the complement of the set Ω_o in the parameter space Ω. In hypothesis test, instead of β, one usually considers the *power of the test* $1 - \beta(\theta)$, and a small probability of type II error is equivalent to large power of the test.

Definition 18.8. Let $H_o : \theta \in \Omega_o$ and $H_a : \theta \notin \Omega_o$ be the null and alternative hypothesis to be tested based on a random sample $X_1, X_2, ..., X_n$ from a population X with density $f(x; \theta)$, where θ is a parameter. The *power function* of a hypothesis test

$$H_o : \theta \in \Omega_o \qquad \text{versus} \qquad H_a : \theta \notin \Omega_o$$

is a function $\pi : \Omega \to [0,1]$ defined by

$$\pi(\theta) = \begin{cases} P(\text{Type I Error}) & \text{if } \mathrm{H_o} \text{ is true} \\ \\ 1 - P(\text{Type II Error}) & \text{if } \mathrm{H_a} \text{ is true.} \end{cases}$$

Example 18.9. A manufacturing firm needs to test the null hypothesis H_o that the probability p of a defective item is 0.1 or less, against the alternative hypothesis $H_a : p > 0.1$. The procedure is to select two items at random. If both are defective, H_o is rejected; otherwise, a third is selected. If the third item is defective H_o is rejected. If all other cases, H_o is accepted, what is the power of the test in terms of p (if H_o is true)?

Answer: Let p be the probability of a defective item. We want to calculate the power of the test at the null hypothesis. The power function of the test is given by

$$\pi(p) = \begin{cases} P(\text{Type I Error}) & \text{if } p \leq 0.1 \\ \\ 1 - P(\text{Type II Error}) & \text{if } p > 0.1. \end{cases}$$

Hence, we have

$\pi(p)$

$= P\,(\text{Reject H}_o\ /\ \text{H}_o \text{ is true})$

$= P\,(\text{Reject H}_o\ /\ \text{H}_o : \text{p} = \text{p})$

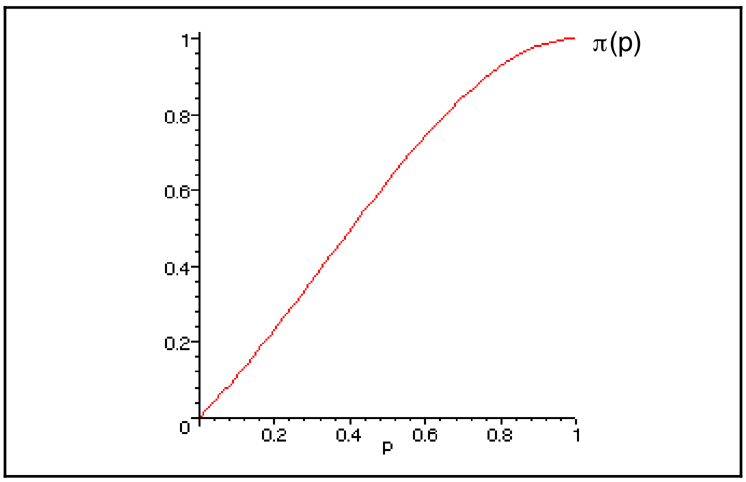
$= P\,(\text{first two items are both defective /p}) +$

$+\, P\,(\text{at least one of the first two items is not defective and third is/p})$

$= p^2 + (1-p)^2\, p + \binom{2}{1}\, p(1-p)p$

$= p + p^2 - p^3.$

The graph of this power function is shown below.



Remark 18.4. If X denotes the number of independent trials needed to obtain the first success, then $X \sim GEO(p)$, and

$$P(X = k) = (1-p)^{k-1}\, p,$$

where $k = 1, 2, 3, ..., \infty$. Further

$$P(X \leq n) = 1 - (1-p)^n$$

since

$$\begin{aligned}\sum_{k=1}^{n} (1-p)^{k-1}\, p &= p\, \sum_{k=1}^{n} (1-p)^{k-1} \\ &= p\, \frac{1-(1-p)^n}{1-(1-p)} \\ &= 1 - (1-p)^n.\end{aligned}$$

Example 18.10. Let X be the number of independent trails required to obtain a success where p is the probability of success on each trial. The hypothesis $H_o : p = 0.1$ is to be tested against the alternative $H_a : p = 0.3$. The hypothesis is rejected if $X \leq 4$. What is the power of the test if H_a is true?

Answer: The power function is given by

$$\pi(p) = \begin{cases} P\,(\text{Type I Error}) & \text{if } p = 0.1 \\ \\ 1 - P\,(\text{Type II Error}) & \text{if p = 0.3.} \end{cases}$$

Hence, we have

$$\begin{aligned}
\alpha &= 1 - P\,(\text{Accept H}_\text{o} \;/\; \text{H}_\text{o} \text{ is false}) \\
&= P\,(\text{Reject H}_\text{o} \;/\; \text{H}_\text{a} \text{ is true}) \\
&= P\,(X \leq 4 \;/\; \text{H}_\text{a} \text{ is true}) \\
&= P\,(X \leq 4 \;/\; p = 0.3) \\
&= \sum_{k=1}^{4} P\,(X = k \;/p = 0.3) \\
&= \sum_{k=1}^{4} (1-p)^{k-1}\, p \qquad \text{(where p = 0.3)} \\
&= \sum_{k=1}^{4} (0.7)^{k-1}\, (0.3) \\
&= 0.3 \;\sum_{k=1}^{4} (0.7)^{k-1} \\
&= 1 - (0.7)^4 \\
&= 0.7599.
\end{aligned}$$

Hence, the power of the test at the alternative is 0.7599.

Example 18.11. Let $X_1, X_2, ..., X_{25}$ be a random sample of size 25 drawn from a normal distribution with unknown mean μ and variance $\sigma^2 = 100$. It is desired to test the null hypothesis $\mu = 4$ against the alternative $\mu = 6$. What is the power at $\mu = 6$ of the test with rejection rule: reject $\mu = 4$ if $\sum_{i=1}^{25} X_i \geq 125$?

Answer: The power of the test at the alternative is

$$\begin{aligned}
\pi(6) &= 1 - P\,(\text{Type II Error}) \\
&= 1 - P\,(\text{Accept H}_\text{o} \;/\; \text{H}_\text{o} \text{ is false}) \\
&= P\,(\text{Reject H}_\text{o} \;/\; \text{H}_\text{a} \text{ is true}) \\
&= P\left(\sum_{i=1}^{25} X_i \geq 125 \;/\; H_a : \mu = 6\right) \\
&= P\left(\overline{X} \geq 5 \;/\; H_a \mu = 6\right) \\
&= P\left(\frac{\overline{X} - 6}{\frac{10}{\sqrt{25}}} \geq \frac{5 - 6}{\frac{10}{\sqrt{25}}}\right) \\
&= P\left(Z \geq -\frac{1}{2}\right) \\
&= 0.6915.
\end{aligned}$$

Example 18.12. A urn contains 7 balls, θ of which are red. A sample of size 2 is drawn without replacement to test $H_o : \theta \leq 1$ against $H_a : \theta > 1$. If the null hypothesis is rejected if one or more red balls are drawn, find the power of the test when $\theta = 2$.

Answer: The power of the test at $\theta = 2$ is given by

$$\begin{aligned}
\pi(2) &= 1 - P\,(\text{Type II Error}) \\
&= 1 - P\,(\text{Accept H}_\text{o} \;/\; \text{H}_\text{o} \text{ is false}) \\
&= 1 - P\,(\text{zero red balls are drawn } /2 \text{ balls were red}) \\
&= 1 - \frac{\binom{5}{2}}{\binom{7}{2}} \\
&= 1 - \frac{10}{21} \\
&= \frac{11}{21} \\
&= 0.524.
\end{aligned}$$

In all of these examples, we have seen that if the rule for rejection of the null hypothesis H_o is given, then one can compute the significance level or power function of the hypothesis test. The rejection rule is given in terms of a statistic $W(X_1, X_2, ..., X_n)$ of the sample $X_1, X_2, ..., X_n$. For instance, in Example 18.5, the rejection rule was: "Reject the null hypothesis H_o if $\sum_{i=1}^{20} X_i \leq 6$." Similarly, in Example 18.7, the rejection rule was: "Reject H_o

if and only if $X_1+X_2+X_3+X_4 < -20$", and so on. The statistic W, used in the statement of the rejection rule, partitioned the set S^n into two subsets, where S denotes the support of the density function of the population X. One subset is called the rejection or critical region and other subset is called the acceptance region. The rejection rule is obtained in such a way that the probability of the type I error is as small as possible and the power of the test at the alternative is as large as possible.

