Proability and Statistics

MA20205

Bibhas Adhikari

Autumn 2022-23, IIT Kharagpur

Lecture 19

November 15, 2022

Hypothesis Testing

Hypothesis is a statement that can be either true or false. A statistical hypothesis is a hypothesis about parameter(s) of some population or process

Hypothesis Testing

Hypothesis is a statement that can be either true or false. A statistical hypothesis is a hypothesis about parameter(s) of some population or process

Suppose a COVID19 vaccine is given to 1000 people and 982 of them have developed antigens. Then with how much confident that we are to say that the vaccine is effective?

Hypothesis Testing

Hypothesis is a statement that can be either true or false. A statistical hypothesis is a hypothesis about parameter(s) of some population or process

Suppose a COVID19 vaccine is given to 1000 people and 982 of them have developed antigens. Then with how much confident that we are to say that the vaccine is effective?

Hypothesis testing - is a statistical procedure to evaluate statements that should be accepted or rejected.

Hypothesis Testing

Hypothesis is a statement that can be either true or false. A statistical hypothesis is a hypothesis about parameter(s) of some population or process

Suppose a COVID19 vaccine is given to 1000 people and 982 of them have developed antigens. Then with how much confident that we are to say that the vaccine is effective?

Hypothesis testing - is a statistical procedure to evaluate statements that should be accepted or rejected.A hypothesis is a statement which requires testing by observation to decide whether it is true or false. For example, such statements could be

- The coin is unbiased

Hypothesis Testing

Hypothesis is a statement that can be either true or false. A statistical hypothesis is a hypothesis about parameter(s) of some population or process

Suppose a COVID19 vaccine is given to 1000 people and 982 of them have developed antigens. Then with how much confident that we are to say that the vaccine is effective?

Hypothesis testing - is a statistical procedure to evaluate statements that should be accepted or rejected.A hypothesis is a statement which requires testing by observation to decide whether it is true or false. For example, such statements could be

- The coin is unbiased
- Students whose registration are accepted for the *Quantum Computing* course have CGPA ≥ 8

Hypothesis Testing

Hypothesis is a statement that can be either true or false. A statistical hypothesis is a hypothesis about parameter(s) of some population or process

Suppose a COVID19 vaccine is given to 1000 people and 982 of them have developed antigens. Then with how much confident that we are to say that the vaccine is effective?

Hypothesis testing - is a statistical procedure to evaluate statements that should be accepted or rejected.A hypothesis is a statement which requires testing by observation to decide whether it is true or false. For example, such statements could be

- The coin is unbiased
- Students whose registration are accepted for the *Quantum Computing* course have CGPA ≥ 8
- An algorithm performs better than another algorithm

Hypothesis Testing

Since a statement will be tested based on a given data, the conclusion (*accepting* / *rejecting*) after testing depends on the statistics of the data and a cutoff threshold, however the truth is unknown.

Hypothesis Testing

Since a statement will be tested based on a given data, the conclusion (*accepting*/*rejecting*) after testing depends on the statistics of the data and a cutoff threshold, however the truth is unknown.

Thus accepting the hypothesis is based on a statistical evidence and it is possible that if we change the data, the conclusion will be different.

Hypothesis Testing

Since a statement will be tested based on a given data, the conclusion (*accepting*/*rejecting*) after testing depends on the statistics of the data and a cutoff threshold, however the truth is unknown.

Thus accepting the hypothesis is based on a statistical evidence and it is possible that if we change the data, the conclusion will be different.

The framework of hypothesis testing usually have two opposite hypotheses:

- H_0 : Null hypothesis (which is the "status quo" i.e. current status)

Hypothesis Testing

Since a statement will be tested based on a given data, the conclusion (*accepting*/*rejecting*) after testing depends on the statistics of the data and a cutoff threshold, however the truth is unknown.

Thus accepting the hypothesis is based on a statistical evidence and it is possible that if we change the data, the conclusion will be different.

The framework of hypothesis testing usually have two opposite hypotheses:

- H_0 : Null hypothesis (which is the "status quo" i.e. current status)
- H_1 : Alternative hypothesis (which is alternative to the null hypothesis)

Hypothesis Testing

Analogue of Hypothesis Testing in courthouse Suppose a person being prosecuted is assumed innocent. This is the null hypothesis. Then the police need to produce sufficient evidence to prove the person guilty. Thus the Hypothesis Testing investigates whether there are enough evidence to reject the null hypothesis.

