MATH 5010 –Foundations of Statistical Theory and Probability

❖ Hypothesis Tests

-Methods of Find Tests

Instructor: He Wang
Department of Mathematics
Northeastern University

Introduction

We need to **test** whether a **hypothesis** is **true** or **false.**

For example, a pharmaceutical company might be interested in knowing if a new drug is effective in treating a disease.

Hypothesis 0: drug is not effective (null hypothesis)

Hypothesis 1: drug is effective (alternative hypothesis)

Hypothesis testing: Decide either to accept H_0 or to reject it, based on observed data.

In introduction to statistics class, we have learned several concrete tests:

- Test for one mean (Z-test, T-test)
- Test for proportion
- Test for variance (Chi square)
- Test for two proportions
- Test for two means
- Test for two variance (F-test)

Here, we will learn the theory behind these tests and more other tests.

Review and Example

You have a coin and you would like to check whether it is fair or not.

let θ be the probability of heads, $\theta = P(H)$. You have two hypotheses:

- H_0 (the **null hypothesis**): The coin is fair, i.e. $\theta = \theta_0 = \frac{1}{2}$.
- H_1 (the **alternative hypothesis**): The coin is not fair, i.e. $\theta \neq \frac{1}{2}$.

We toss the coin 100 times and record the outcomes. (**Data**)

Let $X_1, \dots, X_n$ be a random sample from **Bernoulli**(θ).

Let $X = X_1 + \cdots + X_n$ be the number of heads that we observe.

$$X \sim Binomial\ (100, \theta)$$

If H_0 is true, then $\theta = \theta_0 = \frac{1}{2}$. So, we expect the number of heads to be **close** to 50.

More specifically, we suggest a **threshold** t:

If $|X - 50| \leq t$, accept H_0.

If $|X - 50| > t$, accept H_1.

How do we choose the threshold t?

$$P(type\ I\ error) = P(|X - 50| > t|H_0).$$

We want to have a test

$P(type\ I\ error) \leq \alpha = 0.05$ (**level of significance**)

By CLT, when n is large:

$$Y = \frac{X - n\theta_0}{\sqrt{n\theta_0(1 - \theta_0)}} = \frac{X - 50}{5} \approx \sim Normal(0,1)$$

$$P(type\ I\ error) = P(|X-50| > t|H_0)$$

$$= P\left(\frac{|X-50|}{5} > \frac{t}{5}\middle|H_0\right)$$

$$= P\left(Y > \frac{t}{5}\middle|H_0\right).$$

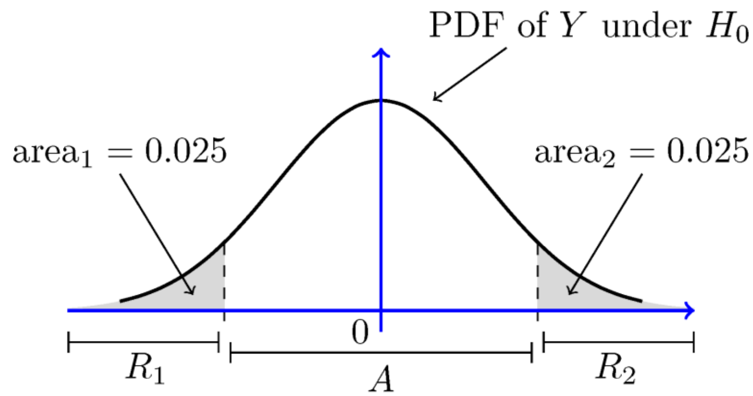
We want to have a test $P(type\ I\ error) \leq \alpha = 0.05$ (level of significance)

So, if $|Y| \leq 1.96$, accept H_0.

On the other hand, $if\ |Y| > 1.96$, accept H_1.

Equivalently,

If $X \in \{41, \dots, 59\}$, accept H_0 $(\text{Fail to reject } H_0)$.
For the rest, reject H_0.