Next, we give two definitions that will lead us to the definition of uniformly most powerful test.

Definition 18.9. Given $0 \leq \delta \leq 1$, a test (or test procedure) T for testing the null hypothesis $H_o : \theta \in \Omega_o$ against the alternative $H_a : \theta \in \Omega_a$ is said to be a *test of level* δ if

$$\max_{\theta \in \Omega_o} \pi(\theta) \leq \delta,$$

where $\pi(\theta)$ denotes the power function of the test T.

Definition 18.10. Given $0 \leq \delta \leq 1$, a test (or test procedure) for testing the null hypothesis $H_o : \theta \in \Omega_o$ against the alternative $H_a : \theta \in \Omega_a$ is said to be a *test of size* δ if

$$\max_{\theta \in \Omega_o} \pi(\theta) = \delta.$$

Definition 18.11. Let T be a test procedure for testing the null hypothesis $H_o : \theta \in \Omega_o$ against the alternative $H_a : \theta \in \Omega_a$. The test (or test procedure) T is said to be the *uniformly most powerful* (UMP) test of level δ if T is of level δ and for any other test W of level δ,

$$\pi_T(\theta) \geq \pi_W(\theta)$$

for all $\theta \in \Omega_a$. Here $\pi_T(\theta)$ and $\pi_W(\theta)$ denote the power functions of tests T and W, respectively.

Remark 18.5. If T is a test procedure for testing $H_o : \theta = \theta_o$ against $H_a : \theta = \theta_a$ based on a sample data $x_1, ..., x_n$ from a population X with a continuous probability density function $f(x; \theta)$, then there is a critical region C associated with the the test procedure T, and power function of T can be computed as

$$\pi_T = \int_C L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n.$$

Similarly, the size of a critical region C, say α, can be given by

$$\alpha = \int_C L(\theta_o, x_1, ..., x_n)\, dx_1 \cdots dx_n.$$

The following famous result tells us which tests are uniformly most powerful if the null hypothesis and the alternative hypothesis are both simple.

Theorem 18.1 (Neyman-Pearson). Let $X_1, X_2, ..., X_n$ be a random sample from a population with probability density function $f(x; \theta)$. Let

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

be the likelihood function of the sample. Then any critical region C of the form

$$C = \left\{ (x_1, x_2, ..., x_n) \;\middle|\; \frac{L(\theta_o, x_1, ..., x_n)}{L(\theta_a, x_1, ..., x_n)} \leq k \right\}$$

for some constant $0 \leq k < \infty$ is best (or uniformly most powerful) of its size for testing $H_o : \theta = \theta_o$ against $H_a : \theta = \theta_a$.

Proof: We assume that the population has a continuous probability density function. If the population has a discrete distribution, the proof can be appropriately modified by replacing integration by summation.

Let C be the critical region of size α as described in the statement of the theorem. Let B be any other critical region of size α. We want to show that the power of C is greater than or equal to that of B. In view of Remark 18.5, we would like to show that

$$\int_C L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n \geq \int_B L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n. \tag{1}$$

Since C and B are both critical regions of size α, we have

$$\int_C L(\theta_o, x_1, ..., x_n)\, dx_1 \cdots dx_n = \int_B L(\theta_o, x_1, ..., x_n)\, dx_1 \cdots dx_n. \tag{2}$$

The last equality (2) can be written as

$$\begin{aligned}&\int_{C \cap B} L(\theta_o, x_1, ..., x_n)\, dx_1 \cdots dx_n + \int_{C \cap B^c} L(\theta_o, x_1, ..., x_n)\, dx_1 \cdots dx_n \\ &= \int_{C \cap B} L(\theta_o, x_1, ..., x_n)\, dx_1 \cdots dx_n + \int_{C^c \cap B} L(\theta_o, x_1, ..., x_n)\, dx_1 \cdots dx_n\end{aligned}$$

since

$$C = (C \cap B) \cup (C \cap B^c) \quad \text{and} \quad B = (C \cap B) \cup (C^c \cap B). \tag{3}$$

Therefore from the last equality, we have

$$\int_{C \cap B^c} L(\theta_o, x_1, ..., x_n)\, dx_1 \cdots dx_n = \int_{C^c \cap B} L(\theta_o, x_1, ..., x_n)\, dx_1 \cdots dx_n. \tag{4}$$

Since

$$C = \left\{ (x_1, x_2, ..., x_n) \;\middle|\; \frac{L(\theta_o, x_1, ..., x_n)}{L(\theta_a, x_1, ..., x_n)} \leq k \right\} \tag{5}$$

we have

$$L(\theta_a, x_1, ..., x_n) \geq \frac{L(\theta_o, x_1, ..., x_n)}{k} \tag{6}$$

on C, and

$$L(\theta_a, x_1, ..., x_n) < \frac{L(\theta_o, x_1, ..., x_n)}{k} \tag{7}$$

on C^c. Therefore from (4), (6) and (7), we have

$$\begin{aligned}
\int_{C \cap B^c} & L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n \\
&\geq \int_{C \cap B^c} \frac{L(\theta_o, x_1, ..., x_n)}{k}\, dx_1 \cdots dx_n \\
&= \int_{C^c \cap B} \frac{L(\theta_o, x_1, ..., x_n)}{k}\, dx_1 \cdots dx_n \\
&\geq \int_{C^c \cap B} L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n.
\end{aligned}$$

Thus, we obtain

$$\int_{C \cap B^c} L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n \geq \int_{C^c \cap B} L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n.$$

From (3) and the last inequality, we see that

$$\begin{aligned}
\int_C & L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n \\
&= \int_{C \cap B} L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n + \int_{C \cap B^c} L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n \\
&\geq \int_{C \cap B} L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n + \int_{C^c \cap B} L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n \\
&\geq \int_B L(\theta_a, x_1, ..., x_n)\, dx_1 \cdots dx_n
\end{aligned}$$

and hence the theorem is proved.

Now we give several examples to illustrate the use of this theorem.

Example 18.13. Let X be a random variable with a density function $f(x)$. What is the critical region for the best test of

$$H_o : f(x) = \begin{cases} \frac{1}{2} & \text{if } -1 < x < 1 \\ \\ 0 & \text{elsewhere,} \end{cases}$$

against

$$H_a : f(x) = \begin{cases} 1 - |x| & \text{if } -1 < x < 1 \\ \\ 0 & \text{elsewhere,} \end{cases}$$

at the significance size $\alpha = 0.10$?

Answer: We assume that the test is performed with a sample of size 1. Using Neyman-Pearson Theorem, the best critical region for the best test at the significance size α is given by

$$\begin{aligned} C &= \left\{ x \in \mathbb{R} \mid \frac{L_o(x)}{L_a(x)} \leq k \right\} \\ &= \left\{ x \in \mathbb{R} \mid \frac{\frac{1}{2}}{1 - |x|} \leq k \right\} \\ &= \left\{ x \in \mathbb{R} \mid |x| \leq 1 - \frac{1}{2k} \right\} \\ &= \left\{ x \in \mathbb{R} \mid \frac{1}{2k} - 1 \leq x \leq 1 - \frac{1}{2k} \right\}. \end{aligned}$$

Since

$$\begin{aligned} 0.1 &= P(C) \\ &= P\left(\frac{L_o(X)}{L_a(X)} \leq k \;/\; H_o \text{ is true} \right) \\ &= P\left(\frac{\frac{1}{2}}{1 - |X|} \leq k \;/\; H_o \text{ is true} \right) \\ &= P\left(\frac{1}{2k} - 1 \leq X \leq 1 - \frac{1}{2k} \;/\; H_o \text{ is true} \right), \\ &= \int_{\frac{1}{2k}-1}^{1-\frac{1}{2k}} \frac{1}{2}\, dx \\ &= 1 - \frac{1}{2k}, \end{aligned}$$

we get the critical region C to be

$$C = \{ x \in \mathbb{R} \mid -0.1 \leq x \leq 0.1 \}.$$

Thus the best critical region is $C = [-0.1,\ 0.1]$ and the best test is: "Reject H_o if $-0.1 \leq X \leq 0.1$".

