Inferential statistics

- Inference involves making a **Generalization** about a larger group of individuals on the basis of a subset or sample.

Null and alternative hypotheses

- In hypotheses testing, a specific hypothesis (Null and alternative Hypothesis) are formulated and tested.

- **The null hypotheses H0 means** : X1=X 2

Or X1-X 2=0

- this means that there is no difference between x1 and x2

- **The alternative hypotheses H1 means**

X1>X2   or   X1< X2

Hypothesis Testing

- The method of assessing the hypotheses testing is known as *significance test*.

- *The significance testing* is a method for assessing whether a result is likely to be due to chance or due to a real effect .

General principles of significance tests

1. set up a null hypothesis and its alternative.

2. find the value of the test statistic.

3. refer the value of the test statistic to a known distribution which it would follow if the null hypothesis was true.

General principles of significance tests

4-conclude that the data are consistent or inconsistent with the null hypothesis.

- If the data are not consistent with the null hypotheses, the difference is said to be statistically significant. If the data are consistent with the null hypotheses it is said that we accept it i.e. statistically insignificant.
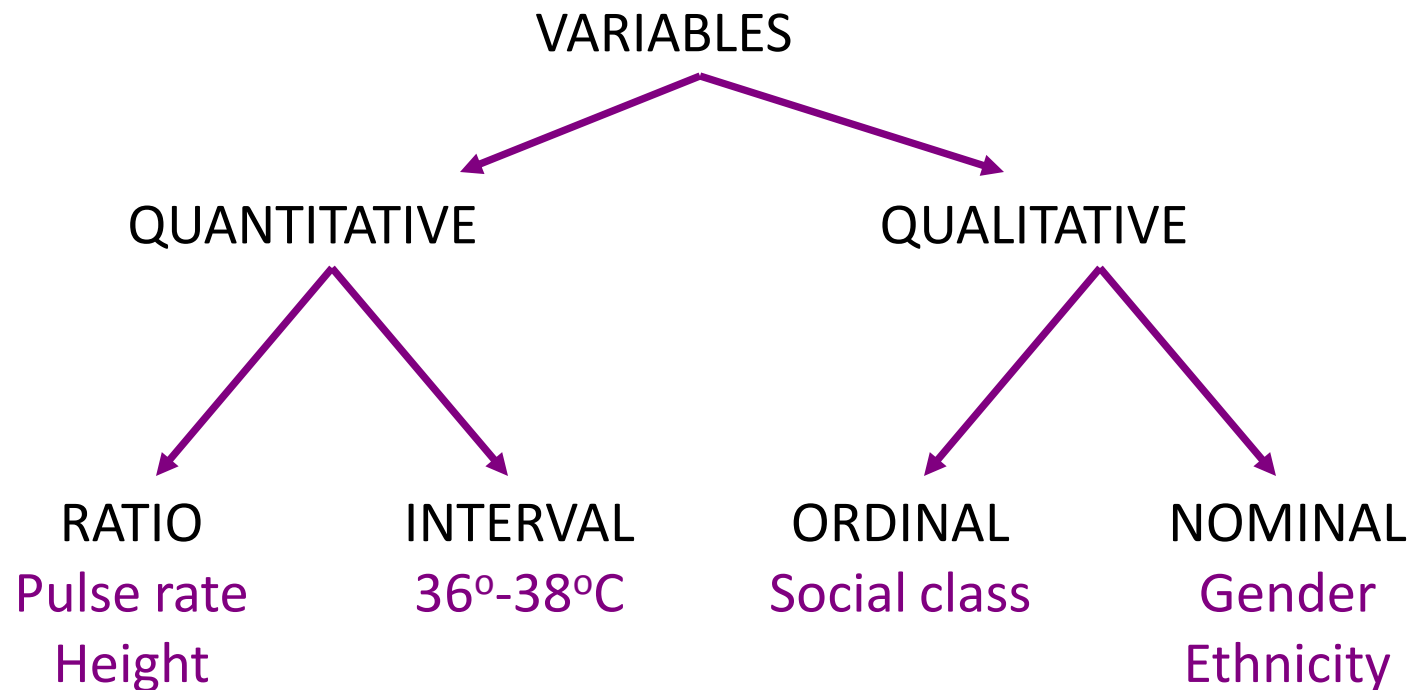
General principles of significance tests P<0.05