A = Acceptance Region

$R = R_1 \cup R_2$ = Rejection Region

$\alpha = \text{P(type I error)} = \text{area}_1 + \text{area}_2 = 0.05$

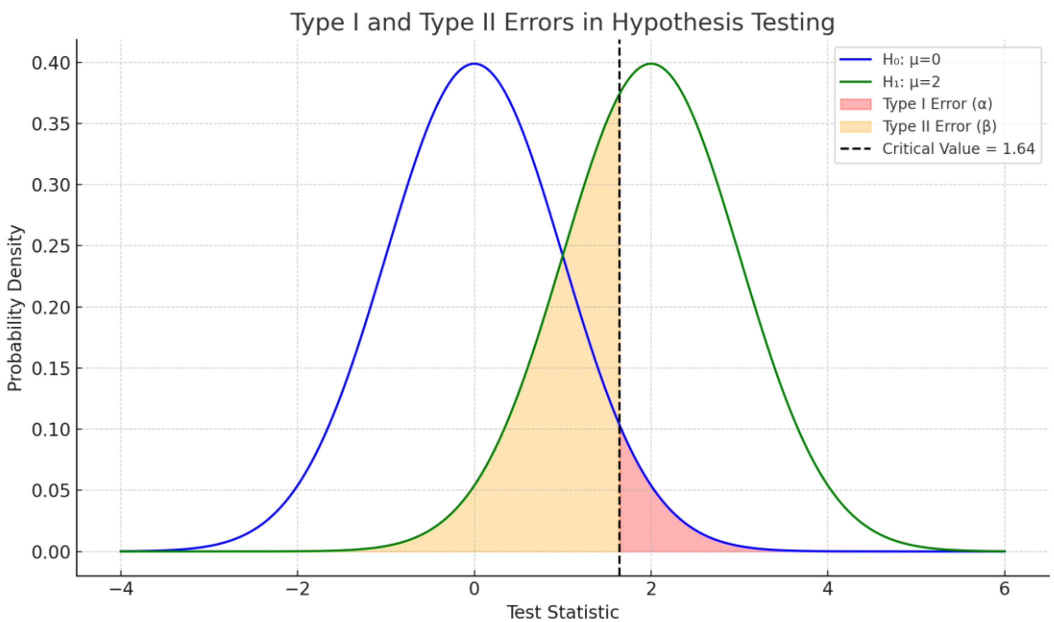
The second possible error that we can make is to accept H_0 when H_0 is false.

$$P(type\ II\ error) := P(Accept\ H_0 |\ \theta = \theta_1).$$

$$\theta_1 \neq \frac{1}{2}$$

Type I & II Errors

Decisions \ Facts	H_0 is True	H_1 is True
Reject H_0	**Type I Error**	Correct Decision
Fail to reject H_0	Correct Decision	**Type II Error**



Example: Radar Aircraft Detection Problem

A radar system uses radio waves to detect whether an aircraft is present.

The received signal is denoted by X.

$$X = \begin{cases} W & \text{if no aircraft is present} \\ 1 + W & \text{if an aircraft is present} \end{cases}$$

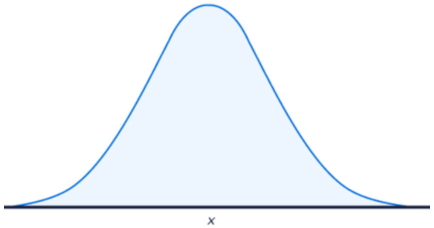
Here, $W \sim Normal(0, \sigma^2)$

So, $X = \theta + W$ with $\theta = \begin{cases} 0 & \text{if no aircraft is present} \\ 1 & \text{if an aircraft is present} \end{cases}$

1. **Hypotheses** (in terms of θ.)

- H_0 (null hypothesis): No aircraft is present. ($\theta = 0$)
- H_1 (alternative hypothesis): An aircraft is present. ($\theta = 1$)

2. Test Design: Construct a test with significance level $\alpha = 0.05$ to decide between H_0 and H_1. (Reject H_0, when $X > c$.)



$$P(type\ I\ error) = P(Reject\ H_0|H_0)$$

$$= P(X > c|\theta = 0)$$

$$= P(W > c) = 0.5$$

$$Z = 3W$$

$$\Longrightarrow c \approx \frac{1.645}{3}$$

3. Type II Error:

$$\beta = P(type\ II\ error) = P(Accept\ H_0|H_1)$$

$$= P(X < c\ |\theta = 1)$$

$$= P(1 + W < c)$$

$$\approx 0.0877$$

4. **Evidence Check**: If an observation $X = 0.6$ is obtained, determine whether there is sufficient evidence to reject H_0 at the significance level $\alpha = 0.01$.

