Lecture 11:
Chapter 11

C C Moxley

UAB Mathematics

5 November 15

§11.1 Analyzing Categorical Data

So far, we have used statistical methods to analyze population parameters - but we knew what types of random variables we were dealing with.

So far, we have used statistical methods to analyze population parameters - but we knew what types of random variables we were dealing with. Now, we move on to making qualitative inferences, namely things like whether or not a sample comes from a particular type of distribution or whether or not two samples are independent.

So far, we have used statistical methods to analyze population parameters - but we knew what types of random variables we were dealing with. Now, we move on to making qualitative inferences, namely things like whether or not a sample comes from a particular type of distribution or whether or not two samples are independent. The first type of test is a **goodness-of-fit test**,

§11.1 Analyzing Categorical Data

So far, we have used statistical methods to analyze population parameters - but we knew what types of random variables we were dealing with. Now, we move on to making qualitative inferences, namely things like whether or not a sample comes from a particular type of distribution or whether or not two samples are independent. The first type of test is a **goodness-of-fit test**, and the second test is done using a is a **contingency test**.

§11.2 Goodness-of-Fit Tests

Definition (Goodness-of-Fit Test)

A **goodness-of-fit test** tests the hypothesis that an observed frequency distribution or a single column of data comes from some claimed distribution.

§11.2 Goodness-of-Fit Tests

Definition (Goodness-of-Fit Test)

A **goodness-of-fit test** tests the hypothesis that an observed frequency distribution or a single column of data comes from some claimed distribution.

There are a few requirements to perform a goodness-of-fit test.

Definition (Goodness-of-Fit Test)

A **goodness-of-fit test** tests the hypothesis that an observed frequency distribution or a single column of data comes from some claimed distribution.

There are a few requirements to perform a goodness-of-fit test.

- The data must be randomly selected.

§11.2 Goodness-of-Fit Tests

Definition (Goodness-of-Fit Test)

A **goodness-of-fit test** tests the hypothesis that an observed frequency distribution or a single column of data comes from some claimed distribution.

There are a few requirements to perform a goodness-of-fit test.

- The data must be randomly selected.
- The data consist of (or can be arranged into) frequency counts for each of the different categories.

Definition (Goodness-of-Fit Test)

A **goodness-of-fit test** tests the hypothesis that an observed frequency distribution or a single column of data comes from some claimed distribution.

There are a few requirements to perform a goodness-of-fit test.

- The data must be randomly selected.
- The data consist of (or can be arranged into) frequency counts for each of the different categories.
- For each category, the expected frequency is at least five.

§11.2 Goodness-of-Fit Tests

We need these variables.

We need these variables.

- O, observed frequency (from the table or data)

We need these variables.

- O, observed frequency (from the table or data)
- E, expected frequency (from the assumption)

§11.2 Goodness-of-Fit Tests

We need these variables.

- O, observed frequency (from the table or data)
- E, expected frequency (from the assumption)
- k, number of different cells or categories

We need these variables.

- O, observed frequency (from the table or data)
- E, expected frequency (from the assumption)
- k, number of different cells or categories
- n, number of trails or observations

§11.2 Goodness-of-Fit Tests

We need these variables.

- O, observed frequency (from the table or data)
- E, expected frequency (from the assumption)
- k, number of different cells or categories
- n, number of trails or observations

We then compute the test statistic $\chi^2 = \sum \frac{(O-E)^2}{E}$.

§11.2 Goodness-of-Fit Tests

We need these variables.

- O, observed frequency (from the table or data)
- E, expected frequency (from the assumption)
- k, number of different cells or categories
- n, number of trails or observations

We then compute the test statistic $\chi^2 = \sum \frac{(O-E)^2}{E}$. And we perform a right-tailed χ^2-test with $k-1$ degrees of freedom.

How to calculate E?

How to calculate E? Well, if the assumed distribution is uniform, then we get that $E = n/k$. Otherwise, we get $E = np$, where p is the probability of falling into a particular category.

§11.2 Goodness-of-Fit Tests

How to calculate E? Well, if the assumed distribution is uniform, then we get that $E = n/k$. Otherwise, we get $E = np$, where p is the probability of falling into a particular category. You can often calculate p by using a "between" calculator on StatCrunch for the appropriate distribution.

§11.2 Example

Example (uniform random variable goodness of fit)

Is the following sample of weights from a population with a continuous uniform distribution where the upper and lower limits are 10 and 60? Use $\alpha = 0.1$.

§11.2 Example

Example (uniform random variable goodness of fit)

Is the following sample of weights from a population with a continuous uniform distribution where the upper and lower limits are 10 and 60? Use $\alpha = 0.1$.

Data Set 1

§11.2 Example

Example (uniform random variable goodness of fit)

Is the following sample of weights from a population with a continuous uniform distribution where the upper and lower limits are 10 and 60? Use $\alpha = 0.1$.

Data Set 1

To check this, let's compute the expected and observed counts in each of the categories below.

§11.2 Example

Example (uniform random variable goodness of fit)

Is the following sample of weights from a population with a continuous uniform distribution where the upper and lower limits are 10 and 60? Use $\alpha = 0.1$.