Example 18.14. Suppose X has the density function

$$f(x;\theta) = \begin{cases} (1+\theta)\, x^{\theta} & \text{if } 0 \leq x \leq 1 \\ \\ 0 & \text{otherwise.} \end{cases}$$

Based on a single observed value of X, find the most powerful critical region of size $\alpha = 0.1$ for testing $H_o : \theta = 1$ against $H_a : \theta = 2$.

Answer: By Neyman-Pearson Theorem, the form of the critical region is given by

$$\begin{aligned} C &= \left\{ x \in \mathbb{R} \mid \frac{L(\theta_o, x)}{L(\theta_a, x)} \leq k \right\} \\ &= \left\{ x \in \mathbb{R} \mid \frac{(1+\theta_o)\, x^{\theta_o}}{(1+\theta_a)\, x^{\theta_a}} \leq k \right\} \\ &= \left\{ x \in \mathbb{R} \mid \frac{2x}{3x^2} \leq k \right\} \\ &= \left\{ x \in \mathbb{R} \mid \frac{1}{x} \leq \frac{3}{2} k \right\} \\ &= \{ x \in \mathbb{R} \mid x \geq a, \} \end{aligned}$$

where a is some constant. Hence the most powerful or best test is of the form: "Reject H_o if $X \geq a$."

Since, the significance level of the test is given to be $\alpha = 0.1$, the constant a can be determined. Now we proceed to find a. Since

$$\begin{aligned} 0.1 &= \alpha \\ &= P\,(\text{Reject } H_o \ / \ H_o \text{ is true}\} \\ &= P\,(X \geq a \ / \ \theta = 1) \\ &= \int_a^1 2x \, dx \\ &= 1 - a^2, \end{aligned}$$

hence

$$a^2 = 1 - 0.1 = 0.9.$$

Therefore

$$a = \sqrt{0.9},$$

since k in Neyman-Pearson Theorem is positive. Hence, the most powerful test is given by "Reject H_o if $X \geq \sqrt{0.9}$".

Example 18.15. Suppose that X is a random variable about which the hypothesis $H_o : X \sim UNIF(0,1)$ against $H_a : X \sim N(0,1)$ is to be tested. What is the most powerful test with a significance level $\alpha = 0.05$ based on one observation of X?

Answer: By Neyman-Pearson Theorem, the form of the critical region is given by

$$\begin{aligned} C &= \left\{ x \in \mathbb{R} \mid \frac{L_o\,(x)}{L_a\,(x)} \leq k \right\} \\ &= \left\{ x \in \mathbb{R} \mid \sqrt{2\pi}\, e^{\frac{1}{2}x^2} \leq k \right\} \\ &= \left\{ x \in \mathbb{R} \mid x^2 \leq 2 \ln \left(\frac{k}{\sqrt{2\pi}} \right) \right\} \\ &= \{ x \in \mathbb{R} \mid x \leq a, \} \end{aligned}$$

where a is some constant. Hence the most powerful or best test is of the form: "Reject H_o if $X \leq a$."

Since, the significance level of the test is given to be $\alpha = 0.05$, the constant a can be determined. Now we proceed to find a. Since

$$\begin{aligned} 0.05 &= \alpha \\ &= P\,(\text{Reject } H_o \ / \ H_o \text{ is true}\} \\ &= P\,(X \leq a \ / \ X \sim UNIF(0,1)) \\ &= \int_0^a dx \\ &= a, \end{aligned}$$

hence $a = 0.05$. Thus, the most powerful critical region is given by

$$C = \{ x \in \mathbb{R} \mid 0 < x \leq 0.05 \}$$

based on the support of the uniform distribution on the open interval $(0,1)$. Since the support of this uniform distribution is the interval $(0,1)$, the acceptance region (or the complement of C in $(0,1)$) is

$$C^c = \{ x \in \mathbb{R} \mid 0.05 < x < 1 \}.$$

However, since the support of the standard normal distribution is $\mathbb{R}$, the actual critical region should be the complement of C^c in $\mathbb{R}$. Therefore, the critical region of this hypothesis test is the set

$$\{x \in \mathbb{R} \,|\, x \leq 0.05 \text{ or } x \geq 1\}.$$

The most powerful test for $\alpha = 0.05$ is: "Reject H_o if $X \leq 0.05$ or $X \geq 1$."

Example 18.16. Let X_1, X_2, X_3 denote three independent observations from a distribution with density

$$f(x;\theta) = \begin{cases} (1+\theta)\, x^{\theta} & \text{if } 0 \leq x \leq 1 \\ \\ 0 & \text{otherwise.} \end{cases}$$

What is the form of the best critical region of size 0.034 for testing $H_o : \theta = 1$ versus $H_a : \theta = 2$?

Answer: By Neyman-Pearson Theorem, the form of the critical region is given by (with $\theta_o = 1$ and $\theta_a = 2$)

$$\begin{aligned} C &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid \frac{L(\theta_o, x_1, x_2, x_3)}{L(\theta_a, x_1, x_2, x_3)} \leq k \right\} \\ &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid \frac{(1+\theta_o)^3 \prod_{i=1}^3 x_i^{\theta_o}}{(1+\theta_a)^3 \prod_{i=1}^3 x_i^{\theta_a}} \leq k \right\} \\ &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid \frac{8 x_1 x_2 x_3}{27 x_1^2 x_2^2 x_3^2} \leq k \right\} \\ &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid \frac{1}{x_1 x_2 x_3} \leq \frac{27}{8} k \right\} \\ &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 x_2 x_3 \geq a, \right\} \end{aligned}$$

where a is some constant. Hence the most powerful or best test is of the form: "Reject H_o if $\displaystyle\prod_{i=1}^{3} X_i \geq a$."

Since, the significance level of the test is given to be $\alpha = 0.034$, the constant a can be determined. To evaluate the constant a, we need the probability distribution of $X_1 X_2 X_3$. The distribution of $X_1 X_2 X_3$ is not easy to get. Hence, we will use Theorem 17.5. There, we have shown that

$-2(1+\theta)\sum_{i=1}^{3}\ln X_i \sim \chi^2(6)$. Now we proceed to find a. Since

$$\begin{aligned}
0.034 &= \alpha \\
&= P\left(\text{Reject } H_o \;/\; H_o \text{ is true}\right\} \\
&= P\left(X_1X_2X_3 \geq a \;/\; \theta = 1\right) \\
&= P\left(\ln(X_1X_2X_3) \geq \ln a \;/\; \theta = 1\right) \\
&= P\left(-2(1+\theta)\ln(X_1X_2X_3) \leq -2(1+\theta)\ln a \;/\; \theta = 1\right) \\
&= P\left(-4\ln(X_1X_2X_3) \leq -4\ln a\right) \\
&= P\left(\chi^2(6) \leq -4\ln a\right)
\end{aligned}$$

hence from chi-square table, we get

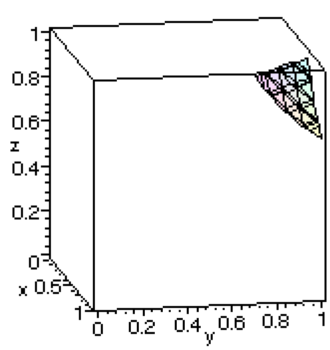
$$-4\ln a = 1.4.$$

Therefore

$$a = e^{-0.35} = 0.7047.$$

Hence, the most powerful test is given by "Reject H_o if $X_1X_2X_3 \geq 0.7047$".

The critical region C is the region above the surface $x_1x_2x_3 = 0.7047$ of the unit cube $[0,1]^3$. The following figure illustrates this region.



Critical region is to the right of the shaded surface

Example 18.17. Let $X_1, X_2, ..., X_{12}$ be a random sample from a normal population with mean zero and variance σ^2. What is the most powerful test of size 0.025 for testing the null hypothesis $H_o : \sigma^2 = 10$ versus $H_a : \sigma^2 = 5$?

