Types of Statistical Inference

Single categorical variable

One-proportion z-interval and test (Chapters 19-21)

Single quantitative variable

One sample t-interval and test (Chapter 23)

Two quantitative variables

Regression inference (Chapter 27)

Two categorical variables

Two categories each:
Two proportion z-interval and test (Chapter 22)

More than two categories each:
Chi-square tests (Chapter 26)

One categorical, one quantitative variable

Two categories:
2-sample t-interval and test (Chapter 24)
Paired t-interval and test (Chapter 25)

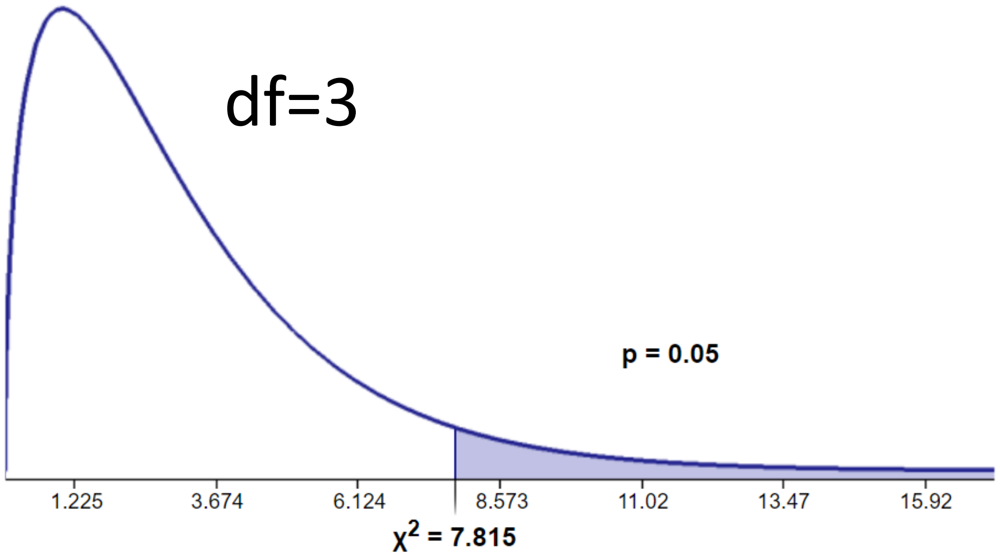
More than two categories:
ANOVA test (Chapter 28)

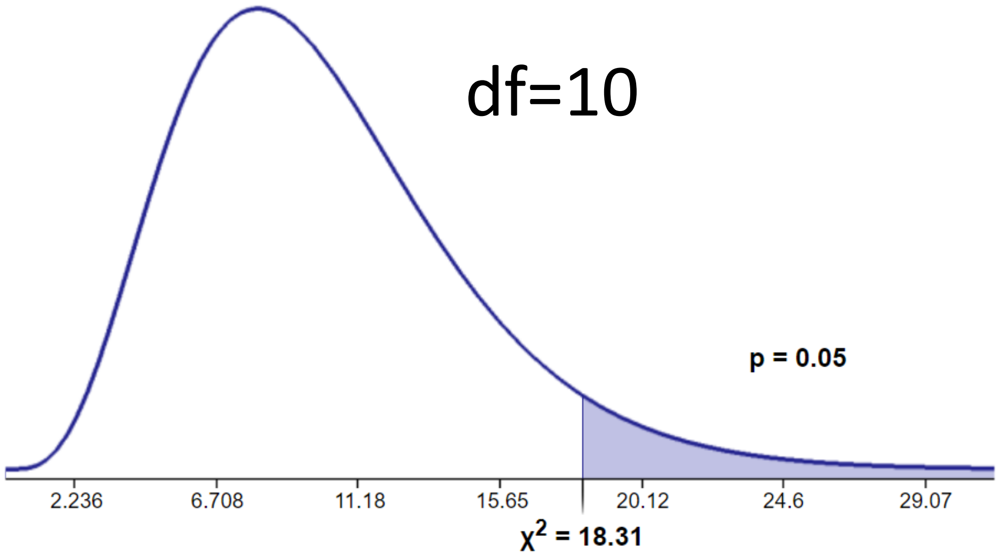
Chi square tests

- Uses the chi-square distributions
- Since the distribution is right-skewed, there's no confidence interval, only a one-sided hypothesis test.

χ

Chi Chai





Types of chi-square tests:

- **Chi-square test for goodness-of-fit:** Compares a list of observed outcomes for a single categorical variable to the expected outcomes given by a model.
(1 categorical variable, 1 sample)

- **Chi-square test for homogeneity:** Compares observed distributions of one categorical variable across multiple samples to test for differences among the populations.
(1 categorical variable, multiple samples)

- **Chi-square test for independence:** Compares observed distributions of two categorical variables to test for an association between the two variables.
(2 categorical variables, 1 sample)
(same technique as homogeneity)

Chi square goodness-of-fit test

1. H_0: The counts in all categories match the expected counts
   H_A: The count in at least one category doesn't match the expected count
2. Find the χ^2-score of the sample.
3. Convert the χ^2-score to a P–value.
4. Compare the P-value to α=.05.
5. Retain the null if the P-value is greater than α, and reject the null hypothesis if the P-value is less than α. Report the P-value of the test.

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

summed over all categories

Degrees of freedom

$$df = number\ of\ categories\ -\ 1$$

Chi square test for homogeneity or independence

1. H_0: The counts in all categories match across different samples/categories
   H_A: The count in at least one category doesn't match
2. Find the χ^2-score of the sample.
3. Convert the χ^2-score to a P–value.
4. Compare the P-value to α=.05.
5. Retain the null if the P-value is greater than α, and reject the null hypothesis if the P-value is less than α. Report the P-value of the test.

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

summed over all categories

$$\text{expected} = \frac{(row\ sum)(column\ sum)}{total\ sum}$$

Degrees of freedom

$$df = (rows - 1)(columns - 1)$$

Chi square conditions

1. **Counted data.** The numbers are counts of cases, not percentages.
2. **Independence.** The cases (people) are independent of each other.
3. **Randomization.** The sample is representative.
4. **10% Condition.** The sample is less than 10% of the population.
5. **Expected cell frequency.** The expected count in each cell in the table is at least 5.