ASDS Statistics, YSU, Fall 2020

Lecture 30

Michael Poghosyan

26 Dec 2020

Contents

t-Test Example

Example: Recall the Etruscans-Italians Problem: Scientists have a data about 84 skull sizes (widths) of adult Etruscans, and the problem was to see if Etruscans were Italians.

Also Scientists believe that the skull size is not changing much through time, and modern adult Italians skull size is in average 132.4mm.

t-Test Example

Example: Recall the Etruscans-Italians Problem: Scientists have a data about 84 skull sizes (widths) of adult Etruscans, and the problem was to see if Etruscans were Italians.

Also Scientists believe that the skull size is not changing much through time, and modern adult Italians skull size is in average 132.4mm.

We make a Hypothesis:

$$\mathcal{H}_0 : \text{Etruscans were Italians} \quad vs \quad \mathcal{H}_1 : \text{They were not}$$

t-Test Example

Example: Recall the Etruscans-Italians Problem: Scientists have a data about 84 skull sizes (widths) of adult Etruscans, and the problem was to see if Etruscans were Italians.

Also Scientists believe that the skull size is not changing much through time, and modern adult Italians skull size is in average 132.4mm.

We make a Hypothesis:

$$\mathcal{H}_0 : \text{Etruscans were Italians} \quad \textit{vs} \quad \mathcal{H}_1 : \text{They were not}$$

In mathematical terms,

$$\mathcal{H}_0 : \mu = 132.4 \qquad \textit{vs} \qquad \mathcal{H}_1 : \mu \neq 132.4$$

t-Test Example

Example: Recall the Etruscans-Italians Problem: Scientists have a data about 84 skull sizes (widths) of adult Etruscans, and the problem was to see if Etruscans were Italians.

Also Scientists believe that the skull size is not changing much through time, and modern adult Italians skull size is in average 132.4mm.

We make a Hypothesis:

$$\mathcal{H}_0 : \text{Etruscans were Italians} \quad \textit{vs} \quad \mathcal{H}_1 : \text{They were not}$$

In mathematical terms,

$$\mathcal{H}_0 : \mu = 132.4 \qquad \textit{vs} \qquad \mathcal{H}_1 : \mu \neq 132.4$$

Now let's test, at 95%, this Hypo in **R**:

t-Test Example

```r
library(Rlab)
data <- etruscan
x <- data$width[data$group == "ancient"]

t.test(x, mu = 132.4)
```

```
##
##  One Sample t-test
##
## data:  x
## t = 17.46, df = 83, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 132.4
## 95 percent confidence interval:
##  142.4781 145.0695
## sample estimates:
## mean of x
##  143.7738
```

Test for the Normal Variance, μ is known

Model: $X_1, X_2, ..., X_n \stackrel{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is known**, the Parameter (our unknown) is σ^2;

Test for the Normal Variance, μ is known

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is known**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0: \ \sigma^2 = \sigma_0^2$

Test for the Normal Variance, μ is known

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is known**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0 : \ \sigma^2 = \sigma_0^2$

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

Test for the Normal Variance, μ is known

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is known**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0: \ \sigma^2 = \sigma_0^2$

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

Test Statistics:

Test for the Normal Variance, μ is known

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is known**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0 : \ \sigma^2 = \sigma_0^2$

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

Test Statistics: $\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma_0} \right)^2$

Test for the Normal Variance, μ is known

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is known**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0: \ \sigma^2 = \sigma_0^2$

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

Test Statistics: $\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma_0} \right)^2$

Distrib of the Test-Statistics Under $\mathcal{H}_0$: $\chi^2 \sim \chi^2(n)$;

Test for the Normal Variance, μ is known

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is known**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0: \ \sigma^2 = \sigma_0^2$

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

Test Statistics: $\chi^2 = \displaystyle\sum_{k=1}^{n} \left(\frac{X_k - \mu}{\sigma_0} \right)^2$

Distrib of the Test-Statistics Under $\mathcal{H}_0$: $\chi^2 \sim \chi^2(n)$;

$\mathcal{H}_1$ is	RR is
$\sigma^2 \neq \sigma_0^2$	$\chi^2 \notin \left[\chi^2_{n,\frac{\alpha}{2}}, \chi^2_{n,1-\frac{\alpha}{2}} \right]$
$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{n,1-\alpha}$
$\sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{n,\alpha}$

Test for the Normal Variance, μ is unknown

Model: $X_1, X_2, ..., X_n \stackrel{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is unknown**, the Parameter (our unknown) is σ^2;

Test for the Normal Variance, μ is unknown

Model: $X_1, X_2, ..., X_n \stackrel{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is unknown**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0: \ \sigma^2 = \sigma_0^2$

Test for the Normal Variance, μ is unknown

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is unknown**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0: \ \sigma^2 = \sigma_0^2$

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

Test for the Normal Variance, μ is unknown

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is unknown**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0 : \ \sigma^2 = \sigma_0^2$