Answer: By Neyman-Pearson Theorem, the form of the critical region is given by (with ${\sigma_o}^2 = 10$ and ${\sigma_a}^2 = 5$)

$$
\begin{aligned}
C &= \left\{ (x_1, x_2, ..., x_{12}) \in \mathbb{R}^{12} \;\middle|\; \frac{L\left({\sigma_o}^2, x_1, x_2, ..., x_{12}\right)}{L\left({\sigma_a}^2, x_1, x_2, ..., x_{12}\right)} \leq k \right\} \\
&= \left\{ (x_1, x_2, ..., x_{12}) \in \mathbb{R}^{12} \;\middle|\; \prod_{i=1}^{12} \frac{\frac{1}{\sqrt{2\pi\sigma_o^2}}\, e^{-\frac{1}{2}\left(\frac{x_i}{\sigma_o}\right)^2}}{\frac{1}{\sqrt{2\pi\sigma_a^2}}\, e^{-\frac{1}{2}\left(\frac{x_i}{\sigma_a}\right)^2}} \leq k \right\} \\
&= \left\{ (x_1, x_2, ..., x_{12}) \in \mathbb{R}^{12} \;\middle|\; \left(\frac{1}{2}\right)^6 e^{\frac{1}{20}\sum_{i=1}^{12} x_i^2} \leq k \right\} \\
&= \left\{ (x_1, x_2, ..., x_{12}) \in \mathbb{R}^{12} \;\middle|\; \sum_{i=1}^{12} x_i^2 \leq a \right\},
\end{aligned}
$$

where a is some constant. Hence the most powerful or best test is of the form: "Reject H_o if $\displaystyle\sum_{i=1}^{12} X_i^2 \leq a$."

Since, the significance level of the test is given to be $\alpha = 0.025$, the constant a can be determined. To evaluate the constant a, we need the probability distribution of $X_1^2 + X_2^2 + \cdots + X_{12}^2$. It can be shown that the distribution of $\sum_{i=1}^{12}\left(\frac{X_i}{\sigma}\right)^2 \sim \chi^2(12)$. Now we proceed to find a. Since

$$
\begin{aligned}
0.025 &= \alpha \\
&= P\left(\text{Reject } H_o \;/\; H_o \text{ is true}\right\} \\
&= P\left(\sum_{i=1}^{12} \left(\frac{X_i}{\sigma}\right)^2 \leq a \;/\; \sigma^2 = 10 \right) \\
&= P\left(\sum_{i=1}^{12} \left(\frac{X_i}{\sqrt{10}}\right)^2 \leq a \;/\; \sigma^2 = 10 \right) \\
&= P\left(\chi^2(12) \leq \frac{a}{10} \right),
\end{aligned}
$$

hence from chi-square table, we get

$$\frac{a}{10} = 4.4.$$

Therefore

$$a = 44.$$

Hence, the most powerful test is given by "Reject H_o if $\sum_{i=1}^{12} X_i^2 \leq 44$." The best critical region of size 0.025 is given by

$$C = \left\{ (x_1, x_2, ..., x_{12}) \in \mathbb{R}^{12} \mid \sum_{i=1}^{12} x_i^2 \leq 44 \right\}.$$

In last five examples, we have found the most powerful tests and corresponding critical regions when the both H_o and H_a are simple hypotheses. If either H_o or H_a is not simple, then it is not always possible to find the most powerful test and corresponding critical region. In this situation, hypothesis test is found by using the likelihood ratio. A test obtained by using likelihood ratio is called the *likelihood ratio test* and the corresponding critical region is called the *likelihood ratio critical region*.

18.4. Some Examples of Likelihood Ratio Tests

In this section, we illustrate, using likelihood ratio, how one can construct hypothesis test when one of the hypotheses is not simple. As pointed out earlier, the test we will construct using the likelihood ratio is not the most powerful test. However, such a test has all the desirable properties of a hypothesis test. To construct the test one has to follow a sequence of steps. These steps are outlined below:

(1) Find the likelihood function $L(\theta, x_1, x_2, ..., x_n)$ for the given sample.

(2) Evaluate $\max_{\theta \in \Omega_o} L(\theta, x_1, x_2, ..., x_n)$.

(3) Find the maximum likelihood estimator $\widehat{\theta}$ of θ.

(4) Compute $\max_{\theta \in \Omega} L(\theta, x_1, x_2, ..., x_n)$ using $L\left(\widehat{\theta}, x_1, x_2, ..., x_n\right)$.

(5) Using steps (2) and (4), find $W(x_1, ..., x_n) = \dfrac{\max_{\theta \in \Omega_o} L(\theta, x_1, x_2, ..., x_n)}{\max_{\theta \in \Omega} L(\theta, x_1, x_2, ..., x_n)}$.

(6) Using step (5) determine $C = \{(x_1, x_2, ..., x_n) \mid W(x_1, ..., x_n) \leq k\}$, where $k \in [0, 1]$.

(7) Reduce $W(x_1, ..., x_n) \leq k$ to an equivalent inequality $\widehat{W}(x_1, ..., x_n) \leq A$.

(8) Determine the distribution of $\widehat{W}(x_1, ..., x_n)$.

(9) Find A such that given α equals $P\left(\widehat{W}(x_1, ..., x_n) \leq A \mid H_o \text{ is true}\right)$.

In the remaining examples, for notational simplicity, we will denote the likelihood function $L(\theta, x_1, x_2, ..., x_n)$ simply as $L(\theta)$.

Example 18.19. Let $X_1, X_2, ..., X_n$ be a random sample from a normal population with mean μ and `known variance` σ^2. What is the likelihood ratio test of size α for testing the null hypothesis $H_o : \mu = \mu_o$ versus the alternative hypothesis $H_a : \mu \neq \mu_o$?

Answer: The likelihood function of the sample is given by

$$\begin{aligned} L(\mu) &= \prod_{i=1}^{n} \left(\frac{1}{\sigma\sqrt{2\pi}} \right) e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} \\ &= \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}. \end{aligned}$$

Since $\Omega_o = \{\mu_o\}$, we obtain

$$\begin{aligned} \max_{\mu\in\Omega_o} L(\mu) &= L(\mu_o) \\ &= \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_o)^2}. \end{aligned}$$

We have seen in Example 15.13 that if $X \sim N(\mu, \sigma^2)$, then the maximum likelihood estimator of μ is $\overline{X}$, that is

$$\widehat{\mu} = \overline{X}.$$

Hence

$$\max_{\mu\in\Omega} L(\mu) = L(\widehat{\mu}) = \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\overline{x})^2}.$$

Now the likelihood ratio statistics $W(x_1, x_2, ..., x_n)$ is given by

$$W(x_1, x_2, ..., x_n) = \frac{\left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_o)^2}}{\left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\overline{x})^2}}$$

which simplifies to

$$W(x_1, x_2, ..., x_n) = e^{-\frac{n}{2\sigma^2}(\overline{x}-\mu_o)^2}.$$

Now the inequality $W(x_1, x_2, ..., x_n) \leq k$ becomes

$$e^{-\frac{n}{2\sigma^2}(\overline{x}-\mu_o)^2} \leq k$$

and which can be rewritten as

$$(\overline{x} - \mu_o)^2 \geq -\frac{2\sigma^2}{n} \ln(k)$$

or

$$|\overline{x} - \mu_o| \geq K$$

where $K = \sqrt{-\frac{2\sigma^2}{n} \ln(k)}$. In view of the above inequality, the critical region can be described as

$$C = \{(x_1, x_2, ..., x_n) \;|\; |\overline{x} - \mu_o| \geq K\}.$$

Since we are given the size of the critical region to be α, we can determine the constant K. Since the size of the critical region is α, we have

$$\alpha = P\left(\left|\overline{X} - \mu_o\right| \geq K\right).$$

For finding K, we need the probability density function of the statistic $\overline{X} - \mu_o$ when the population X is $N(\mu, \sigma^2)$ and the null hypothesis $H_o : \mu = \mu_o$ is true. Since σ^2 is known and $X_i \sim N(\mu, \sigma^2)$,

$$\frac{\overline{X} - \mu_o}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

and

$$\begin{aligned}
\alpha &= P\left(\left|\overline{X} - \mu_o\right| \geq K\right) \\
&= P\left(\left|\frac{\overline{X} - \mu_o}{\frac{\sigma}{\sqrt{n}}}\right| \geq K\frac{\sqrt{n}}{\sigma}\right) \\
&= P\left(|Z| \geq K\frac{\sqrt{n}}{\sigma}\right) \qquad \text{where} \qquad Z = \frac{\overline{X} - \mu_o}{\frac{\sigma}{\sqrt{n}}} \\
&= 1 - P\left(-K\,\frac{\sqrt{n}}{\sigma} \leq Z \leq K\,\frac{\sqrt{n}}{\sigma}\right)
\end{aligned}$$

we get

$$z_{\frac{\alpha}{2}} = K \, \frac{\sqrt{n}}{\sigma}$$

which is

$$K = z_{\frac{\alpha}{2}} \, \frac{\sigma}{\sqrt{n}},$$

where $z_{\frac{\alpha}{2}}$ is a real number such that the integral of the standard normal density from $z_{\frac{\alpha}{2}}$ to ∞ equals $\frac{\alpha}{2}$.

