

Using the Chi-Square Test to Test for Independence

When you want to test whether variables within a sample are independent or related to each other.

e.g. Of all the students planning to go to a four year college, are males and females different in their preference to go to a “local” school, a distant school, or have no preference?

It could turn out that males and females have the same preference which would indicate they are related, or have different preferences which would indicate they are independent.

Use the same 5 step process with one alteration.

$$\text{Degrees of Freedom} = (\#Rows - 1)(\#Columns - 1)$$

E.g. Suppose you invent a radical new surgical procedure at your hospital. You want to know if doctors and nurses feel the same about the procedure or differently. What you will be testing is whether opinion is dependent on profession or not.

So you take a survey of 400 doctors and nurses.

<u>Group</u>	<u>pref. New</u>	<u>Pref. Old</u>	<u>No pref.</u>
Nurses	100	80	20
Doctors	50	120	20

Now, using $\alpha = 0.05$, do the chi-square test for independence.

1. H_0 : The opinion about the surgery is **independent** of profession
 H_1 : The opinion about the surgery is **dependent** on profession (claim)

2. $\alpha = 0.05$, d.f. = $(2 - 1)(3 - 1) = 2$,
from table: cv = 5.991

Now we need to calculate the expected values.
Start with the data table.

Sum the rows and columns

<u>Group</u>	<u>pref. New</u>	<u>pref. Old</u>	<u>No pref.</u>	
Nurses	100	75	80	100
Doctors	50	75	120	100
	150		200	50
				n=400

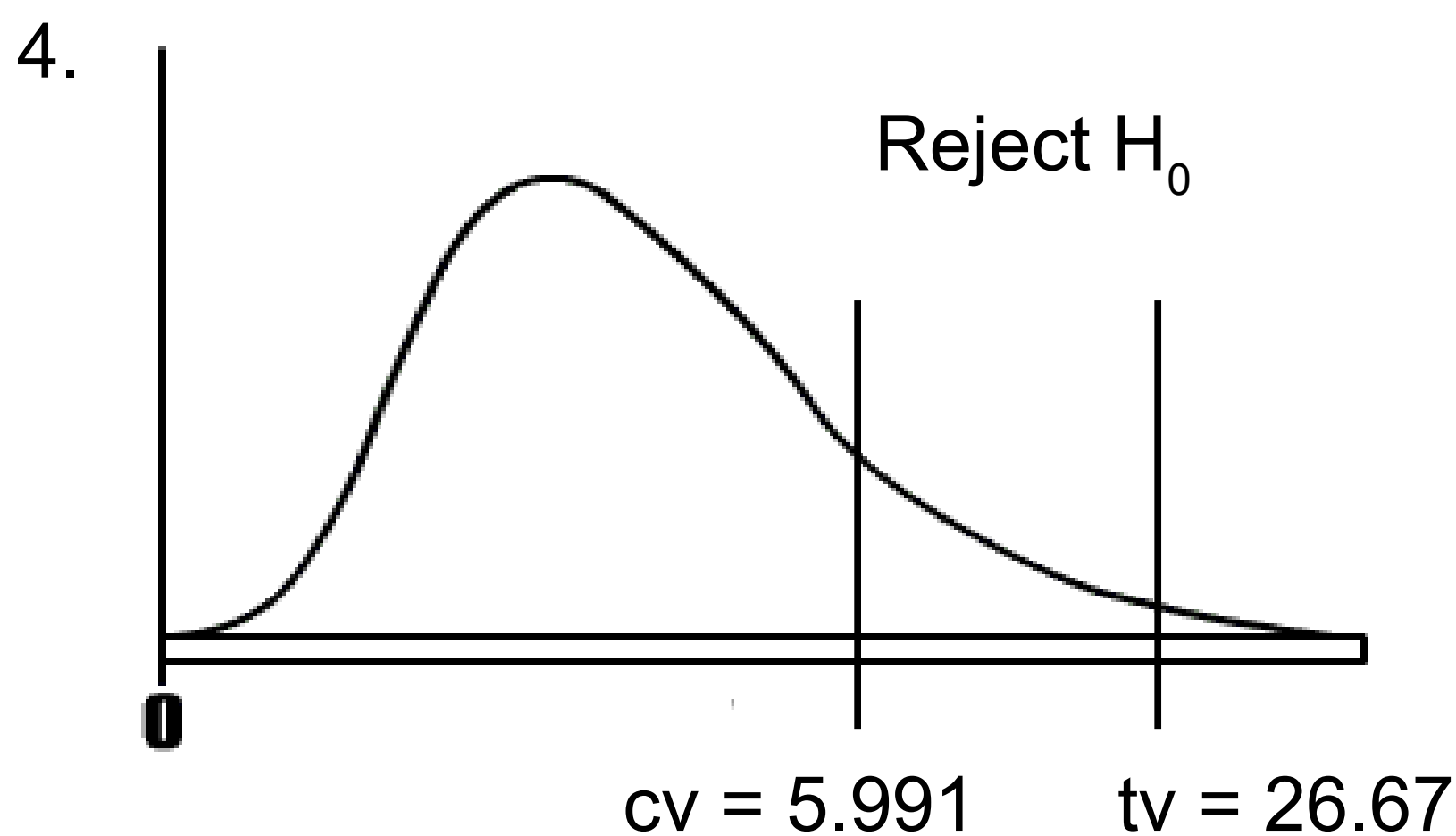
Now calculate the expected value for each cell.

