

AUA CS 108, Statistics, Fall 2019

Lecture 41

Michael Poghosyan

YSU, AUA

michael@ysu.am, mpoghosyan@aua.am

27 Nov 2019

Contents

- ▶ Likelihood Ratio Tests
- ▶ Neyman-Pearson Lemma
- ▶ p -Values

Last Lecture ReCap

Last Lecture ReCap

- ▶ What were we doing last time?

Notes

Last time we have talked about 2 Sample Tests for Normally Distributed Samples (variances known/unknown, equal/unequal, paired/unpaired).

Notes

Last time we have talked about 2 Sample Tests for Normally Distributed Samples (variances known/unknown, equal/unequal, paired/unpaired).

Note: An important note is that you can perform 2 Sample t -Test even the DataSets are not Normally Distributed, but n and m are large (say, $n, m \geq 30$).

Notes

Last time we have talked about 2 Sample Tests for Normally Distributed Samples (variances known/unknown, equal/unequal, paired/unpaired).

Note: An important note is that you can perform 2 Sample t -Test even the DataSets are not Normally Distributed, but n and m are large (say, $n, m \geq 30$). This is because the t -Statistics will have, approximately (in the limit), a Standard Normal Distribution, so we can use Asymptotic Tests, and replace the quantiles of Standard Normal by t -quantiles (as we have talked in the previous lecture).

Notes

Note: To perform a two-Sample t -Test in **R**, we can use the same `t.test`.

Notes

Note: To perform a two-Sample t -Test in **R**, we can use the same `t.test`. When the input of `t.test` consists of two Datasets, it performs a two-Sample Test.

Notes

Note: To perform a two-Sample t -Test in **R**, we can use the same `t.test`. When the input of `t.test` consists of two Datasets, it performs a two-Sample Test. Some appropriate parameters are:

- ▶ `paired` : if TRUE, then it is doing a Paired t -Test. The default value is FALSE;

Notes

Note: To perform a two-Sample t -Test in **R**, we can use the same `t.test`. When the input of `t.test` consists of two Datasets, it performs a two-Sample Test. Some appropriate parameters are:

- ▶ `paired` : if TRUE, then it is doing a Paired t -Test. The default value is FALSE;
- ▶ `mu` : our μ_0 , in two-Sample Test, the value of $\mu_X - \mu_Y$ from the Null.

Notes

Note: To perform a two-Sample t -Test in **R**, we can use the same `t.test`. When the input of `t.test` consists of two Datasets, it performs a two-Sample Test. Some appropriate parameters are:

- ▶ `paired` : if TRUE, then it is doing a Paired t -Test. The default value is FALSE;
- ▶ `mu` : our μ_0 , in two-Sample Test, the value of $\mu_X - \mu_Y$ from the Null. The default value is 0, so by default,

$$\mathcal{H}_0 : \mu_X = \mu_Y$$

Notes

Note: To perform a two-Sample t -Test in **R**, we can use the same `t.test`. When the input of `t.test` consists of two Datasets, it performs a two-Sample Test. Some appropriate parameters are:

- ▶ `paired` : if TRUE, then it is doing a Paired t -Test. The default value is FALSE;
- ▶ `mu` : our μ_0 , in two-Sample Test, the value of $\mu_X - \mu_Y$ from the Null. The default value is 0, so by default,

$$\mathcal{H}_0 : \mu_X = \mu_Y$$

- ▶ `var.equal`: if TRUE, then we are dealing with Equal Variances case, so Pooled Standard Deviation is calculated. Otherwise, **R** is doing non-equal Variance Test.

Notes

Note: To perform a two-Sample t -Test in **R**, we can use the same `t.test`. When the input of `t.test` consists of two Datasets, it performs a two-Sample Test. Some appropriate parameters are:

- ▶ `paired` : if TRUE, then it is doing a Paired t -Test. The default value is FALSE;
- ▶ `mu` : our μ_0 , in two-Sample Test, the value of $\mu_X - \mu_Y$ from the Null. The default value is 0, so by default,

$$\mathcal{H}_0 : \mu_X = \mu_Y$$

- ▶ `var.equal`: if TRUE, then we are dealing with Equal Variances case, so Pooled Standard Deviation is calculated. Otherwise, **R** is doing non-equal Variance Test. By default, the value is FALSE.

Note: To perform a two-Sample Proportion Test in **R**, we can use `prop.test` command.

Likelihood Ratio Test

Likelihood Ratio Test (LRT)

LRT is one of a general methods to Design a Test.

Likelihood Ratio Test (LRT)

LRT is one of a general methods to Design a Test.

Model: $X_1, X_2, \dots, X_n \stackrel{IID}{\sim} \mathcal{F}_\theta, \theta \in \Theta;$

Likelihood Ratio Test (LRT)

LRT is one of a general methods to Design a Test.

Model: $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{F}_\theta, \theta \in \Theta$; We assume the Parametric Family \mathcal{F}_θ is given by its PD(M)Fs $f(x|\theta)$.

Likelihood Ratio Test (LRT)

LRT is one of a general methods to Design a Test.

