FST 6-5 Notes
Topic: Contingency Tables
<u>GOAL</u>
This lesson shows how to compute relative frequencies and probabilities from contingency tables.

SPUR Objectives

H Use a contingency table to compute percentages involving categorical variables.

M Represent information about relative frequencies or frequencies in a contingency table.

Vocabulary

contingency tables

Simpson's Paradox

Contingency tables – are tables that divide outcomes among two or more categorical variables.

Warm-up
Willie Fielder hurt himself two games in the 2047 baseball season and only batted 6 times with 1 hit, for a batting average of 0.167. Scott ("Scruffy") Scrub played the entire season but was a second-stringer, so was up only 100 times and got 19 hits, for a batting average of 0.190, better than Willie's average. In the 2048 season, Willie was well and got 201 hits in 600 at-bats. Scott still remained a second-stringer and was up only 100 times again, but got 35 hits.

a. What was Willie's batting average for the 2048 season?

$$\frac{201}{600} = 0.335$$

b. What was Scott's batting average for the 2048 season?

$$\frac{35}{100} = 0.350$$

c. What was Willie's combined batting average for the 2047 and 2048 seasons?

$$\frac{1}{6} + \frac{201}{600} = \frac{202}{606} = 0.333$$

d. What was Scott's combined batting average for the 2047 and 2048 seasons?

$$\frac{19}{100} + \frac{35}{100} = \frac{54}{200} = 0.270$$

Player	2047		2048	
	At Bats	Hits	At Bats	Hits
Fielder	6	1	600	201
Scrub	100	19	100	35

Titanic Table 1 below lists the number of passengers and crew who survived and died (the possible outcomes) in the sinking of the Titanic, categorized by status (first-class, second-class, third-class, and crew).

Titanic Table 1: Status and Survival

	First	Second	Third	Crew	Totals
Survived	203	118	178	212	711
Died	122	167	528	673	1490
Totals	325	285	706	885	2201

Source: British Wreck Commissioner's Inquiry Report

Example 1: Use the table above.

a. Out of all the people on the ship, what percent died?

$$\frac{1490}{2201} = 67.7\%$$

b. What percent of passengers in third class died?

$$\frac{528}{706} = 74.8\%$$

c. What percent of passengers in first or second class died?

$$\frac{122 + 167}{325 + 285} = \frac{289}{610} = 47.4\%$$

Example 2:

A 2001 study by the University of Texas Southwestern Medical Center examined 626 patients to see if there was a connection between getting a tattoo and infection with Hepatitis C (HCV). The results are in the contingency table below.

	Tattoo Done in Commercial Tattoo Parlor	Tattoo Done Elsewhere	No Tattoo	Totals
Has Hepatitis C	17	8	18	43
No Hepatitis C	35	53	495	583
Totals	52	61	513	626

Source: Haley RW, Fischer RP in Medicine, March 2001

a. What percent of people in the study did not have a tattoo?

$$513/626 = 81.9\%$$

b. What percent of people in the study with no tattoo had Hepatitis C?

$$18/513 = 3.5\%$$

c. What percent of people in the study with a tattoo had Hepatitis C?

$$\frac{17 + 8}{52 + 61} = \frac{25}{113} = 22.1\%$$

d. What can you conclude from answers to Parts b and c?

People w/ tattoo about 6 times as likely to have Hep C.

Example 3:

Fifth-grade students in a school were surveyed about their favorite book series.
The results are reported in the contingency table below.

	Harry Potter	Animorphs	Lemony Snicket	Lord of the Rings	Other
Boys	42%	15%	23%	11%	9%
Girls	51%	8%	28%	5%	8%

(handwritten: 125 Boys, 117 Girls)

Suppose the 5th-grade class contains 117 girls and 125 boys. Did more boys than girls prefer Lemony Snicket?

Boys: $(0.23)(125) = 28.75 \approx 29$

Girls: $(0.28)(117) = 32.76 \approx 33$

$33 > 29$

More girls than boys prefer Lemony Snicket.