

**Goodness-Of-Fit**

**&**

**$\chi^2$ GOF-Test**

What is  $\chi^2$ -Test?

---

A  $\chi^2$ -Test is a statistical method used to analyze categorical data to determine if observed frequencies are significantly different from expected frequencies.

What is the purpose of  $\chi^2$ -Test ?

---

It helps determining if differences between observed and expected data are due to chance or a significant relationship.

---

What are the types of  $\chi^2$ -Test?

---

There are two main types:

① **Chi-Square Test of Independence:**

Determines if categorical variables are dependent or independent.

② **Chi-Square Goodness-Of-Fit Test:**

Determines if a sample distribution fits a known distribution.

---

## How do we perform the **Goodness-Of-Fit Test**?

Here are the six steps to perform this test:

- 1 Set up  $H_0$  and  $H_1$ .

$H_0$  : Observed Frequencies = Expected Frequencies

$H_1$  : Observed Frequencies  $\neq$  Expected Frequencies, RTT

- 2 Set up the **Test Table** and find the **Computed Test Statistic** using the formula below.

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

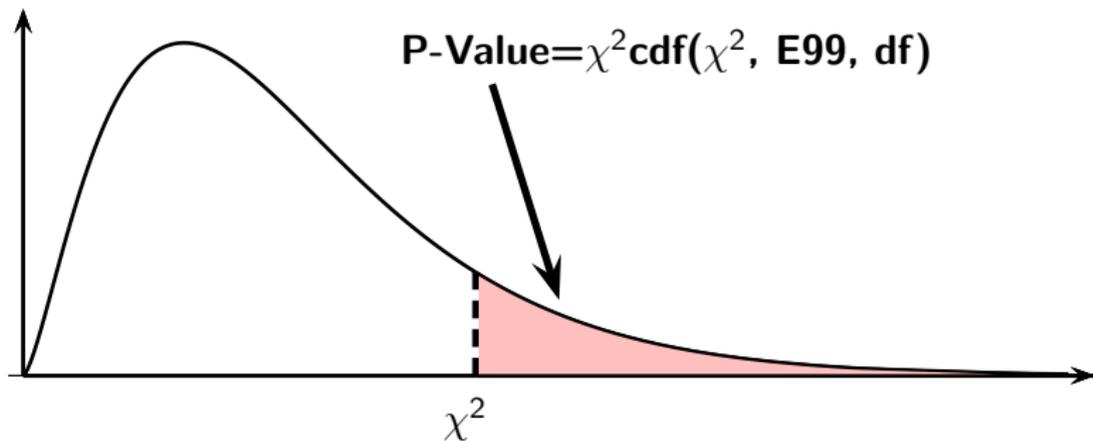
where O is the observed value and E is the expected frequency for each category.

## How do we perform the **Goodness-Of-Fit Test**?

- ③ Find the **P-value** using the degrees of freedom below.

$$df = k - 1$$

where  $k$  is the number of categories.



## How do we perform the **Goodness-Of-Fit Test**?

---

- ④ Determine the validity of  $H_0$  and  $H_1$ .

If  $P$ -Value  $> \alpha \implies H_0$  valid &  $H_1$  invalid

If  $P$ -Value  $\leq \alpha \implies H_0$  invalid &  $H_1$  valid

- ⑤ Draw the final conclusion.

If  $H_0$  valid  $\implies$  Observed Frequencies = Expected Frequencies

If  $H_1$  valid  $\implies$  Observed Frequencies  $\neq$  Expected Frequencies

---

## What is the **Test Table**?

A **Test Table** is intended to compare observed frequencies  $O$  with expected frequencies  $E$  across categories that leads to calculating the **Computed Test Statistic**.

It typically includes columns for Category, Observed Values  $O$ , Expected Values  $E$ , and  $\frac{(O - E)^2}{E}$  for calculation process for the **Computed Test Statistic**  $\chi^2$ .

Category	$O$	$E$	$\frac{(O - E)^2}{E}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
Total	$\sum O$	$\sum E$	$\chi^2 = \sum \frac{(O - E)^2}{E}$

*Example:*

Suppose a survey of 100 students conducted on which political issue concerns them the most right now. Here is the result of the survey while it was expected equal frequency for all categories.

Category	Observed $O$	Expected $E$
Climate Change	15	25
Education Cost	45	25
Healthcare	30	25
Immigration	10	25

Make a complete **Test Table**.

---

Solution:

Here is the complete **Test Table**.