Hypothesis Testing

Analogue of Hypothesis Testing in courthouse Suppose a person being prosecuted is assumed innocent. This is the null hypothesis. Then the police need to produce sufficient evidence to prove the person guilty. Thus the Hypothesis Testing investigates whether there are enough evidence to reject the null hypothesis.

Another example Null and Alternative hypotheses for the statement regarding whether a coin is unbiased

Hypothesis Testing

Analogue of Hypothesis Testing in courthouse Suppose a person being prosecuted is assumed innocent. This is the null hypothesis. Then the police need to produce sufficient evidence to prove the person guilty. Thus the Hypothesis Testing investigates whether there are enough evidence to reject the null hypothesis.

Another example Null and Alternative hypotheses for the statement regarding whether a coin is unbiasedLet θ be the probability of getting head. Then

- $H_0 : \theta = 0.5$ and $H_1 : \theta > 0.5$

Hypothesis Testing

Analogue of Hypothesis Testing in courthouse Suppose a person being prosecuted is assumed innocent. This is the null hypothesis. Then the police need to produce sufficient evidence to prove the person guilty. Thus the Hypothesis Testing investigates whether there are enough evidence to reject the null hypothesis.

Another example Null and Alternative hypotheses for the statement regarding whether a coin is unbiasedLet θ be the probability of getting head. Then

- $H_0 : \theta = 0.5$ and $H_1 : \theta > 0.5$
- $H_0 : \theta = 0.5$ and $H_1 : \theta < 0.5$

Hypothesis Testing

Analogue of Hypothesis Testing in courthouse Suppose a person being prosecuted is assumed innocent. This is the null hypothesis. Then the police need to produce sufficient evidence to prove the person guilty. Thus the Hypothesis Testing investigates whether there are enough evidence to reject the null hypothesis.

Another example Null and Alternative hypotheses for the statement regarding whether a coin is unbiasedLet θ be the probability of getting head. Then

- $H_0 : \theta = 0.5$ and $H_1 : \theta > 0.5$
- $H_0 : \theta = 0.5$ and $H_1 : \theta < 0.5$
- $H_0 : \theta = 0.5$ and $H_1 : \theta \neq 0.5$

Hypothesis Testing

Analogue of Hypothesis Testing in courthouse Suppose a person being prosecuted is assumed innocent. This is the null hypothesis. Then the police need to produce sufficient evidence to prove the person guilty. Thus the Hypothesis Testing investigates whether there are enough evidence to reject the null hypothesis.

Another example Null and Alternative hypotheses for the statement regarding whether a coin is unbiasedLet θ be the probability of getting head. Then

- $H_0 : \theta = 0.5$ and $H_1 : \theta > 0.5$
- $H_0 : \theta = 0.5$ and $H_1 : \theta < 0.5$
- $H_0 : \theta = 0.5$ and $H_1 : \theta \neq 0.5$

Two important approaches for Hypothesis Testing:

- critical-value test
- p-value test

Hypothesis Testing

Critical-value test Consider a toy example. Suppose we have a four-sided die and we want to test whether the die is unbiased. Define

- $H_0 : \theta = 0.25$

Hypothesis Testing

Critical-value test Consider a toy example. Suppose we have a four-sided die and we want to test whether the die is unbiased. Define

- $H_0 : \theta = 0.25$
- $H_1 : \theta > 0.25$

where θ is the unbiasedness measure

Hypothesis Testing

Critical-value test Consider a toy example. Suppose we have a four-sided die and we want to test whether the die is unbiased. Define

- $H_0 : \theta = 0.25$
- $H_1 : \theta > 0.25$

where θ is the unbiasedness measure
Suppose

- the die is drawn $n = 1000$ times
- 3 appears 290 times

Hypothesis Testing

Critical-value test Consider a toy example. Suppose we have a four-sided die and we want to test whether the die is unbiased. Define

- $H_0 : \theta = 0.25$
- $H_1 : \theta > 0.25$

where θ is the unbiasedness measure
Suppose

- the die is drawn $n = 1000$ times
- 3 appears 290 times

Let $X_1, X_2, \ldots, X_n$ denote the n copies of the Bernoulli random variable with success means 3 occurs and false means 3 does not occur.