Model: $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{F}_\theta, \theta \in \Theta$; We assume the Parametric Family \mathcal{F}_θ is given by its PD(M)Fs $f(x|\theta)$.

Hypothesis: $\mathcal{H}_0 : \theta \in \Theta_0$ vs $\mathcal{H}_1 : \theta \in \Theta_0^c$

Likelihood Ratio Test (LRT)

LRT is one of a general methods to Design a Test.

Model: $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{F}_\theta, \theta \in \Theta$; We assume the Parametric Family \mathcal{F}_θ is given by its PD(M)Fs $f(x|\theta)$.

Hypothesis: $\mathcal{H}_0 : \theta \in \Theta_0$ vs $\mathcal{H}_1 : \theta \in \Theta_0^c$

Significance Level: $\alpha \in (0, 1)$;

Likelihood Ratio Test (LRT)

LRT is one of a general methods to Design a Test.

Model: $X_1, X_2, \dots, X_n \stackrel{IID}{\sim} \mathcal{F}_\theta, \theta \in \Theta$; We assume the Parametric Family \mathcal{F}_θ is given by its PD(M)Fs $f(x|\theta)$.

Hypothesis: $\mathcal{H}_0 : \theta \in \Theta_0$ vs $\mathcal{H}_1 : \theta \in \Theta_0^c$

Significance Level: $\alpha \in (0, 1)$;

Rejection Region: we define the **Likelihood Ratio** by

$$LR = \frac{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta)}{\sup_{\theta \in \Theta} \mathcal{L}(\theta)}$$

Likelihood Ratio Test (LRT)

LRT is one of a general methods to Design a Test.

Model: $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{F}_\theta, \theta \in \Theta$; We assume the Parametric Family \mathcal{F}_θ is given by its PD(M)Fs $f(x|\theta)$.

Hypothesis: $\mathcal{H}_0 : \theta \in \Theta_0$ vs $\mathcal{H}_1 : \theta \in \Theta_0^c$

Significance Level: $\alpha \in (0, 1)$;

Rejection Region: we define the **Likelihood Ratio** by

$$LR = \frac{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta)}{\sup_{\theta \in \Theta} \mathcal{L}(\theta)}$$

where $\mathcal{L}(\theta)$ is our GOF *Likelihood Function*:

$$\mathcal{L}(\theta) = f(X_1|\theta) \cdot f(X_2|\theta) \cdot \dots \cdot f(X_n|\theta).$$

Likelihood Ratio Test (LRT), Cont'd

Now, our RR is of the form:

$$RR = \{LR \leq c\}$$

Likelihood Ratio Test (LRT), Cont'd

Now, our RR is of the form:

$$RR = \{LR \leq c\}$$

where we choose $c \in (0, 1)$ from the requirement to have a Significance Level α :

$$\sup_{\Theta_0} \mathbb{P}(\text{Type I Error}) = \sup_{\theta_0 \in \Theta_0} \mathbb{P}(LR \leq c \mid \theta = \theta_0) \leq \alpha.$$

Likelihood Ratio Test (LRT), Cont'd

Now, our RR is of the form:

$$RR = \{LR \leq c\}$$

where we choose $c \in (0, 1)$ from the requirement to have a Significance Level α :

$$\sup_{\Theta_0} \mathbb{P}(\text{Type I Error}) = \sup_{\theta_0 \in \Theta_0} \mathbb{P}(LR \leq c \mid \theta = \theta_0) \leq \alpha.$$

Tests, obtained by this method, are called LRTs.

Likelihood Ratio Test (LRT), Cont'd

Now, our RR is of the form:

$$RR = \{LR \leq c\}$$

where we choose $c \in (0, 1)$ from the requirement to have a Significance Level α :

$$\sup_{\Theta_0} \mathbb{P}(\text{Type I Error}) = \sup_{\theta_0 \in \Theta_0} \mathbb{P}(LR \leq c \mid \theta = \theta_0) \leq \alpha.$$

Tests, obtained by this method, are called LRTs.

Idea: Clearly, $0 \leq LR \leq 1$.

Likelihood Ratio Test (LRT), Cont'd

Now, our RR is of the form:

$$RR = \{LR \leq c\}$$

where we choose $c \in (0, 1)$ from the requirement to have a Significance Level α :

$$\sup_{\Theta_0} \mathbb{P}(\text{Type I Error}) = \sup_{\theta_0 \in \Theta_0} \mathbb{P}(LR \leq c \mid \theta = \theta_0) \leq \alpha.$$

Tests, obtained by this method, are called LRTs.

Idea: Clearly, $0 \leq LR \leq 1$. If $LR \approx 0$, then the Likelihood of our Data for $\theta \in \Theta_0$ is pretty small compared to the one with $\theta \notin \Theta_0$.

Likelihood Ratio Test (LRT), Cont'd

Now, our RR is of the form:

$$RR = \{LR \leq c\}$$

where we choose $c \in (0, 1)$ from the requirement to have a Significance Level α :

$$\sup_{\Theta_0} \mathbb{P}(\text{Type I Error}) = \sup_{\theta_0 \in \Theta_0} \mathbb{P}(LR \leq c \mid \theta = \theta_0) \leq \alpha.$$

Tests, obtained by this method, are called LRTs.

