Chapter 5: Hypothesis Testing

STAT2602A Probability and statistics II
(2024-2025 1st Semester)

Contents

Introduction

5.1 Basic concepts

- Power function
- P-value

5.2 Most powerful tests

5.3 Generalized likelihood ratio tests: One-sample case

- 5.3.1 Testing for the mean: Variance is known
- 5.3.2 Testing for the mean: Variance is unknown
- 5.3.3 Testing for the variance
- 5.3.4 Test and interval estimation

5.4 Generalized likelihood ratio tests: Two-sample case

- 5.4.1 Testing for the mean: Variance is known
- 5.4.2 Testing for the mean: Variance is unknown
- 5.4.3 Testing for the variance

5.5 Generalized likelihood ratio tests: Large samples

- 5.5.1 Goodness-of-fit tests
- 5.5.2 Pearson Chi-squared test of independence

Introduction

Hypothesis Testing

A procedure used to determine (make a decision) whether a hypothesis should be rejected (declared false) or not

5.1 Basic concepts

- Hypothesis: a statement (or claim) about a population.
- Hypothesis test: a rule that leads to a decision to or not to reject a hypothesis.
- Simple hypothesis: a hypothesis that completely specifies the distribution of the population.
- Composite hypothesis: a hypothesis that does not completely specify the distribution of the populatoin.
- Null hypothesis (H_0): a hypothesis that is assumed to be true before it can be rejected.
- Alternative hypothesis (H_1 or H_a): a hypothesis that will be accepted if the null hypothesis is rejected.

Hypothesis

Two competing hypotheses: $H_0: \ \theta \in \Omega_0$ versus $H_1: \ \theta \in \Omega_1$,

- Ω_0 and Ω_1 are disjoint sets of possible values of the parameter θ.

Example 5.1.
Suppose that the score of one course follows $N(\theta, 10)$, and we want to know whether the theoretical mean $\theta = 80$. In this case,

$H_0: \ \theta = 80$ versus $H_1: \ \theta \neq 80$.

Example 5.2.
Suppose that the score of one course follows $N(\theta, 10)$, and we want to know whether the theoretical mean $\theta = 80$ or 70. In this case, $H_0: \ \theta = 80$ versus $H_1: \ \theta = 70$.

5.1 Basic concepts

- **Test statistic:** the statistic upon which the statistical decision will be based.
- **Rejection region or critical region:** the set of values of the test statistic for which the null hypothesis is rejected.
- **Acceptance region:** the set of values of the test statistic for which the null hypothesis is not rejected (is accepted).
- **Type I error:** rejection of the null hypothesis when it is true.
- **Type II error:** acceptance of the null hypothesis when it is false.

5.1 Basic concepts

(Def 5.1)

Power function

The power function $\pi(\theta)$ is the probability of rejecting H_0 when the true value of the parameter is θ, i.e.,

$$\pi(\theta) := \mathrm{P}_\theta(W(\mathbf{X}) \in R).$$

Let $\alpha(\theta)$ and $\beta(\theta)$ be probabilities of committing a type I and type II error respectively when the true value of the parameter is θ.

$$\begin{aligned} \alpha(\theta) &= \mathrm{P}_\theta(W(\mathbf{X}) \in R) \quad \text{for } \theta \in \Omega_0; \\ \beta(\theta) &= \mathrm{P}_\theta(W(\mathbf{X}) \in R^c) \quad \text{for } \theta \in \Omega_1. \end{aligned}$$

From $\alpha(\theta)$ and $\beta(\theta)$, we know that

$$\pi(\theta) = \begin{cases} \alpha(\theta), & \text{for } \theta \in \Omega_0; \\ 1 - \beta(\theta), & \text{for } \theta \in \Omega_1. \end{cases}$$

5.1 Basic concepts

Power

The power of a hypothesis test is the probability that the test correctly rejects the null hypothesis H_0 when a specific alternative hypothesis H_1 is true.

$\beta =$ probability of a Type II error, known as a "false negative".
$1 - \beta =$ probability of a "true positive", i.e., correctly rejecting the null hypothesis, is also known as the power of the test.
$\alpha =$ probability of a Type I error, known as a "false positive".
$1 - \alpha =$ probability of a "true negative", i.e., correctly not rejecting the null hypothesis.

5.1 Basic concepts

(Example 5.3)

A manufacturer of drugs has to decide whether 90% of all patients given a new drug will recover from a certain disease. Suppose
(a) the alternative hypothesis is that 60% of all patients given the new drug will recover;
(b) the test statistic is W, the observed number of recoveries in 20 trials;
(c) he will accept the null hypothesis when $W > 14$ and reject it otherwise.

- Find the power function of W.

5.1 Basic concepts

(Example 5.3)
Let $p = \mathrm{P}(\text{"recovery"})$. The hypotheses are

$$H_0 : p = 0.9 \text{ versus } H_1 : p = 0.6.$$

The test statistic W follows a binomial distribution $\mathrm{B}(n, p)$ with parameters $n = 20$ and p. The rejection region is $\{W \leq 14\}$.

$$\begin{aligned}
\pi(p) &= \mathrm{P}_p(W \leq 14) \\
&= 1 - \mathrm{P}_p(W > 14) \\
&= 1 - \sum_{k=15}^{20} \binom{20}{k} p^k (1-p)^{20-k} \\
&\approx \begin{cases} 0.0113, & \text{for } p = 0.9; \\ 0.8744, & \text{for } p = 0.6. \end{cases}
\end{aligned}$$

This implies that the probability of committing a type I and type II error are 0.0113 and 0.1256, respectively.

5.1 Basic concepts

(Example 5.4)
Let $\mathbf{X}$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$, where σ^2 is known. Consider a test statistic $W = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ for hypotheses $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$. Assume that the rejection region is $\{W \geq K\}$. Then, the power function is

$$\begin{aligned} \pi(\mu) &= \mathrm{P}_\mu(W \geq K) \\ &= \mathrm{P}_\mu\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq K + \frac{\mu_0-\mu}{\sigma/\sqrt{n}}\right) \\ &= \mathrm{P}\left(Z \geq K + \frac{\mu_0-\mu}{\sigma/\sqrt{n}}\right), \end{aligned}$$

where $Z \sim \mathrm{N}(0,1)$. It is easy to see that

$$\lim_{\mu\to-\infty} \pi(\mu) = 0, \quad \lim_{\mu\to\infty} \pi(\mu) = 1, \quad \text{and} \quad \pi(\mu_0) = \alpha \text{ if } \mathrm{P}(Z \geq K) = \alpha.$$

$\square$

5.1 Basic concepts

(Def 5.2)

Size

For $\alpha \in [0,1]$, a test with power function $\pi(\theta)$ is a size α test if

$$\max_{\theta \in \Omega_0} \pi(\theta) = \alpha.$$

α is also called the *level of significance* or *significance level*. If H_0 is a simple hypothesis $\theta = \theta_0$, then $\alpha = \pi(\theta_0)$.

5.1 Basic concepts

(Example 5.5)

Suppose that we want to test the null hypothesis that the mean of a normal population with $\sigma^2 = 1$ is μ_0 against the alternative hypothesis that it is μ_1, where $\mu_1 > \mu_0$.

(a) Find the value of K such that $\{\overline{X} \geq K\}$ provides a rejection region with the level of significance $\alpha = 0.05$ for a random sample of size n.

(b) For the rejection region found in (a), if $\mu_0 = 10$, $\mu_1 = 11$ and we need the type II probability $\beta \leq 0.06$, what should n be?