Hypothesis Testing

Critical-value test Consider a toy example. Suppose we have a four-sided die and we want to test whether the die is unbiased. Define

- $H_0 : \theta = 0.25$
- $H_1 : \theta > 0.25$

where θ is the unbiasedness measure
Suppose

- the die is drawn $n = 1000$ times
- 3 appears 290 times

Let $X_1, X_2, \ldots, X_n$ denote the n copies of the Bernoulli random variable with success means 3 occurs and false means 3 does not occur. If the true probability is $\theta = 0.25$ then we should have

$$P(X_i = 3) = \theta = 0.25 \text{ and } P(X_i \neq 3) = 1 - \theta = 0.75$$

Hypothesis Testing

Since we do not have the knowledge of true probability, let us consider an estimator:

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Hypothesis Testing

Since we do not have the knowledge of true probability, let us consider an estimator:

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Then from the experiment,

$$\widehat{\theta} = \frac{290}{1000} = 0.29$$

Hypothesis Testing

Since we do not have the knowledge of true probability, let us consider an estimator:

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Then from the experiment,

$$\widehat{\theta} = \frac{290}{1000} = 0.29$$

Further, $Var(X_i) = \theta(1-\theta) = 0.25(1-0.25) = 0.1875.$

Hypothesis Testing

Since we do not have the knowledge of true probability, let us consider an estimator:

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Then from the experiment,

$$\widehat{\theta} = \frac{290}{1000} = 0.29$$

Further, $Var(X_i) = \theta(1-\theta) = 0.25(1-0.25) = 0.1875.$

Then the question for Hypothesis testing is: How far is $\widehat{\theta} = 0.29$ from $\theta = 0.25$? Note that as per the data H_0 should be rejected!

Hypothesis Testing

However, we should do a theoretical analysis over the sample size before rejecting H_0. If n is large, from central limit theorem,

$$\widehat{\theta} \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{n}\right)$$

Recall that $\widehat{\theta}$ is an unbiased estimator!!

Hypothesis Testing

However, we should do a theoretical analysis over the sample size before rejecting H_0. If n is large, from central limit theorem,

$$\widehat{\theta} \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{n}\right)$$

Recall that $\widehat{\theta}$ is an unbiased estimator!!

Further,

$$\widehat{Z} = \frac{\widehat{\theta} - \theta}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Hypothesis Testing

However, we should do a theoretical analysis over the sample size before rejecting H_0. If n is large, from central limit theorem,

$$\widehat{\theta} \sim \mathcal{N}\left(\theta, \frac{\sigma^2}{n}\right)$$

Recall that $\widehat{\theta}$ is an unbiased estimator!!

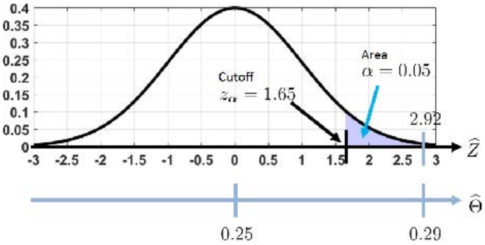
Further,

$$\widehat{Z} = \frac{\widehat{\theta} - \theta}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

Therefore, $\widehat{\theta} = 0.29$ is equivalent to $\widehat{Z} = 2.92$, and $\widehat{\theta} = 0.25$ is equivalent to $\widehat{Z} = 0$.

Hypothesis Testing

The mapping between $\widehat{\theta}$ and $\widehat{Z}$ is plotted below:



Critical level α is chosen a small value, exp. $\alpha = 0.05$ such that the corresponding cutoff is given by

$$z_\alpha = \text{cutoff location where the area under the curve is } \alpha$$

Thus

$$\Phi(z_\alpha) = 1 - \alpha.$$

Hypothesis testing

For the example above, the cutoff $z_\alpha = z_{0.05} = 1.65$

Hypothesis testing

For the example above, the cutoff $z_\alpha = z_{0.05} = 1.65$

Critical-value test

- Set a critical value z_α, and compute $\widehat{Z} = \frac{\widehat{\theta} - \theta}{\sigma/\sqrt{n}}$
- If $\widehat{Z} \geq z_\alpha$ then reject H_0
- If $\widehat{Z} < z_\alpha$ then keep H_0

Hypothesis Testing

p-value test

Instead of looking at the cutoff value z_α, here we inspect the probability of obtaining our hypothesis if H_0 is true.

Hypothesis Testing

p-value test

Instead of looking at the cutoff value z_α, here we inspect the probability of obtaining our hypothesis if H_0 is true.

Consider the example of tossing a coin, where the probability of getting head is unknown. Let

- $H_0 : \theta = 0.9$
- $H_1 : \theta < 0.9$

Hypothesis Testing

p-value test

Instead of looking at the cutoff value z_α, here we inspect the probability of obtaining our hypothesis if H_0 is true.

