Unit 12: Bivariate Data

**Lesson 2: Two way frequency and relative frequency**

**Objectives:**

- I can analyze two way frequency table for two categorical data.

- I can construct relative frequency tables for two way frequency tables.

- I can analyze data by calculating the percentages and relative frequencies of different events in the study

**Agenda:**

* Discoveries
* Practice

**Focus Questions**:

* How can I examine data in relative frequency table to study two variables?
* How can I identify if there is an association between the variables studied?

**Vocabulary:**

* Contingency table, relative frequency table, conditional relative frequency by row or column

**Homework: HW12-2**

Web support:

<https://www.youtube.com/watch?v=l5MrtV7ZN88>

<https://www.youtube.com/watch?v=k8xFH6fCIWs>

Web Practice:

<https://www.khanacademy.org/math/probability/two-way-tables-categorical-data-a1/two-way-frequency-tables/e/two-way-frequency-tables>

**What is a Contingency Table?**

**A contingency table** is also called a **two-way frequency table** which list bivariate data or data for two variables. These tables examine the relationships between the two variables and often used to compare and analyze survey results. The table cells that show the frequency of both variables are called **joint frequency** and the table cells that show the total of each category is called **marginal frequency**.

**Do Now:**

1. Daniel is processing a large document on a computer. This scatter plot shows how many pages he produced each hour. Using this information, what is the best prediction of the number of pages Daniel can produce in 10 hours?

1] 30 2] 45

3] 15 4] 40

2)Which of the following correlation coefficients would indicate no significant linear relationship for the independent and dependent variables in a data set?

a) – 1 b) – 0.52 c) 0.15 d) 0.90

**Example 1:** A group of students were surveyed in New Paltz High School’s Cafeteria to identify which type of lunch they mostly prefer. The following two-way frequency table identifies the number of 9th and 10th grade students who prefer either hot or cold lunch. Analyze the table and answer the following questions.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Cold Lunch | Hot Lunch | Total |
| 9th graders | 32 | 33 | 65 |
| 10th graders | 25 | 10 | 35 |
| Total | 57 | 43 | 100 |

1. How many students were surveyed in the cafeteria?
2. How many 10th grade students were surveyed?
3. How many 9th grade students prefer hot lunch?
4. How many 10th grade students prefer cold lunch?
5. How many students prefer cold lunch?

**Example 2:** Fill the two-way frequency table and answer the following questions:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Sport Utility (SUV) | Sports Car | Total |
| Male | 60 |  | 140 |
| Female | 100 | 20 | 120 |
| Total |  | 100 |  |

1. How many people responded to the survey? \_\_\_\_\_\_\_
2. How many males responded to the survey? \_\_\_\_\_\_\_
3. How many people chose an SUV? \_\_\_\_\_\_\_
4. How many females chose a sports car? \_\_\_\_\_\_\_
5. How many males chose an SUV? \_\_\_\_\_\_\_\_\_

**Relative Frequency Table:** A two-way frequency table in which all frequencies are divided by the total (Also called **probability).** Relative frequency tables can help us find the number of times an event occurs compare to the total number of events. We can change the relative frequency to a percentage as well.

**Example 3:** Fill the relative frequency table and answer the questions:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Cold Lunch | Hot Lunch | Total |
| 9th graders |  |  |  |
| 10th graders |  |  |  |
| Total |  |  |  |

1. What is the relative frequency of 9th grade students who were surveyed?
2. What is the probability **of all students** who are 9th graders and refer cold lunch?
3. What is the percentage **of all students** who are 10th graders and prefer hot lunch?
4. What is the percentage **of all students** who prefer hot lunch?
5. Which has a higher percentage, **students** who are 9th graders and prefer hot lunch or **students** who are 10th graders and preferred cold lunch

***Exercise* 4:** A class of 20 students recorded their hair color and eye color which are shown in the **two-way frequency table** below.

1. How many students have blond hair and blue eyes?
2. How many students have red hair?
3. How many students have brown eyes?

Construct a relative frequency table for the data above

1. What is the relative frequency of students who have black hair and green eyes?
2. ****What is the probability that a student has red hair and green eyes?
3. What is the percentage of students who have blond hair and green eyes?