Hence, the likelihood ratio test is given by "Reject H_o if

$$\left|\overline{X} - \mu_o\right| \geq z_{\frac{\alpha}{2}} \, \frac{\sigma}{\sqrt{n}}."$$

If we denote

$$z = \frac{\overline{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}}$$

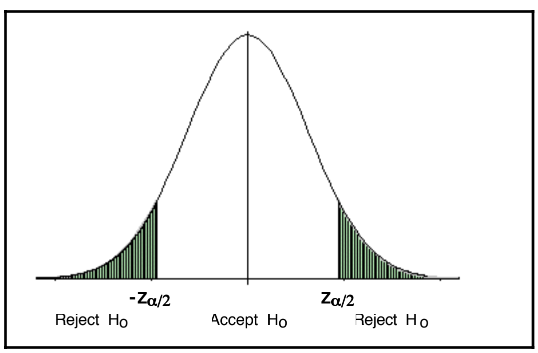
then the above inequality becomes

$$|Z| \geq z_{\frac{\alpha}{2}}.$$

Thus critical region is given by

$$C = \left\{(x_1, x_2, ..., x_n) \mid \; |z| \geq z_{\frac{\alpha}{2}} \right\}.$$

This tells us that the null hypothesis must be rejected when the absolute value of z takes on a value greater than or equal to $z_{\frac{\alpha}{2}}$.



Remark 18.6. The hypothesis $H_a : \mu \neq \mu_o$ is called a two-sided alternative hypothesis. An alternative hypothesis of the form $H_a : \mu > \mu_o$ is called a right-sided alternative. Similarly, $H_a : \mu < \mu_o$ is called the a left-sided

alternative. In the above example, if we had a right-sided alternative, that is $H_a : \mu > \mu_o$, then the critical region would have been

$$C = \{(x_1, x_2, ..., x_n) \mid \; z \geq z_\alpha \}.$$

Similarly, if the alternative would have been left-sided, that is $H_a : \mu < \mu_o$, then the critical region would have been

$$C = \{(x_1, x_2, ..., x_n) \mid \; z \leq -z_\alpha \}.$$

We summarize the three cases of hypotheses test of the mean (of the normal population with known variance) in the following table.

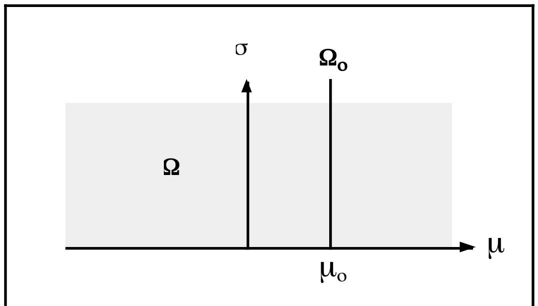
H_o	H_a	Critical Region (or Test)
$\mu = \mu_o$	$\mu > \mu_o$	$z = \frac{\overline{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}} \geq z_\alpha$
$\mu = \mu_o$	$\mu < \mu_o$	$z = \frac{\overline{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}} \leq -z_\alpha$
$\mu = \mu_o$	$\mu \neq \mu_o$	$\lvert z \rvert = \left\lvert \frac{\overline{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}} \right\rvert \geq z_{\frac{\alpha}{2}}$

Example 18.20. Let $X_1, X_2, ..., X_n$ be a random sample from a normal population with mean μ and `unknown variance` σ^2. What is the likelihood ratio test of size α for testing the null hypothesis $H_o : \mu = \mu_o$ versus the alternative hypothesis $H_a : \mu \neq \mu_o$?

Answer: In this example,

$$\begin{aligned}
\Omega &= \left\{ \left(\mu, \sigma^2\right) \in \mathbb{R}^2 \mid \; -\infty < \mu < \infty, \; \sigma^2 > 0 \right\}, \\
\Omega_o &= \left\{ \left(\mu_o, \sigma^2\right) \in \mathbb{R}^2 \mid \sigma^2 > 0 \right\}, \\
\Omega_a &= \left\{ \left(\mu, \sigma^2\right) \in \mathbb{R}^2 \mid \mu \neq \mu_o, \; \sigma^2 > 0 \right\}.
\end{aligned}$$

These sets are illustrated below.



Graphs of Ω and Ωo

The likelihood function is given by

$$\begin{aligned} L\left(\mu, \sigma^2\right) &= \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}. \end{aligned}$$

Next, we find the maximum of $L\left(\mu, \sigma^2\right)$ on the set Ω_o. Since the set Ω_o is equal to $\left\{\left(\mu_o, \sigma^2\right) \in \mathbb{R}^2 \mid 0 < \sigma < \infty\right\}$, we have

$$\max_{(\mu,\sigma^2)\in\Omega_o} L\left(\mu, \sigma^2\right) = \max_{\sigma^2>0} L\left(\mu_o, \sigma^2\right).$$

Since $L\left(\mu_o, \sigma^2\right)$ and $\ln L\left(\mu_o, \sigma^2\right)$ achieve the maximum at the same σ value, we determine the value of σ where $\ln L\left(\mu_o, \sigma^2\right)$ achieves the maximum. Taking the natural logarithm of the likelihood function, we get

$$\ln\left(L\left(\mu, \sigma^2\right)\right) = -\frac{n}{2}\ln(\sigma^2) - \frac{n}{2}\ln(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_o)^2.$$

Differentiating $\ln L\left(\mu_o, \sigma^2\right)$ with respect to σ^2, we get from the last equality

$$\frac{d}{d\sigma^2}\ln\left(L\left(\mu, \sigma^2\right)\right) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu_o)^2.$$

Setting this derivative to zero and solving for σ, we obtain

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_o)^2}.$$

Thus $\ln\left(L\left(\mu,\sigma^2\right)\right)$ attains maximum at $\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_o)^2}$. Since this value of σ is also yield maximum value of $L\left(\mu,\sigma^2\right)$, we have

$$\max_{\sigma^2>0} L\left(\mu_o,\sigma^2\right) = \left(2\pi\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_o)^2\right)^{-\frac{n}{2}} e^{-\frac{n}{2}}.$$

Next, we determine the maximum of $L\left(\mu,\sigma^2\right)$ on the set Ω. As before, we consider $\ln L\left(\mu,\sigma^2\right)$ to determine where $L\left(\mu,\sigma^2\right)$ achieves maximum. Taking the natural logarithm of $L\left(\mu,\sigma^2\right)$, we obtain

$$\ln\left(L\left(\mu,\sigma^2\right)\right) = -\frac{n}{2}\ln(\sigma^2) - \frac{n}{2}\ln(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.$$

Taking the partial derivatives of $\ln L\left(\mu,\sigma^2\right)$ first with respect to μ and then with respect to σ^2, we get

$$\frac{\partial}{\partial\mu}\ln L\left(\mu,\sigma^2\right) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu),$$

and

$$\frac{\partial}{\partial\sigma^2}\ln L\left(\mu,\sigma^2\right) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2,$$

respectively. Setting these partial derivatives to zero and solving for μ and σ, we obtain

$$\mu = \overline{x} \qquad \text{and} \qquad \sigma^2 = \frac{n-1}{n}\, s^2,$$

where $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2$ is the sample variance.