Category	$O$	$E$	$\frac{(O - E)^2}{E}$
Climate Change	15	25	$\frac{(15 - 25)^2}{25} = 4$
Education Cost	45	25	$\frac{(45 - 25)^2}{25} = 16$
Healthcare	30	25	$\frac{(30 - 25)^2}{25} = 1$
Immigration	10	25	$\frac{(10 - 25)^2}{25} = 9$
Total	100	100	$\chi^2 = 30$

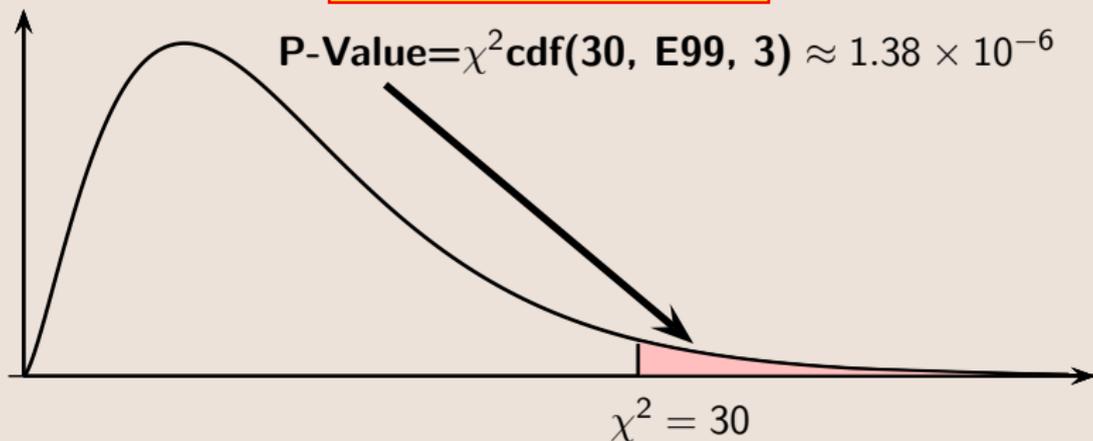
*Example:*

Find the corresponding **P-Value** for our computed test statistic  $\chi^2 = 30$ .

**Solution:**

Since we have  $k = 4$  categories, then the degrees of freedom is

$$df = k - 1 = 4 - 1 = 3$$



$\chi^2$ GOF-Test & TI

First we store the **Observed Values** in  $L1$ , and the **Expected Frequencies** in  $L2$ . Press **STAT** followed by option **1** from the **EDIT** menu, then store the values in  $L1$  and  $L2$ .

```

2001 CALC TESTS
1:Edit...
2:SortA(
3:SortD(
4:ClrList
5:SetUpEditor

```

L1	L2	L3	2
15	25	-----	
45	25		
30	25		
10	25		
-----	-----		
L2(5) =			

$\chi^2$ GOF-Test & TI

Now press **STAT** followed by **→** to get to the **TESTS** menu, then use the **↓** to go to  **$\chi^2$ GOF-Test** option, and press **ENTER** to select it.

```
EDIT CALC TESTS
B1:2-PropZInt...
C:  $\chi^2$ -Test...
D:  $\chi^2$ GOF-Test...
E: 2-SampFTTest...
F: LinRegTTest...
G: LinRegTInt...
H: ANOVA(
```

$\chi^2$ GOF-Test & TI

Make sure you follow the steps on the screen and update the degrees of freedom, then press **Calculate** to arrive at the result screen.

```
χ²GOF-Test
Observed:L1
Expected:L2
df:3
Calculate Draw
```

```
χ²GOF-Test
χ²=30
P=1.380057E-6
df=3
CNTRB={4 16 1 ...
█
```

**CNTRB** stands for Contribution in a  $\chi^2$ GOF-Test.

What is the purpose of **CNTRB** ?

It represents how much each individual category contributes to the  $\chi^2$  test statistic.

Higher CNTRB values indicate categories that differ most from expected values.

By pressing  you can view all values of **CNTRB**.

```
χ²GOF-Test  
χ²=30  
P=1.380057E-6  
df=3  
CNTRB=(4 16 1 ...  
█
```

```
χ²GOF-Test  
χ²=30  
P=1.380057E-6  
df=3  
CNTRB=... 16 1 9)  
█
```

*Example:*

Suppose a survey of 120 students conducted on how students study for exams in a math class. Here is the result of the survey.