5.1 Basic concepts

(Example 5.5)
(a)

$H_0 : \mu = \mu_0$, $H_1 : \mu = \mu_1$, the rejection region is $\{\overline{X} \geq K\}$.

$$
\begin{aligned}
\alpha = \pi(\mu_0) &= \mathrm{P}_{\mu_0}(\overline{X} \geq K) \\
&= \mathrm{P}_{\mu_0}\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \geq \frac{K - \mu_0}{\sigma/\sqrt{n}}\right) \\
&= \mathrm{P}\left(Z \geq \frac{K - \mu_0}{1/\sqrt{n}}\right),
\end{aligned}
$$

where $Z \sim \mathrm{N}(0, 1)$.

When $\alpha = 0.05$, $0.05 = \mathrm{P}\left(Z \geq \frac{K-\mu_0}{1/\sqrt{n}}\right)$

$$
\frac{K - \mu_0}{1/\sqrt{n}} = z_{0.05} \approx 1.645 \quad \text{or} \quad K \approx \mu_0 + \frac{1.645}{\sqrt{n}}.
$$

5.1 Basic concepts

(Example 5.5)

(b)

$$\begin{aligned}
\beta = 1 - \pi(\mu_1) &= \mathrm{P}_{\mu_1}(\overline{X} < K) \\
&= \mathrm{P}_{\mu_1}\left(\frac{\overline{X} - \mu_1}{\sigma/\sqrt{n}} < \frac{K - \mu_1}{\sigma/\sqrt{n}}\right) \\
&= \mathrm{P}\left(Z < \frac{K - \mu_1}{\sigma/\sqrt{n}}\right) \\
&\approx \mathrm{P}\left(Z < \frac{\mu_0 + \frac{1.645}{\sqrt{n}} - \mu_1}{\sigma/\sqrt{n}}\right).
\end{aligned}$$

With $\mu_0 = 10$, $\mu_1 = 11$, $\sigma^2 = 1$,

$$\beta \approx \mathrm{P}\left(Z < \sqrt{n}(\mu_0 - \mu_1) + 1.645\right) = \mathrm{P}\left(Z < -\sqrt{n} + 1.645\right).$$

$$\begin{aligned}
\beta \leq 0.06 &\iff -\sqrt{n} + 1.645 \leq -z_{0.06} \approx -1.555 \\
&\iff n \geq (1.645 + 1.555)^2 \approx 10.24,\ n \geq 11.
\end{aligned}$$

5.1 Basic concepts

Steps to perform a hypothesis test

- State the null and alternative hypotheses and the level of significance α
- Choose a test statistic
- Determine the rejection region
- Calculate the value of the test statistic according to the particular sample drawn
- Make a decision: reject H_0 if and only if the value of the test statistic falls in the rejection region

5.1 Basic concepts

P-value

We can also use P-value to make a decision.

(Def 3.1)
Let $W(\mathbf{x})$ be the observed value of the test statistic $W(\mathbf{X})$.
Case 1: The rejection region is $\{W(\mathbf{X}) \leq K\}$, then

$$p\text{-value} = \max_{\theta \in \Omega_0} \mathrm{P}_\theta(W(\mathbf{X}) \leq W(\mathbf{x}));$$

Case 2: The rejection region is $\{W(\mathbf{X}) \geq K\}$, then

$$p\text{-value} = \max_{\theta \in \Omega_0} \mathrm{P}_\theta(W(\mathbf{X}) \geq W(\mathbf{x}));$$

Case 3: The rejection region is $\{|W(\mathbf{X})| \geq K\}$, then

$$p\text{-value} = \max_{\theta \in \Omega_0} \mathrm{P}_\theta(|W(\mathbf{X})| \geq |W(\mathbf{x})|).$$

5.1 Basic concepts

P-value

(Property 5.1)

For a test statistic $W(\mathbf{X})$,

$$H_0 \text{ is rejected at the significance level } \alpha \iff p\text{-value} \leq \alpha.$$

(Example 5.3)

A manufacturer of drugs has to decide whether 90% of all patients given a new drug will recover from a certain disease. Suppose
(a) the alternative hypothesis is that 60% of all patients given the new drug will recover;
(b) the test statistic is W, the observed number of recoveries in 20 trials;
(c) he will accept the null hypothesis when $W > 14$ and reject it otherwise.

- Find the power function of W.

5.1 Basic concepts

P-value

(Example 5.3)

<u>Method 2</u>

If the observed value of W is 12, then

$$\begin{aligned} p\text{-value} &= \mathrm{P}_p(W \le 12) \text{ for } p = 0.9 \\ &= \sum_{k=0}^{12} \binom{20}{k} (0.9)^k (0.1)^{20-k} \approx 0.0004. \end{aligned}$$

At the significance level $\alpha = 0.05$, the null hypothesis is rejected.

5.1 Basic concepts

(Example 5.5)

Suppose that we want to test the null hypothesis that the mean of a normal population with $\sigma^2 = 1$ is μ_0 against the alternative hypothesis that it is μ_1, where $\mu_1 > \mu_0$.

(a) Find the value of K such that $\{\overline{X} \geq K\}$ provides a rejection region with the level of significance $\alpha = 0.05$ for a random sample of size n.

(b) For the rejection region found in (a), if $\mu_0 = 10$, $\mu_1 = 11$ and we need the type II probability $\beta \leq 0.06$, what should n be?

5.1 Basic concepts

P-value

(Example 5.5)

Method 2

If the observed value of $\overline{X}$ is 10.417, then

$$\begin{aligned} p\text{-value} &= \mathrm{P}_\mu(\overline{X} \geq 10.417) \text{ for } \mu = 10 \\ &= \mathrm{P}_\mu\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{10.417-\mu}{\sigma/\sqrt{n}}\right) \text{ for } \mu = 10 \\ &= \mathrm{P}\left(Z \geq \frac{10.417-10}{1/\sqrt{11}}\right) \\ &\approx 0.0833. \end{aligned}$$

At the significance level $\alpha = 0.05$, the null hypothesis $H_0 : \mu = 10$ is not rejected.

5.2 Most powerful tests

(Def 5.4)

Most powerful tests

A test concerning a simple null hypothesis $\theta = \theta_0$ against a simple alternative hypothesis $\theta = \theta_1$ is said to be most powerful if the power of the test at $\theta = \theta_1$ is a maximum.

Method:

Consider the **likelihood ratio** $\frac{L(\theta_0)}{L(\theta_1)}$.

5.2 Most powerful tests

(Theorem 5.1)

Neyman-Pearson Lemma

Suppose $X_1, X_2, \ldots, X_n$ constitute a random sample of size n from a population with exactly one unknown parameter θ. Suppose that there is a positive constant k and a region C such that

$$(i)\ \mathrm{P}_\theta\left\{(X_1, X_2, \ldots, X_n) \in C\right\} = \alpha \quad \text{for } \theta = \theta_0,$$

and

$$\begin{aligned}(ii)\ \frac{\mathbf{f}(x_1, x_2, \ldots, x_n; \theta_0)}{\mathbf{f}(x_1, x_2, \ldots, x_n; \theta_1)} &\leq k \qquad \text{when } (x_1, x_2, \ldots, x_n) \in C,\\ (iii)\ \frac{\mathbf{f}(x_1, x_2, \ldots, x_n; \theta_0)}{\mathbf{f}(x_1, x_2, \ldots, x_n; \theta_1)} &\geq k \qquad \text{when } (x_1, x_2, \ldots, x_n) \notin C.\end{aligned}$$

5.2 Most powerful tests

(Theorem 5.1)

Neyman-Pearson Lemma

Construct a test, called the **likelihood ratio test**, which rejects $H_0 : \theta = \theta_0$ and accepts $H_1 : \theta = \theta_1$ if and only if $(X_1, X_2, \ldots, X_n) \in C$.
Then any other test which has significance level $\alpha^* \leq \alpha$ has power not more than that of this likelihood ratio test. In other words, the likelihood ratio test is **most powerful** among all tests having significance level $\alpha^* \leq \alpha$.

