

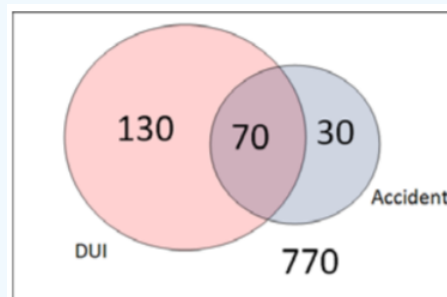
## 5.7: Contingency (Two-way) Tables

Contingency Tables, also known as cross tabulations, crosstabs or two-way tables, is a method of displaying the counts of the responses of two categorical variables from data.

### Example: Accidents and DUI

1000 drivers were asked if they were involved in an accident in the last year. They were also asked if during this time, they were DUI, driving under the influence of alcohol or drugs. The totals are summarized in a contingency table:

	Accident	No Accident	Total
DUI	70	130	200
Non-DUI	30	770	800
Total	100	900	1000



### Solution

In the table, each column represents a choice for the accident question and each row represents a choice for the DUI question.

Marginal Probabilities can be determined from the contingency table by using the outside total values for each event divided by the total sample size.

- Probability a driver had an accident =  $P(A) = 100/1000 = 0.10$
- Probability a driver was not DUI =  $P(D') = 1 - P(D) = 1 - 200/1000 = 0.80$

Joint Probabilities can be determined from the contingency table by using the inside values of the table divided by the total sample size.

- Probability a driver had an accident **and** was DUI =  $P(A \text{ and } D) = 70/1000 = 0.07$
- Probability a driver had an accident **or** was DUI =  $P(A \text{ or } D) = (100+200-70)/1000 = 0.23$

Conditional Probabilities can be determined from the contingency table by using the inside values of the table divided by the outside total value of the conditional event.

- Probability a driver was DUI **given** the driver had an accident =  $P(D|A) = 70/100 = 0.70$
- Probability a DUI driver had an accident =  $P(A|D) = 70/200 = 0.35$

### Creating a two-table from reported probabilities

We can create a hypothetical two-way table from reported cross tabulated probabilities, such as the CNN exit poll for the 2016 presidential election:

gender			
	clinton	trump	other/no answer
male 47%	41%	52%	7%
female 53%	54%	41%	5%

national map			
popular vote			
candidate	%		votes
● trump	46.1%	<div></div>	62,979,879
● clinton	48.2%	<div></div>	65,844,954

Step 1: Choose a convenient total number. (This is called the radix of the table).

GENDER			
VOTED FOR	Female	Male	Total
Trump			
Clinton			
Other			
Total			10000

Radix chosen = 10000 random voters

Step 2: Determine the outside values of the table by multiplying the radix times the marginal probabilities for gender.

GENDER			
VOTED FOR	Female	Male	Total
Trump			
Clinton			
Other			
Total	5300	4700	10000

Total Female =  $(0.53)(10000) = 5300$

Total Male =  $(0.47)(10000) = 4700$

Step 3: Determine the inside values of the table by multiplying the appropriate gender total times the conditional probabilities from the exit polls.

GENDER			
VOTED FOR	Female	Male	Total
Trump	2173	2444	
Clinton	2862	1927	
Other	265	329	
Total	5300	4700	10000

$$\text{Trump Female} = (0.41)(5300) = 2173$$

$$\text{Clinton Female} = (0.54)(5300) = 2862$$

$$\text{Other Female} = (0.05)(5300) = 265$$

$$\text{Trump Male} = (0.52)(4700) = 2444$$

$$\text{Clinton Male} = (0.41)(4700) = 1927$$

$$\text{Other Male} = (0.057)(4700) = 329$$

Step 4: Add each row to get the row totals.

GENDER			
VOTED FOR	Female	Male	Total
Trump	2173	2444	4617
Clinton	2862	1927	4789
Other	265	329	594
Total	5300	4700	10000

$$\text{Trump} = 2173 + 2444 = 4617$$

$$\text{Clinton} = 2862 + 1927 = 4789$$

$$\text{Other} = 265 + 329 = 594$$

From the last column, we can now get the marginal probabilities (which are slightly off from the actual vote due to rounding in the exit polls): Donald Trump received 46%, Hillary Clinton received 48% and other candidates received 6% of the total vote.

---

This page titled [5.7: Contingency \(Two-way\) Tables](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Maurice A. Geraghty](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.