Redo the part 2 with $\alpha = 0.01$ and get $c = 0.775$

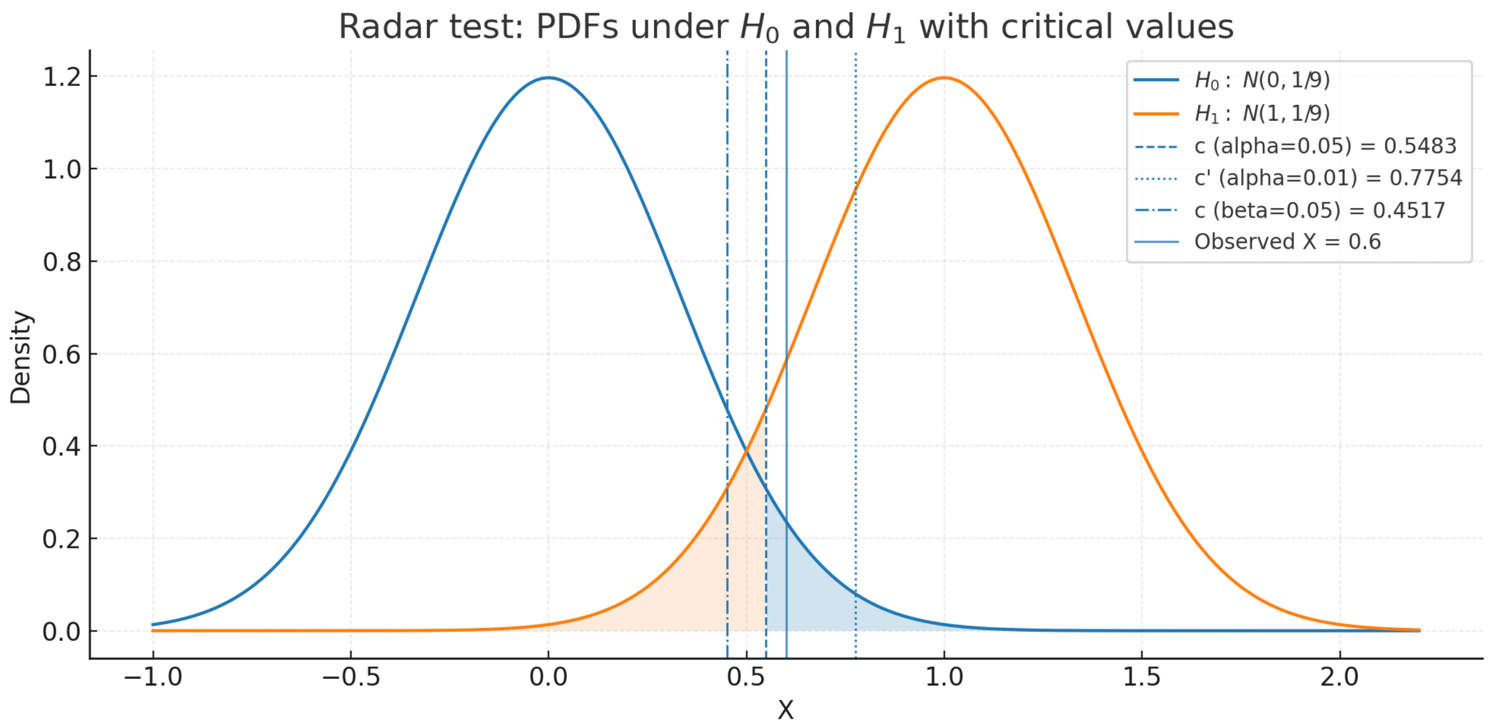
$X = 0.6 < c$. So, we do not reject H_0 at the 0.01 level.

5. Power Constraint: If we want the probability of missing a present aircraft to be less than 5% ($\beta < 0.05$), find the **smallest significance level** α that can be achieved.

$$0.05 = \beta = P(type\ II\ error) = P(1 + W < c) \qquad \Longrightarrow c \approx 0.45167$$

$$\alpha = P(type\ I\ error) = P(W > c) = P(Z > 3c) \approx 0.0877$$

Trade-off Between α *and* β



Shaded areas show the rejection region under H_0 (right tail) and the corresponding type II region under H_1.

P-value

P-value is the lowest significance level α that results in rejecting the null hypothesis.

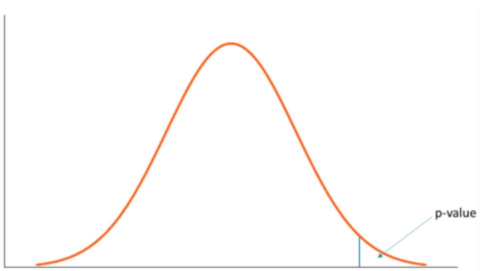
Formally,

P (a test statistic more **extreme** than what we observed | H_0 is true)

$P(W(X_1, \ldots, X_n) \geq W(x_1, \ldots, x_n) \mid H_0 \text{ is true})$ Right-side

$P(W(X_1, \ldots, X_n) \leq W(x_1, \ldots, x_n) \mid H_0 \text{ is true})$ Left-Side

P-value indicates how **close** the decision was in accept or reject decision.



Continue the example:

6. p-value. The observed value is $X_o = 0.6$, compute the **p-value** for this test under H_0.

