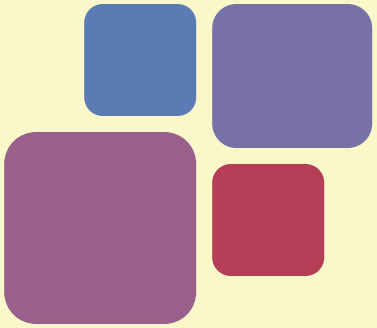
Contingency Tables; Joint & Marginal Probabilities

Section 4.4

Statistics, MATH 181

Mr. Keltner



Univariate Data

- Data from one variable of a population are called ***univariate data***.
  - This is what we focused on in Chapter 2 when constructing a frequency distribution for grouped data.
    - **Example**: Making a frequency distribution for the mileage on the odometer of the cars in the school parking lot (in thousands).
  - In this section, we are working to group and analyze data from ***two*** variables of a population.

Bivariate Data

- Data from two variables of a population are called ***bivariate data***.
  - A frequency distribution for bivariate data is called a ***contingency table***, or a ***two-way table***.
    - **Example**: Showing the breakdown of hair color and gender for the population of students at EHS.

Bivariate Data: Football Fans

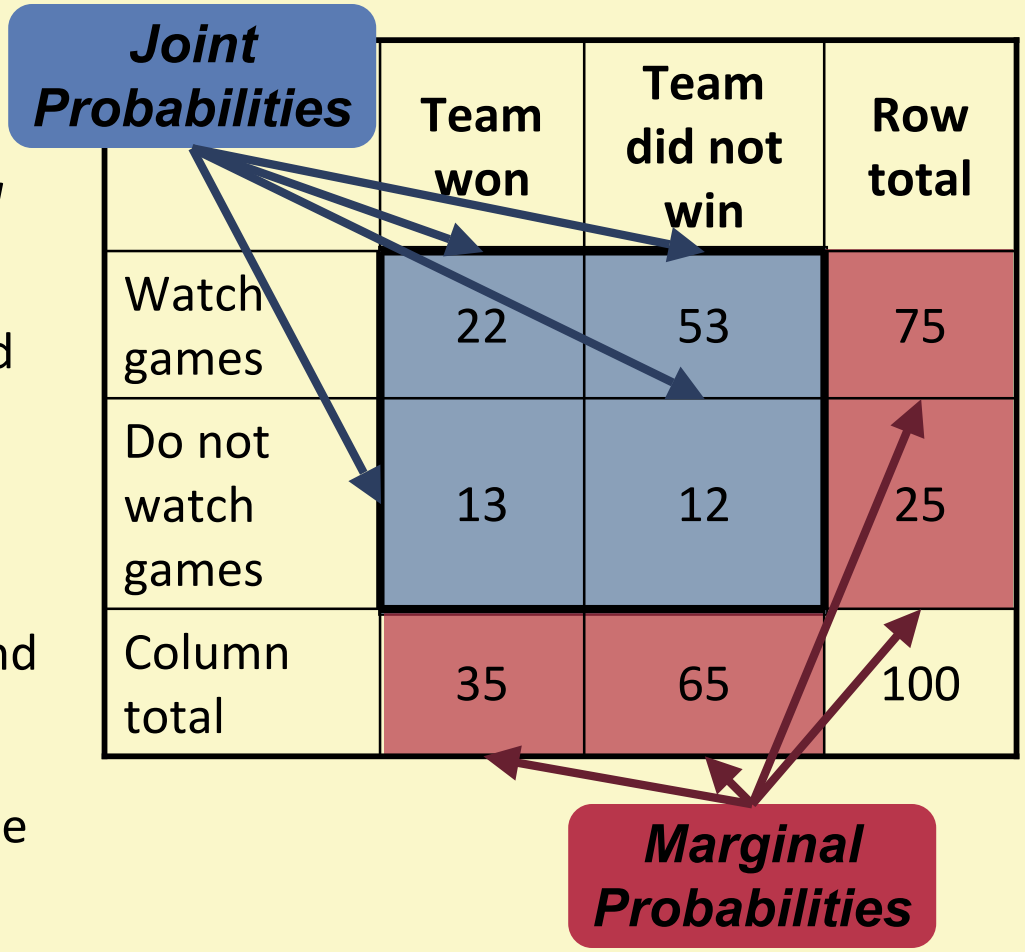
- A random sample of 100 adults were surveyed.
  - They were asked if they regularly watch NFL football games.
  - They were also asked if their favorite team had ever won the Super Bowl.
  - The results are in the table below.

	Team won	Team did not win	Row total
Watch games	22	53	75
Do not watch games	13	12	25
Column total	35	65	100

The small boxes inside the rectangle formed by the heavy lines are called ***cells.***

Marginal & Joint Probabilities

- The *contingency table*, or *two-way table*, introduces the concepts of ***joint probabilities*** and ***marginal probabilities***.
  - Probabilities that correspond to events represented in the margin of the contingency table are ***marginal probabilities.***
  - ***Joint probabilities*** correspond to events that are the combination of the two variables that are listed in the contingency table.



Example: Football Fans

- For a person selected at random from the sample:
  - Find P(favorite team won the Super Bowl).
  - Find P(watch NFL games).
  - Find P(watch NFL games and favorite team has not won Super Bowl).
  - Find P(favorite team won *and* watch NFL games).

	Team won	**Team did not win**	**Row total**
Watch games	22	53	75
Do not watch games	13	12	25
Column total	35	65	100

Example 2: Complete the Contingency Table

- A joint frequency distribution for number of farms, by acreage and tenure of operator, is shown below. Frequencies are in thousands.

Tenure of Operator

	Full Owner	Part Owner	Tenant	Total
Under 50	*639*	64	41	*744*
50→180	487	131	41	659
180→500	203	*153*	*33*	389
500→1000	54	91	17	162
1000 & Over	46	112	18	176
Total	1429	551	*150*	*2130*

- The sum of the joint probabilities in a row or column must equal the marginal probability for that row or column.
  - Any discrepancy in values is chalked up to rounding error.
- Use this information to fill in the missing entries of the table at the left.

Assessment

Pgs. 186-187:

#’s 4.75 - .79