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

Test Statistics:

Test for the Normal Variance, μ is unknown

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is unknown**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0 : \ \sigma^2 = \sigma_0^2$

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

Test Statistics: $\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \overline{X}}{\sigma_0} \right)^2 = \frac{(n-1) \cdot S^2}{\sigma_0^2}$

Test for the Normal Variance, μ is unknown

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is unknown**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0: \ \sigma^2 = \sigma_0^2$

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

Test Statistics: $\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \overline{X}}{\sigma_0} \right)^2 = \frac{(n-1) \cdot S^2}{\sigma_0^2}$

Distrib of the Test-Statistics Under $\mathcal{H}_0$: $\chi^2 \sim \chi^2(n-1)$;

Test for the Normal Variance, μ is unknown

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ **is unknown**, the Parameter (our unknown) is σ^2;

Null Hypothesis: $\mathcal{H}_0: \ \sigma^2 = \sigma_0^2$

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

Test Statistics: $\chi^2 = \sum_{k=1}^{n} \left(\frac{X_k - \overline{X}}{\sigma_0} \right)^2 = \frac{(n-1) \cdot S^2}{\sigma_0^2}$

Distrib of the Test-Statistics Under $\mathcal{H}_0$: $\chi^2 \sim \chi^2(n-1)$;

$\mathcal{H}_1$ is	RR is
$\sigma^2 \neq \sigma_0^2$	$\chi^2 \notin \left[\chi^2_{n-1,\frac{\alpha}{2}}, \chi^2_{n-1,1-\frac{\alpha}{2}} \right]$
$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{n-1,1-\alpha}$
$\sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{n-1,\alpha}$

Large Sample Hypothesis Testing

aka

Asymptotic Testing

Asymptotic Test for the Mean of General Distribution

Model: $X_1, X_2, ..., X_n$ are IID from some Distribution with Mean μ and Variance σ^2. σ^2 is not known. The Parameter is μ;

Asymptotic Test for the Mean of General Distribution

Model: $X_1, X_2, ..., X_n$ are IID from some Distribution with Mean μ and Variance σ^2. σ^2 is not known. The Parameter is μ;

Null Hypothesis: $\mathcal{H}_0: \ \mu = \mu_0$

Asymptotic Test for the Mean of General Distribution

Model: $X_1, X_2, ..., X_n$ are IID from some Distribution with Mean μ and Variance σ^2. σ^2 is not known. The Parameter is μ;

Null Hypothesis: $\mathcal{H}_0: \ \mu = \mu_0$

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

Asymptotic Test for the Mean of General Distribution

Model: $X_1, X_2, ..., X_n$ are IID from some Distribution with Mean μ and Variance σ^2. σ^2 is not known. The Parameter is μ;

Null Hypothesis: $\mathcal{H}_0 : \ \mu = \mu_0$

Asymptotic Significance Level: $\alpha \in (0, 1)$;This means that we want to have

$$\mathbb{P}(\text{Type I Error}) = \mathbb{P}(\text{Reject } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ is True}) \to \alpha, \quad \textit{as} \quad n \to +\infty$$

Asymptotic Test for the Mean of General Distribution

Model: $X_1, X_2, ..., X_n$ are IID from some Distribution with Mean μ and Variance σ^2. σ^2 is not known. The Parameter is μ;

Null Hypothesis: $\mathcal{H}_0: \ \mu = \mu_0$

Asymptotic Significance Level: $\alpha \in (0,1)$;This means that we want to have

$$\mathbb{P}(\text{Type I Error}) = \mathbb{P}(\text{Reject } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ is True}) \to \alpha, \quad \textit{as} \quad n \to +\infty$$

Test Statistics: $t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$, where $S^2 = \dfrac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}$.

Asymptotic Test for the Mean of General Distribution

Model: $X_1, X_2, ..., X_n$ are IID from some Distribution with Mean μ and Variance σ^2. σ^2 is not known. The Parameter is μ;

Null Hypothesis: $\mathcal{H}_0: \ \mu = \mu_0$

Asymptotic Significance Level: $\alpha \in (0, 1)$;This means that we want to have

$$\mathbb{P}(\text{Type I Error}) = \mathbb{P}(\text{Reject } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ is True}) \to \alpha, \quad \textit{as} \quad n \to +\infty$$

Test Statistics: $t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$, where $S^2 = \dfrac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}$.

Asymptotic Distrib of the TS Under $\mathcal{H}_0$: $t \xrightarrow{D} \mathcal{N}(0, 1)$;

Asymptotic Test for the Mean of General Distribution

Model: $X_1, X_2, ..., X_n$ are IID from some Distribution with Mean μ and Variance σ^2. σ^2 is not known. The Parameter is μ;

Null Hypothesis: $\mathcal{H}_0 : \ \mu = \mu_0$

Asymptotic Significance Level: $\alpha \in (0,1)$;This means that we want to have

$$\mathbb{P}(\text{Type I Error}) = \mathbb{P}(\text{Reject } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ is True}) \to \alpha, \quad as \quad n \to +\infty$$

Test Statistics: $t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$, where $S^2 = \dfrac{\sum_{k=1}^{n}(X_k - \overline{X})^2}{n-1}$.