In medicine, we usually consider that differences are significant if the probability is less than 0.05. This means that if the null hypothesis is true, we shall make a wrong decision less than 5 in a hundred times

The t-test

Inferences about Population Means

TYPES OF DATA

VARIABLES

QUANTITATIVE   QUALITATIVE

RATIO   INTERVAL   ORDINAL   NOMINAL
Pulse rate 36°-38^oC   Social class   Gender
Height   Ethnicity

Background

- The t statistic allows researchers to use sample data to test hypotheses about an unknown population.

- Developed by Gossett for the quality control of beer.

- Comes in 3 varieties:

- Single sample t test , independent samples t test , and dependent samples t test (paired t-test).

The T Distribution:

1- It has mean of zero.

2- It is symmetric about the

mean.

3- It ranges from -∞ to ∞.
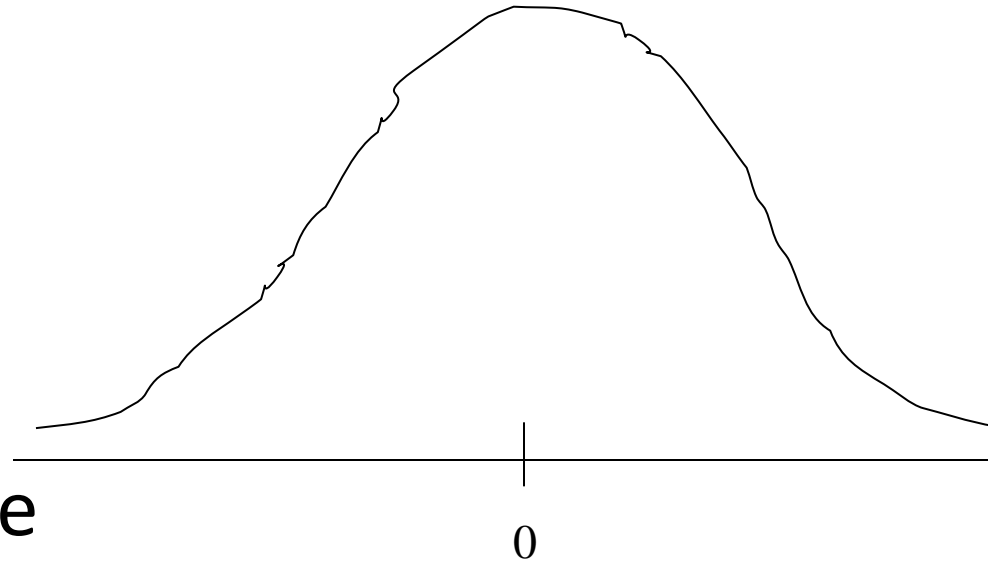
Assumptions of t-Test

- Dependent variables are interval or ratio.
- The population from which samples are drawn is normally distributed.
- Samples are randomly selected.
- For very large samples, the t-distribution approximates the standard normal (z) distribution.
- In practice, it is best to use t-distributions any time the population standard deviation is not known.

Important guidelines

- critical *t*-values for common alpha levels (0.10, 0.05, 0.01, and so forth) are usually given in a single table.

- Values in the *t*-table are not actually listed by sample size but by degrees of freedom *(df)*.

- The number of degrees of freedom for a problem involving the *t*-distribution for sample size *n* is simply $n - 1$ for a one-sample mean problem.

- The *t*-distribution is particularly useful for tests with small samples ($n < 30$).

One-Sample t-test

- **Requirements**: Normally distributed population, σ is unknown
- **Test for population mean**
- **Hypothesis test**
- **Formula**:

$$t = \frac{\bar{x} - \Delta}{\frac{s}{\sqrt{n}}}$$

- where x is the sample mean, Δ is a specified value to be tested, s is the sample standard deviation, and n is the size of the sample.