5.2 Most powerful tests

Neyman-Pearson Lemma says that to test $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, the rejection region for the likelihood ratio test is

$$\frac{L(\theta_0)}{L(\theta_1)} \leq k \Longleftrightarrow (X_1, X_2, \cdots, X_n) \in C \Longleftrightarrow W(\mathbf{X}) \in R,$$

where the interval R is chosen so that the test has the significance level α.

5.2 Most powerful tests

(Example 5.6)
A random sample $\{X_1, X_2, \ldots, X_n\}$ from a normal population $\mathrm{N}(\mu, \sigma^2)$, where $\sigma^2 = \sigma_0^2$ is known, is to be used to test the null hypothesis $\mu = \mu_0$ against the alternative hypothesis $\mu = \mu_1$, where $\mu_1 > \mu_0$.
Use the Neyman-Pearson Lemma to construct the most powerful test.

5.2 Most powerful tests

(Example 5.6)

$$L(\mu) = \left(\frac{1}{\sigma_0\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(X_i-\mu)^2\right].$$

The likelihood ratio test rejects the null hypothesis $\mu = \mu_0$ if and only if $\frac{L(\mu_0)}{L(\mu_1)} \le k$,

$$\exp\left\{\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\left[(X_i-\mu_1)^2-(X_i-\mu_0)^2\right]\right\} \le k$$

$$\iff \quad \sum_{i=1}^{n}(-2\mu_1 X_i + \mu_1^2 + 2\mu_0 X_i - \mu_0^2) \le 2\sigma_0^2 \log k$$

$$\iff \quad n(\mu_1^2-\mu_0^2) + 2(\mu_0-\mu_1)\sum_{i=1}^{n} X_i \le 2\sigma_0^2 \log k$$

$$\iff \quad \overline{X} \ge \frac{2\sigma_0^2 \log k - n(\mu_1^2-\mu_0^2)}{2n(\mu_0-\mu_1)} \qquad \text{(since } \mu_1 > \mu_0\text{)}.$$

5.2 Most powerful tests

(Example 5.6)
Therefore, in order that the level of significance is α, we should choose a constant K such that $\mathrm{P}_\mu(\overline{X} \geq K) = \alpha$ for $\mu = \mu_0$,

$$\mathrm{P}\left(\frac{\overline{X} - \mu_0}{\sigma_0/\sqrt{n}} \geq \frac{K - \mu_0}{\sigma_0/\sqrt{n}}\right) = \alpha \Longleftrightarrow \frac{K - \mu_0}{\sigma_0/\sqrt{n}} = z_\alpha \Longleftrightarrow K = \mu_0 + \frac{\sigma_0 z_\alpha}{\sqrt{n}}.$$

Therefore, the most powerful test having significance level $\alpha^* \leq \alpha$ is the one which has the rejection region

$$\left\{\overline{X} \geq \mu_0 + \frac{\sigma_0 z_\alpha}{\sqrt{n}}\right\} \quad \text{or} \quad \left\{\frac{\overline{X} - \mu_0}{\sigma_0/\sqrt{n}} \geq z_\alpha\right\}.$$

5.2 Most powerful tests

(Example 5.7)
Suppose $X_1, X_2, \ldots, X_n$ constitute a random sample of size n from a population given by a density

$$f(x) = \theta x^{\theta-1} \mathrm{I}(0 \le x \le 1).$$

If $0 \le X_i \le 1$ for $i = 1, 2, \ldots, n$, find the form of the most powerful test for testing

$$H_0 : \theta = 2 \text{ versus } H_1 : \theta = 1.$$

5.2 Most powerful tests

(Example 5.7)
The likelihood function of the sample is

$$L(\theta) = \theta^n \left(\prod_{i=1}^{n} X_i \right)^{\theta - 1} \prod_{i=1}^{n} \mathrm{I}(0 \leq X_i \leq 1) = \theta^n \left(\prod_{i=1}^{n} X_i \right)^{\theta - 1}.$$

Hence, the likelihood ratio is

$$\frac{L(2)}{L(1)} = 2^n \prod_{i=1}^{n} X_i.$$

The likelihood ratio test rejects H_0 if and only if $2^n \prod_{i=1}^{n} X_i \leq K$ where K is a positive constant (or, equivalently, $\prod_{i=1}^{n} X_i \leq k$ where k is a positive constant).

5.3 Generalized likelihood ratio tests: One-sample case

Suppose that $\theta \in \Omega$, where Ω is the parametric space. Consider the following hypotheses:

$$H_0 : \theta \in \Omega_0 \quad \text{versus} \quad H_1 : \theta \in \Omega_1,$$

where Ω_1 is the complement of Ω_0 with respect to Ω (i.e., $\Omega_1 = \Omega/\Omega_0$). Let

$$L(\Omega_0) = \max_{\theta \in \Omega_0} L(\theta) \quad \text{and} \quad L(\Omega) = \max_{\theta \in \Omega} L(\theta).$$

The **generalized likelihood ratio** is defined as

$$\Lambda = \frac{L(\Omega_0)}{L(\Omega)}.$$

The null hypothesis H_0 is rejected if and only if Λ falls in a rejection region of the form $\Lambda \leq k$, where $0 \leq k \leq 1$.

5.3.1 Testing for the mean: Variance is known

(Example 5.8)
Find the generalized likelihood ratio test for testing

$$H_0 : \mu = \mu_0 \text{ versus } H_1 : \mu \neq \mu_0$$

on the basis of a random sample of size n from $\mathrm{N}(\mu, \sigma^2)$, where $\sigma^2 = \sigma_0^2$ is known.

5.3.1 Testing for the mean: Variance is known

(Example 5.8)

Since Ω_0 contains only μ_0,

$$L(\Omega_0) = \left(\frac{1}{\sigma_0\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(X_i - \mu_0)^2\right].$$

Since the maximum likelihood estimator of μ is $\overline{X}$,

$$L(\Omega) = \left(\frac{1}{\sigma_0\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(X_i - \overline{X})^2\right].$$

$$\begin{aligned}\Lambda = \frac{L(\Omega_0)}{L(\Omega)} &= \exp\left\{-\frac{1}{2\sigma_0^2}\left[\sum_{i=1}^{n}(X_i - \mu_0)^2 - \sum_{i=1}^{n}(X_i - \overline{X})^2\right]\right\} \\ &= \exp\left[-\frac{n(\overline{X} - \mu_0)^2}{2\sigma_0^2}\right].\end{aligned}$$

5.3.1 Testing for the mean: Variance is known

(Example 5.8)

The rejection region is $\left\{\left|\overline{X}-\mu_0\right| \geq K\right\}$. In order that the level of significance is α, that is,

$$\mathrm{P}_\mu\left(\left|\overline{X}-\mu_0\right| \geq K\right)=\alpha \quad \text{for } \mu=\mu_0,$$

we should let $K=z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}$, so that

$$\begin{aligned}\mathrm{P}_\mu\left(\left|\overline{X}-\mu_0\right| \geq K\right) &= \mathrm{P}_\mu\left(\left|\overline{X}-\mu_0\right| \geq z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right)\\ &= \mathrm{P}\left(Z \geq z_{\alpha/2}\right)+\mathrm{P}\left(Z \leq -z_{\alpha/2}\right)\\ &= \frac{\alpha}{2}+\frac{\alpha}{2}=\alpha\end{aligned}$$

for $\mu=\mu_0$. So, the generalized likelihood ratio test has the rejection region

$$\left\{\frac{\left|\overline{X}-\mu_0\right|}{\sigma_0/\sqrt{n}} \geq z_{\alpha/2}\right\}$$

at the significance level α.