Letting these optimal values of μ and σ into $L\left(\mu,\sigma^2\right)$, we obtain

$$\max_{(\mu,\sigma^2)\in\Omega} L\left(\mu,\sigma^2\right) = \left(2\pi\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2\right)^{-\frac{n}{2}} e^{-\frac{n}{2}}.$$

Hence

$$\frac{\max\limits_{(\mu,\sigma^2)\in\Omega_o} L\left(\mu,\sigma^2\right)}{\max\limits_{(\mu,\sigma^2)\in\Omega} L\left(\mu,\sigma^2\right)} = \frac{\left(2\pi\frac{1}{n}\sum\limits_{i=1}^{n}(x_i-\mu_o)^2\right)^{-\frac{n}{2}} e^{-\frac{n}{2}}}{\left(2\pi\frac{1}{n}\sum\limits_{i=1}^{n}(x_i-\overline{x})^2\right)^{-\frac{n}{2}} e^{-\frac{n}{2}}} = \left(\frac{\sum\limits_{i=1}^{n}(x_i-\mu_o)^2}{\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}\right)^{-\frac{n}{2}}.$$

Since

$$\sum_{i=1}^{n}(x_i - \overline{x})^2 = (n-1)\, s^2$$

and

$$\sum_{i=1}^{n}(x_i - \mu)^2 = \sum_{i=1}^{n}(x_i - \overline{x})^2 + n\,(\overline{x} - \mu_o)^2\,,$$

we get

$$W(x_1, x_2, ..., x_n) = \frac{\displaystyle\max_{(\mu,\sigma^2)\in\Omega_o} L\left(\mu,\sigma^2\right)}{\displaystyle\max_{(\mu,\sigma^2)\in\Omega} L\left(\mu,\sigma^2\right)} = \left(1 + \frac{n}{n-1}\frac{(\overline{x}-\mu_o)^2}{s^2}\right)^{-\frac{n}{2}}.$$

Now the inequality $W(x_1, x_2, ..., x_n) \leq k$ becomes

$$\left(1 + \frac{n}{n-1}\frac{(\overline{x}-\mu_o)^2}{s^2}\right)^{-\frac{n}{2}} \leq k$$

and which can be rewritten as

$$\left(\frac{\overline{x}-\mu_o}{s}\right)^2 \geq \frac{n-1}{n}\left(k^{-\frac{2}{n}} - 1\right)$$

or

$$\left|\frac{\overline{x}-\mu_o}{\frac{s}{\sqrt{n}}}\right| \geq K$$

where $K = \sqrt{(n-1)\left[k^{-\frac{2}{n}} - 1\right]}$. In view of the above inequality, the critical region can be described as

$$C = \left\{(x_1, x_2, ..., x_n) \mid \left|\frac{\overline{x}-\mu_o}{\frac{s}{\sqrt{n}}}\right| \geq K\right\}$$

and the best likelihood ratio test is: "Reject H_o if $\left|\frac{\overline{x}-\mu_o}{\frac{s}{\sqrt{n}}}\right| \geq K$". Since we are given the size of the critical region to be α, we can find the constant K. For finding K, we need the probability density function of the statistic $\frac{\overline{x}-\mu_o}{\frac{s}{\sqrt{n}}}$ when the population X is $N(\mu, \sigma^2)$ and the null hypothesis $H_o : \mu = \mu_o$ is true.

Since the population is normal with mean μ and variance σ^2,

$$\frac{\overline{X}-\mu_o}{\frac{S}{\sqrt{n}}} \sim t(n-1),$$

where S^2 is the sample variance and equals to $\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$. Hence

$$K = t_{\frac{\alpha}{2}}(n-1)\,\frac{s}{\sqrt{n}},$$

where $t_{\frac{\alpha}{2}}(n-1)$ is a real number such that the integral of the t-distribution with $n-1$ degrees of freedom from $t_{\frac{\alpha}{2}}(n-1)$ to ∞ equals $\frac{\alpha}{2}$.

Therefore, the likelihood ratio test is given by "Reject $H_o : \mu = \mu_o$ if

$$\left|\overline{X} - \mu_o\right| \geq t_{\frac{\alpha}{2}}(n-1)\,\frac{S}{\sqrt{n}}."$$

If we denote

$$t = \frac{\overline{x} - \mu_o}{\frac{s}{\sqrt{n}}}$$

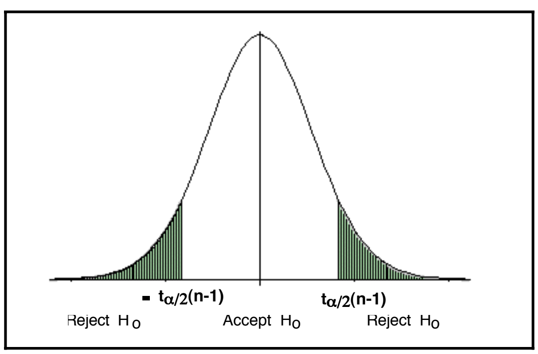
then the above inequality becomes

$$|T| \geq t_{\frac{\alpha}{2}}(n-1).$$

Thus critical region is given by

$$C = \left\{(x_1, x_2, ..., x_n) \mid \; |t| \geq t_{\frac{\alpha}{2}}(n-1)\right\}.$$

This tells us that the null hypothesis must be rejected when the absolute value of t takes on a value greater than or equal to $t_{\frac{\alpha}{2}}(n-1)$.



Remark 18.7. In the above example, if we had a right-sided alternative, that is $H_a : \mu > \mu_o$, then the critical region would have been

$$C = \left\{(x_1, x_2, ..., x_n) \mid \; t \geq t_{\alpha}(n-1)\right\}.$$

Similarly, if the alternative would have been left-sided, that is $H_a : \mu < \mu_o$, then the critical region would have been

$$C = \{(x_1, x_2, ..., x_n) \mid \; t \leq -t_\alpha(n-1)\}.$$

We summarize the three cases of hypotheses test of the mean (of the normal population with unknown variance) in the following table.

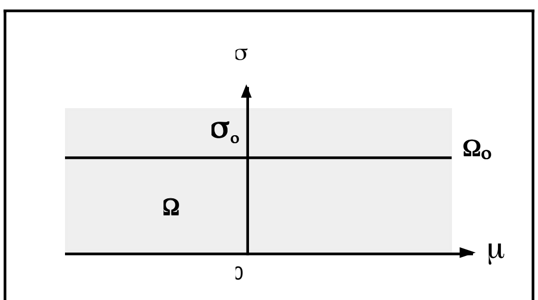
H_o	H_a	Critical Region (or Test)
$\mu = \mu_o$	$\mu > \mu_o$	$t = \frac{\overline{x}-\mu_o}{\frac{s}{\sqrt{n}}} \geq t_\alpha(n-1)$
$\mu = \mu_o$	$\mu < \mu_o$	$t = \frac{\overline{x}-\mu_o}{\frac{s}{\sqrt{n}}} \leq -t_\alpha(n-1)$
$\mu = \mu_o$	$\mu \neq \mu_o$	$\lvert t \rvert = \left\lvert \frac{\overline{x}-\mu_o}{\frac{s}{\sqrt{n}}} \right\rvert \geq t_{\frac{\alpha}{2}}(n-1)$

Example 18.21. Let $X_1, X_2, ..., X_n$ be a random sample from a normal population with mean μ and variance σ^2. What is the likelihood ratio test of significance of size α for testing the null hypothesis $H_o : \sigma^2 = \sigma_o^2$ versus $H_a : \sigma^2 \neq \sigma_o^2$?

Answer: In this example,

$$\Omega = \left\{ \left(\mu, \sigma^2\right) \in \mathbb{R}^2 \mid \; -\infty < \mu < \infty, \; \sigma^2 > 0 \right\},$$
$$\Omega_o = \left\{ \left(\mu, \sigma_o^2\right) \in \mathbb{R}^2 \mid \; -\infty < \mu < \infty \right\},$$
$$\Omega_a = \left\{ \left(\mu, \sigma^2\right) \in \mathbb{R}^2 \mid \; -\infty < \mu < \infty, \; \sigma \neq \sigma_o \right\}.$$

These sets are illustrated below.