Note

- Note that the formula for the one-sample t-test is the same as the z-test, except that the t-test substitutes the sample standard deviation s for the population standard deviation σ and takes critical values from the t-distribution instead of the z-distribution.

Example 1

- A professor wants to know if her introductory statistics class has a good grasp of basic math. Six students are chosen at random from the class and given a math proficiency test. The professor wants the class to be able to score above 70 on the test. The six students get scores of 62, 92, 75, 68, 83, and 95. Can the professor have 90 percent confidence that the mean score for the class on the test would be above 70?

- **null hypothesis**: H_0: $\mu = 70$

- **alternative hypothesis**: H_a: $\mu > 70$

One sample t-test

- Values in the *t*-table are not actually listed by sample size but by degrees of freedom*(df)*. The number of degrees of freedom for a problem involving the *t*-distribution for sample size *n* is simply *n* – 1 for a one-sample mean problem.

Example 1

- First, compute the sample mean and standard deviation:

$$62$$
$$92$$
$$75$$
$$68$$
$$83$$
$$+\ 95$$
$$\overline{}$$
$$475$$

$$\bar{x} = \frac{475}{6} = 79.17$$

$$s = 13.17$$

Example 1

- Next, compute the *t*-value:

$$t = \frac{\bar{x} - \Delta}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{79.17 - 70}{\frac{13.17}{\sqrt{6}}} = \frac{9.17}{5.38} = 1.71$$

Example 1

- To test the hypothesis, the computed t-value of 1.71 will be compared to the critical value in the t-table.

- Because the sample mean is in the numerator, the larger it is, the larger the resulting t-value will be.

- The higher sample mean will make it more likely that the professor will conclude that the math proficiency of the class is satisfactory and that the null hypothesis of less-than-satisfactory class math knowledge can be rejected.

- The larger the computed t-value, the greater the chance that the null hypothesis can be rejected. It follows, then, that if the computed t-value is larger than the critical t-value from the table, the null hypothesis can be rejected.

Student's t-Test Table

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05
two-tails	1.00	0.50	0.40	0.30	0.20	0.10
df						
1	0.000	1.000	1.376	1.963	3.078	6.314
2	0.000	0.816	1.061	1.386	1.886	2.920
3	0.000	0.765	0.978	1.250	1.638	2.353
4	0.000	0.741	0.941	1.190	1.533	2.132
5	0.000	0.727	0.920	1.156	1.476	2.015
6	0.000	0.718	0.906	1.134	1.440	1.943
7	0.000	0.711	0.896	1.119	1.415	1.895
8	0.000	0.706	0.889	1.108	1.397	1.860
9	0.000	0.703	0.883	1.100	1.383	1.833
10	0.000	0.700	0.879	1.093	1.372	1.812

Example 1

- A 90 percent confidence level is equivalent to an alpha level of 0.10. Because extreme values in one rather than two directions will lead to rejection of the null hypothesis, this is a one-tailed test, and you do not divide the alpha level by 2.

- The number of degrees of freedom for the problem is 6 − 1 = 5. The value in the *t*-table for $t_{.10,5}$ is 1.476.

- Because the computed *t*-value of 1.71 is larger than the critical value in the table, the null hypothesis can be rejected, and the professor has evidence that the class mean on the math test would be at least 70.

Example 2

- A Little League baseball coach wants to know if his team is representative of other teams in scoring runs. Nationally, the average number of runs scored by a Little League team in a game is 5.7. He chooses five games at random in which his team scored 5 , 9, 4, 11, and 8 runs. Is it likely that his team's scores could have come from the national distribution?