5.3.1 Testing for the mean: Variance is known

(Example 5.9)

The standard deviation of the annual incomes of government employees is \$1400. The mean is claimed to be \$35,000. Now a sample of 49 employees has been drawn and their average income is \$35,600. At the 5% significance level, can you conclude that the mean annual income of all government employees is not \$35,000?

5.3.1 Testing for the mean: Variance is known

(Example 5.9)

Solution 1

$$H_0 : \mu = 35000 \quad \text{versus} \quad H_1 : \mu \neq 35000.$$

The test statistic is

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\overline{X} - 35000}{1400/\sqrt{49}} = \frac{\overline{X} - 35000}{200},$$

which follows $N(0,1)$ under H_0.
At the significance level $\alpha = 5\%$, the rejection region is

$$\left\{|Z| \geq z_{\alpha/2}\right\} \approx \{|Z| \geq 1.960\}.$$

Since $\overline{x} = 35600$, the value of the test statistic is

$$\frac{35600 - 35000}{200} = 3.$$

Since $|3| \geq 1.960$, we reject H_0 and accept H_1.

5.3.1 Testing for the mean: Variance is known

(Example 5.9)

Solution 2

$$H_0 : \mu = 35000 \quad \text{versus} \quad H_1 : \mu \neq 35000.$$

The test statistic is

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\overline{X} - 35000}{1400/\sqrt{49}} = \frac{\overline{X} - 35000}{200},$$

which follows $N(0,1)$ under H_0. Since $\overline{x} = 35600$, the value of the test statistic is

$$\frac{35600 - 35000}{200} = 3.$$

$$p\text{-value} \quad = \quad \mathrm{P}\left(|Z| \geq |3|\right) = 2\mathrm{P}\left(Z \geq 3\right) \approx 2(0.5 - 0.4987) = 0.0026,$$

where Z follows $N(0,1)$. Since $0.0026 \leq 0.05 = \alpha$, we reject H_0 and accept H_1.

Therefore we conclude that the mean annual income of all government employees is not \$35,000.

5.3.1 Testing for the mean: Variance is known

(Example 5.10)
The chief financial officer in FedEx believes that including a stamped self-addressed envelope in the monthly invoice sent to customers will reduce the amount of time it takes for customers to pay their monthly bills. Currently, customers return their payments in 24 days on average, with a standard deviation of 6 days. It was calculated that an improvement of two days on average would cover the costs of the envelopes (because cheques can be deposited earlier). A random sample of 220 customers was selected and stamped self-addressed envelopes were included in their invoice packs. The amounts of time taken for these customers to pay their bills were recorded and their mean is 21.63 days. Assume that the corresponding population standard deviation is still 6 days.

- Can the chief financial officer conclude that the plan will be profitable at the 10% significance level?

5.3.1 Testing for the mean: Variance is known

(Example 5.10)

<u>Solution 1</u>

$$H_0 : \mu \geq 22 \quad \text{versus} \quad H_1 : \mu < 22.$$

$$\frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{21.63 - 22}{6/\sqrt{220}} \approx -0.9147.$$

Since $-0.9147 > -1.282 \approx -z_{0.1}$, H_0 should not be rejected.

The chief financial officer cannot conclude that the plan is profitable at the 10% significance level.

5.3.1 Testing for the mean: Variance is known

(Example 5.10)
Solution 2 Consider

$$p\text{-value} \approx \mathrm{P}(Z \leq -0.9147) \approx 0.1814 > 0.1 = \alpha,$$

where Z follows $N(0,1)$.
Therefore, H_0 should not be rejected.

The chief financial officer cannot conclude that the plan is profitable at the 10% significance level.

5.3.2 Testing for the mean: Variance is unknown

(Example 5.11)
Find the generalized likelihood ratio test for testing

$$H_0 : \mu = \mu_0 \text{ versus } H_1 : \mu > \mu_0$$

on the basis of a random sample of size n from $\mathrm{N}(\mu, \sigma^2)$.

5.3.2 Testing for the mean: Variance is unknown

(Example 5.11)

$$\begin{aligned} \Omega &= \{(\mu, \sigma) : \mu \geq \mu_0, \sigma > 0\}, \\ \Omega_0 &= \{(\mu, \sigma) : \mu = \mu_0, \sigma > 0\}, \\ \Omega_1 &= \{(\mu, \sigma) : \mu > \mu_0, \sigma > 0\}. \end{aligned}$$

The likelihood function of the sample is

$$L(\mu, \sigma) \quad = \quad = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2\right],$$

and hence,

$$\begin{aligned} \frac{\partial \ln L(\mu, \sigma)}{\partial \mu} &= \frac{n}{\sigma^2}(\overline{X} - \mu), \\ \frac{\partial \ln L(\mu, \sigma)}{\partial \sigma} &= -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(X_i - \mu)^2. \end{aligned}$$

5.3.2 Testing for the mean: Variance is unknown

(Example 5.11)

On Ω_0, the maximum value of $L(\mu, \sigma)$ is $L(\mu_0, \tilde{\sigma})$, where $\tilde{\sigma}$ satisfies

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2.$$

This is because $\tilde{\sigma}$ is the maximum value of $\ln L(\mu_0, \sigma)$, by noting that for all $\mu > 0$,

$$\sigma < \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2} \iff \frac{\partial \ln L(\mu, \sigma)}{\partial \sigma} > 0,$$

$$\sigma > \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2} \iff \frac{\partial \ln L(\mu, \sigma)}{\partial \sigma} < 0.$$

Therefore,

$$L(\Omega_0) = L(\mu_0, \tilde{\sigma}) = \left(\frac{1}{\tilde{\sigma}\sqrt{2\pi}}\right)^n \exp\left(-\frac{n}{2}\right).$$

5.3.2 Testing for the mean: Variance is unknown

(Example 5.11)
On Ω, the maximum value of $L(\mu, \sigma)$ is $L(\hat{\mu}, \hat{\sigma})$, where (*noting that* $L(\mu, \sigma)$ *decreases with respect to* μ *when* $\mu > \overline{X}$ *and increases with respect to* μ *when* $\mu < \overline{X}$)

$$\hat{\mu} = \begin{cases} \mu_0, & \text{if } \overline{X} \leq \mu_0; \\ \overline{X}, & \text{if } \overline{X} > \mu_0, \end{cases}$$

and

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \hat{\mu})^2.$$

Therefore,

$$L(\Omega) = L(\hat{\mu}, \hat{\sigma}) = \left(\frac{1}{\hat{\sigma}\sqrt{2\pi}} \right)^n \exp\left(-\frac{n}{2} \right).$$

5.3.2 Testing for the mean: Variance is unknown

(Example 5.11)

Thus, we have

$$\Lambda = \frac{L(\Omega_0)}{L(\Omega)} = \left(\frac{\hat{\sigma}}{\tilde{\sigma}}\right)^n = \left(\frac{\hat{\sigma}^2}{\tilde{\sigma}^2}\right)^{n/2} = \begin{cases} 1, & \text{if } \overline{X} \le \mu_0; \\ \left[\dfrac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sum_{i=1}^{n}(X_i - \mu_0)^2}\right]^{n/2}, & \text{if } \overline{X} > \mu_0. \end{cases}$$