Graphs of Ω and Ωo

The likelihood function is given by

$$\begin{aligned} L\left(\mu,\sigma^2\right) &= \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}. \end{aligned}$$

Next, we find the maximum of $L\left(\mu,\sigma^2\right)$ on the set Ω_o. Since the set Ω_o is equal to $\left\{\left(\mu,\sigma_o^2\right) \in \mathbb{R}^2 \mid -\infty < \mu < \infty\right\}$, we have

$$\max_{(\mu,\sigma^2)\in\Omega_o} L\left(\mu,\sigma^2\right) = \max_{-\infty<\mu<\infty} L\left(\mu,\sigma_o^2\right).$$

Since $L\left(\mu,\sigma_o^2\right)$ and $\ln L\left(\mu,\sigma_o^2\right)$ achieve the maximum at the same μ value, we determine the value of μ where $\ln L\left(\mu,\sigma_o^2\right)$ achieves the maximum. Taking the natural logarithm of the likelihood function, we get

$$\ln\left(L\left(\mu,\sigma_o^2\right)\right) = -\frac{n}{2}\ln(\sigma_o^2) - \frac{n}{2}\ln(2\pi) - \frac{1}{2\sigma_o^2}\sum_{i=1}^{n}(x_i-\mu)^2.$$

Differentiating $\ln L\left(\mu,\sigma_o^2\right)$ with respect to μ, we get from the last equality

$$\frac{d}{d\mu}\ln\left(L\left(\mu,\sigma^2\right)\right) = \frac{1}{\sigma_o^2}\sum_{i=1}^{n}(x_i-\mu).$$

Setting this derivative to zero and solving for μ, we obtain

$$\mu = \overline{x}.$$

Hence, we obtain

$$\max_{-\infty<\mu<\infty} L\left(\mu,\sigma^2\right) = \left(\frac{1}{2\pi\sigma_o^2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

Next, we determine the maximum of $L\left(\mu,\sigma^2\right)$ on the set Ω. As before, we consider $\ln L\left(\mu,\sigma^2\right)$ to determine where $L\left(\mu,\sigma^2\right)$ achieves maximum. Taking the natural logarithm of $L\left(\mu,\sigma^2\right)$, we obtain

$$\ln\left(L\left(\mu,\sigma^2\right)\right) = -n\ln(\sigma) - \frac{n}{2}\ln(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2.$$

Taking the partial derivatives of $\ln L\left(\mu, \sigma^2\right)$ first with respect to μ and then with respect to σ^2, we get

$$\frac{\partial}{\partial \mu} \ln L\left(\mu, \sigma^2\right) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu),$$

and

$$\frac{\partial}{\partial \sigma^2} \ln L\left(\mu, \sigma^2\right) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2,$$

respectively. Setting these partial derivatives to zero and solving for μ and σ, we obtain

$$\mu = \overline{x} \qquad \text{and} \qquad \sigma^2 = \frac{n-1}{n}\, s^2,$$

where $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ is the sample variance.

Letting these optimal values of μ and σ into $L\left(\mu, \sigma^2\right)$, we obtain

$$\max_{(\mu,\sigma^2)\in\Omega} L\left(\mu, \sigma^2\right) = \left(\frac{n}{2\pi(n-1)s^2}\right)^{\frac{n}{2}} e^{-\frac{n}{2(n-1)s^2}\sum_{i=1}^{n}(x_i - \overline{x})^2}.$$

Therefore

$$\begin{aligned} W(x_1, x_2, ..., x_n) &= \frac{\max\limits_{(\mu,\sigma^2)\in\Omega_o} L\left(\mu, \sigma^2\right)}{\max\limits_{(\mu,\sigma^2)\in\Omega} L\left(\mu, \sigma^2\right)} \\ &= \frac{\left(\frac{1}{2\pi\sigma_o^2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma_o^2}\sum_{i=1}^{n}(x_i - \overline{x})^2}}{\left(\frac{n}{2\pi(n-1)s^2}\right)^{\frac{n}{2}} e^{-\frac{n}{2(n-1)s^2}\sum_{i=1}^{n}(x_i - \overline{x})^2}} \\ &= n^{-\frac{n}{2}}\, e^{\frac{n}{2}} \left(\frac{(n-1)s^2}{\sigma_o^2}\right)^{\frac{n}{2}} e^{-\frac{(n-1)s^2}{2\sigma_o^2}}. \end{aligned}$$

Now the inequality $W(x_1, x_2, ..., x_n) \leq k$ becomes

$$n^{-\frac{n}{2}}\, e^{\frac{n}{2}} \left(\frac{(n-1)s^2}{\sigma_o^2}\right)^{\frac{n}{2}} e^{-\frac{(n-1)s^2}{2\sigma_o^2}} \leq k$$

which is equivalent to

$$\left(\frac{(n-1)s^2}{\sigma_o^2}\right)^{n} e^{-\frac{(n-1)s^2}{\sigma_o^2}} \leq \left(k\left(\frac{n}{e}\right)^{\frac{n}{2}}\right)^2 := K_o,$$

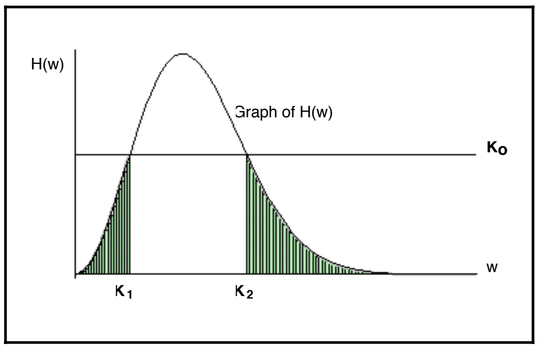
where K_o is a constant. Let H be a function defined by

$$H(w) = w^n \, e^{-w}.$$

Using this, we see that the above inequality becomes

$$H\left(\frac{(n-1)s^2}{\sigma_o^2}\right) \leq K_o.$$

The figure below illustrates this inequality.



From this it follows that

$$\frac{(n-1)s^2}{\sigma_o^2} \leq K_1 \qquad \text{or} \qquad \frac{(n-1)s^2}{\sigma_o^2} \geq K_2.$$

In view of these inequalities, the critical region can be described as

$$C = \left\{ (x_1, x_2, ..., x_n) \;\middle|\; \frac{(n-1)s^2}{\sigma_o^2} \leq K_1 \;\text{ or }\; \frac{(n-1)s^2}{\sigma_o^2} \geq K_2 \right\},$$

and the best likelihood ratio test is: "Reject H_o if

$$\frac{(n-1)S^2}{\sigma_o^2} \leq K_1 \;\text{ or }\; \frac{(n-1)S^2}{\sigma_o^2} \geq K_2."$$

Since we are given the size of the critical region to be α, we can determine the constants K_1 and K_2. As the sample $X_1, X_2, ..., X_n$ is taken from a normal distribution with mean μ and variance σ^2, we get

$$\frac{(n-1)S^2}{\sigma_o^2} \sim \chi^2(n-1)$$

when the null hypothesis $H_o : \sigma^2 = \sigma_o^2$ is true.

Therefore, the likelihood ratio critical region C becomes

$$\left\{ (x_1, x_2, ..., x_n) \;\middle|\; \frac{(n-1)s^2}{\sigma_o^2} \leq \chi^2_{\frac{\alpha}{2}}(n-1) \;\text{ or }\; \frac{(n-1)s^2}{\sigma_o^2} \geq \chi^2_{1-\frac{\alpha}{2}}(n-1) \right\}$$

and the likelihood ratio test is: "Reject $H_o : \sigma^2 = \sigma_o^2$ if

$$\frac{(n-1)S^2}{\sigma_o^2} \leq \chi^2_{\frac{\alpha}{2}}(n-1) \;\text{ or }\; \frac{(n-1)S^2}{\sigma_o^2} \geq \chi^2_{1-\frac{\alpha}{2}}(n-1)\text{"}$$

where $\chi^2_{\frac{\alpha}{2}}(n-1)$ is a real number such that the integral of the chi-square density function with $(n-1)$ degrees of freedom from 0 to $\chi^2_{\frac{\alpha}{2}}(n-1)$ is $\frac{\alpha}{2}$. Further, $\chi^2_{1-\frac{\alpha}{2}}(n-1)$ denotes the real number such that the integral of the chi-square density function with $(n-1)$ degrees of freedom from $\chi^2_{1-\frac{\alpha}{2}}(n-1)$ to ∞ is $\frac{\alpha}{2}$.

Remark 18.8. We summarize the three cases of hypotheses test of the variance (of the normal population with unknown mean) in the following table.

H_o	H_a	Critical Region (or Test)
$\sigma^2 = \sigma_o^2$	$\sigma^2 > \sigma_o^2$	$\chi^2 = \frac{(n-1)s^2}{\sigma_o^2} \geq \chi^2_{1-\alpha}(n-1)$
$\sigma^2 = \sigma_o^2$	$\sigma^2 < \sigma_o^2$	$\chi^2 = \frac{(n-1)s^2}{\sigma_o^2} \leq \chi^2_{\alpha}(n-1)$
$\sigma^2 = \sigma_o^2$	$\sigma^2 \neq \sigma_o^2$	$\chi^2 = \frac{(n-1)s^2}{\sigma_o^2} \geq \chi^2_{1-\alpha/2}(n-1)$ or $\chi^2 = \frac{(n-1)s^2}{\sigma_o^2} \leq \chi^2_{\alpha/2}(n-1)$

18.5. Review Exercises

1. Five trials $X_1, X_2, ..., X_5$ of a Bernoulli experiment were conducted to test $H_o : p = \frac{1}{2}$ against $H_a : p = \frac{3}{4}$. The null hypothesis H_o will be rejected if $\sum_{i=1}^{5} X_i = 5$. Find the probability of Type I and Type II errors.