Idea: Clearly, $0 \leq LR \leq 1$. If $LR \approx 0$, then the Likelihood of our Data for $\theta \in \Theta_0$ is pretty small compared to the one with $\theta \notin \Theta_0$. So we do not believe that $\theta \in \Theta_0$, and we reject the Null.

Neyman-Pearson Lemma

Assume $X_1, \dots, X_n \sim \mathcal{F}_\theta$, and we want to Test, at α -level, a Simple Hypothesis vs Simple Hypothesis:

$$\mathcal{H}_0 : \theta = \theta_0 \quad \text{vs} \quad \mathcal{H}_1 : \theta = \theta_1.$$

Neyman-Pearson Lemma

Assume $X_1, \dots, X_n \sim \mathcal{F}_\theta$, and we want to Test, at α -level, a Simple Hypothesis vs Simple Hypothesis:

$$\mathcal{H}_0 : \theta = \theta_0 \quad \text{vs} \quad \mathcal{H}_1 : \theta = \theta_1.$$

Let $\mathcal{L}(\theta)$ be the Likelihood Function for X_1, \dots, X_n , and let

$$LR = \frac{\mathcal{L}(\theta_0)}{\mathcal{L}(\theta_1)}$$

be the Likelihood Ratio for our Test.

Neyman-Pearson Lemma

Assume $X_1, \dots, X_n \sim \mathcal{F}_\theta$, and we want to Test, at α -level, a Simple Hypothesis vs Simple Hypothesis:

$$\mathcal{H}_0 : \theta = \theta_0 \quad \text{vs} \quad \mathcal{H}_1 : \theta = \theta_1.$$

Let $\mathcal{L}(\theta)$ be the Likelihood Function for X_1, \dots, X_n , and let

$$LR = \frac{\mathcal{L}(\theta_0)}{\mathcal{L}(\theta_1)}$$

be the Likelihood Ratio for our Test. Assume there exists a number c such that for the Rejection Region

$$RR = \{LR \leq c\},$$

the significance level is α :

$$\mathbb{P}(LR \leq c \mid \mathcal{H}_0 \text{ is true}) = \alpha.$$

Neyman-Pearson Lemma

Assume $X_1, \dots, X_n \sim \mathcal{F}_\theta$, and we want to Test, at α -level, a Simple Hypothesis vs Simple Hypothesis:

$$\mathcal{H}_0 : \theta = \theta_0 \quad \text{vs} \quad \mathcal{H}_1 : \theta = \theta_1.$$

Let $\mathcal{L}(\theta)$ be the Likelihood Function for X_1, \dots, X_n , and let

$$LR = \frac{\mathcal{L}(\theta_0)}{\mathcal{L}(\theta_1)}$$

be the Likelihood Ratio for our Test. Assume there exists a number c such that for the Rejection Region

$$RR = \{LR \leq c\},$$

the significance level is α :

$$\mathbb{P}(LR \leq c \mid \mathcal{H}_0 \text{ is true}) = \alpha.$$

Then this Test has the highest Power among all other Tests of Significance Level α .

Example: LRT

Assume $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, and σ^2 is **known**.

Example: LRT

Assume $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, and σ^2 is **known**. We want to construct the LRT, for Significance Level α , for

$$\mathcal{H}_0 : \mu = \mu_0 \quad \text{vs} \quad \mathcal{H}_1 : \mu \neq \mu_0.$$

Example: LRT

Assume $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, and σ^2 is **known**. We want to construct the LRT, for Significance Level α , for

$$\mathcal{H}_0 : \mu = \mu_0 \quad \text{vs} \quad \mathcal{H}_1 : \mu \neq \mu_0.$$

Step 1: We calculate the Likelihood:

Example: LRT

Assume $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, and σ^2 is **known**. We want to construct the LRT, for Significance Level α , for

$$\mathcal{H}_0 : \mu = \mu_0 \quad vs \quad \mathcal{H}_1 : \mu \neq \mu_0.$$

Step 1: We calculate the Likelihood: the PDF is

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}},$$

Example: LRT

Assume $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, and σ^2 is **known**. We want to construct the LRT, for Significance Level α , for

$$\mathcal{H}_0 : \mu = \mu_0 \quad \text{vs} \quad \mathcal{H}_1 : \mu \neq \mu_0.$$

Step 1: We calculate the Likelihood: the PDF is

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}},$$

and the Likelihood Function is:

$$\mathcal{L}(\mu) = f(X_1|\mu, \sigma^2) \cdot f(X_2|\mu, \sigma^2) \cdot \dots \cdot f(X_n|\mu, \sigma^2)$$

Example: LRT

Assume $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, and σ^2 is **known**. We want to construct the LRT, for Significance Level α , for

$$\mathcal{H}_0 : \mu = \mu_0 \quad \text{vs} \quad \mathcal{H}_1 : \mu \neq \mu_0.$$