Example 2

- Assume an alpha level of 0.05.
- Because the team's scoring rate could be either higher than or lower than the national average, the problem calls for a two-tailed test. First, state the null and alternative hypotheses:
- **null hypothesis**: $H_0: \mu = 5.7$
- **alternative hypothesis**: $H_a: \mu \neq 5.7$

Example 2

- Next, the *t*-value:

$$t = \frac{\bar{x} - \Delta}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{7.4 - 5.7}{\frac{2.88}{\sqrt{5}}} = \frac{1.7}{1.29} = 1.32$$

Example 2

- Now, look up the critical value from the *t*-table
- You need to know two things in order to do this: the degrees of freedom and the desired alpha level.
- The degrees of freedom is $5 - 1 = 4$. The overall alpha level is 0.05, but because this is a two-tailed test, the alpha level must be divided by two, which yields 0.025.
- The tabled value for $t_{.025,4}$ is 2.776.
- The computed *t* of 1.32 is smaller, so you cannot reject the null hypothesis that the mean of this team is equal to the population mean.
- The coach cannot conclude that his team is different from the national distribution on runs scored.

Student's t-Test Table

(1 tail)	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
(2 tail)	0.1	0.05	0.02	0.01	0.005	0.002	0.001
df							
1	6.3138	12.7065	31.8193	63.6551	127.3447	318.493	636.405
2	2.92	4.3026	6.9646	9.9247	14.0887	22.3276	31.5989
3	2.3534	3.1824	4.5407	5.8408	7.4534	10.2145	12.9242
4	2.1319	2.7764	3.747	4.6041	5.5967	7.1732	8.6103
5	2.015	2.5706	3.365	4.0322	4.7734	5.8934	6.8688
6	1.9432	2.4469	3.1426	3.7074	4.3168	5.2076	5.9589
7	1.8946	2.3646	2.998	3.4995	4.0294	4.7852	5.4079
8	1.8595	2.306	2.8965	3.3554	3.8325	4.5008	5.0414
9	1.8331	2.2621	2.8214	3.2498	3.8325	4.2969	4.7809
10	1.8124	2.2282	2.7638	3.1693	3.6896	4.2969	4.5869
11	1.7959	2.201	2.7181	3.0158	3.4966	4.0247	4.4369
12	1.7823	2.1788	2.681	3.0545	3.4284	3.9296	4.3178
13	1.7709	2.1604	2.6503	3.0123	3.3725	3.852	4.2208
14	1.7613	2.1448	2.6245	2.9768	3.3257	3.7874	4.1404
15	1.753	2.1314	2.6025	2.9467	3.286	3.7328	4.0728
16	1.7459	2.1199	2.5835	2.9208	3.252	3.6861	4.015
17	1.7396	2.1098	2.5669	2.8983	3.2224	3.6458	3.9651
18	1.7341	2.1009	2.5524	2.8784	3.1966	3.6105	3.9216
19	1.7291	2.093	2.5395	2.8609	3.1737	3.5794	3.8834
20	1.7247	2.086	2.528	2.854	3.1534	3.5518	3.8495

The t Statistic

The t statistic is used when a researcher wants to determine whether or not a treatment causes a change in a population mean. In this case you must know the value of μ for the original, untreated population. A sample is obtained from the population and the treatment is administered to the sample. If the resulting sample mean is significantly different from the original population mean, you can conclude that the treatment has a significant effect.

The t Statistic

- The particular advantage of the t statistic, is that it can be used to test hypotheses about a *completely unknown* population; that is, both μ and σ are unknown, and the only available information about the population comes from the sample.

The t Statistic

If the population mean is unknown . A sample is then obtained from the population and the t statistic is used to compare the actual sample mean with the hypothesized population mean. A significant difference indicates that the hypothesized value for μ should be rejected.

The t Statistic (cont.)

- All that is required for a hypothesis test with t is a sample and a reasonable hypothesis about the population mean.

The Estimated Standard Error and the t Statistic

- Whenever a sample is obtained from a population you expect to find some discrepancy or "error" between the sample mean and the population mean.

- This general phenomenon is known as **sampling error**.

- The goal for a hypothesis test is to evaluate the *significance* of the observed discrepancy between a sample mean and the population mean.

The Estimated Standard Error and the t Statistic (cont.)

The hypothesis test attempts to decide between the following two alternatives:

1. Is it reasonable that the discrepancy between M and μ is simply due to sampling error and not the result of a treatment effect?

2. Is the discrepancy between M and μ more than would be expected by sampling error alone? That is, is the sample mean significantly different from the population mean?

The t Distributions and Degrees of Freedom (cont.)