2. A manufacturer of car batteries claims that the life of his batteries is normally distributed with a standard deviation equal to 0.9 year. If a random sample of 10 of these batteries has a standard deviation of 1.2 years, do you think that $\sigma > 0.9$ year? Use a 0.05 level of significance.

3. Let $X_1, X_2, ..., X_8$ be a random sample of size 8 from a Poisson distribution with parameter λ. Reject the null hypothesis $H_o : \lambda = 0.5$ if the observed sum $\sum_{i=1}^{8} x_i \geq 8$. First, compute the significance level α of the test. Second, find the power function $\beta(\lambda)$ of the test as a sum of Poisson probabilities when H_a is true.

4. Suppose X has the density function

$$f(x) = \begin{cases} \frac{1}{\theta} & \text{for } 0 < x < \theta \\ \\ 0 & \text{otherwise.} \end{cases}$$

If one observation of X is taken, what are the probabilities of Type I and Type II errors in testing the null hypothesis $H_o : \theta = 1$ against the alternative hypothesis $H_a : \theta = 2$, if H_o is rejected for $X > 0.92$.

5. Let X have the density function

$$f(x) = \begin{cases} (\theta + 1)\, x^{\theta} & \text{for } 0 < x < 1 \text{ where } \theta > 0 \\ \\ 0 & \text{otherwise.} \end{cases}$$

The hypothesis $H_o : \theta = 1$ is to be rejected in favor of $H_1 : \theta = 2$ if $X > 0.90$. What is the probability of Type I error?

6. Let $X_1, X_2, ..., X_6$ be a random sample from a distribution with density function

$$f(x) = \begin{cases} \theta\, x^{\theta - 1} & \text{for } 0 < x < 1 \text{ where } \theta > 0 \\ \\ 0 & \text{otherwise.} \end{cases}$$

The null hypothesis $H_o : \theta = 1$ is to be rejected in favor of the alternative $H_a : \theta > 1$ if and only if *at least* 5 of the sample observations are larger than 0.7. What is the significance level of the test?

7. A researcher wants to test $H_o : \theta = 0$ versus $H_a : \theta = 1$, where θ is a parameter of a population of interest. The statistic W, based on a random sample of the population, is used to test the hypothesis. Suppose that under H_o, W has a normal distribution with mean 0 and variance 1, and under H_a, W has a normal distribution with mean 4 and variance 1. If H_o is rejected when $W > 1.50$, then what are the probabilities of a Type I or Type II error respectively?

8. Let X_1 and X_2 be a random sample of size 2 from a normal distribution $N(\mu,\ 1)$. Find the *likelihood ratio critical region* of size 0.005 for testing the null hypothesis $H_o : \mu = 0$ against the composite alternative $H_a : \mu \neq 0$?

9. Let $X_1, X_2, ..., X_{10}$ be a random sample from a Poisson distribution with mean θ. What is the most powerful (or best) critical region of size 0.08 for testing the null hypothesis $H_0 : \theta = 0.1$ against $H_a : \theta = 0.5$?

10. Let X be a random sample of size 1 from a distribution with probability density function

$$f(x,\theta) = \begin{cases} (1 - \frac{\theta}{2}) + \theta\, x & \text{if } 0 \leq x \leq 1 \\ \\ 0 & \text{otherwise.} \end{cases}$$

For a significance level $\alpha = 0.1$, what is the *best (or uniformly most powerful) critical region* for testing the null hypothesis $H_o : \theta = -1$ against $H_a : \theta = 1$?

11. Let X_1, X_2 be a random sample of size 2 from a distribution with probability density function

$$f(x,\theta) = \begin{cases} \frac{\theta^x\, e^{-\theta}}{x!} & \text{if } x = 0, 1, 2, 3, ... \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta \geq 0$. For a significance level $\alpha = 0.053$, what is the *best critical region* for testing the null hypothesis $H_o : \theta = 1$ against $H_a : \theta = 2$? Sketch the graph of the best critical region.

12. Let $X_1, X_2, ..., X_8$ be a random sample of size 8 from a distribution with probability density function

$$f(x,\theta) = \begin{cases} \frac{\theta^x\, e^{-\theta}}{x!} & \text{if } x = 0, 1, 2, 3, ... \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta \geq 0$. What is the *likelihood ratio critical region* for testing the null hypothesis $H_o : \theta = 1$ against $H_a : \theta \neq 1$? If $\alpha = 0.1$ can you determine the best likelihood ratio critical region?

13. Let $X_1, X_2, ..., X_n$ be a random sample of size n from a distribution with probability density function

$$f(x,\theta) = \begin{cases} \frac{x^6\, e^{-\frac{x}{\beta}}}{\Gamma(7)\beta^7}, & \text{if } x > 0 \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $\beta \geq 0$. What is the *likelihood ratio critical region* for testing the null hypothesis $H_o : \beta = 5$ against $H_a : \beta \neq 5$? What is the *most powerful test ?*

14. Let $X_1, X_2, ..., X_5$ denote a random sample of size 5 from a population X with probability density function

$$f(x; \theta) = \begin{cases} (1-\theta)^{x-1}\,\theta & \text{if } x = 1, 2, 3, ..., \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta < 1$ is a parameter. What is the *likelihood ratio critical region* of size 0.05 for testing $H_o : \theta = 0.5$ versus $H_a : \theta \neq 0.5$?

15. Let X_1, X_2, X_3 denote a random sample of size 3 from a population X with probability density function

$$f(x; \mu) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2}} \qquad\qquad -\infty < x < \infty,$$

where $-\infty < \mu < \infty$ is a parameter. What is the *likelihood ratio critical region* of size 0.05 for testing $H_o : \mu = 3$ versus $H_a : \mu \neq 3$?

16. Let X_1, X_2, X_3 denote a random sample of size 3 from a population X with probability density function

$$f(x; \theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & \text{if } 0 < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta < \infty$ is a parameter. What is the *likelihood ratio critical region* for testing $H_o : \theta = 3$ versus $H_a : \theta \neq 3$?

17. Let X_1, X_2, X_3 denote a random sample of size 3 from a population X with probability density function

$$f(x; \theta) = \begin{cases} \frac{e^{-\theta}\,\theta^x}{x!} & \text{if } x = 0, 1, 2, 3, ..., \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta < \infty$ is a parameter. What is the *likelihood ratio critical region* for testing $H_o : \theta = 0.1$ versus $H_a : \theta \neq 0.1$?

18. A box contains 4 marbles, θ of which are white and the rest are black. A sample of size 2 is drawn to test $H_o : \theta = 2$ versus $H_a : \theta \neq 2$. If the null hypothesis is rejected if both marbles are the same color, find the significance level of the test.

19. Let X_1, X_2, X_3 denote a random sample of size 3 from a population X with probability density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & \text{for } 0 \leq x \leq \theta \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta < \infty$ is a parameter. What is the *likelihood ratio critical region* of size $\frac{117}{125}$ for testing $H_o : \theta = 5$ versus $H_a : \theta \neq 5$?

20. Let X_1, X_2 and X_3 denote three independent observations from a distribution with density

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

where $0 < \beta < \infty$ is a parameter. What is the *best (or uniformly most powerful* critical region for testing $H_o : \beta = 5$ versus $H_a : \beta = 10$?

21. Suppose X has the density function

$$f(x) = \begin{cases} \frac{1}{\theta} & \text{for } 0 < x < \theta \\ \\ 0 & \text{otherwise.} \end{cases}$$

If X_1, X_2, X_3, X_4 is a random sample of size 4 taken from X, what are the probabilities of Type I and Type II errors in testing the null hypothesis $H_o : \theta = 1$ against the alternative hypothesis $H_a : \theta = 2$, if H_o is rejected for $\max\{X_1, X_2, X_3, X_4\} \leq \frac{1}{2}$.

22. Let X_1, X_2, X_3 denote a random sample of size 3 from a population X with probability density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & \text{if } 0 < x < \infty \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \theta < \infty$ is a parameter. The null hypothesis $H_o : \theta = 3$ is to be rejected in favor of the alternative $H_a : \theta \neq 3$ if and only if $\overline{X} > 6.296$. What is the significance level of the test?