Step 1: We calculate the Likelihood: the PDF is

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}},$$

and the Likelihood Function is:

$$\mathcal{L}(\mu) = f(X_1|\mu, \sigma^2) \cdot f(X_2|\mu, \sigma^2) \cdot \dots \cdot f(X_n|\mu, \sigma^2)$$

i.e.

$$\mathcal{L}(\mu) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \cdot e^{-\frac{1}{2\sigma^2} \cdot \sum_{k=1}^n (X_k - \mu)^2}$$

Example: LRT, Cont'd

Step 2: Now, we calculate the LR:

$$LR = \frac{\sup_{\mu=\mu_0} \mathcal{L}(\mu)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)} =$$

Example: LRT, Cont'd

Step 2: Now, we calculate the LR:

$$LR = \frac{\sup_{\mu=\mu_0} \mathcal{L}(\mu)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)} = \frac{\mathcal{L}(\mu_0)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)}$$

Example: LRT, Cont'd

Step 2: Now, we calculate the LR:

$$LR = \frac{\sup_{\mu=\mu_0} \mathcal{L}(\mu)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)} = \frac{\mathcal{L}(\mu_0)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)}$$

We can write

$$\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu) = \mathcal{L}(\hat{\mu}^{MLE}),$$

since, by the very definition, $\hat{\mu}^{MLE}$ is the global max point of $\mathcal{L}(\mu)$.

Example: LRT, Cont'd

Step 2: Now, we calculate the LR:

$$LR = \frac{\sup_{\mu=\mu_0} \mathcal{L}(\mu)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)} = \frac{\mathcal{L}(\mu_0)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)}$$

We can write

$$\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu) = \mathcal{L}(\hat{\mu}^{MLE}),$$

since, by the very definition, $\hat{\mu}^{MLE}$ is the global max point of $\mathcal{L}(\mu)$. And, of course, we know that in this model

$$\hat{\mu}^{MLE} =$$

Example: LRT, Cont'd

Step 2: Now, we calculate the LR:

$$LR = \frac{\sup_{\mu=\mu_0} \mathcal{L}(\mu)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)} = \frac{\mathcal{L}(\mu_0)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)}$$

We can write

$$\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu) = \mathcal{L}(\hat{\mu}^{MLE}),$$

since, by the very definition, $\hat{\mu}^{MLE}$ is the global max point of $\mathcal{L}(\mu)$. And, of course, we know that in this model

$$\hat{\mu}^{MLE} = \overline{X}.$$

Example: LRT, Cont'd

Step 2: Now, we calculate the LR:

$$LR = \frac{\sup_{\mu=\mu_0} \mathcal{L}(\mu)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)} = \frac{\mathcal{L}(\mu_0)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)}$$

We can write

$$\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu) = \mathcal{L}(\hat{\mu}^{MLE}),$$

since, by the very definition, $\hat{\mu}^{MLE}$ is the global max point of $\mathcal{L}(\mu)$. And, of course, we know that in this model

$$\hat{\mu}^{MLE} = \overline{X}.$$

Plugging the obtained into the LR formula, we will obtain:

$$LR = \frac{\mathcal{L}(\mu_0)}{\mathcal{L}(\overline{X})} =$$

Example: LRT, Cont'd

Step 2: Now, we calculate the LR:

$$LR = \frac{\sup_{\mu=\mu_0} \mathcal{L}(\mu)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)} = \frac{\mathcal{L}(\mu_0)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)}$$

We can write

$$\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu) = \mathcal{L}(\hat{\mu}^{MLE}),$$

since, by the very definition, $\hat{\mu}^{MLE}$ is the global max point of $\mathcal{L}(\mu)$. And, of course, we know that in this model

$$\hat{\mu}^{MLE} = \bar{X}.$$

Plugging the obtained into the LR formula, we will obtain:

$$LR = \frac{\mathcal{L}(\mu_0)}{\mathcal{L}(\bar{X})} = \frac{e^{-\frac{1}{2\sigma^2} \cdot \sum_{k=1}^n (X_k - \mu_0)^2}}{e^{-\frac{1}{2\sigma^2} \cdot \sum_{k=1}^n (X_k - \bar{X})^2}} =$$

Example: LRT, Cont'd

Step 2: Now, we calculate the LR:

$$LR = \frac{\sup_{\mu=\mu_0} \mathcal{L}(\mu)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)} = \frac{\mathcal{L}(\mu_0)}{\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu)}$$

We can write

$$\sup_{\mu \in \mathbb{R}} \mathcal{L}(\mu) = \mathcal{L}(\hat{\mu}^{MLE}),$$

since, by the very definition, $\hat{\mu}^{MLE}$ is the global max point of $\mathcal{L}(\mu)$. And, of course, we know that in this model

$$\hat{\mu}^{MLE} = \bar{X}.$$

Plugging the obtained into the LR formula, we will obtain:

$$LR = \frac{\mathcal{L}(\mu_0)}{\mathcal{L}(\bar{X})} = \frac{e^{-\frac{1}{2\sigma^2} \cdot \sum_{k=1}^n (X_k - \mu_0)^2}}{e^{-\frac{1}{2\sigma^2} \cdot \sum_{k=1}^n (X_k - \bar{X})^2}} = e^{-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2}$$