- The value of **degrees of freedom**, df = n - 1, is used to determine how well the distribution of t approximates a normal distribution.

- For large values of df, the **t distribution** will be nearly normal, but with small values for df, the t distribution will be flatter and more spread out than a normal distribution.

The t Distributions and Degrees of Freedom (cont.)

- To evaluate the t statistic from a hypothesis test, you must select an α level, find the value of df for the t statistic, and consult the t distribution table.

- If the obtained t statistic is larger than the critical value from the table, you can reject the null hypothesis.

- In this case, you have demonstrated that the obtained difference between the data and the hypothesis (numerator of the ratio) is significantly larger than the difference that would be expected if there was no treatment effect (the standard error in the denominator).

Student's t-Test Table

(1 tail)	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
(2 tail)	0.1	0.05	0.02	0.01	0.005	0.002	0.001
df							
1	6.3138	12.7065	31.8193	63.6551	127.3447	318.493	636.405
2	2.92	4.3026	6.9646	9.9247	14.0887	22.3276	31.5989
3	2.3534	3.1824	4.5407	5.8408	7.4534	10.2145	12.9242
4	2.1319	2.7764	3.747	4.6041	5.5967	7.1732	8.6103
5	2.015	2.5706	3.365	4.0322	4.7734	5.8934	6.8688
6	1.9432	2.4469	3.1426	3.7074	4.3168	5.2076	5.9589
7	1.8946	2.3646	2.998	3.4995	4.0294	4.7852	5.4079
8	1.8595	2.306	2.8965	3.3554	3.8325	4.5008	5.0414
9	1.8331	2.2621	2.8214	3.2498	3.8325	4.2969	4.7809
10	1.8124	2.2282	2.7638	3.1693	3.6896	4.2969	4.5869
11	1.7959	2.201	2.7181	3.0158	3.4966	4.0247	4.4369
12	1.7823	2.1788	2.681	3.0545	3.4284	3.9296	4.3178
13	1.7709	2.1604	2.6503	3.0123	3.3725	3.852	4.2208
14	1.7613	2.1448	2.6245	2.9768	3.3257	3.7874	4.1404
15	1.753	2.1314	2.6025	2.9467	3.286	3.7328	4.0728
16	1.7459	2.1199	2.5835	2.9208	3.252	3.6861	4.015
17	1.7396	2.1098	2.5669	2.8983	3.2224	3.6458	3.9651
18	1.7341	2.1009	2.5524	2.8784	3.1966	3.6105	3.9216
19	1.7291	2.093	2.5395	2.8609	3.1737	3.5794	3.8834
20	1.7247	2.086	2.528	2.854	3.1534	3.5518	3.8495

Two Independent Samples t test for Comparing Two Means

- **Requirements**: Two normally distributed but independent populations, σ is unknown.

- Degrees of freedom: is the smaller of $n_1 - 1$ and $n_2 - 1$.

Hypothesis testing

- **Formula**:

$$t = \frac{\overline{x}_1 - \overline{x}_2 - \Delta}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

- where $\overline{x}_1$ and $\overline{x}_2$ are the means of the two samples, Δ is the hypothesized difference between the population means (0 if testing for equal means), s_1 and s_2 are the standard deviations of the two samples, and n_1 and n_2 are the sizes of the two samples.

Example 3

- An experiment is conducted to determine whether intensive tutoring (covering a great deal of material in a fixed amount of time) is more effective than paced tutoring (covering less material in the same amount of time). Two randomly chosen groups are tutored separately and then administered proficiency tests. Use a significance level of $\alpha < 0.05$.

- Let μ_1 represent the population mean for the intensive tutoring group and μ_2 represent the population mean for the paced tutoring group.