Asymptotic Distrib of the TS Under $\mathcal{H}_0$: $t \xrightarrow{D} \mathcal{N}(0,1)$;

$\mathcal{H}_1$ is	RR is
$\mu \neq \mu_0$	$\lvert t\rvert > z_{1-\frac{\alpha}{2}}$
$\mu > \mu_0$	$t > z_{1-\alpha}$
$\mu < \mu_0$	$t < z_{\alpha}$

Asymptotic Tests based on MLEs

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{F}_\theta$;

Asymptotic Tests based on MLEs

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{F}_\theta$;

Null Hypothesis: $\mathcal{H}_0: \ \theta = \theta_0$

Asymptotic Tests based on MLEs

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{F}_\theta$;

Null Hypothesis: $\mathcal{H}_0 : \ \theta = \theta_0$

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

Asymptotic Tests based on MLEs

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{F}_\theta$;

Null Hypothesis: $\mathcal{H}_0 : \ \theta = \theta_0$

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

Test Statistics: $W = \dfrac{\hat{\theta}^{MLE} - \theta_0}{1/\sqrt{n \cdot \mathcal{I}(\theta_0)}}$ or $W = \dfrac{\hat{\theta}^{MLE} - \theta_0}{1/\sqrt{n \cdot \mathcal{I}\left(\hat{\theta}^{MLE}\right)}}$

Asymptotic Tests based on MLEs

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{F}_\theta$;

Null Hypothesis: $\mathcal{H}_0 : \ \theta = \theta_0$

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

Test Statistics: $W = \dfrac{\hat{\theta}^{MLE} - \theta_0}{1/\sqrt{n \cdot \mathcal{I}(\theta_0)}}$ or $W = \dfrac{\hat{\theta}^{MLE} - \theta_0}{1/\sqrt{n \cdot \mathcal{I}\left(\hat{\theta}^{MLE}\right)}}$

Asymptotic Distrib of the TS Under $\mathcal{H}_0$: $W \xrightarrow{D} \mathcal{N}(0, 1)$;

Asymptotic Tests based on MLEs

Model: $X_1, X_2, ..., X_n \overset{IID}{\sim} \mathcal{F}_\theta$;

Null Hypothesis: $\mathcal{H}_0 : \ \theta = \theta_0$

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

Test Statistics: $W = \dfrac{\hat{\theta}^{MLE} - \theta_0}{1/\sqrt{n \cdot \mathcal{I}(\theta_0)}}$ or $W = \dfrac{\hat{\theta}^{MLE} - \theta_0}{1/\sqrt{n \cdot \mathcal{I}\left(\hat{\theta}^{MLE}\right)}}$

Asymptotic Distrib of the TS Under $\mathcal{H}_0$: $W \xrightarrow{D} \mathcal{N}(0, 1)$;

$\mathcal{H}_1$ is	RR is
$\theta \neq \theta_0$	$\lvert W \rvert > z_{1-\frac{\alpha}{2}}$
$\theta > \theta_0$	$W > z_{1-\alpha}$
$\theta < \theta_0$	$W < z_\alpha$

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, othewise, 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, othewise, 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, othewise, 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, othewise, 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

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1$$

based on the Test Statistics TS.

p-Values

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

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1$$

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

$$\mathcal{H}_0 \qquad vs \qquad \mathcal{H}_1$$

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.

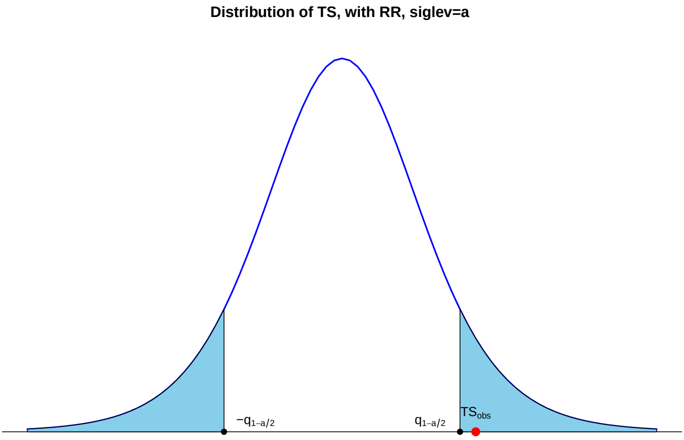
Now, assume the Distribution of TS, our Test Statistics, **under** $\mathcal{H}_0$, is given like this (I am drawing for Z- or t-Statistics, for Two Tailed Test, the other cases can be considered in a similar way):