**On to Conditional Relative Frequency: Web Support:**

<https://www.youtube.com/watch?v=Kbyl5oUZG0I>

<https://www.youtube.com/watch?v=xy_Bz06IZVc>

We often need to exam the table by separate categories (rows or columns). When a relative frequency is determined based upon a row or column (specific condition), it is called a "conditional" relative frequency. To obtain a conditional relative frequency for Rows, the entries are divided by the total for that Row. **This is useful when we answer a question like: What percentage of 9th graders prefer hot lunch?**

**Example 1:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Cold Lunch | Hot Lunch | Total |
| 9th graders |  |  |  |
| 10th graders |  |  |  |
| Total |  |  |  |

1. What is the conditional relative frequency **of 9th graders** who prefer hot lunch?

You can also see this question written as: What is the probability that a student prefer hot lunch **given that** he is a 9th grader?

1. What is the conditional relative frequency **of 9th graders** who prefer cold lunch? **Or** what is the probability that a student prefer cold lunch **given that** he is a 9th grader?
2. What is the percentage **of 10th graders** who prefer hot lunch?
3. What is the percentage **of 10th graders** who prefer cold lunch?

**Conditional relative frequency is also important when we look for an Association between the two variables?** In this situation grade level and preferred type of lunch.

There is an "**association"** between two categorical variables, if the row (or column) conditional relative frequencies are **different** for the rows (or columns) of the table. The bigger the differences in the conditional relative frequencies, the stronger the association between the variables. If the conditional relative frequencies are nearly equal for all categories, there may be no association between the variables. Such variables are said to be **independent.**

***Practice:*** Look at the following conditional relative frequency table.



1. What are the two variables examined in this table?
2. What is the percentage of students who passed the test given that they completed the review sheet?
3. What is the percentage of students who failed the test given that they did not complete the review sheet?
4. What is the percentage of students who passed the test?
5. How many students were surveyed?
6. A deli offers two types of bread and three types of meat for sandwiches. The two-way table of data below shows the relative frequencies of some sandwich types the deli sold one afternoon.

Complete the frequency table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Turkey | Ham | Salami | Totals |
| White Bread |  |  |  | 44 |
| Wheat Bread |  |  |  | 36 |
| Totals | 28 | 32 | 20 | 80 |

1. What is the number of salami sandwiches on wheat bread that were sold?

14 c. 20

28 d. 10

1. What is the relative frequency of sandwiches with turkey on white bread that were sold?

0.175 c. 0.25

0.35 d. 0.075

* 1. What is the percentage of sandwiches with ham on wheat bread that were sold?

17.5% c. 15%

35% d. 40%

***Name; \_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_ Homework 12-2***

***1*:** Juniors and seniors were asked if they plan to attend college immediately after graduation, seek full-time employment, or choose some other option. A random sample of 100 students was selected from those who completed the survey. Complete the calculations of the relative frequencies for each of the blank cells. Round your answers to the nearest thousandth.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Plan to attend College | Plan to seek full-time employment | Other options | Totals |
| Seniors |  |  |  |  |
| Juniors |  |  |  |  |
| Totals |  |  |  |  |

1. How many students are juniors?
2. What is the probability that a student is seeking a full time employment?
3. What is the relative frequency that a student is a senior and looking for other options?
4. What is the probability that a student is a senior and planning to attend college?
5. Philip surveyed students in his school about whether they play a musical instrument or play sport. The results of his survey are shown in the two way frequency table.

a. Complete the table:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Play a musical instrument | Do not play a musical instrument | Totals |
| Play sports | 9 |  | 27 |
| Do not play sports |  | 27 |  |
| Totals | 15 |  | 60 |

b. How many students play sports and a musical instruments?

c. How many students Do not play sports?

d. How many students were surveyed?

1. Construct a relative frequency table :

|  |  |  |  |
| --- | --- | --- | --- |
|  | Play a musical instrument | Do not play a musical instrument | Totals |
| Play sports |  |  |  |
| Do not play sports |  |  |  |
| Totals |  |  |  |

1. What is the percentage of students who do not play sports?
2. What is the joint frequency of students who play an instrument and sports?
3. What is the percentages of students who play an instrument?
4. What is the marginal frequency of students who do not play sports?

Have you worked on Regents Review 10??