Compute the test statistic (standardized):

$$Z = \frac{X}{\sigma} = 3X$$

For the observed value is $X_o = 0.6$, we have $Z_o = 3(0.6) = 1.8$

$$\text{p} - \text{value} = P(X \geq 0.6 | H_0) = P(Z \geq 1.8) = 1 - \phi(1.8) \approx 0.0359$$

Decision,

At $\alpha = 0.05$: $0.0359 < 0.05$, so **reject $\boldsymbol{H_0}$**
At $\alpha = 0.01$: $0.0359 > 0.01$, so **do not reject $\boldsymbol{H_0}$**.

Statistical Hypothesis (Mathematical Theory)

A **hypothesis** is a statement about a **population parameter** $\boldsymbol{\theta}$.

- **Null Hypothesis** $(\boldsymbol{H_0})$: $\theta \in \Theta_0$

Represents the **default**, **no effect**, or **no difference**.

Example: H_0: $\mu = 100$, or H_0: $\mu \geq 100$.

- **Alternative Hypothesis** $(\boldsymbol{H}_1)$: $\theta \in \Theta_0^c$

Represents the **new claim**, **effect**, or **difference**.

Example: H_1: $\mu \neq 100$, or H_0: $\mu < 100$.

Goal of a hypothesis test: Based on a **sample** $(x_1, \ldots, x_n)$ from the population, **decide** which of two complementary hypotheses is true.

A **Hypothesis Test** is a rule that specifies:

- For which sample values the decision is made to accept H_0 as true.
- For which sample values H_0 is rejected and H_1 is accepted as true.

Terminology: Some people will not use accept H_0, but only claim: there's not enough evidence to reject H_0. (Fail to reject H_0)

A hypothesis test is specified in terms of a **test statistic** $\boldsymbol{W(\vec{X}) = W(X_1, \dots, X_n)}$

Example: H_0 is rejected if sample mean $\bar{X} > 100$.

$$\text{Reject region} = \{(x_1, \dots, x_n) \mid \bar{x} > 100\}$$

Example: Hypothesis Testing for the Mean

Suppose H_0 is $\boldsymbol{true}$ $(i.e., \mu = \mu_0)$.

By CLT, when n is large,

$$\bar{X} \sim Normal(\mu_0, \sigma^2/n)$$

Equivalently,

$$W(X_1, \ldots, X_n) \coloneqq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad \sim Normal(0,1)$$

Given data $x_1, \ldots, x_n$, we can calculate

$$P(\bar{x} > 100 | H_0)$$

Methods of Finding Tests

1. **Likelihood Ratio Test (LRT)**

Compare the **maximum likelihood** under the null hypothesis H_0 with the **maximum likelihood** over the entire parameter space (null + alternative).

2. **Bayes Tests**

Use **Bayesian decision theory** to minimize the **expected loss** (risk) under a prior distribution over parameters.

3. **Union-Intersection Test (UIT)**

Test H_0 against a **composite alternative** by testing all **simple hypotheses** in the alternative space and **rejecting if any one of them leads to rejection**.

4. Intersection-Union Test (IUT)

5. Neyman–Pearson Lemma *(for simple hypotheses)*

❖ 1. Likelihood Ratio Test (LRT)

$\{X_1, \dots, X_n\}$ is a sample from a population distribution with pdf $f(x|\vec{\theta})$

Recall from Maximum Likelihood Estimation, given an observed sample $x_1, \dots, x_n$, **the likelihood function** is

$$L(\vec{\theta}|\vec{x}) := L(\theta_1, \dots, \theta_k | x_1, \dots, x_n) = f(x_1, \dots, x_n | \theta_1, \dots, \theta_k) = \prod_{i=1}^{n} f(x_i|\vec{\theta})$$

The **likelihood ratio test statistic** for testing $(\boldsymbol{H_0}): \theta \in \Theta_0 \; v.s. \; (\boldsymbol{H}_1): \; \theta \in \Theta_0^c$ is

$$\lambda(\vec{x}) := \frac{\sup_{\Theta_0} L(\vec{\theta}|\vec{x})}{\sup_{\Theta} L(\vec{\theta}|\vec{x})} = \frac{L(\vec{\theta}_0|\vec{x})}{L\left(\hat{\vec{\theta}}_{MLE}\middle|\vec{x}\right)}$$

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

$$\{\vec{x}|\lambda(\vec{x}) \le c\}$$

where c is any number satisfying $0 \le c \le 1$.

Likelihood Ratio Test for Simple Hypotheses

Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with a parameter θ.

Observed $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$.

Decide between two simple hypotheses

$$H_0: \theta = \theta_0$$
$$H_1: \theta = \theta_1$$

Define the **statistics**

$$\lambda(\vec{x}) := \frac{L(\theta_0|\vec{x})}{L\,(\theta_1|\vec{x})}$$

Likelihood ratio test: Choose a constant c.