p-Values

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

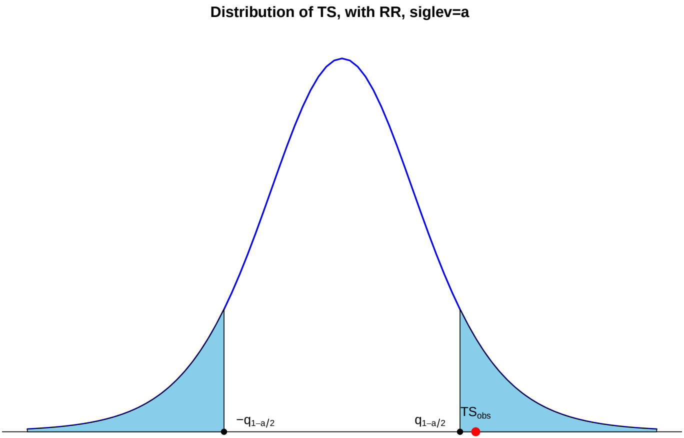
p-Values

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



p-Values

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



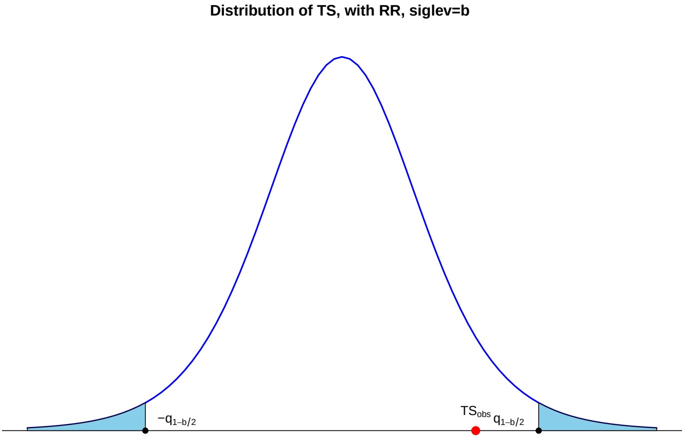
We Reject $\mathcal{H}_0$ at the level a

p-Values

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

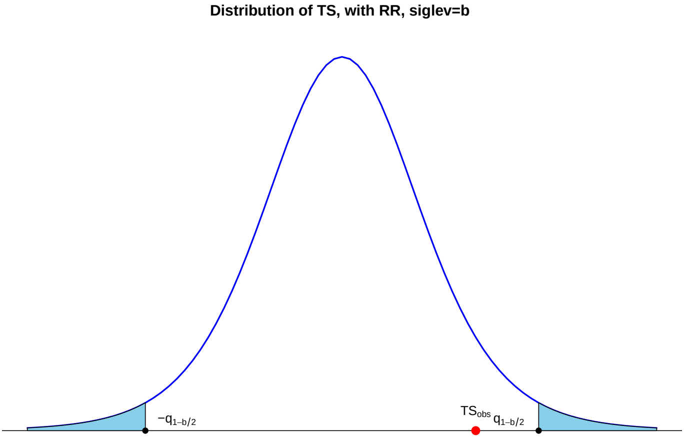
p-Values

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



p-Values

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



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

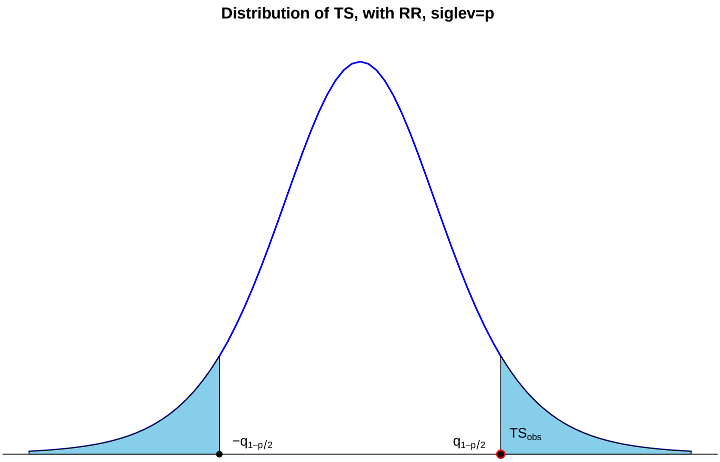
- We Reject $\mathcal{H}_0$ for any $\alpha > \alpha^*$
- We Fail to Reject $\mathcal{H}_0$ for any $\alpha < \alpha^*$