Example: LRT, Cont'd

Step 3: Our RR is (c will be calculated later)

$$RR = \{LR \leq c\} =$$

Example: LRT, Cont'd

Step 3: Our RR is (c will be calculated later)

$$RR = \{LR \leq c\} = \{e^{-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2} \leq c\}$$

Example: LRT, Cont'd

Step 3: Our RR is (c will be calculated later)

$$RR = \{LR \leq c\} = \{e^{-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2} \leq c\}$$

which is equivalent to

$$-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2 \leq \ln c$$

Example: LRT, Cont'd

Step 3: Our RR is (c will be calculated later)

$$RR = \{LR \leq c\} = \{e^{-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2} \leq c\}$$

which is equivalent to

$$-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2 \leq \ln c$$

This is equivalent to

$$\left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \geq \sqrt{-2 \ln c} =: C$$

Example: LRT, Cont'd

Step 3: Our RR is (c will be calculated later)

$$RR = \{LR \leq c\} = \{e^{-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2} \leq c\}$$

which is equivalent to

$$-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2 \leq \ln c$$

This is equivalent to

$$\left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \geq \sqrt{-2 \ln c} =: C$$

But this is exactly

$$|Z| \geq C$$

where Z is our Z -Test Statistics. So we have arrived at our GOF Z -Test!

Example: LRT, Cont'd

Step 3: Our RR is (c will be calculated later)

$$RR = \{LR \leq c\} = \{e^{-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2} \leq c\}$$

which is equivalent to

$$-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2 \leq \ln c$$

This is equivalent to

$$\left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \geq \sqrt{-2 \ln c} =: C$$

But this is exactly

$$|Z| \geq C$$

where Z is our Z -Test Statistics. So we have arrived at our GOF Z -Test! And, in particular, we know that value of C : $C =$

Example: LRT, Cont'd

Step 3: Our RR is (c will be calculated later)

$$RR = \{LR \leq c\} = \{e^{-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2} \leq c\}$$

which is equivalent to

$$-\frac{n}{2\sigma^2} \cdot (\bar{X} - \mu_0)^2 \leq \ln c$$

This is equivalent to

$$\left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \geq \sqrt{-2 \ln c} =: C$$

But this is exactly

$$|Z| \geq C$$

where Z is our Z -Test Statistics. So we have arrived at our GOF Z -Test! And, in particular, we know that value of C : $C = z_{1-\alpha/2}$.

p-Values

3 Methods of Decision Making in Testing

When Testing a Hypothesis, at a Significance Level α , we can make the Decision in 3 ways:

3 Methods of Decision Making in Testing

When Testing a Hypothesis, at a Significance Level α , we can make the Decision in 3 ways:

- ▶ Based on the Test Statistics TS and the Rejection Region RR :
if $TS \in RR$, then we Reject Null Hypothesis, otherwise, we Fail to Reject Null;

3 Methods of Decision Making in Testing

When Testing a Hypothesis, at a Significance Level α , we can make the Decision in 3 ways:

- ▶ Based on the Test Statistics TS and the Rejection Region RR : if $TS \in RR$, then we Reject Null Hypothesis, otherwise, we Fail to Reject Null;
- ▶ Based on the Confidence Interval CI for the Parameter: if θ is our Parameter, (L, U) is a CI of $(1 - \alpha)$ -level for θ , and our Null is $\mathcal{H}_0 : \theta = \theta_0$, then we Reject Null if and only if $\theta_0 \notin (L, U)$;

3 Methods of Decision Making in Testing

When Testing a Hypothesis, at a Significance Level α , we can make the Decision in 3 ways:

- ▶ Based on the Test Statistics TS and the Rejection Region RR : if $TS \in RR$, then we Reject Null Hypothesis, otherwise, we Fail to Reject Null;
- ▶ Based on the Confidence Interval CI for the Parameter: if θ is our Parameter, (L, U) is a CI of $(1 - \alpha)$ -level for θ , and our Null is $\mathcal{H}_0 : \theta = \theta_0$, then we Reject Null if and only if $\theta_0 \notin (L, U)$;
- ▶ Based on the p -Value: if $p < \alpha$, Reject Null, otherwise, Fail to Reject.

3 Methods of Decision Making in Testing

When Testing a Hypothesis, at a Significance Level α , we can make the Decision in 3 ways:

- ▶ Based on the Test Statistics TS and the Rejection Region RR : if $TS \in RR$, then we Reject Null Hypothesis, otherwise, we Fail to Reject Null;
- ▶ Based on the Confidence Interval CI for the Parameter: if θ is our Parameter, (L, U) is a CI of $(1 - \alpha)$ -level for θ , and our Null is $\mathcal{H}_0 : \theta = \theta_0$, then we Reject Null if and only if $\theta_0 \notin (L, U)$;
- ▶ Based on the p -Value: if $p < \alpha$, Reject Null, otherwise, Fail to Reject.