Data Set 1

To check this, let's compute the expected and observed counts in each of the categories below. (Note: We'll use five equal categories to keep the expected counts 5 or more.)

§11.2 Example

Example (uniform random variable goodness of fit)

Is the following sample of weights from a population with a continuous uniform distribution where the upper and lower limits are 10 and 60? Use $\alpha = 0.1$.

Data Set 1

To check this, let's compute the expected and observed counts in each of the categories below. (Note: We'll use five equal categories to keep the expected counts 5 or more.)

Category	10-20	20-30	30-40	40-50	50-60
Expected Count	5	5	5	5	5
Observed Count	4	4	3	4	10

§11.2 Example

Example (uniform random variable goodness of fit, con't.)

Category	10-20	20-30	30-40	40-50	50-60
Expected Count	5	5	5	5	5
Observed Count	4	4	3	4	10

Thus, the test statistic is

Example (uniform random variable goodness of fit, con't.)

Category	10-20	20-30	30-40	40-50	50-60
Expected Count	5	5	5	5	5
Observed Count	4	4	3	4	10

Thus, the test statistic is

$$X^2 = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{4}{5} + \frac{25}{5} = 6.4.$$

Example (uniform random variable goodness of fit, con't.)

Category	10-20	20-30	30-40	40-50	50-60
Expected Count	5	5	5	5	5
Observed Count	4	4	3	4	10

Thus, the test statistic is

$$X^2 = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{4}{5} + \frac{25}{5} = 6.4.$$

And so the P-value is $P(\chi^2 > 6.4) = 0.1712$.

§11.2 Example

Example (uniform random variable goodness of fit, con't.)

Category	10-20	20-30	30-40	40-50	50-60
Expected Count	5	5	5	5	5
Observed Count	4	4	3	4	10

Thus, the test statistic is

$$X^2 = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{4}{5} + \frac{25}{5} = 6.4.$$

And so the P-value is $P(\chi^2 > 6.4) = 0.1712$. And we fail to reject our null hypothesis.

§11.2 Example

Example (uniform random variable goodness of fit, con't.)

Category	10-20	20-30	30-40	40-50	50-60
Expected Count	5	5	5	5	5
Observed Count	4	4	3	4	10

Thus, the test statistic is

$$X^2 = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{4}{5} + \frac{25}{5} = 6.4.$$

And so the P-value is $P(\chi^2 > 6.4) = 0.1712$. And we fail to reject our null hypothesis. Our data is not significant enough to reject the claim that the data came from a population with a continuous uniform distribution with bound 10 and 60.

§11.2 Example

A salesman wants to make a pitch on a day a company is most likely to make major purchases. He obtained data about the company's previous purchases and tested (using $\alpha = 0.05$) to see if the purchases were uniformly distributed throughout the week (Mon-Sun). He obtained the following.

- $k = 7$,

§11.2 Example

A salesman wants to make a pitch on a day a company is most likely to make major purchases. He obtained data about the company's previous purchases and tested (using $\alpha = 0.05$) to see if the purchases were uniformly distributed throughout the week (Mon-Sun). He obtained the following.

- $k = 7$, $n = 20$,

§11.2 Example

A salesman wants to make a pitch on a day a company is most likely to make major purchases. He obtained data about the company's previous purchases and tested (using $\alpha = 0.05$) to see if the purchases were uniformly distributed throughout the week (Mon-Sun). He obtained the following.

- $k = 7$, $n = 20$, $E = 2.86$,

§11.2 Example

A salesman wants to make a pitch on a day a company is most likely to make major purchases. He obtained data about the company's previous purchases and tested (using $\alpha = 0.05$) to see if the purchases were uniformly distributed throughout the week (Mon-Sun). He obtained the following.

- $k = 7$, $n = 20$, $E = 2.86$, test statistic $\chi^2 = 15.85$, critical value 15.507, P-value 0.045

What's the null hypothesis in this case?

§11.2 Example

A salesman wants to make a pitch on a day a company is most likely to make major purchases. He obtained data about the company's previous purchases and tested (using $\alpha = 0.05$) to see if the purchases were uniformly distributed throughout the week (Mon-Sun). He obtained the following.

- $k = 7$, $n = 20$, $E = 2.86$, test statistic $\chi^2 = 15.85$, critical value 15.507, P-value 0.045

What's the null hypothesis in this case? Do we accept or reject it?

§11.2 Example

A salesman wants to make a pitch on a day a company is most likely to make major purchases. He obtained data about the company's previous purchases and tested (using $\alpha = 0.05$) to see if the purchases were uniformly distributed throughout the week (Mon-Sun). He obtained the following.

- $k = 7$, $n = 20$, $E = 2.86$, test statistic $\chi^2 = 15.85$, critical value 15.507, P-value 0.045

What's the null hypothesis in this case? Do we accept or reject it? What does this mean?