The rejection region is $\{\Lambda \le k\}$ for some nonnegative constant $k < 1$ (since we do not want α to be 1). Then $\{\Lambda \le k\} \subseteq \{\overline{X} > \mu_0\}$ and $\Lambda \le k$ is equivalent to

$$k^{2/n} \ge \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sum_{i=1}^{n}(X_i - \mu_0)^2} = \frac{1}{1 + \dfrac{n(\overline{X} - \mu_0)^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2}}$$

5.3.2 Testing for the mean: Variance is unknown

(Example 5.11)

$$\frac{(\overline{X}-\mu_0)^2}{S^2} = \frac{n(\overline{X}-\mu_0)^2}{\sum_{i=1}^{n}(X_i-\overline{X})^2} \geq k^{-2/n}-1,$$

In order that the level of significance is α, that is,

$$\mathrm{P}_{(\mu,\sigma)}\left(\frac{\overline{X}-\mu_0}{S/\sqrt{n-1}} \geq c\right) = \alpha \quad \text{for } \mu=\mu_0,$$

we should let $c = t_{\alpha,n-1}$, since

$$\mathrm{P}_{(\mu,\sigma)}\left(\frac{\overline{X}-\mu_0}{S/\sqrt{n-1}} \geq t_{\alpha,n-1}\right) = \mathrm{P}\left(t_{n-1} \geq t_{\alpha,n-1}\right) \quad \text{for } \mu=\mu_0$$

by Property 4.1(iii). So, the generalized likelihood ratio test has the rejection region $\left\{\frac{\overline{X}-\mu_0}{S/\sqrt{n-1}} \geq t_{\alpha,n-1}\right\}$ at the significance level α.

5.3.2 Testing for the mean: Variance is unknown

(Example 5.12)

What will happen if we change H_0 in the previous example(5.11) to be $\mu \leq \mu_0$?

5.3.2 Testing for the mean: Variance is unknown

(Example 5.12)

$$\begin{aligned}
\Omega &= \{(\mu, \sigma) : -\infty < \mu < \infty, \sigma > 0\}, \\
\Omega_0 &= \{(\mu, \sigma) : \mu \le \mu_0, \sigma > 0\}, \\
\Omega_1 &= \{(\mu, \sigma) : \mu > \mu_0, \sigma > 0\}.
\end{aligned}$$

On Ω_0, the maximum value of $L(\mu, \sigma)$ is $L(\tilde{\mu}, \tilde{\sigma})$ where

$$\tilde{\mu} = \begin{cases} \overline{X}, & \text{if } \overline{X} < \mu_0; \\ \mu_0, & \text{if } \overline{X} \ge \mu_0, \end{cases}$$

and

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \tilde{\mu})^2.$$

Therefore,

$$L(\Omega_0) = L(\tilde{\mu}, \tilde{\sigma}) = \left(\frac{1}{\tilde{\sigma}\sqrt{2\pi}} \right)^n \exp\left(-\frac{n}{2} \right).$$

5.3.2 Testing for the mean: Variance is unknown

(Example 5.12)

$$\Lambda = \frac{L(\Omega_0)}{L(\Omega)} = \left(\frac{\hat{\sigma}}{\tilde{\sigma}}\right)^n = \left(\frac{\hat{\sigma}^2}{\tilde{\sigma}^2}\right)^{n/2} = \begin{cases} 1, & \text{if } \overline{X} \le \mu_0; \\ \left[\dfrac{\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2}{\displaystyle\sum_{i=1}^{n}(X_i - \mu_0)^2}\right]^{n/2}, & \text{if } \overline{X} > \mu_0. \end{cases}$$

So the generalized likelihood ratio test is the same as that in the previous example(5.11).

5.3.2 Testing for the mean: Variance is unknown

(Example 5.13)

According to the last census in a city, the mean family annual income was 316 thousand dollars. A random sample of 900 families taken this year produced a mean family annual income of 313 thousand dollars and a standard deviation of 70 thousand dollars.

At the 2.5% significance level, can we conclude that the mean family annual income has declined since the last census?

5.3.2 Testing for the mean: Variance is unknown

(Example 5.13)

Consider hypothesis

$$H_0 : \mu \geq 316 \quad \text{versus} \quad H_1 : \mu < 316.$$

The value of the test statistic is

$$\frac{\overline{x} - \mu_0}{s/\sqrt{n-1}} = \frac{313 - 316}{70/\sqrt{900-1}} \approx -1.286.$$

Since $-1.286 > -1.963 = -t_{0.025,899}$, we do not reject H_0. Thus we cannot conclude that the mean family annual income has declined since the last census at the 2.5% level of significance.

5.3.3 Testing for the variance

(Example 5.14)

Given a random sample of size n from a normal population with unknown mean and variance,

find the generalized likelihood ratio test for testing the null hypothesis $\sigma = \sigma_0$ ($\sigma_0 > 0$) against the alternative hypothesis $\sigma \neq \sigma_0$.

5.3.3 Testing for the variance

(Example 5.14)

Note that

$$\begin{aligned}\Omega &= \{(\mu,\sigma) : -\infty < \mu < \infty, \sigma > 0\}, \\ \Omega_0 &= \{(\mu,\sigma) : -\infty < \mu < \infty, \sigma = \sigma_0\}, \\ \Omega_1 &= \{(\mu,\sigma) : -\infty < \mu < \infty, \sigma > 0, \sigma \neq \sigma_0\},\end{aligned}$$

and the likelihood function of the sample $\{X_1, \cdots, X_n\}$ is

$$L(\mu,\sigma) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-\mu)^2\right].$$

On Ω_0, the maximum value of $L(\mu,\sigma)$ is $L(\tilde{\mu},\sigma_0)$ where $\tilde{\mu} = \overline{X}$. Therefore,

$$\begin{aligned}L(\Omega_0) &= L(\tilde{\mu},\sigma_0) \\ &= \left(\frac{1}{\sigma_0\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^n (X_i-\overline{X})^2\right].\end{aligned}$$

5.3.3 Testing for the variance

(Example 5.14)
On Ω, the maximum value of $L(\mu, \sigma)$ is $L(\hat{\mu}, \hat{\sigma})$ where $\hat{\mu} = \overline{X}$ and $\hat{\sigma}^2 = S^2$. Therefore,

$$L(\Omega) = L(\hat{\mu}, \hat{\sigma}) = \left(\frac{1}{\hat{\sigma}\sqrt{2\pi}}\right)^n \exp\left(-\frac{n}{2}\right).$$

Thus, we have

$$\begin{aligned}
\Lambda &= \frac{L(\Omega_0)}{L(\Omega)} = \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} \exp\left[-\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{2\sigma_0^2} + \frac{n}{2}\right] \\
&= \left[\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n\sigma_0^2}\right]^{n/2} \exp\left[-\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{2\sigma_0^2} + \frac{n}{2}\right].
\end{aligned}$$

5.3.3 Testing for the variance

(Example 5.14)

The rejection region is $\{\Lambda \le k\}$ for some positive constant $k < 1$ (since we do not want α to be 1). Letting $Y = \frac{1}{n\sigma_0^2}\sum_{i=1}^{n}(X_i - \overline{X})^2$,

$$\begin{aligned}
\Lambda \le k &\iff Y^{n/2}\exp\left(-\frac{nY}{2} + \frac{n}{2}\right) \le k, \\
&\iff Y\exp(-Y+1) \le k^{2/n}, \\
&\iff Y\exp(-Y) \le \frac{k^{2/n}}{e}.
\end{aligned}$$

For $y > 0$ define a function $g(y) = ye^{-y}$. Then,

$$\frac{\mathrm{d}g(y)}{\mathrm{d}y} = e^{-y} - ye^{-y} = (1-y)e^{-y}.$$

Since

$$y < 1 \iff \frac{\mathrm{d}g(y)}{\mathrm{d}y} > 0 \text{ and } y > 1 \iff \frac{\mathrm{d}g(y)}{\mathrm{d}y} < 0,$$

$g(y)$ will be small when y is close to zero or very large.