Hypothesis testing

- **null hypothesis**: H_0: $\mu_1 = \mu_2$
- or H_0: $\mu_1 - \mu_2 = 0$
- **alternative hypothesis**: H_a: $\mu_1 > \mu_2$
- or: H_a: $\mu_1 - \mu_2 > 0$

Let us organize the inputs we have

Group	Method	n	$\bar{x}$	s
1	Intensive	12	46.31	6.44
2	Paced	10	42.79	7.52

Calculating t value

$$t = \frac{46.31 - 42.79 - 0}{\sqrt{\frac{6.44^2}{12} + \frac{7.52^2}{10}}} = \frac{3.52}{\sqrt{3.46 + 5.66}} = 1.166$$

Testing the hypothesis

- The degrees of freedom parameter is the smaller of (12 – 1) and (10 – 1), or 9.

- Because this is a one-tailed test, the alpha level (0.05) is not divided by two.

- The next step is to look up $t_{.05,9}$ in the *t*-table, which gives a critical value of 1.833. The computed *t* of 1.166 does not exceed the tabled value, so the null hypothesis cannot be rejected.

- Our conclusion is: This test has not provided statistically significant evidence that intensive tutoring is superior to paced tutoring.

Paired Difference t-test

- **Requirements**: A set of paired observations from a normal population

- This t-test compares one set of measurements with a second set from the same sample. It is often used to compare "before" and "after" scores in experiments to determine whether significant change has occurred.

Hypothesis testing

- **Formula**:

$$t = \frac{\overline{x} - \Delta}{\frac{s}{\sqrt{n}}}$$

- where $\overline{x}$ is the mean of the change scores, Δ is the hypothesized difference (0 if testing for equal means), s is the sample standard deviation of the differences, and n is the sample size. The number of degrees of freedom for the problem is $n - 1$.

Example 4

 A farmer decides to try out a new fertilizer on a test plot containing 10 stalks of corn. Before applying the fertilizer, he measures the height of each stalk. Two weeks later, he measures the stalks again, being careful to match each stalk's new height to its previous one. The stalks would have grown an average of 6 inches during that time even without the fertilizer. Did the fertilizer help? Use a significance level of 0.05.

Hypothesis testing

- **null hypothesis**: H_0: μ = 6
- **alternative hypothesis**: H_a : μ > 6

Stalk	1	2	3	4	5	6	7	8	9	10
Before height	35.5	31.7	31.2	36.3	22.8	28.0	24.6	26.1	34.5	27.7
After height	45.3	36.0	38.6	44.7	31.4	33.5	28.8	35.8	42.9	35.0

Steps of calculation of t

- Subtract each stalk's "before" height from its "after" height to get the change score for each stalk; then compute the mean and standard deviation of the change scores and insert these into the formula.

$$45.3 - 35.5 = 9.8$$

$$36.0 - 31.7 = 4.3$$

$$38.6 - 31.2 = 7.4$$

$$44.7 - 36.3 = 8.4$$

$$31.4 - 22.8 = 8.6$$

$$33.5 - 28.0 = 5.5$$

$$28.8 - 24.6 = 4.2$$

$$35.8 - 26.1 = 9.7$$

$$42.9 - 34.5 = 8.4$$

$$35.0 - 27.7 = \underline{7.3}$$

$$\frac{73.6}{10} = 7.36$$

$$s^2 = \frac{(9.8 - 7.36)^2 + (4.3 - 7.36)^2 + \ldots + (7.3 - 7.36)^2}{10 - 1}$$

$$= 4.216$$

$$s = \sqrt{4.216}$$

$$= 2.05$$

$$t = \frac{7.36 - 6}{\frac{2.05}{\sqrt{10}}} = \frac{1.36}{0.65} = 2.09$$

Hyopothesis testing

- The problem has $n - 1$, or $10 - 1 = 9$ degrees of freedom.

- The test is one-tailed because you are asking only whether the fertilizer increases growth, not reduces it. The critical value from the t-table for $t_{.05,9}$ is 1.833.

- Because the computed t-value of 2.098 is larger than 1.833, the null hypothesis can be rejected.

- **Conclusion: The test has provided evidence that the fertilizer caused the corn to grow more than if it had not been fertilized. The amount of actual increase was not large (1.36 inches over normal growth), but it was statistically significant.**

Chi-Square (X2)

- The statistical procedures that we have reviewed thus far are appropriate only for numerical variables.

- The **chi-square** (χ^2) test can be used to evaluate a relationship between two categorical variables. It is one example of a **nonparametric test**.