We reject H_0 if $\lambda < c$ and accept it if $\lambda \geq c$.

Radar Example:

Random Variable $X = \theta + W$ with $W \sim Normal\left(0, \sigma^2 = \frac{1}{9}\right)$ and

$$\theta = \begin{cases} 0 & \text{if no aircraft is present} \\ 1 & \text{if an aircraft is present} \end{cases}$$

Let $X = x$. Design a level 0.05 test $(\alpha = 0.05)$ to decide between $H_0\ and\ H_1$

$$H_0: \theta = \theta_0 = 0$$
$$H_1: \theta = \theta_1 = 1$$

$$L(\theta_0|\vec{x}) = \frac{3}{\sqrt{2\pi}} e^{-\frac{9x^2}{2}}$$

$$L(\theta_1|\vec{x}) = \frac{3}{\sqrt{2\pi}} e^{-\frac{9(x-1)^2}{2}}$$

Likelihood ratio test statistics

$$\lambda(\vec{x}) := \frac{L(\theta_0|\vec{x})}{L\,(\theta_1|\vec{x})} = \exp\left(-\frac{9x^2}{2} + \frac{9(x-1)^2}{2}\right)$$

$$= \exp\left(\frac{9(1-2x)}{2}\right)$$

Set threshold c. We reject H_0 if

$$\exp\left(\frac{9(1-2x)}{2}\right) < c$$

Equivalently, We reject H_0 if

$$x > \frac{1}{2}\left(1 - \frac{2}{9}\ln c\right) := c'$$

The choice of c' (based on $\alpha = 0.05$)

$$\alpha = P(type\ I\ error) = P(Reject\ H_0|H_0)$$

$$= P(X > c'|\theta = \theta_0)$$

$$= P(X > c') \qquad X \sim Normal\left(0, \sigma^2 = \frac{1}{9}\right)$$

$$= 1 - \phi(3c')$$

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Solve $c' = \phi^{-1}(1 - \alpha) \approx 0.548$

In this case, the likelihood ratio test is exactly the same test that we obtained

Likelihood Ratio Test for normal means with known variance

$\{X_1, \dots, X_n\}$ is a sample from $Normal(\mu, \sigma^2)$ with known σ^2.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Test: $H_0: \mu = \mu_0 \;\; v.s. \;\; H_1: \mu \neq \mu_1$

$$\lambda(\vec{x}) \coloneqq \frac{\sup_{\Theta_0} L(\vec{\theta}|\vec{x})}{\sup_{\Theta} L(\vec{\theta}|\vec{x})} = \frac{L(\vec{\theta}_0|\vec{x})}{L\left(\hat{\vec{\theta}}_{MLE}\middle|\vec{x}\right)}$$

Step 1. Likelihood Function

$$L(\mu|\vec{x}) = \prod_{i=1}^{n} f(x_i) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right)$$

Step 2 – Likelihood Ratio Test statistics

Maximum Likelihood Estimates:

- Under $\boldsymbol{H_0}$, μ is fixed at μ_0, so the MLE is simply μ_0.
- Under **no restriction** (H_1), the MLE of μ is: $\hat{\mu}_{MLE} = \bar{X}$

Likelihood Ratio Test statistics

$$\lambda(\vec{x}) = \frac{L(\mu_0|\vec{x})}{L(\bar{X}|\vec{x})}$$

where

$$L(\mu_0|\vec{x}) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0)^2\right)$$

$$L(\bar{X}|\vec{x}) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)$$

So, $$\lambda(\vec{x}) = \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i - \mu_0)^2 - (x_i - \bar{x})^2\right)\right)$$

$$= \exp\left(-\frac{n(\bar{x} - \mu_0)^2}{2\sigma^2}\right)$$

Since $\lambda(\vec{x})$ is decreasing of $|\bar{x} - \mu_0|$, the rejection region $\{\vec{x}|\lambda(\vec{x}) \leq c\}$

$$\{\vec{x}||\bar{x} - \mu_0| \geq \sqrt{-2\sigma^2(\log c)/n}\}$$

Equivalent, we can define **statistics**

$$Z = \frac{(\bar{X} - \mu_0)}{\sigma/\sqrt{n}} \sim Normal(0,1)$$

Reject H_0 if $|Z| > z_{\alpha/2}$ $\qquad c = \exp(-z_{\alpha/2}^2/2)$

The LRT for normal means with known σ^2 reduces to the **classical two-sided Z-test**.

Likelihood Ratio Test for normal means with unknown variance

$\{X_1, \dots, X_n\}$ is a sample from $Normal(\mu, \sigma^2)$ with unknown σ^2.