Note: When doing Tests, say, with `t.test`, **R** is calculating the p -Value, and sometimes also the CI. So, to decide whether to Reject Null or Not, using **R**, you can use the 2nd and 3rd Methods.

p -Values

Now, about the p -Value.

p -Values

Now, about the p -Value. Assume we are Testing a Hypothesis

$$\mathcal{H}_0 \quad \text{vs} \quad \mathcal{H}_1$$

based on the Test Statistics TS .

p -Values

Now, about the p -Value. Assume we are Testing a Hypothesis

$$\mathcal{H}_0 \quad \text{vs} \quad \mathcal{H}_1$$

based on the Test Statistics TS . Assume we already have Observations, and we calculate the value of TS , let us denote that by TS_{obs} (this is just a number).

p -Values

Now, about the p -Value. Assume we are Testing a Hypothesis

$$\mathcal{H}_0 \quad \text{vs} \quad \mathcal{H}_1$$

based on the Test Statistics TS . Assume we already have Observations, and we calculate the value of TS , let us denote that by TS_{obs} (this is just a number). We know that, for a given Significance Level α , we will Reject \mathcal{H}_0 , iff TS_{obs} will be in the RR .

Now, assume the Distribution of TS , our Test Statistics, **under** \mathcal{H}_0 , is given like this (I am drawing for Z - or t -Statistics, for Two Tailed Test, the other cases can be considered in a similar way):

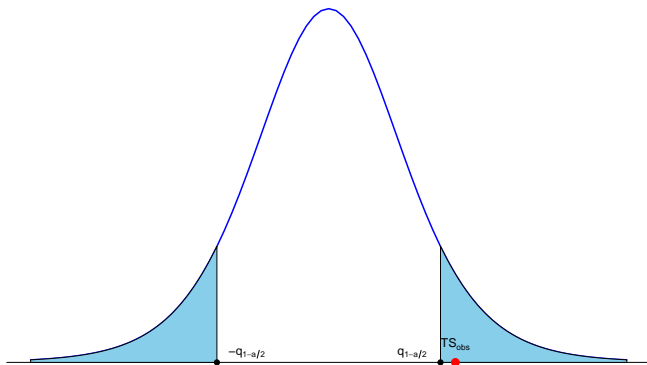
p -Values

Let us take a Significance Level $\alpha \in (0, 1)$:

p -Values

Let us take a Significance Level $\alpha \in (0, 1)$:

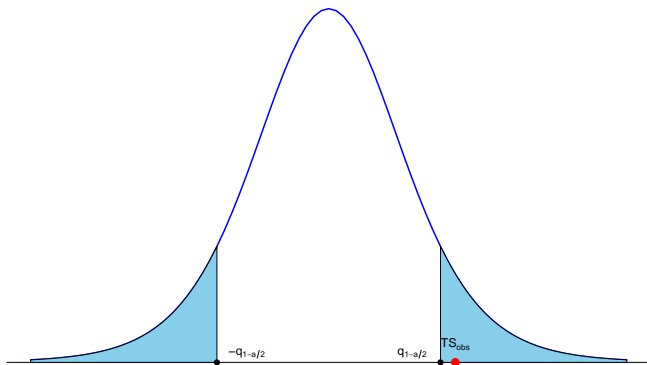
Distribution of TS, with RR, siglev= α



p -Values

Let us take a Significance Level $\alpha \in (0, 1)$:

Distribution of TS, with RR, siglev= α



We Reject \mathcal{H}_0 at the level α

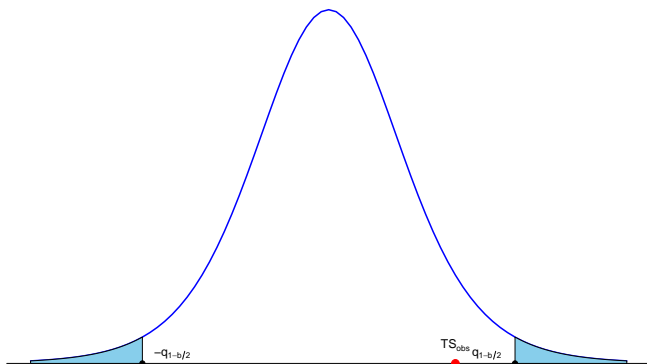
p -Values

Now, let us change our Significance Level to $b < a$:

p -Values

Now, let us change our Significance Level to $b < a$:

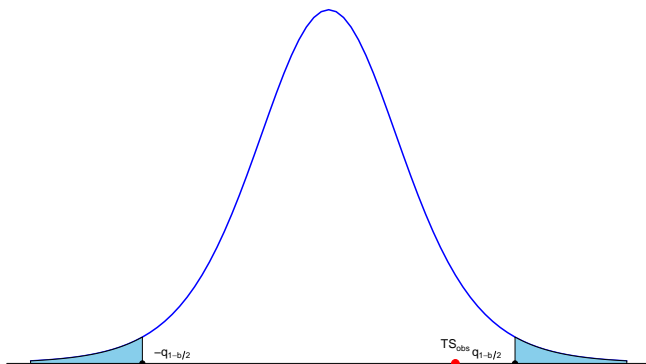
Distribution of TS, with RR, siglev= b



p -Values

Now, let us change our Significance Level to $b < a$:

Distribution of TS, with RR, siglev= b



We Do Not Reject \mathcal{H}_0 at the level b

p -Values

Now, it is clear that, having our Observed Test Statistics TS_{obs} , we will Reject our Null for some large values of α , but Fail to Reject for very small values of α .

p -Values

Now, it is clear that, having our Observed Test Statistics TS_{obs} , we will Reject our Null for some large values of α , but Fail to Reject for very small values of α .