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

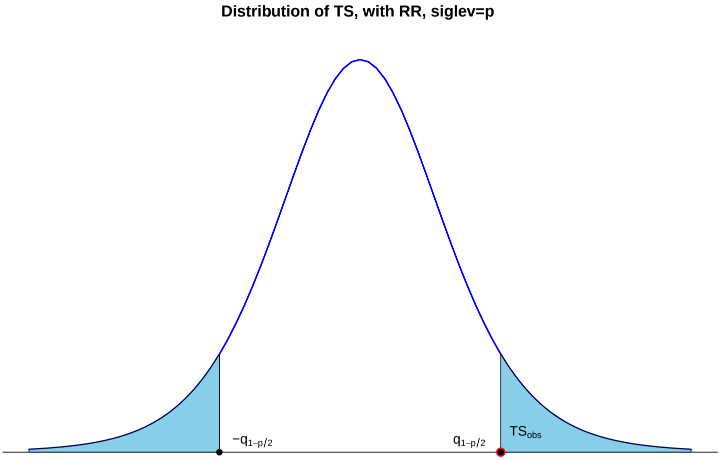
Now, we denote $p = \alpha^*$ and call it the *p*-**Value of the Test:**

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

p-Values



p-Values



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 unprobable, 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 unprobable, very unbelievable, under $\mathcal{H}_0$, so we safely Reject $\mathcal{H}_0$.

To Remember:

- If p-Value$< \alpha$, then we Reject $\mathcal{H}_0$
- If p-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]))
```

Example

Example: Assume we have an observation from $\mathcal{N}(\mu, 3^2)$, and we want to Test

$$\mathcal{H}_0: \ \mu = 1.2 \qquad \textit{vs} \qquad \mathcal{H}_1: \ \mu \neq 1.2.$$

Example

Example: Assume we have an observation from $\mathcal{N}(\mu, 3^2)$, and we want to Test

$$\mathcal{H}_0: \ \mu = 1.2 \qquad \textit{vs} \qquad \mathcal{H}_1: \ \mu \neq 1.2.$$

So we will do a Z-Test.

Example

Example: Assume we have an observation from $\mathcal{N}(\mu, 3^2)$, and we want to Test

$$\mathcal{H}_0: \ \mu = 1.2 \qquad vs \qquad \mathcal{H}_1: \ \mu \neq 1.2.$$

So we will do a Z-Test. Assume that after plugging our observations into the Z-Statistics, we get

$$Z_{obs} = 1.72.$$

Example

Example: Assume we have an observation from $\mathcal{N}(\mu, 3^2)$, and we want to Test

$$\mathcal{H}_0: \ \mu = 1.2 \qquad vs \qquad \mathcal{H}_1: \ \mu \neq 1.2.$$

So we will do a Z-Test. Assume that after plugging our observations into the Z-Statistics, we get

$$Z_{obs} = 1.72.$$

Then, the p-Value of the Test is

$$p\text{-Value} = \mathbb{P}(|Z| > Z_{obs} \mid \text{Null is True}) = \mathbb{P}\Big(|Z| > 1.72 \mid Z \sim \mathcal{N}(0,1)\Big)$$

Example

Example: Assume we have an observation from $\mathcal{N}(\mu, 3^2)$, and we want to Test

$$\mathcal{H}_0: \ \mu = 1.2 \qquad vs \qquad \mathcal{H}_1: \ \mu \neq 1.2.$$

So we will do a Z-Test. Assume that after plugging our observations into the Z-Statistics, we get

$$Z_{obs} = 1.72.$$

Then, the p-Value of the Test is

$$p\text{-Value} = \mathbb{P}(|Z| > Z_{obs} \mid \text{Null is True}) = \mathbb{P}\Big(|Z| > 1.72 \mid Z \sim \mathcal{N}(0,1)\Big)$$

and the value of p-Value is ⌣̈