$$\frac{(\text{Row Sum})(\text{Col Sum})}{n}$$

3. Calculate the test value.

$$tv = \frac{(100-75)^2}{75} + \frac{(80-100)^2}{100} + \frac{(20-25)^2}{25} + \frac{(50-75)^2}{75} + \frac{(120-100)^2}{100} + \frac{(30-25)^2}{25}$$

$$tv = 26.67$$



5. *There is enough evidence to reject H_0 at $\alpha=0.05$, so there is enough evidence to support the claim that opinion is related to profession. Nurse and Doctors differ in their opinion about the new procedure.*

Your turn.

Let us find out if there's a relationship between your gender and GPA in our Prob & Stats class.

Gender	GPA		
	<3.0	3.0 to 3.5	>3.5
Male	7	13	4
Female	3	11	6

What is the first step?

1. H_0 : GPA is independent of the student's gender.
 H_1 : GPA is dependent on the student's gender.

Is there a relationship between your gender and GPA in our Prob & Stats class.

Gender	GPA		
	<3.0	3.0 to 3.5	>3.5
Male	7	13	4
Female	3	11	6

1. H_0 : GPA is independent of the student's gender.
 H_1 : GPA is dependent on the student's gender.

What is the second step?

2. Critical Value

d.f. = $(2 - 1)(3 - 1) = 2$ and let's use $\alpha = 0.10$

From table you should get $cv = 4.605$

Your turn. Is there a relationship between your gender and GPA in our Prob & Stats class.

GPA

Gender	<3.0	3.0 to 3.5	>3.5
Male	7	13	4
Female	3	11	6
	10	24	10
			n=44

1. H_0 : GPA is independent of the student's gender.

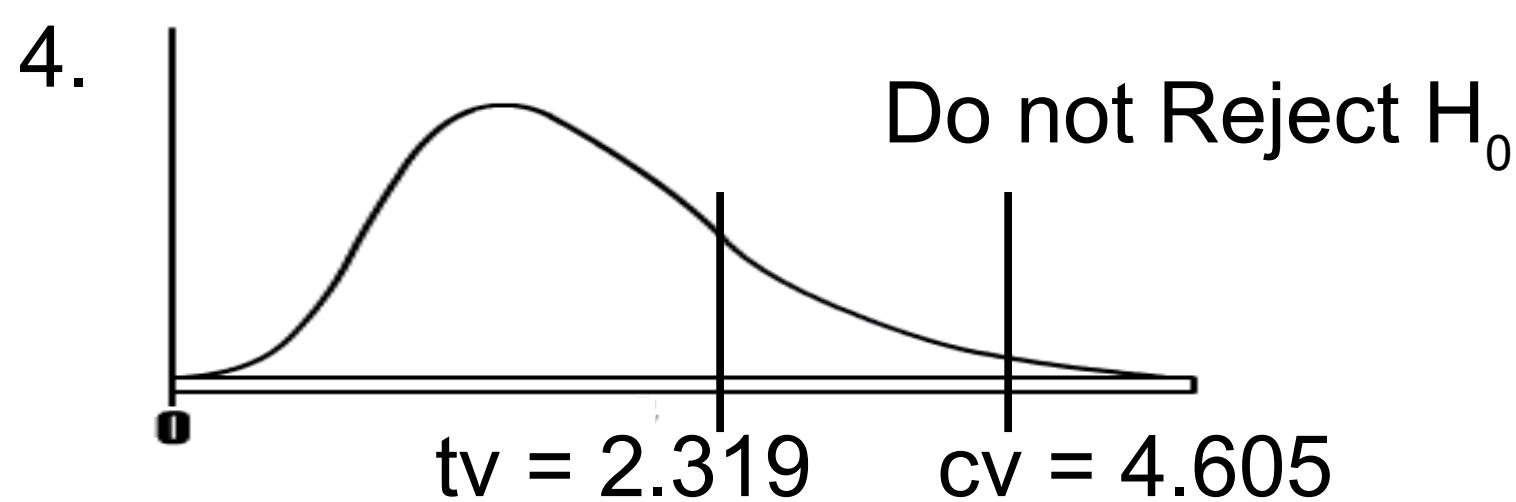
H_1 : GPA is dependent on the student's gender.

2. $cv = 4.605$

What is the third step?

3. Calculate the test value. Add **rows & columns** and calculate **expected values** first.

$tv = 2.319$



5. Depending on your claim....

If H_0 is the claim: *There is not enough evidence to reject H_0 at $\alpha=0.10$, so there is not enough evidence to reject the claim that GPA is independent of Gender.*

If H_1 is the claim: *There is not enough evidence to reject H_0 at $\alpha=0.10$, so there is not enough evidence to support the claim that GPA is dependent Gender.*