Also, clearly, if we are Rejecting Null at the level α , then we will Reject also at any level $\alpha' \geq \alpha$.

p -Values

Now, it is clear that, having our Observed Test Statistics TS_{obs} , we will Reject our Null for some large values of α , but Fail to Reject for very small values of α .

Also, clearly, if we are Rejecting Null at the level α , then we will Reject also at any level $\alpha' \geq \alpha$. And, similarly, if we Fail to Reject at the level β , then we also will Fail to Reject at any level $\beta' \leq \beta$.

p -Values

Now, it is clear that, having our Observed Test Statistics TS_{obs} , we will Reject our Null for some large values of α , but Fail to Reject for very small values of α .

Also, clearly, if we are Rejecting Null at the level α , then we will Reject also at any level $\alpha' \geq \alpha$. And, similarly, if we Fail to Reject at the level β , then we also will Fail to Reject at any level $\beta' \leq \beta$.

Then we will have a point $\alpha^* \in (0, 1)$ such that

p -Values

Now, it is clear that, having our Observed Test Statistics TS_{obs} , we will Reject our Null for some large values of α , but Fail to Reject for very small values of α .

Also, clearly, if we are Rejecting Null at the level α , then we will Reject also at any level $\alpha' \geq \alpha$. And, similarly, if we Fail to Reject at the level β , then we also will Fail to Reject at any level $\beta' \leq \beta$.

Then we will have a point $\alpha^* \in (0, 1)$ such that

- ▶ We Reject \mathcal{H}_0 for any $\alpha > \alpha^*$

p -Values

Now, it is clear that, having our Observed Test Statistics TS_{obs} , we will Reject our Null for some large values of α , but Fail to Reject for very small values of α .

Also, clearly, if we are Rejecting Null at the level α , then we will Reject also at any level $\alpha' \geq \alpha$. And, similarly, if we Fail to Reject at the level β , then we also will Fail to Reject at any level $\beta' \leq \beta$.

Then we will have a point $\alpha^* \in (0, 1)$ such that

- ▶ We Reject \mathcal{H}_0 for any $\alpha > \alpha^*$
- ▶ We Fail to Reject \mathcal{H}_0 for any $\alpha < \alpha^*$

p -Values

Now, it is clear that, having our Observed Test Statistics TS_{obs} , we will Reject our Null for some large values of α , but Fail to Reject for very small values of α .

Also, clearly, if we are Rejecting Null at the level α , then we will Reject also at any level $\alpha' \geq \alpha$. And, similarly, if we Fail to Reject at the level β , then we also will Fail to Reject at any level $\beta' \leq \beta$.

Then we will have a point $\alpha^* \in (0, 1)$ such that

- ▶ We Reject \mathcal{H}_0 for any $\alpha > \alpha^*$
- ▶ We Fail to Reject \mathcal{H}_0 for any $\alpha < \alpha^*$

Note: Give here the real line with picture, MP!

p -Values

Now, it is clear that, having our Observed Test Statistics TS_{obs} , we will Reject our Null for some large values of α , but Fail to Reject for very small values of α .

Also, clearly, if we are Rejecting Null at the level α , then we will Reject also at any level $\alpha' \geq \alpha$. And, similarly, if we Fail to Reject at the level β , then we also will Fail to Reject at any level $\beta' \leq \beta$.

Then we will have a point $\alpha^* \in (0, 1)$ such that

- ▶ We Reject \mathcal{H}_0 for any $\alpha > \alpha^*$
- ▶ We Fail to Reject \mathcal{H}_0 for any $\alpha < \alpha^*$

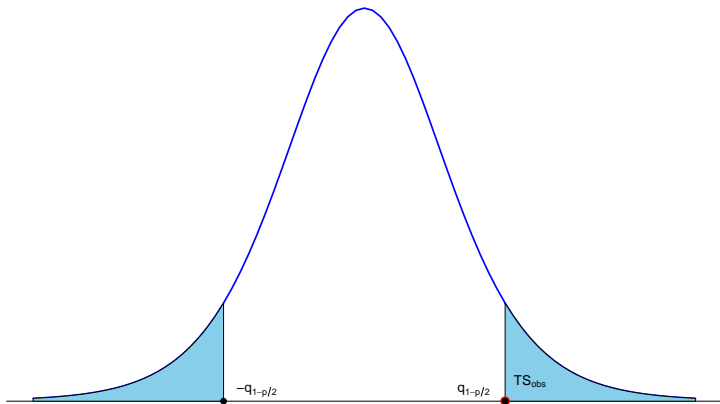
Note: Give here the real line with picture, MP!