```
1-(pnorm(1.72,mean=0,sd=1)-pnorm(-1.72,mean=0,sd=1))
```

```
## [1] 0.08543244
```

Goodness-of-Fit Tests

Intro to GoF Tests

Here, we have a DataSet $x_1, x_2, ..., x_n$, and a Statistical Model, and we want to see how good our Model is fitting the Data.

Intro to GoF Tests

Here, we have a DataSet $x_1, x_2, ..., x_n$, and a Statistical Model, and we want to see how good our Model is fitting the Data.

First we consider Pearson's χ^2-Test: a famous GoF Test for the Multinomial Distribution.

Goodness-of-Fit Tests: Pearson's χ^2 Test

Model: Here we assume that the result of an Experiment can be one of the $A_1, ..., A_m$ (different classes), with Probabilities

$$p_1 = \mathbb{P}(A_1), ..., p_m = \mathbb{P}(A_m).$$

Goodness-of-Fit Tests: Pearson's χ^2 Test

Model: Here we assume that the result of an Experiment can be one of the $A_1, ..., A_m$ (different classes), with Probabilities

$$p_1 = \mathbb{P}(A_1), ..., p_m = \mathbb{P}(A_m).$$

Data: We have the results of a repetition of the previous Experiment: The results are: the number of A_1 shown is X_1, the number of A_2 shown is X_2, . . . , the number of A_m shown is X_m;

Goodness-of-Fit Tests: Pearson's χ^2 Test

Model: Here we assume that the result of an Experiment can be one of the $A_1, ..., A_m$ (different classes), with Probabilities

$$p_1 = \mathbb{P}(A_1), ..., p_m = \mathbb{P}(A_m).$$

Data: We have the results of a repetition of the previous Experiment: The results are: the number of A_1 shown is X_1, the number of A_2 shown is X_2, . . . , the number of A_m shown is X_m;

Null Hypothesis:

$$\mathcal{H}_0 : \text{ The Actual Probabilities are } p_1, p_2, ..., p_m.$$

vs

$$\mathcal{H}_1 : \mathcal{H}_0 \text{ is not correct.}$$

Goodness-of-Fit Tests: Pearson's χ^2 Test

Model: Here we assume that the result of an Experiment can be one of the $A_1, ..., A_m$ (different classes), with Probabilities

$$p_1 = \mathbb{P}(A_1), ..., p_m = \mathbb{P}(A_m).$$

Data: We have the results of a repetition of the previous Experiment: The results are: the number of A_1 shown is X_1, the number of A_2 shown is X_2, . . . , the number of A_m shown is X_m;

Null Hypothesis:

$$\mathcal{H}_0: \text{ The Actual Probabilities are } p_1, p_2, ..., p_m.$$

vs

$$\mathcal{H}_1: \ \mathcal{H}_0 \text{ is not correct.}$$

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

Goodness-of-Fit Tests: Pearson's χ^2 Test

Test Statistics:

Goodness-of-Fit Tests: Pearson's χ^2 Test

Test Statistics: $\chi^2 = \sum_{k=1}^{m} \frac{(X_k - n \cdot p_k)^2}{n \cdot p_k}$

Goodness-of-Fit Tests: Pearson's χ^2 Test

Test Statistics: $\chi^2 = \sum_{k=1}^{m} \frac{(X_k - n \cdot p_k)^2}{n \cdot p_k} = \sum_{k=1}^{m} \frac{(O_k - E_k)^2}{E_k}$

Goodness-of-Fit Tests: Pearson's χ^2 Test

Test Statistics: $\chi^2 = \sum_{k=1}^{m} \frac{(X_k - n \cdot p_k)^2}{n \cdot p_k} = \sum_{k=1}^{m} \frac{(O_k - E_k)^2}{E_k}$

Here usually one constructs the following χ^2-Table:

	A_1	A_2	...	A_m
Observed Freq., O_k	X_1	X_2	...	X_m
Expected Freq., E_k	$n \cdot p_1$	$n \cdot p_2$	...	$n \cdot p_m$

Goodness-of-Fit Tests: Pearson's χ^2 Test

Test Statistics: $\chi^2 = \sum_{k=1}^{m} \frac{(X_k - n \cdot p_k)^2}{n \cdot p_k} = \sum_{k=1}^{m} \frac{(O_k - E_k)^2}{E_k}$

Here usually one constructs the following χ^2-Table:

	A_1	A_2	...	A_m
Observed Freq., O_k	X_1	X_2	...	X_m
Expected Freq., E_k	$n \cdot p_1$	$n \cdot p_2$	...	$n \cdot p_m$

Assumption: We assume that $n \cdot p_k \geq 5$ for any k;

Goodness-of-Fit Tests: Pearson's χ^2 Test

Test Statistics: $\chi^2 = \sum_{k=1}^{m} \frac{(X_k - n \cdot p_k)^2}{n \cdot p_k} = \sum_{k=1}^{m} \frac{(O_k - E_k)^2}{E_k}$

Here usually one constructs the following χ^2-Table:

	A_1	A_2	...	A_m
Observed Freq., O_k	X_1	X_2	...	X_m
Expected Freq., E_k	$n \cdot p_1$	$n \cdot p_2$	...	$n \cdot p_m$

Assumption: We assume that $n \cdot p_k \geq 5$ for any k;

Distrib of the Test-Statistics Under $\mathcal{H}_0$: $\chi^2 \approx \chi^2(m-1)$;

Goodness-of-Fit Tests: Pearson's χ^2 Test

Test Statistics: $\chi^2 = \sum_{k=1}^{m} \frac{(X_k - n \cdot p_k)^2}{n \cdot p_k} = \sum_{k=1}^{m} \frac{(O_k - E_k)^2}{E_k}$

Here usually one constructs the following χ^2-Table:

	A_1	A_2	...	A_m
Observed Freq., O_k	X_1	X_2	...	X_m
Expected Freq., E_k	$n \cdot p_1$	$n \cdot p_2$	...	$n \cdot p_m$

Assumption: We assume that $n \cdot p_k \geq 5$ for any k;

Distrib of the Test-Statistics Under $\mathcal{H}_0$: $\chi^2 \approx \chi^2(m-1)$;

Rejection Region: $\chi^2 > \chi^2_{m-1,1-\alpha}$

Example

Example: I am claiming that, for my Stat courses, the percentage of A-grade students is 15%, of B-grade students is 25%, of C-grades are 20%, for D I have 15%, and all others are Failing the course.

Example

Example: I am claiming that, for my Stat courses, the percentage of A-grade students is 15%, of B-grade students is 25%, of C-grades are 20%, for D I have 15%, and all others are Failing the course. Now, one of my courses finished with the following result:

$$\#A = 27,\ \#B = 22,\ \#C = 10,\ \#D = 10,\ \#F = 12.$$

Is this Data supporting my claim?

Solution:

Example

Example: I am claiming that, for my Stat courses, the percentage of A-grade students is 15%, of B-grade students is 25%, of C-grades are 20%, for D I have 15%, and all others are Failing the course. Now, one of my courses finished with the following result:

$$\#A = 27, \ \#B = 22, \ \#C = 10, \ \#D = 10, \ \#F = 12.$$

Is this Data supporting my claim?

Solution: We have $n = 27 + 22 + 10 + 10 + 12 = 81$. Next, we make the Table:

	A	B	C	D	E
Obs. Frq., O_k	27	22	10	10	12
Exp. Frq., E_k	$81 \cdot 0.15$	$81 \cdot 0.25$	$81 \cdot 0.2$	$81 \cdot 0.15$	$81 \cdot 0.25$

Example

Example: I am claiming that, for my Stat courses, the percentage of A-grade students is 15%, of B-grade students is 25%, of C-grades are 20%, for D I have 15%, and all others are Failing the course. Now, one of my courses finished with the following result:

$$\#A = 27,\ \#B = 22,\ \#C = 10,\ \#D = 10,\ \#F = 12.$$

Is this Data supporting my claim?

Solution: We have $n = 27 + 22 + 10 + 10 + 12 = 81$. Next, we make the Table:

	A	B	C	D	E
Obs. Frq., O_k	27	22	10	10	12
Exp. Frq., E_k	$81 \cdot 0.15$	$81 \cdot 0.25$	$81 \cdot 0.2$	$81 \cdot 0.15$	$81 \cdot 0.25$

Now, we can calculate the TS:

$$\chi^2 = \sum_{k=1}^{5} \frac{(O_k - E_k)^2}{E_k} = \frac{(27 - 81 \cdot 0.15)^2}{81 \cdot 0.15} + \ldots + \frac{(12 - 81 \cdot 0.15)^2}{81 \cdot 0.15}$$

Example, Cont'd

The rest is in **R**:

```
obsd <-  c(27, 22, 10, 10, 12)
expd <- 81* c(0.15, 0.25, 0.2, 0.15, 0.25)
xi2 <- sum((obsd-expd)^2/expd)
xi2
```

```
## [1] 24.41564
```

```
q <- qchisq(1-0.05, df = length(obsd)-1)
q
```

```
## [1] 9.487729
```

```
xi2 > q
```

```
## [1] TRUE
```

Example, Cont'd

```r
obsd <-  c(27, 22, 10, 10, 12)
p <-  c(0.15, 0.25, 0.2, 0.15, 0.25)
chisq.test(obsd, p = p)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  obsd
## X-squared = 24.416, df = 4, p-value = 6.592e-05
```

Kolmogorov-Smirnov Test

```
x <- rnorm(50, mean = 3, sd = 1)
ks.test(x, y = "pnorm", mean = 0, sd = 1)
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  x
## D = 0.87562, p-value = 8.882e-16
## alternative hypothesis: two-sided
```

Kolmogorov-Smirnov Test

```r
x <- rnorm(50, mean = 3, sd = 1)
ks.test(x, y = "pnorm", mean = 0, sd = 1)
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  x
## D = 0.87562, p-value = 8.882e-16
## alternative hypothesis: two-sided
```

```r
x <- rexp(50, rate = 3.1)
ks.test(x, y = "pnorm", mean = 0, sd = 1)
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  x
## D = 0.50228, p-value = 3.382e-12
## alternative hypothesis: two-sided
```

Example

```r
x <- runif(40)
y <- rexp(30)
ks.test(x,y)
```

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  x and y
## D = 0.31667, p-value = 0.05272
## alternative hypothesis: two-sided
```

Some Important Tests

Here are some important Tests in **R** to know:

Some Important Tests

Here are some important Tests in **R** to know:

- **Shapiro-Wilk Normality Test:** Tests if a vector `x` is coming from a Normal Distribution;

  usage: `shapiro.test(x)`

Some Important Tests

Here are some important Tests in **R** to know:

- **Shapiro-Wilk Normality Test:** Tests if a vector `x` is coming from a Normal Distribution;

  usage: `shapiro.test(x)`

- **One Sample Kolmogorov-Smirnov Test:** Tests if a vector `x` is coming from a given continuous Distribution,

  usage: `ks.test(x, y = "CDF Name", CDF Parameters)`

Some Important Tests

Here are some important Tests in **R** to know:

- **Shapiro-Wilk Normality Test:** Tests if a vector `x` is coming from a Normal Distribution;

  usage: `shapiro.test(x)`

- **One Sample Kolmogorov-Smirnov Test:** Tests if a vector `x` is coming from a given continuous Distribution,

  usage: `ks.test(x, y = "CDF Name", CDF Parameters)`

- **Two Sample Kolmogorov-Smirnov Test:** Tests if two vectors `x` and `y` are coming from the same Distribution

  usage: `ks.test(x,y)`

Some Important Tests

Here are some important Tests in **R** to know:

- **Shapiro-Wilk Normality Test:** Tests if a vector `x` is coming from a Normal Distribution;

  usage: `shapiro.test(x)`

- **One Sample Kolmogorov-Smirnov Test:** Tests if a vector `x` is coming from a given continuous Distribution,

  usage: `ks.test(x, y = "CDF Name", CDF Parameters)`

- **Two Sample Kolmogorov-Smirnov Test:** Tests if two vectors `x` and `y` are coming from the same Distribution

  usage: `ks.test(x,y)`

- **F-Test of equality of variances:** Given 2 Samples `x, y` from Normal Distributions, it tests if the variances are equal.

  usage: `var.test(x,y)`

Some Important Tests

- **Runs Test:** given a vector `x` performs a test if that vector is random

  usage: `DescTools::RunsTest(x)`

Some Important Tests

- **Runs Test:** given a vector `x` performs a test if that vector is random

  usage: `DescTools::RunsTest(x)`

- **Bartel's Test of Randomness:** Given a vector `x` it tests if `x` is random (versus there is a trend or systematic oscillations)

  usage: `DescTools::BartelsRankTest(x)`

Some Important Tests

- **Runs Test:** given a vector `x` performs a test if that vector is random

  usage: `DescTools::RunsTest(x)`

- **Bartel's Test of Randomness:** Given a vector `x` it tests if `x` is random (versus there is a trend or systematic oscillations)

  usage: `DescTools::BartelsRankTest(x)`

- **Two Sample t-Test:** Given two vectors `x` and `y` coming from Normal Distributions, it tests whether their means are equal

  usage: `t.test(x,y)`

Some Important Tests

- **Runs Test:** given a vector `x` performs a test if that vector is random

  usage: `DescTools::RunsTest(x)`

- **Bartel's Test of Randomness:** Given a vector `x` it tests if `x` is random (versus there is a trend or systematic oscillations)

  usage: `DescTools::BartelsRankTest(x)`

- **Two Sample *t*-Test:** Given two vectors `x` and `y` coming from Normal Distributions, it tests whether their means are equal

  usage: `t.test(x,y)`

- **One Sample Proportions Test:** Given a Bernoulli Experiment, it tests if the probability of success is the given value

  usage: `prop.test(no.of.successes, no.of.trials, p)` or `binom.test(no.of.successes, no.of.trials, p)`

Some Important Tests

- **Two Sample Proportions Test:** Given two Bernoulli Experiments, it tests if the probabilities of successes are equal to each other

  usage: `prop.test(vec.of.no.of.successes, vec.of.no.of.trials)`

Some Important Tests

- **Two Sample Proportions Test:** Given two Bernoulli Experiments, it tests if the probabilities of successes are equal to each other

  usage: `prop.test(vec.of.no.of.successes, vec.of.no.of.trials)`

- **Pearson's χ^2 Goodness of Fit Test:** Given a Multinomial Experiment, it Tests if the probabilities are given ones

  usage: `chisq.test(vec.of.obs, vec.of.probs)`

Some Important Tests

- **Two Sample Proportions Test:** Given two Bernoulli Experiments, it tests if the probabilities of successes are equal to each other

  usage: `prop.test(vec.of.no.of.successes, vec.of.no.of.trials)`

- **Pearson's χ^2 Goodness of Fit Test:** Given a Multinomial Experiment, it Tests if the probabilities are given ones

  usage: `chisq.test(vec.of.obs, vec.of.probs)`

- **Pearson's χ^2 Independence Test:** Given 2 count vectors `x` and `y`, it Tests if the variables `x` and `y` are independent

  usage: `chisq.test(x, y)`

Some Important Tests

- **Two Sample Proportions Test:** Given two Bernoulli Experiments, it tests if the probabilities of successes are equal to each other

  usage: `prop.test(vec.of.no.of.successes, vec.of.no.of.trials)`

- **Pearson's χ^2 Goodness of Fit Test:** Given a Multinomial Experiment, it Tests if the probabilities are given ones

  usage: `chisq.test(vec.of.obs, vec.of.probs)`

- **Pearson's χ^2 Independence Test:** Given 2 count vectors `x` and `y`, it Tests if the variables `x` and `y` are independent

  usage: `chisq.test(x, y)`

- **Location Tests:** See, for example, Here