- Nonparametric tests are used when assumptions about normal distribution in the population cannot be met. These tests are less powerful than parametric tests.

Example 5

• Suppose that 125 children are shown three television commercials for breakfast cereal and are asked to pick which they liked shown in Table 1:

Table 1. Commercial Preference for Boys and Girls

	A	B	C	Totals
Boys	30	29	16	75
Girls	12	33	5	50
Totals	42	62	21	125

Calculation

- You would like to know if the choice of favorite commercial was related to whether the child was a boy or a girl or if these two variables are independent.
- The totals in the margins will allow you to determine the overall probability of (1) liking commercial A, B, or C, regardless of gender, and (2) being either a boy or a girl, regardless of favorite commercial.

Table 1. Commercial Preference for Boys and Girls

	A	B	C	Totals
Boys	30	29	16	75
Girls	12	33	5	50
Totals	42	62	21	125

Calculation

- If the two variables are independent, then you should be able to use these probabilities to predict approximately how many children should be in each cell.

- If the actual count is very different from the count that you would expect if the probabilities are independent, the two variables must be related.

Table 1. Commercial Preference for Boys and Girls

	A	B	C	Totals
Boys	30	29	16	75
Girls	12	33	5	50
Totals	42	62	21	125

- The overall probability of a child in the sample being a boy is 75 ÷ 125 = 0.6.
- The overall probability of liking Commercial A is 42 ÷ 125 = 0.336.
- The multiplication rule states that the probability of both of two independent events occurring is the product of their two probabilities.
- Therefore, the probability of a child both being a boy and liking Commercial A is 0.6 × 0.336 = 0.202.
- The expected number of children in this cell, then, is 0.202 × 125 = 25.2.

Table 1. Commercial Preference for Boys and Girls

	A	B	C	Totals
Boys	30	29	16	75
Girls	12	33	5	50
Totals	42	62	21	125

- There is a faster way of computing the expected count for each cell:
- Multiply the row total by the column total and divide by *n*. The expected count for the first cell is, therefore, (75 × 42) ÷ 125 = 25.2.
- If you perform this operation for each cell, you get the expected counts (in parentheses) shown in Table 2.

Table 2. Chi-Square Results for Table

	A	B	C	Totals
Boys	30 (25.2)	29 (37.2)	16 (12.6)	75
Girls	12 (16.8)	33 (24.8)	5 (8.4)	50
Totals	42	62	21	125

- The formula for χ^2, compares each cell's actual count to its expected count:

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

- The formula describes an operation that is performed on each cell and which yields a number. When all the numbers are summed, the result is χ^2. Now, compute it for the six cells in the example:

$$\chi^2 = \frac{(30-25.2)^2}{25.2} + \frac{(29-37.2)^2}{37.2} + \frac{(16-12.6)^2}{12.6} +$$

$$\frac{(12-16.8)^2}{16.8} + \frac{(33-24.8)^2}{24.8} + \frac{(5-8.4)^2}{8.4}$$

$$= 0.914 + 1.808 + 0.917 + 1.371 + 2.711 + 1.376$$

$$= 9.098$$

Table 2. Chi-Square Results for Table

	A	B	C	Totals
Boys	30 (25.2)	29 (37.2)	16 (12.6)	75
Girls	12 (16.8)	33 (24.8)	5 (8.4)	50
Totals	42	62	21	125

Example 5

- The larger χ^2, the more likely that the variables are related; note that the cells that contribute the most to the resulting statistic are those in which the expected count is very different from the actual count.

- Chi-square has a probability distribution, the critical values for which are listed in χ^2 Statistics Table.

- " As with the *t*-distribution, χ^2 has a degrees-of-freedom parameter, the formula for which is (number of rows – 1) × (number of columns – 1)

- or in your example: (2 – l) × (3 – 1) = 1 × 2 = 2

Conclusion

- Based on the table," a chi-square of 9.097 with two degrees of freedom falls between the commonly used significance levels of 0.05 and 0.01.

- If you had specified an alpha of 0.05 for the test, you could, therefore, reject the null hypothesis that gender and favorite commercial are independent.