5.3.3 Testing for the variance

(Example 5.14)
Thus we reject the null hypothesis $\sigma = \sigma_0$ when the value of Y (or nY) is large or small, that is, the rejection region of our generalized likelihood ratio test has the rejection region:

$$\{nY \leq K_1\} \cup \{nY \geq K_2\}.$$

Note that $nY = \dfrac{nS^2}{\sigma_0^2}$.

5.3.3 Testing for the variance

(Example 5.14)
In order that the level of significance is α, that is,

$$\mathrm{P}_{(\mu,\sigma)}\left(\frac{nS^2}{\sigma_0^2} \le K_1\right) + \mathrm{P}\left(\frac{nS^2}{\sigma_0^2} \ge K_2\right) = \alpha \quad \text{for } \sigma = \sigma_0,$$

we should let $K_1 = \chi^2_{1-\alpha/2,n-1}$ and $K_2 = \chi^2_{\alpha/2,n-1}$, since

$$\mathrm{P}_{(\mu,\sigma)}\left(\frac{nS^2}{\sigma_0^2} \le K_1\right) = \mathrm{P}\left(\chi^2_{n-1} \le \chi^2_{1-\alpha/2,n-1}\right) = \frac{\alpha}{2}$$

and

$$\mathrm{P}_{(\mu,\sigma)}\left(\frac{nS^2}{\sigma_0^2} \ge K_2\right) = \mathrm{P}\left(\chi^2_{n-1} \ge \chi^2_{\alpha/2,n-1}\right) = \frac{\alpha}{2}$$

for $\sigma = \sigma_0$ by using the fact that $nY \sim \chi^2_{n-1}$ from Property 4.1(ii).

5.3.3 Testing for the variance

(Example 5.15)

One important factor in inventory control is the variance of the daily demand for the product. A manager has developed the optimal order quantity and reorder point, assuming that the variance is equal to 250. Recently, the company has experienced some inventory problems, which induced the operations manager to doubt the assumption. To examine the problem, the manager took a sample of 25 daily demands and found that $s^2 = 270.58$.

Do these data provide sufficient evidence at the 5% significance level to infer that the management scientist's assumption about the variance is wrong?

5.3.3 Testing for the variance

(Example 5.15)

Consider hypothesis

$$H_0 : \sigma^2 = 250 \quad \text{vesus} \quad H_1 : \sigma^2 \neq 250.$$

The value of test statistic is

$$\frac{ns^2}{\sigma_0^2} = \frac{25 \times 270.58}{250} \approx 25.976.$$

Since $\chi^2_{1-0.05/2,25-1} \approx 12.401 \leq 25.976 \leq 39.364 \approx \chi^2_{0.05/2,25-1}$, we do not reject H_0.

There is not sufficient evidence at the 5% significance level to infer that the management scientist's assumption about the variance is wrong.

5.3.4 Test and interval estimation

We can obtain the interval estimation by using the two-tailed hypothesis testing. For example, consider hypotheses

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad \mu \neq \mu_0.$$

If the variance is known, the acceptance region is

$$\left\{ \left| \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \right| < z_{\alpha/2} \right\} \Longleftrightarrow \overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu_0 < \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

at the significance level α. As H_0 is accepted, $\mu = \mu_0$ hence

$$\mathrm{P}\left(\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right) = 1 - \alpha.$$

That is, $\left[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right]$ is the $1 - \alpha$ confidence interval of μ.

5.3.4 Test and interval estimation

Similarly, we can find the confidence interval of μ when the variance is unknown, and σ^2 by using the two-tailed hypothesis testing.

5.4 Generalized likelihood ratio tests: Two-sample case

In this section, we assume that there are two populations following $\mathrm{N}(\mu_1, \sigma_1^2)$ and $\mathrm{N}(\mu_2, \sigma_2^2)$ respectively.

A sample $\{X_i,\ i = 1, 2, \ldots, n_1\}$ is taken from the population $\mathrm{N}(\mu_1, \sigma_1^2)$ and a sample $\{Y_j,\ j = 1, 2, \ldots, n_2\}$ is taken from the population $\mathrm{N}(\mu_2, \sigma_2^2)$.

Assume that these two samples are independent (that is, $X_1, X_2, \ldots, X_{n_1},\ Y_1, Y_2, \ldots, Y_{n_2}$ are independent).

5.4.1 Testing for the mean: Variance is known

We first consider the hypothesis testing for $\mu_1 - \mu_2$ when σ_1 and σ_2 are known.

(Example 5.16)

Assume that σ_1 and σ_2 are known.
Find the generalized likelihood ratio for testing

$$H_0 : \mu_1 - \mu_2 = \delta \qquad \text{versus} \qquad H_1 : \mu_1 - \mu_2 \neq \delta.$$

5.4.1 Testing for the mean: Variance is known

(Example 5.16)

Note that

$$\begin{aligned}
\Omega_0 &= \{(\mu_1, \mu_2) : \mu_1 - \mu_2 = \delta\}, \\
\Omega_1 &= \{(\mu_1, \mu_2) : \mu_1 - \mu_2 \neq \delta\}, \\
\Omega &= \Omega_0 \cup \Omega_1 = \{(\mu_1, \mu_2) : -\infty < \mu_1 < \infty, \ -\infty < \mu_2 < \infty\}.
\end{aligned}$$

The likelihood function of the two samples is

$$\begin{aligned}
L(\mu_1, \mu_2) &= \left(\frac{1}{\sigma_1\sqrt{2\pi}}\right)^{n_1} \exp\left[-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n_1}(X_i - \mu_1)^2\right] \\
&\quad \times \left(\frac{1}{\sigma_2\sqrt{2\pi}}\right)^{n_2} \exp\left[-\frac{1}{2\sigma_2^2}\sum_{j=1}^{n_2}(Y_j - \mu_2)^2\right].
\end{aligned}$$

5.4.1 Testing for the mean: Variance is known

(Example 5.16)

On Ω_0, we have

$$\ln L(\mu_1, \mu_2) = C - \frac{1}{2\sigma_1^2}\sum_{i=1}^{n_1}(X_i - \mu_1)^2 - \frac{1}{2\sigma_2^2}\sum_{j=1}^{n_2}(Y_j - \mu_1 + \delta)^2,$$

where C depends on neither μ_1 nor μ_2.

$$\begin{aligned}
\frac{\partial}{\partial \mu_1}\ln L(\mu_1, \mu_1 - \delta) &= \frac{1}{\sigma_1^2}\sum_{i=1}^{n_1}(X_i - \mu_1) + \frac{1}{\sigma_2^2}\sum_{j=1}^{n_2}(Y_j - \mu_1 + \delta) \\
&= \frac{n_1(\overline{X} - \mu_1)}{\sigma_1^2} + \frac{n_2(\overline{Y} - \mu_1 + \delta)}{\sigma_2^2} \\
&= \frac{n_1\overline{X}}{\sigma_1^2} + \frac{n_2(\overline{Y} + \delta)}{\sigma_2^2} - \left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)\mu_1.
\end{aligned}$$

