Categorical vs Numerical Data

Data Type	Question Type	Possible Responses
Categorical	What is your sex?	male or female
Numerical	Disrete- How many cars do you own?	two or three
Numerical	Continuous - How tall are you?	72 inches

Notice that discrete data arise fom a counting process, while continuous data arise from a measuring process.

CHI SQUARE OF INDEPENDENCE - 2 X 2 CONTINGENCY TABLE

Variable 2	Data type 1	Data type 2	Totals
Category 1	a	b	a + b
Category 2	c	d	c + d
Total	a + c	b + d	a + b + c + d = N

CHI SQUARE OF INDEPENDENCE - 3 X 3 CONTINGENCY TABLE

	Category I	Category II	Category III	Row Totals
Sample A	a	b	c	a+b+c
Sample B	d	e	f	d+e+f
Sample C	g	h	i	g+h+i
Column Totals	a+d+g	b+e+h	c+f+i	a+b+c+d+e+f+g+h+i=N

Ho: The two categorical variables are independent.

Ha: The two categorical variables are related.

	Blue	Red	Yellow
Extroverted	5	20	5
Introverted	10	5	5

Observed Frequencies

Is there a relationship between personality type and color preference? Our hypotheses will state exactly that, with the null as usual stating that there is no effect or no relationship.

H_1: There is a relationship between color preference and personality type (variables are not independent).

H_0: There is no relationship between color preference and personality type (variables are independent).

Since we have two variables, our degrees of freedom will change.

df = (R - 1) (C- 1) ← where R is the number of rows and C is the number of columns in our table. There are two rows going across and three rows going down. So, degrees of freedom for this example are:

df = (2 - 1) (3- 1) = (1)(2) = 2