Likelihood Ratio Test: $H_0: \mu = \mu_0 \;\; v.s. \;\; H_1: \mu \neq \mu_1$

Maximum Likelihood Estimates (MLE):

- Under $\boldsymbol{H_0}$, $\widehat{\mu_0} = \mu_0$ and $\hat{\sigma}_0 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2$

- Under **no restriction** (H_1), $\hat{\mu} = \bar{X}$

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

Likelihood Ratio Test statistics

$$L(\mu_0|\vec{x}) = (2\pi\hat{\sigma}_0^2)^{-n/2} \exp\left(-\frac{1}{2\hat{\sigma}_0^2}\sum_{i=1}^{n}(x_i - \mu_0)^2\right)$$

$$L(\bar{X}|\vec{x}) = (2\pi\hat{\sigma}^2)^{-n/2} \exp\left(-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)$$

$$\lambda(\vec{x}) = \frac{L(\mu_0|\vec{x})}{L(\bar{X}|\vec{x})} = \frac{(2\pi\hat{\sigma}_0^2)^{-n/2}}{(2\pi\hat{\sigma}^2)^{-n/2}} = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)^{-n/2} = \cdots$$

$$= 1 + \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Connection to the t-statistic

Recall the **t-statistic**:

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}, \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Rewrite:

$$\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} = 1 + \frac{t^2}{n-1}$$

The LRT rejects H_0 for **small** $\lambda(\vec{x}) \Leftrightarrow$ **large** | t |.

$$\text{Reject } H_0 \quad \text{if} \quad |t| > t_{n-1,1-\alpha/2}$$

The LRT for normal means with unknown σ^2 reduces to the **classical two-sided Student's t-test with $n-1$** degrees of freedom.

Remark:

The classical Z, t, chi-square, and proportion tests can all be seen as **special cases or asymptotic results of the likelihood ratio test**. But historically, they were discovered first for practical problems.

The two sample tests, pooled t test, two-proportion, F test are also LRT.

Sufficient statistic

Suppose $T(X)$ is a sufficient statistic for a parameter θ.

Theorem: The LRT statistic $\lambda^*(t)$ based on the sufficient statistic T equals the LRT statistic $\lambda\,(x)$ based on the original data X.

$$\lambda^*\left(T(\vec{x})\right) = \lambda(\vec{x})$$

Example: LRT for normal means with known variance

$T(X) = \bar{X}$ is a sufficient statistics for μ

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \qquad f(t) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left(-\frac{n(t-\mu)^2}{2\sigma^2}\right)$$

Test: $H_0\colon \mu = \mu_0 \ \ v.s.\ \ H_1\colon \mu \neq \mu_1$

$$\lambda(t) \coloneqq \frac{\sup_{\Theta_0} L(\vec{\theta}|t)}{\sup_{\Theta} L(\vec{\theta}|t)} = \frac{L(\vec{\theta}_0|t)}{L\left(\hat{\vec{\theta}}_{MLE}\middle|t\right)}$$

Likelihood Function

$$L(\mu|t) = f(t) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left(-\frac{n(t-\mu)^2}{2\sigma^2}\right)$$

Step 2 – Likelihood Ratio Test statistics

Maximum Likelihood Estimates:

- Under $\boldsymbol{H_0}$, μ is fixed at μ_0, so the MLE is simply μ_0.
- Under **no restriction** (H_1), the MLE of μ is: $\hat{\mu}_{MLE} = \bar{X} = \mathrm{t}$

Likelihood Ratio Test statistics

$$\lambda(\vec{x}) = \frac{L(\mu_0|\vec{x})}{L(\bar{X}|\vec{x})}$$

where

$$L(\mu_0|t) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left(-\frac{n(t-\mu_0)^2}{2\sigma^2}\right)$$

$$L(\bar{X}|t) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left(-\frac{n(t-t)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi\sigma^2/n}}$$

So,

$$\lambda(t) = \exp\left(-\frac{n(t-\mu_0)^2}{2\sigma^2}\right)$$

This verifies

$$\lambda^*\big(T(\vec{x})\big) = \exp\left(-\frac{n(\bar{x}-\mu_0)^2}{2\sigma^2}\right) = \lambda(\vec{x})$$

❖ 2. Bayesian Tests

Random Sample $X_1, \ldots, X_n$ from a population distribution with pdf $f(x|\theta)$

In the **Bayesian** approach, random variable θ has **prior distribution** $\pi(\theta)$.