5.4.1 Testing for the mean: Variance is known

(Example 5.16)
Therefore, the maximum likelihood estimator of μ_1 is

$$\tilde{\mu}_1 = \frac{\dfrac{n_1 \overline{X}}{\sigma_1^2} + \dfrac{n_2(\overline{Y} + \delta)}{\sigma_2^2}}{\dfrac{n_1}{\sigma_1^2} + \dfrac{n_2}{\sigma_2^2}},$$

since

$$\begin{aligned}
\mu_1 < \tilde{\mu}_1 \quad &\Longleftrightarrow \quad \frac{\partial}{\partial \mu_1} \ln L(\mu_1, \mu_1 - \delta) > 0, \\
\mu_1 > \tilde{\mu}_1 \quad &\Longleftrightarrow \quad \frac{\partial}{\partial \mu_1} \ln L(\mu_1, \mu_1 - \delta) < 0.
\end{aligned}$$

On Ω, it is easy to see that the maximum likelihood estimator of μ_1 is $\overline{X}$ and that of μ_2 is $\overline{Y}$,

5.4.1 Testing for the mean: Variance is known

(Example 5.16)
since $\mu_1 = \overline{X}$ maximizes

$$\left(\frac{1}{\sigma_1\sqrt{2\pi}}\right)^{n_1} \exp\left[-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n_1}(X_i-\mu_1)^2\right],$$

and $\mu_2 = \overline{Y}$ maximizes

$$\left(\frac{1}{\sigma_2\sqrt{2\pi}}\right)^{n_2} \exp\left[-\frac{1}{2\sigma_2^2}\sum_{j=1}^{n_2}(Y_j-\mu_2)^2\right].$$

Thus, the generalized likelihood ratio is

$$\begin{aligned}
\Lambda &= \frac{L(\Omega_0)}{L(\Omega)} \\
&= \exp\left[-\frac{n_1(\overline{X}-\tilde{\mu}_1)^2}{2\sigma_1^2} - \frac{n_2(\overline{Y}-\tilde{\mu}_1+\delta)^2}{2\sigma_2^2}\right] \\
&= \exp\left[C'(\overline{X}-\overline{Y}-\delta)^2\right],
\end{aligned}$$

where C' is negative and does not depend on the samples

5.4.1 Testing for the mean: Variance is known

(Example 5.16)

Because

$$\overline{X} - \tilde{\mu}_1 = \frac{\frac{n_2}{\sigma_2^2}(\overline{X} - \overline{Y} - \delta)}{\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}} \quad \text{and} \quad \overline{Y} - \tilde{\mu}_1 + \delta = \frac{\frac{n_1}{\sigma_1^2}(\overline{Y} + \delta - \overline{X})}{\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}}.$$

Therefore the rejection region should be $\left\{|\overline{X} - \overline{Y} - \delta| \geq K\right\}$.

Under H_0, we have

$$\begin{cases} \overline{X} \text{ follows } \mathrm{N}\left(\mu_1, \frac{\sigma_1^2}{n_1}\right), \\ \overline{Y} \text{ follows } \mathrm{N}\left(\mu_1 - \delta, \frac{\sigma_2^2}{n_2}\right), \end{cases}$$

and thus (by the independence between the two samples)

$$\overline{X} - \overline{Y} \text{ follows } \mathrm{N}\left(\delta, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

5.4.1 Testing for the mean: Variance is known

(Example 5.16)
Therefore, the rejection region is

$$\left\{ \frac{|\overline{X} - \overline{Y} - \delta|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \geq z_{\alpha/2} \right\},$$

where the test statistic is $\dfrac{\overline{X} - \overline{Y} - \delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$.

5.4.2 Testing for the mean: Variance is unknown

In this section, we consider the hypothesis testing for $\mu_1 - \mu_2$ when σ_1 and σ_2 are unknown but equal.

5.4.2 Testing for the mean: Variance is unknown

(Example 5.17)
Assume that σ_1 and σ_2 are unknown but equal to σ. Find the generalized likelihood ratio for testing

$$H_0: \mu_1 - \mu_2 = \delta \qquad \text{versus} \qquad H_1: \mu_1 - \mu_2 \neq \delta.$$

5.4.2 Testing for the mean: Variance is unknown

(Example 5.17)
The generalized likelihood ratio is

$$\Lambda = \frac{L(\Omega_0)}{L(\Omega)} = \frac{\tilde{\sigma}^{-(n_1+n_2)}}{\hat{\sigma}^{-(n_1+n_2)}} = \left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2}\right)^{-(n_1+n_2)/2}.$$

The rejection region is $\left\{|W| \geq t_{\alpha/2, n_1+n_2-2}\right\}$,

The test statistic is $W = \dfrac{\overline{X} - \overline{Y} - \delta}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$.

5.4.2 Testing for the mean: Variance is unknown

(Example 5.18)

A consumer agency wanted to estimate the difference in the mean amounts of caffeine in two brands of coffee. The agency took a sample of 15 500-gramme jars of Brand I coffee that showed the mean amount of caffeine in these jars to be 80 mg per jar and the standard deviation to be 5 mg. Another sample of 12 500-gramme jars of Brand II coffee gave a mean amount of caffeine equal to 77 mg per jar and a standard deviation of 6 mg. Assuming that the two populations are normally distributed with equal variances, check at the 5% significance level whether the mean amount of caffeine in 500-gramme jars is greater for Brand 1 than for Brand 2.

5.4.2 Testing for the mean: Variance is unknown

(Example 5.18)
Let the amounts of caffeine in jars of Brand I be referred to as population 1 and those of Brand II be referred to as population 2. We consider the hypotheses:

$$H_0 : \mu_1 \le \mu_2 \text{ versus } H_1 : \mu_1 > \mu_2,$$

where μ_1 and μ_2 are the mean of population 1 and population 2, respectively.
Note that

$$n_1 = 15,\ \bar{x}_1 = 80,\ s_1 = 5,$$

and

$$n_2 = 12,\ \bar{x}_2 = 77,\ s_2 = 6,\ \alpha = 0.05.$$

Hence, $\bar{x}_1 - \bar{x}_2 = 80 - 77 = 3,\ t_{\alpha, n_1+n_2-2} = t_{0.05,25} \approx 1.708$,
$\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \approx 0.3873$, and

$$s_p = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{15 * 5^2 + 12 * 6^2}{15 + 12 - 2}} \approx \ 5.4626.$$

5.4.2 Testing for the mean: Variance is unknown

(Example 5.18)
Therefore, the observed value of the test statistic is

$$w = \frac{\overline{x}_1 - \overline{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{3}{5.4626 * 0.3873} \approx 1.42.$$

As $1.42 < 1.708$, we can not reject H_0.

Thus, we conclude that the mean amount of caffeine in 500-gramme jars is not greater for Brand 1 than for Brand 2 at the 5% significance level.

5.4.3 Testing for the variance

In this section, we perform tests comparing σ_1 and σ_2.

5.4.3 Testing for the variance

(Example 5.19)

Find the generalized likelihood ratio test for hypotheses

$$H_0 : \sigma_1 = \sigma_2 \;\; \text{versus} \;\; H_1 : \sigma_1 \neq \sigma_2.$$

5.4.3 Testing for the variance

(Example 5.19)

The generalized likelihood ratio is $\dfrac{C\left(\dfrac{S_1^2}{S_2^2}\right)^{n_1/2}}{\left[n_1\dfrac{S_1^2}{S_2^2}+n_2\right]^{(n_1+n_2)/2}}$,

where C is a constant.

For $w > 0$ define the function

$$G(w) = \frac{w^{n_1/2}}{[n_1 w + n_2]^{(n_1+n_2)/2}}.$$

$$\begin{aligned} \ln G(w) &= \frac{n_1}{2}\ln w - \frac{n_1+n_2}{2}\ln\left[n_1 w + n_2\right], \\ \frac{\mathrm{d}}{\mathrm{d}w}\ln G(w) &= \frac{n_1 n_2(1-w)}{2w\left[n_1 w + n_2\right]}, \end{aligned}$$

which is negative when $w > 1$ and is positive when $w < 1$.