Consider the example of tossing a coin, where the probability of getting head is unknown. Let

- $H_0 : \theta = 0.9$
- $H_1 : \theta < 0.9$

Suppose the given data is as follows: $n = 150$ tosses, and number of heads in those tosses is 128. Then

$$\widehat{\theta} = \frac{128}{150} = 0.853.$$

Hypothesis Testing

Then we have

$$\widehat{Z} = \frac{\widehat{\theta} - \theta}{\sigma/\sqrt{n}} = \frac{0.853 - 0.9}{\sqrt{\frac{0.9(1-0.9)}{150}}} = -1.92$$

Hypothesis Testing

Then we have

$$\widehat{Z} = \frac{\widehat{\theta} - \theta}{\sigma/\sqrt{n}} = \frac{0.853 - 0.9}{\sqrt{\frac{0.9(1-0.9)}{150}}} = -1.92$$

Now we find the the probability under the curve if we integrate the pdf of $\widehat{Z}$ from $-\infty$ to -1.92. Since $\widehat{Z} \sim \mathcal{N}(0,1)$, we have

$$P\left(\widehat{Z} \leq -1.92\right) = 0.0274 \text{ (}p\text{-value)}$$

Hypothesis Testing

Then we have

$$\widehat{Z} = \frac{\widehat{\theta} - \theta}{\sigma/\sqrt{n}} = \frac{0.853 - 0.9}{\sqrt{\frac{0.9(1-0.9)}{150}}} = -1.92$$

Now we find the the probability under the curve if we integrate the pdf of $\widehat{Z}$ from $-\infty$ to -1.92. Since $\widehat{Z} \sim \mathcal{N}(0,1)$, we have

$$P\left(\widehat{Z} \leq -1.92\right) = 0.0274 \text{ (}p\text{-value)}$$

Since this value is less than the critical level α (approximately 5%,) we reject the null hypothesis.

Hypothesis Testing

Then we have

$$\widehat{Z} = \frac{\widehat{\theta} - \theta}{\sigma/\sqrt{n}} = \frac{0.853 - 0.9}{\sqrt{\frac{0.9(1-0.9)}{150}}} = -1.92$$

Now we find the the probability under the curve if we integrate the pdf of $\widehat{Z}$ from $-\infty$ to -1.92. Since $\widehat{Z} \sim \mathcal{N}(0, 1)$, we have

$$P\left(\widehat{Z} \leq -1.92\right) = 0.0274 \text{ (}p\text{-value)}$$

Since this value is less than the critical level α (approximately 5%,) we reject the null hypothesis.

p-value

$$p - \text{value} = P(\widehat{Z} \leq z)$$

where z is the random realization of $\widehat{Z}$ estimated from the data

Hypothesis Testing

Decision rule from the p-value:

Hypothesis Testing

Decision rule from the p-value:

Reject H_0 if p-value $< \alpha$, otherwise accept H_0.

Hypothesis Testing

Decision rule from the p-value:

Reject H_0 if p-value $< \alpha$, otherwise accept H_0.

Relation between critical-value and p-value

There is a one-one correspondence. In the p-value test, if $\widehat{Z}$ is normal then

$$p\text{-value} = P(\widehat{Z} \leq z) = \Phi(z)$$

where Φ is the CDF of $\mathcal{N}(0, 1)$. Then taking the inverse

$$z = \Phi^{-1}(p\text{-value})$$

Hypothesis Testing

To test *p*-value, we compare it with the critical level α by checking

$$p\text{-value} < \alpha.$$

Hypothesis Testing

To test p-value, we compare it with the critical level α by checking

$$p\text{-value} < \alpha.$$

Taking inverse both sides,

$$\Phi^{-1}(p\text{-value}) < \Phi^{-1}(\alpha) \text{ i.e. } z < z_{\alpha}$$

Hypothesis Testing

To test p-value, we compare it with the critical level α by checking

$$p\text{-value} < \alpha.$$

Taking inverse both sides,

$$\Phi^{-1}(p\text{-value}) < \Phi^{-1}(\alpha) \text{ i.e. } z < z_{\alpha}$$

Thus if the test statistic fails for the p-value, it will also fail in the critical-value test, and vice-versa.

Hypothesis Testing

Difference between critical-value test and p-value test

- critical-value test: compare with respect to critical value, which is the cutoff on the z-axis

Hypothesis Testing

Difference between critical-value test and *p*-value test

- critical-value test: compare with respect to critical value, which is the cutoff on the *z*-axis
- *p*-value test: compare with respect to α, which is the probability

Hypothesis Testing

Difference between critical-value test and *p*-value test

- critical-value test: compare with respect to critical value, which is the cutoff on the *z*-axis
- *p*-value test: compare with respect to α, which is the probability
- both give the same conclusion

Hypothesis Testing

In general, we will now discuss the method of Hypothesis Testing for estimating parameters of the distribution of a population X.

