

Section 4.4

Contingency Tables and Association

Objectives

- Compute the marginal distribution of a variable
- Use the conditional distribution to identify association among categorical data

Titanic - statistics

	Men	Women	Boys	Girls
Survived	334	318	29	27
Died	1360	104	35	18

The table is referred to as a **contingency table**, or **two-way table**, because it relates two categories of data

A **marginal distribution** of a variable is a frequency or relative frequency distribution of either the row or column variable in the contingency table.

		Class				Total	
		First	Second	Third	Crew		
Survival	Alive	203	118	178	212	711	.32
	Dead	122	167	528	673	1490	.68
	Total	325	285	706	885	2201	

$\frac{711}{2201}$ $\frac{325}{2201}$

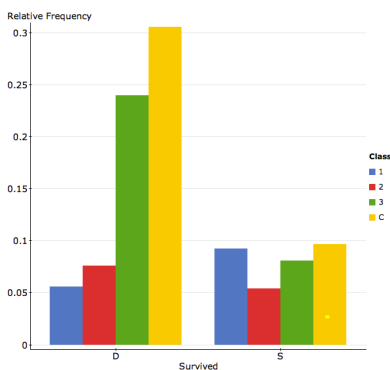
Construct the relative frequency marginal distribution.

A **conditional distribution** lists the relative frequency of each category of the response variable, given a specific value of the explanatory variable in the contingency table.

		Class				Total
		First	Second	Third	Crew	
Survival	Alive	203	118	178	212	711
	Dead	122	167	528	673	1490
	Total	325	285	706	885	2201

$\frac{203}{325}$ $\frac{528}{706}$

Construct the conditional distribution of Survival by Class. *Given*

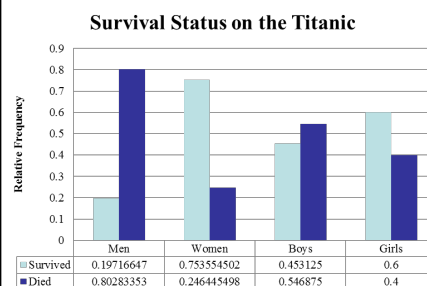


The following contingency table shows the survival status and demographics of passengers on the ill-fated Titanic.

	Men	Women	Boys	Girls
Survived	334	318	29	27
Died	1360	104	35	18

$\frac{334}{1694}$ $\frac{104}{422}$

Construct a conditional distribution of Survival by gender. *Given*



		Class				Total
		First	Second	Third	Crew	
Survival	Alive	203	118	178	212	711
	Dead	122	167	528	673	1490
	Total	325	285	706	885	2201

Find the proportion of crew who died on the titanic.

$$\frac{673}{885} = .76$$

Among those that survived what proportion were from first class.

$$\frac{203}{711} = .29$$

Was class a determine factor in whether a person survived or not?

yes