5.4.3 Testing for the variance

(Example 5.19)

Therefore, the value of $G(w)$ will be small when w is very large or very small.

Therefore H_0 should be rejected when $\frac{S_1^2}{S_2^2}$ is large or small.

When H_0 is true, $\frac{n_1(n_2-1)S_1^2}{n_2(n_1-1)S_2^2} = \frac{\frac{n_1 S_1^2}{(n_1-1)\sigma_1^2}}{\frac{n_2 S_2^2}{(n_2-1)\sigma_2^2}}$ follows F_{n_1-1,n_2-1} by Property 4.3.

Thus, we let the test statistic be $W = \frac{n_1(n_2-1)S_1^2}{n_2(n_1-1)S_2^2}$,

and the rejection region is

$\{W \leq F_{1-\alpha/2,n_1-1,n_2-1}\} \cup \{W \geq F_{\alpha/2,n_1-1,n_2-1}\}$.

5.4.3 Testing for the variance

(Example 5.20)
A study involves the number of absences per year among union and non-union workers. A sample of 16 union workers has a sample standard deviation of 3.0 days. A sample of 10 non-union workers has a sample standard deviation of 2.5 days. At the 10% significance level, can we conclude that the variance of the number of days absent for union workers is different from that for nonunion workers?

5.4.3 Testing for the variance

(Example 5.20)

Let all union workers be referred to as population 1 and all non-union workers be referred to as population 2.

$$H_0 : \sigma_1 = \sigma_2 \text{ versus } H_1 : \sigma_1 \neq \sigma_2,$$

where σ_1^2 and σ_2^2 are the variance of population 1 and population 2, respectively.

Note that $n_1 = 16$, $s_1 = 3$, $n_2 = 10$, and $s_2 = 2.5$.

Hence, the value of the test statistic is

$$\frac{n_1(n_2 - 1)}{n_2(n_1 - 1)} \frac{s_1^2}{s_2^2} = 0.96 * \frac{3.0^2}{2.5^2} = 1.3824.$$

Since

$$\frac{1}{f_{0.05,9,15}} < 1 < 1.3824 < 3.006 \approx f_{0.05,15,9},$$

we cannot reject H_0.

Thus we conclude that the data do not indicate that the variance of the number of days absent for union workers is different from that for non-union workers at the 10% significance level.

5.5 Generalized likelihood ratio tests: Large samples

(Theorem 5.2)

Suppose that we are testing

$$H_0 : \theta_i = \theta_{i,0} \text{ for all } i = 1, 2, \ldots, d$$

versus

$$H_1 : \theta_i \neq \theta_{i,0} \text{ for at least one } i = 1, 2, \ldots, d$$

and that Λ is the generalized likelihood ratio. Then, under very general conditions, when H_0 is true,

$$-2 \ln \Lambda \to_d \chi^2_d \text{ as } n \to \infty.$$

5.5.1 Goodness-of-fit tests

(Example 5.21)

A journal reported that, in a bag of m&m's chocolate peanut candies, there are 30% brown, 30% yellow, 10% blue, 10% red, 10% green and 10% orange candies. Suppose you purchase a bag of m&m's chocolate peanut candies at a nearby store and find 17 brown, 20 yellow, 13 blue, 7 red, 6 green and 9 orange candies, for a total of 72 candies. At the 0.1 level of significance, does the bag purchased agree with the distribution suggested by the journal?

5.5.1 Goodness-of-fit tests

(Example 5.21)
H_0: the bag purchased agrees with the distribution suggested by the journal,
versus
H_1: the bag purchased does not agree with the distribution suggested by the journal.

Colour	O_i	E_i	$O_i - E_i$
Brown	17	$72 \times 30\% = 21.6$	-4.6
Yellow	20	$72 \times 30\% = 21.6$	-1.6
Blue	13	$72 \times 10\% = 7.2$	5.8
Red	7	$72 \times 10\% = 7.2$	-0.2
Green	6	$72 \times 10\% = 7.2$	-1.2
Orange	9	$72 \times 10\% = 7.2$	1.8
Total	72	72	0

All expected frequencies are at least 5.

5.5.1 Goodness-of-fit tests

(Example 5.21)
Therefore, as the sample is large enough,

$$
\begin{aligned}
-2\ln\Lambda &\approx \sum_{i=1}^{6}\frac{O_i^2}{E_i} - n \\
&= \frac{17^2+20^2}{21.6} + \frac{13^2+7^2+6^2+9^2}{7.2} - 72 \\
&\approx 6.426 < 9.236 \approx \chi^2_{0.1,6-1}.
\end{aligned}
$$

Hence we should not reject H_0.

At the significance level 10%, we cannot conclude that the bag purchased does not agree with the distribution suggested by the journal.

5.5.1 Goodness-of-fit tests

(Example 5.22)
A traffic engineer wishes to study whether drivers have a preference for certain tollbooths at a bridge during non-rush hours. The number of automobiles passing through each tollbooth lane was counted during a randomly selected 15-minute interval. The sample information is as follows.

Tollbooth Lane	1	2	3	4	5	Total
Number of Cars observed	171	224	211	180	214	100

Can we conclude that there are differences in the numbers of cars selecting respectively each of the lanes? Test at the 5% significance level.

5.5.1 Goodness-of-fit tests

(Example 5.22)
H_0: there is no preference among the five lanes,
versus
H_1: there is a preference among the five lanes.
All the five expected frequencies equal $1000 \div 5 = 200$, which is not less than 5. Therefore, as the sample is large enough,

$$\begin{aligned} -2\ln\Lambda &\approx \sum_{i=1}^{5} \frac{O_i^2}{E_i} - n \\ &= \frac{171^2 + 224^2 + 211^2 + 180^2 + 214^2}{200} - 1000 \\ &\approx 10.67 \geq 9.488 \approx \chi^2_{0.05,5-1}. \end{aligned}$$

Hence, H_0 should be rejected.
At the significance level 5%, we can conclude that there are differences in the numbers of cars selecting respectively each of the lanes.

5.5.2 Pearson Chi-squared test of independence

(Example 5.23)
Suppose we draw a sample of 360 students and obtain the following information. At the 0.01 level of significance, test whether a student's ability in mathematics is independent of the student's interest in statistics.

		Ability in Math			sum
		Low	Average	High	
Interest in Statistics	Low	63	42	15	120
	Average	58	61	31	150
	High	14	47	29	90
Sum		135	150	75	360

5.5.2 Pearson Chi-squared test of independence

(Example 5.23)
H_0: ability in mathematics and interest in statistics are independent,
versus
H_1: ability in mathematics and interest in statistics are not independent.
The table below shows the expected frequencies

		Ability in Math			sum
		Low	Average	High	
Interest in Statistics	Low	45	50	25	120
	Average	56.25	62.5	31.25	150
	High	33.75	37.5	18.75	90
Sum		135	150	75	360

All expected frequencies are at least 5.

5.5.2 Pearson Chi-squared test of independence

(Example 5.23)
Therefore, as the sample is large enough,

$$
\begin{aligned}
& n\left(\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{O_{i,j}^2}{n_{i\cdot}n_{\cdot j}}-1\right) \\
= \quad & 360\left(\frac{63^2}{120\times 135}+\frac{42^2}{120\times 150}+\cdots+\frac{29^2}{90\times 75}-1\right) \\
\approx \quad & 32.140 \geq 13.277 \approx \chi^2_{0.01,(3-1)(3-1)}.
\end{aligned}
$$

Hence, at the significance level 1%, we reject H_0 and conclude

that there is a relationship between a student's ability in mathematics and the student's interest in statistics. in

The end of Chapter 5