- At a = 0.01, however, you could not reject the null hypothesis.

Accept
Hypothesis

Reject
Hypothesis

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09

Last

- The χ^2 test does not allow you to conclude anything more specific than that there is some relationship in your sample between gender and commercial liked (at $\alpha = 0.05$). Examining the observed versus expected counts in each cell might give you a clue as to the nature of the relationship and which levels of the variables are involved. For example, Commercial B appears to have been liked more by girls than boys. But χ^2 tests only the very general null hypothesis that the two variables are independent.

-

chi² test

- **Used to test strength of association between <u>qualitative</u> variables**
- **Used for categorical data**

Requirements

- **Data should be in form of frequency**
- **Total number of observed must exceed 20**
- **Expected frequency in one category or in any cell must be >5 (When 1 of the cells have <5 in observed yats correction) or if (When 1 of the cells have <5 in expected fischer exact)**

Hypothesis Testing

- Alpha (α) – level of probability needed to reject H_O; usually set at .05.

- p value – from computer; the probability that you could have gotten the data that you have if H_O is true.

- If p < α, then reject H_O.

Hypothesis Testing

Truth Table for Null Hypothesis:

	H_O True	H_O False
Accept H_O	Correct Decision	Type II error (β)
Reject H_O	Type I error (α)	Correct Decision

Hypothesis Testing

- α = probability of making a Type I error (if H_O is true).

  - If H_O is true, just due to chance, we could get values that fall in the region of rejection.
  - The probability of this happening (if H_O is true) = α.

- Type I error – incorrectly rejecting H_O, when H_O is true.

- Type II error – incorrectly accepting H_O when H_O is false.

- β = probability of making a Type II error when H_O is false.

- Power = probability of rejecting H_O when H_O is false.

Hypothesis Testing

- Tails – where to put the region of rejection?
  - Two-tailed test: hypothesis is non-directional, i.e., we make no claim that one condition is better then another, only that they are different; 50% of α goes in each tail.
  - One-tailed test: hypothesis is directional, i.e., we test that one condition is better then the other; all of α goes in one end of the distribution.

Student's t-Test Table

(1 tail)	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
(2 tail)	0.1	0.05	0.02	0.01	0.005	0.002	0.001
df							
1	6.3138	12.7065	31.8193	63.6551	127.3447	318.493	636.405
2	2.92	4.3026	6.9646	9.9247	14.0887	22.3276	31.5989
3	2.3534	3.1824	4.5407	5.8408	7.4534	10.2145	12.9242
4	2.1319	2.7764	3.747	4.6041	5.5967	7.1732	8.6103
5	2.015	2.5706	3.365	4.0322	4.7734	5.8934	6.8688
6	1.9432	2.4469	3.1426	3.7074	4.3168	5.2076	5.9589
7	1.8946	2.3646	2.998	3.4995	4.0294	4.7852	5.4079
8	1.8595	2.306	2.8965	3.3554	3.8325	4.5008	5.0414
9	1.8331	2.2621	2.8214	3.2498	3.8325	4.2969	4.7809
10	1.8124	2.2282	2.7638	3.1693	3.6896	4.2969	4.5869
11	1.7959	2.201	2.7181	3.0158	3.4966	4.0247	4.4369
12	1.7823	2.1788	2.681	3.0545	3.4284	3.9296	4.3178
13	1.7709	2.1604	2.6503	3.0123	3.3725	3.852	4.2208
14	1.7613	2.1448	2.6245	2.9768	3.3257	3.7874	4.1404
15	1.753	2.1314	2.6025	2.9467	3.286	3.7328	4.0728
16	1.7459	2.1199	2.5835	2.9208	3.252	3.6861	4.015
17	1.7396	2.1098	2.5669	2.8983	3.2224	3.6458	3.9651
18	1.7341	2.1009	2.5524	2.8784	3.1966	3.6105	3.9216
19	1.7291	2.093	2.5395	2.8609	3.1737	3.5794	3.8834
20	1.7247	2.086	2.528	2.854	3.1534	3.5518	3.8495