Hypothesis Testing

In general, we will now discuss the method of Hypothesis Testing for estimating parameters of the distribution of a population X.

Suppose $X_1, \ldots, X_n$ is a random sample from a population with pdf

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

where $\theta > 0$ is an unknown parameter.

Hypothesis Testing

In general, we will now discuss the method of Hypothesis Testing for estimating parameters of the distribution of a population X.

Suppose $X_1, \ldots, X_n$ is a random sample from a population with pdf

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

where $\theta > 0$ is an unknown parameter.

Let $n = 4$ and $x_1 = 0.92$, $x_2 = 0.75$, $x_3 = 0.85$, $x_4 = 0.8$ is a sample data from the above distribution. Applying ML method, the estimator $\widehat{\theta}$ of θ is

$$\widehat{\theta} = -1 - \frac{4}{\ln(X_1) + \ln(X_2) + \ln(X_3) + \ln(X_4)}$$

Hypothesis Testing

In general, we will now discuss the method of Hypothesis Testing for estimating parameters of the distribution of a population X.

Suppose $X_1, \ldots, X_n$ is a random sample from a population with pdf

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

where $\theta > 0$ is an unknown parameter.

Let $n = 4$ and $x_1 = 0.92$, $x_2 = 0.75$, $x_3 = 0.85$, $x_4 = 0.8$ is a sample data from the above distribution. Applying ML method, the estimator $\widehat{\theta}$ of θ is

$$\widehat{\theta} = -1 - \frac{4}{\ln(X_1) + \ln(X_2) + \ln(X_3) + \ln(X_4)}$$

hence the ML estimate of θ is

$$\widehat{\theta} = -1 - \frac{4}{\ln(0.92) + \ln(0.75) + \ln(0.85) + \ln(0.80)} = 4.2861.$$

Hypothesis Testing

We we denote the null hypothesis and alternative hypothesis as

$$H_o : \theta \in \Omega_0 \text{ and } H_a : \theta \in \Omega_a$$

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

$$\Omega_o \cap \Omega_a = \emptyset \text{ and } \Omega_o \cup \Omega_a \subset \Omega$$

Hypothesis Testing

We we denote the null hypothesis and alternative hypothesis as

$$H_o : \theta \in \Omega_0 \text{ and } H_a : \theta \in \Omega_a$$

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

$$\Omega_o \cap \Omega_a = \emptyset \text{ and } \Omega_o \cup \Omega_a \subset \Omega$$

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

$$W(x_1, \ldots, x_n) = \frac{\max_{\theta \in \Omega_o} L(\theta, x_1, x_2, \ldots, x_n)}{\max_{\theta \in \Omega} L(\theta, x_1, x_2, \ldots, x_n)}$$

where

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

Hypothesis Testing

The critical region C (a Borel set in $\mathbb{R}^n$) of a test statistic $W(X_1, X_2, \ldots, X_n)$ is such that if $W(x_1, \ldots, x_n)$ is an element of C then we decide to accept H_a; otherwise we accept H_o.

Hypothesis Testing

The critical region C (a Borel set in $\mathbb{R}^n$) of a test statistic $W(X_1, X_2, \ldots, X_n)$ is such that if $W(x_1, \ldots, x_n)$ is an element of C then we decide to accept H_a; otherwise we accept H_o.

In hypothesis test, the problem is to decide whether the null hypothesis is true based on a random sample. There are four possible situations that determine whether the decision is correct or in error.

Hypothesis Testing

The critical region C (a Borel set in $\mathbb{R}^n$) of a test statistic $W(X_1, X_2, \ldots, X_n)$ is such that if $W(x_1, \ldots, x_n)$ is an element of C then we decide to accept H_a; otherwise we accept H_o.

In hypothesis test, the problem is to decide whether the null hypothesis is true based on a random sample. There are four possible situations that determine whether the decision is correct or in error.

	H_o is true	H_o is false
Accept H_0	Correct decision	Type II Error
Reject H_o	Type I Error	Correct Decision

Hypothesis testing

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

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

denoted by α, is defined as

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

Hypothesis testing

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

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

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,

$$\alpha = P(\text{Reject } H_o | H_o \text{ is true})$$

which is equivalent to

$$\alpha = P(\text{Accept } H_a | H_o \text{ is true})$$

Hypothesis Testing

The *probability of type II error* is defined as

$$\beta = P(\text{Accept } H_o | H_o \text{ is false})$$

which is equivalent to

$$\alpha = P(\text{Accept } H_o | H_a \text{ is true})$$