The **posterior distribution** $\pi(\theta|\vec{x})$ based on the sample $\vec{x} = (x_1, \ldots, x_n)$:

By Bayes' Rule:

$$\pi(\theta|\vec{x}) = \frac{f(\vec{x}|\theta)\pi(\theta)}{m(\vec{x})}$$

$m(\vec{x})$ is the marginal distribution of $\vec{x}$:

All inference about θ is based on the posterior distribution.

Test

- Null Hypothesis (H_0): $\theta \in \Theta_0$
- Alternative Hypothesis (H_1): $\theta \in \Theta_0^c$

Using posterior distribution $\pi(\theta|\vec{x})$, we have Posterior Probabilities:

$$P(\theta \in \Theta_0|\vec{x}) = P(H_0\ is\ true\ |\vec{x})$$

$$P(\theta \in \Theta_0^c|\vec{x}) = P(H_1\ is\ true\ |\vec{x})$$

So, we can decide to reject H_0 if $P(\theta \in \Theta_0|\vec{x}) < P(\theta \in \Theta_0^c|\vec{x})$

Rejection region is $\left\{\vec{x}\ \middle| P(\theta \in \Theta_0|\vec{x}) < \frac{1}{2}\right\}$

If we want to avoid falsely rejecting H_0, we can choose

Rejection region is $\{\vec{x}\ |P(\theta \in \Theta_0|\vec{x}) < 0.05\}$

Bayesian Test (Normal mean)

Suppose we have iid data $\mathcal{D} = \{x^{(1)}, \dots, x^{(n)}\}$ observed from normal distribution $N(\mu, \sigma^2)$ with known σ.

$$p\left(x^{(i)}|\mu\right) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}\left(x^{(i)} - \mu\right)^2\right)$$

The **prior** distribution of μ is $Normal(\theta, \tau^2)$

$$p(\mu) = \frac{1}{\sqrt{2\pi}\tau} \exp\left(-\frac{1}{2\tau^2}(\mu - \theta)^2\right)$$

The **posterior** of $\mu|\vec{x}$ is normal with mean and variance:

$$E(\mu|\vec{x}) = \frac{\tau^2\left(\sum_{i=1}^{N} x^{(i)}\right) + \sigma^2\theta}{n\tau^2 + \sigma^2} \qquad Var(\mu|\vec{x}) = \frac{\sigma^2\tau^2}{\sigma^2 + n\tau^2}$$

Test

- Null Hypothesis (H_0): $\mu \leq \mu_0$
- Alternative Hypothesis (H_1): $\mu > \mu_0$

Suppose we decide to reject H_0 when $P(H_0\ true|\vec{x}) < P(H_1\ true\ |\vec{x})$

$$P(\mu \leq \mu_0|\vec{x}) < P(\mu > \mu_0|\vec{x})$$

$$P(\mu \leq \mu_0|\vec{x}) < \frac{1}{2}$$

Since $\mu|\vec{x}$ is normal, by symmetry, we reject H_0 if

$$E(\mu|\vec{x}) = \frac{\tau^2\left(\sum_{i=1}^{N} x^{(i)}\right) + \sigma^2\theta}{n\tau^2 + \sigma^2} > \mu_0$$

That is

$$\bar{x} \leq \mu_0 + \frac{\sigma^2(\mu_0 - \theta)}{n\tau^2}$$

3. **Union-Intersection Test (UIT)**

Instead of constructing one big test for the whole alternative, we:

Express H_0 as an intersection:

$$H_0: \theta \in \Theta_0 = \bigcap_{\gamma \in \Gamma} \Theta_\gamma$$

By De Morgan's law:

$$H_1: \theta \in \Theta_0^c = \bigcup_{\gamma \in \Gamma} \Theta_\gamma^c$$

For each sub-alternative, test $H_{0\gamma}: \theta \in \Theta_\gamma \ v.s.\ H_{1\gamma}: \theta \in \Theta_\gamma^c$.

Obtain **reject** region

$$R_\gamma = \{\vec{x}: T_\gamma(\vec{x}) > c\}$$

where $T_\gamma(\vec{x})$ is the test statistic.

Reject the null if any of the sub-tests rejects:

$$R = \bigcup_{\gamma\in\Gamma} R_\gamma$$

$$= \bigcup_{\gamma\in\Gamma} \{\vec{x}: T_\gamma(\vec{x}) > c\}$$

$$= \{\vec{x}: \sup_{\gamma} T_\gamma(\vec{x}) > c\}$$

The test statistics for H_0 v.s. H_1 is

$$T(\vec{x}) = \sup_{\gamma} T_\gamma(\vec{x})$$

Example: (UIT for Normal)

Suppose we have iid data $\mathcal{D} = \{x^{(1)}, \ldots, x^{(n)}\}$ observed from normal distribution $N(\mu, \sigma^2)$ with known σ.