Now, we denote $p = \alpha^*$ and call it the **p -Value of the Test:**

$$p\text{-Value} = p = \inf\{\alpha : \text{we Reject } \mathcal{H}_0 \text{ at level } \alpha\}.$$

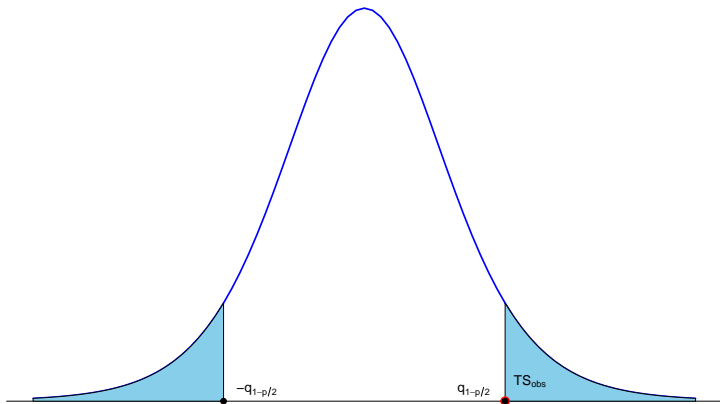
p -Values

Distribution of TS, with RR, siglev= p



p -Values

Distribution of TS, with RR, siglev= p



p -Value, the inf value of α at which we Reject \mathcal{H}_0

p -Values

It is clear from the Figure above that

$$p\text{-Value} = p = \mathbb{P}(|TS| > TS_{obs} \mid \mathcal{H}_0 \text{ is True}).$$

p -Values

It is clear from the Figure above that

$$p\text{-Value} = p = \mathbb{P}(|TS| > TS_{obs} \mid \mathcal{H}_0 \text{ is True}).$$

In words, p -Value is the Probability, under the Null Hypothesis, that we will have a value of Test Statistics TS as extreme as the Observed one.

p -Values

It is clear from the Figure above that

$$p\text{-Value} = p = \mathbb{P}(|TS| > TS_{obs} \mid \mathcal{H}_0 \text{ is True}).$$

In words, p -Value is the Probability, under the Null Hypothesis, that we will have a value of Test Statistics TS as extreme as the Observed one. I.e., p -Value is measuring somehow *how extreme is our Observed Test Statistics TS_{obs}* under \mathcal{H}_0 .

p -Values

It is clear from the Figure above that

$$p\text{-Value} = p = \mathbb{P}(|TS| > TS_{obs} \mid \mathcal{H}_0 \text{ is True}).$$

In words, p -Value is the Probability, under the Null Hypothesis, that we will have a value of Test Statistics TS as extreme as the Observed one. I.e., p -Value is measuring somehow *how extreme is our Observed Test Statistics TS_{obs}* under \mathcal{H}_0 .

And, if p -Value is small, then TS_{obs} is very improbable, very unbelievable, under \mathcal{H}_0 , so we safely Reject \mathcal{H}_0 .

p -Values

It is clear from the Figure above that

$$p\text{-Value} = p = \mathbb{P}(|TS| > TS_{obs} \mid \mathcal{H}_0 \text{ is True}).$$

In words, p -Value is the Probability, under the Null Hypothesis, that we will have a value of Test Statistics TS as extreme as the Observed one. I.e., p -Value is measuring somehow *how extreme is our Observed Test Statistics TS_{obs}* under \mathcal{H}_0 .

And, if p -Value is small, then TS_{obs} is very improbable, very unbelievable, under \mathcal{H}_0 , so we safely Reject \mathcal{H}_0 .

To Remember:

- ▶ If $p\text{-Value} < \alpha$, then we Reject \mathcal{H}_0
- ▶ If $p\text{-Value} \geq \alpha$, then we Fail to Reject \mathcal{H}_0

R Code for the Graphics

```
df <- 8;
x <- seq(-4,4,0.1); y <- dt(x, df = df)
plot.new()
plot.window(xlim = c(-4, 4), ylim = c(-0.05,0.4))
plot(x,y, type="l",col="blue",lwd=2,xaxt="n",yaxt="n",
      bty="n",xlab="",ylab="")
abline(h=0)
title("Distribution of TS, with RR, siglev=a ")
qpoint <- 1.5; tspoint <- 1.7
cord.x <- c(qpoint,seq(qpoint,4,0.01),4)
cord.y <- c(0,dt(seq(qpoint,4,0.01), df=df),0)
polygon(cord.x,cord.y,col='skyblue')
points(c(qpoint), c(0), pch=20, cex=1.4)
text(c(qpoint-0.38),c(0.01),labels=expression("q"[1-a/2]))
cord.x1 <- c(-4,seq(-4,-qpoint,0.01),-qpoint)
cord.y1 <- c(0,dt(seq(-4,-qpoint,0.01), df=df),0)
polygon(cord.x1,cord.y1,col='skyblue')
points(c(-qpoint), c(0), pch=20, cex=1.4)
text(c(-qpoint+0.4),c(0.01),labels=expression("-q"[1-a/2]))
points(c(tspoint), c(0), col="red", pch=19, cex=1.4)
text(c(tspoint), c(0.02), labels = expression("TS"[obs]))
```