§11.2 Example (with frequency table)

A casino wants to determine if a 10-sided die is loaded (with significance $\alpha = 0.01$). They roll the die 64 times and get the following data:

§11.2 Example (with frequency table)

A casino wants to determine if a 10-sided die is loaded (with significance $\alpha = 0.01$). They roll the die 64 times and get the following data:

Side	1	2	3	4	5	6	7	8	9	10
Rolls	6	6	5	7	7	7	7	7	5	7

§11.2 Example (with frequency table)

A casino wants to determine if a 10-sided die is loaded (with significance $\alpha = 0.01$). They roll the die 64 times and get the following data:

Side	1	2	3	4	5	6	7	8	9	10
Rolls	6	6	5	7	7	7	7	7	5	7

We end up getting that the *P*-value for this test is 0.9994, so we **fail to reject** the null hypothesis.

§11.2 Example (with frequency table)

A casino wants to determine if a 10-sided die is loaded (with significance $\alpha = 0.01$). They roll the die 64 times and get the following data:

Side	1	2	3	4	5	6	7	8	9	10
Rolls	6	6	5	7	7	7	7	7	5	7

We end up getting that the P-value for this test is 0.9994, so we **fail to reject** the null hypothesis. This means that the data does not support the claim that the die is loaded.

§11.2 Example (with frequency table)

A casino wants to determine if a 10-sided die is loaded (with significance $\alpha = 0.01$). They roll the die 64 times and get the following data:

Side	1	2	3	4	5	6	7	8	9	10
Rolls	6	6	5	7	7	7	7	7	5	7

We end up getting that the *P*-value for this test is 0.9994, so we **fail to reject** the null hypothesis. This means that the data does not support the claim that the die is loaded. We can conduct the same test with a critical value. We'd see that our test statistic was 1 and our critical value was 21.666, telling us that we should reject our null hypothesis as well.

§11.2 Example (with frequency table)

A casino wants to determine if a 10-sided die is loaded (with significance $\alpha = 0.01$). They roll the die 64 times and get the following data:

Side	1	2	3	4	5	6	7	8	9	10
Rolls	6	6	5	7	7	7	7	7	5	7

We end up getting that the P-value for this test is 0.9994, so we **fail to reject** the null hypothesis. This means that the data does not support the claim that the die is loaded. We can conduct the same test with a critical value. We'd see that our test statistic was 1 and our critical value was 21.666, telling us that we should reject our null hypothesis as well. Use the Goodness-of-Fit tab in StatCrunch.

§11.3 Contingency (Two-Way Frequency) Tables

Definition (Contingency Table)

A **contingency table** is a table consisting of frequency counts of categorical data corresponding to two different variables.

§11.3 Contingency (Two-Way Frequency) Tables

Definition (Contingency Table)

A **contingency table** is a table consisting of frequency counts of categorical data corresponding to two different variables.

Definition (Test of Independence)

A **test of independence** tests the claim that the row and column variables are independent.

§11.3 Contingency (Two-Way Frequency) Tables

Definition (Contingency Table)

A **contingency table** is a table consisting of frequency counts of categorical data corresponding to two different variables.

Definition (Test of Independence)

A **test of independence** tests the claim that the row and column variables are independent.

A test of independence is a χ^2 test, and its degrees of freedom is computed by finding the product of one fewer than the number of columns and one fewer than the number of rows in the contingency table: $df = (c - 1)(r - 1)$.

§11.3 Contingency (Two-Way Frequency) Tables

The requirements for a test of independence are below.

- The sample data must be randomly selected.

§11.3 Contingency (Two-Way Frequency) Tables

The requirements for a test of independence are below.

- The sample data must be randomly selected.
- The sample data are frequency counts in a two-way table.

§11.3 Contingency (Two-Way Frequency) Tables

The requirements for a test of independence are below.

- The sample data must be randomly selected.
- The sample data are frequency counts in a two-way table.
- Each cell in the **expected** table has a value at least 5. (The expected table is the table we'd get if the columns and rows were independent.)

§11.3 Contingency (Two-Way Frequency) Tables

The requirements for a test of independence are below.

- The sample data must be randomly selected.
- The sample data are frequency counts in a two-way table.
- Each cell in the **expected** table has a value at least 5. (The expected table is the table we'd get if the columns and rows were independent.)

You can read more details about the test of independence on page 578 of your text.

§11.3 Example

A hospital wants to verify that a test for diabetes is effective by seeing if the result of the test and the status of the patient are dependent with significance $\alpha = 0.01$. They have the following data.

	Negative	Positive
Tested Negative	17	7
Tested Positive	8	15

§11.3 Example

A hospital wants to verify that a test for diabetes is effective by seeing if the result of the test and the status of the patient are dependent with significance $\alpha = 0.01$. They have the following data.

	Negative	Positive
Tested Negative	17	7
Tested Positive	8	15

Use the contingency table with summary tab in StatCrunch to see that the P-value is 0.0133. So, we fail to reject the null hypothesis! This means that the data supports the claim that the result of the test is independent of the status of the patient, meaning that the diabetes test is not effective.

Note: We would reach the same conclusion if we used a critical value test because the test is right-tailed and the critical value would be 6.635 with a test statistic of 6.131.