Test

- Null Hypothesis (H_0): $\mu = \mu_0$
- Alternative Hypothesis (H_1): $\mu \neq \mu_0$

H_0 can be written as intersection:

$$H_0: \{\mu | \mu \leq \mu_0\} \cap \{\mu | \mu \geq \mu_0\}$$

The LRT of $H_{0L}: \mu \leq \mu_0$ $v.s$ $H_{1L}: \mu > \mu_0$

Reject H_{0L} if $\dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq z_L$

The LRT of $H_{0R}: \mu \geq \mu_0$ $v.s$ $H_{1R}: \mu < \mu_0$

Reject H_{0R} if $\dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq z_R$

The Union-Intersection test of $(H_0): \mu = \mu_0 \; v.s. \; (H_1): \mu \neq \mu_0$

Reject H_0 if $\dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq z_L$ or $\dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq z_R$

Equivalently, if $z_L = -z_R \geq 0$

Reject H_0 if $\dfrac{|\bar{X} - \mu_0|}{S/\sqrt{n}} \geq z_L$

This is the two-sided $z - test$.

Similarly, we have the the two-sided t $- \, test$ for unknown σ^2.

4. Intersection-Union Test (IUT)

Express H_0 as an intersection:

$$H_0 : \theta \in \Theta_0 = \bigcup_{\gamma \in \Gamma} \Theta_\gamma$$

By De Morgan's law:

$$H_1 : \theta \in \Theta_0^c = \bigcap_{\gamma \in \Gamma} \Theta_\gamma^c$$

For each sub-alternative, test $H_{0\gamma} : \theta \in \Theta_\gamma \; v.s. H_{1\gamma} : \theta \in \Theta_\gamma^c$.

Obtain **reject** region

$$R_\gamma = \{\vec{x} : T_\gamma(\vec{x}) > c\}$$

where $T_\gamma(\vec{x})$ is the test statistic.

Reject the null if any of the sub-tests rejects:

$$R = \bigcap_{\gamma \in \Gamma} R_\gamma$$

$$= \bigcap_{\gamma \in \Gamma} \{\vec{x}: T_\gamma(\vec{x}) > c\}$$

$$= \{\vec{x}: \inf_\gamma T_\gamma(\vec{x}) > c\}$$

The test statistics for H_0 v.s. H_1 is

$$T(\vec{x}) = \inf_\gamma T_\gamma(\vec{x})$$

Example: (acceptance sampling)

We have two **quality parameters** for upholstery fabric:

- θ_1: mean breaking strength (should be > 50 pounds)
- θ_2: probability of passing a flammability test (should be > 0.95)

We **accept** a batch **only if** both requirements are met.

Hypotheses:

$$H_0: \{\theta_1 \leq 50 \; or \; \theta_2 \leq 0.95\}$$ (at least one standard fails)

$$H_1: \{\theta_1 > 50 \; and \; \theta_2 > 0.95\}$$ (both pass standards)

H_0 can be written as intersection:

$$H_0: \{\theta_1 \leq 50\} \cup \{\theta_2 \leq 0.95\}$$

$$H_1: \{\theta_1 > 50\} \cap \{\theta_2 > 0.95\}$$

Data Collection

$X_1, \dots, X_n$: breaking strength measurements ($Normal(\theta_1, \sigma^2)$)

$Y_1, \dots, Y_m$: flammability pass/fail indicators (Bernoulli(θ_2)) $\quad Y_i = \begin{cases} 1 \text{ if pass} \\ 0 \text{ if fail.} \end{cases}$

Testing Each Component using Likelihood Ratio Tests (LRT):

Strength Test ($H_{01}: \theta_1 \leq 50$):

$$\text{Reject if } \frac{\bar{X} - 50}{S/\sqrt{n}} > t$$

Flammability Test ($H_{02}: \theta_2 \leq 0.95$)

$$\text{Reject if } \sum_{i=1}^{m} Y_i > b$$

Intersection–Union Test Rule

Reject H_0 iff both component nulls are rejected.

$$\frac{\bar{x} - 50}{s/\sqrt{n}} > t \quad \text{AND} \quad \sum_{i=1}^{m} y_i > b$$

Only if **both** pass do we declare the batch acceptable.

If more than two parameters define a product's quality, individual tests for each parameter can be combined to yield an overall test of the product's quality.

References:

- **Book 1. [CB] Statistical Inference**, by Casella, George, Berger, Roger L, 2nd edition
- **Book 2. [W]: All of Statistics: Larry Wasserman**
- https://www.probabilitycourse.com/

Online books and courses:

- https://online.stat.psu.edu/stat415/
- https://stat110.hsites.harvard.edu/
- https://bookdown.org/egarpor/inference/