Category	Observed $O$
Group Study	44
Review Notes	18
Review Textbook	17
Practice Problems	21
Making Flashcards	20

Use the significance level 0.02 to test the claim that it is twice as likely for students to join **Group Study** as any other methods to prepare for the exam in a math class.

---

## Solution:

Let  $P$  be the probability of any category except the Group Study. Since the Group Study is twice as likely as other categories, then its probability is  $2P$ .

Category	Probability
Group Study	$2P$
Review Notes	$P$
Review Textbook	$P$
Practice Problems	$P$
Making Flashcards	$P$

We know that the sum of all probabilities is **Always** equal to 1.

$$2P + P + P + P + P = 1$$

$$6P = 1$$

$$P = \frac{1}{6}$$

## Solution Continued:

So the probability of any category except the Group Study is  $\frac{1}{6}$ .  
while the probability of the Group Study is  $2p = 2 \cdot \frac{1}{6} = \frac{1}{3}$ .

Category	Probability
Group Study	1/3
Review Notes	1/6
Review Textbook	1/6
Practice Problems	1/6
Making Flashcards	1/6

$H_0 : P_1 = 1/3, P_2 = P_3 = P_4 = P_5 = 1/6$  Claim

$H_1 : At least one of these proportions is different. RTT$

## Solution Continued:

Now we can find the expected values for each category for the 120 randomly selected students.

Category	Probability	Expected Values
Group Study	$\frac{1}{3}$	$120 \cdot \frac{1}{3} = 40$
Review Notes	$\frac{1}{6}$	$120 \cdot \frac{1}{6} = 20$
Review Textbook	$\frac{1}{6}$	$120 \cdot \frac{1}{6} = 20$
Practice Problems	$\frac{1}{6}$	$120 \cdot \frac{1}{6} = 20$
Making Flashcards	$\frac{1}{6}$	$120 \cdot \frac{1}{6} = 20$

## Solution Continued:

Now we have the expected frequencies for each category for 120 randomly selected students.

Category	Observed $O$	Expected $E$
Group Study	44	40
Review Notes	18	20
Review Textbook	17	20
Practice Problems	21	20
Making Flashcards	20	20

## Solution Continued:

Now we store the **Observed Values** in  $L1$ , and the **Expected Frequencies** in  $L2$ . Press **STAT** followed by option **1** from the **EDIT** menu, then store the values in  $L1$  and  $L2$ .

```

EDIT  CALC TESTS
1:Edit...
2:SortA(
3:SortD(
4:ClrList
5:SetUPEditor
  
```

L1	L2	L3	2
44	40	-----	
18	20		
17	20		
21	20		
20	20		
-----	-----		
L2(6) =			

## Solution Continued:

Now press **STAT** followed by  $\rightarrow$  to get to the **TESTS** menu, then use the  $\downarrow$  to go to  **$\chi^2$ GOF-Test** option, and press **ENTER** to select it.

```
 $\chi^2$ GOF-Test  
Observed: L1  
Expected: L2  
df: 4  
Calculate Draw
```

Since we have 5 categories, then the degrees of freedom is

$$df = k - 1 = 5 - 1 = 4$$

Make sure to update the degrees of freedom before you press

**Calculate**

## Solution Continued:

Here is the result screen.

```

      χ²GOF-Test
χ²=1.1
P=.8942722061
df=4
CNTRB={.4 .2 ...
█

```

```

      χ²GOF-Test
χ²=1.1
P=.8942722061
df=4
CNTRB={...5 .05 0}

```

Since  $p\text{-value}=0.894$  is greater than the given significance level  $\alpha = 0.02$ , therefore  $H_0$  is valid and  $H_1$  is invalid.

We have a valid claim, therefore we **Fail-To-Reject** the claim.

---

## Solution Continued:

If we wish to find the computed test statistic manually, we complete **Test Table**.

Category	$O$	$E$	$\frac{(O - E)^2}{E}$
Group Study	44	40	$\frac{(44 - 40)^2}{40} = 0.4$
Review Notes	18	20	$\frac{(18 - 20)^2}{20} = 0.2$
Review Textbook	17	20	$\frac{(17 - 20)^2}{20} = 0.45$
Practice Problems	21	20	$\frac{(21 - 20)^2}{20} = 0.05$
Making Flashcards	20	20	$\frac{(20 - 20)^2}{20} = 0$
Total	120	120	$\chi^2 = 1.1$

## Solution Continued:

And finally to find the corresponding  **$p$ -value** with the degrees of freedom 4, we draw the  **$\chi^2$  curve**, clearly label, and shade.

Use the TI command to find the  **$p$ -value**.

