Worksheet 1—Normal curves

Aim
To explore the effects of changing the mean and standard deviation on the shape of the normal distribution.

Activity 1: Changing the mean
• Use Samsim to draw a normal curve with a mean ($\mu$) = 10 and a standard deviation ($\sigma$) = 2.
  1. Select Normal Distributions from the Main Menu.
  2. Select Normal Curves from the Normal Distributions Menu.
  3. You will now have on your screen two sliders labelled mean and standard deviation. The value for the mean is set at 10.0 and the value for the standard deviation is set at 2.0. Press the Enter key and a Normal Curve will be drawn using these values.
  4. Sketch the graph displayed in the section provided. Make sure you include the mean value along the X axis.

• Learn how to adjust the value on the mean slider.
  1. Press the Enter key and the sliders with again be displayed. You can change the values in the sliders using the following keys: Home, End, Page Up, Page Down, Left Arrow (←) and Right Arrow (→).
  2. Press the Home key and record below the value displayed in the mean window.
3. Press the End key and record the value displayed.

4. Press the Page Down key and circle below the number of units by which the value changes.
   10 5 1 .5 1 0.1 0.05 0.01

5. Move the sliding bar until the value displayed is 10.0. Press the Page Up key twice and record the value displayed.

6. Press the left arrow key and circle the number of units by which the value displayed changes.
   10 5 1 .5 1 0.1 0.05 0.01

7. Press the right arrow key and circle the number of units by which the value displayed changes.
   10 5 1 .5 1 0.1 0.05 0.01

8. If your computer has a mouse installed move the mouse pointer until it is approximately over the number 15. Click the mouse. If the value displayed is not 15.0 use the arrow keys to adjust the value.

**Summary**

Use the Home and End keys to move to the lowest and highest values use the arrow keys to adjust the value by a small amount, use Page Up and Page Down for large adjustments or use a combination of the mouse and the arrow keys.

• Use the slider to change the mean to 15. Leave $\sigma$ at 2.

1. If the value displayed on the mean slider is not 15 use the appropriate keys to select that value. Leave the standard deviation at 2. Press the Enter key and a normal curve will be drawn using these values.

2. Sketch the resulting graph on the same axis as the previous graph. Make sure you include the mean value along the X axis.

• Change the mean to 5. Leave $\sigma$ at 2.

1. Press Enter to display the sliders.

2. Change the value on the mean slider to 5. Press the Enter key and a normal curve will be drawn using these values.

3. Sketch the resulting graph on the same axis as the previous graphs. Include the mean value.

• What effect does changing the mean have on the curve? Does it change the shape of the normal distribution?
Clear the graphs from the screen by pressing the F5 button.

**Activity 2: Changing the standard deviation**

- Use Samsim to draw a normal curve with a mean \((\mu) = 10\) and a standard deviation \((\sigma) = 2\).

1. You will now have on your screen two sliders labelled mean and standard deviation. The value for the mean is set at 10.0 and the value for the standard deviation is set at 2.0. Press the Enter key and a Normal Curve will be drawn using these values.

2. Sketch the graph displayed in the section provided. Make sure you include the mean along the X axis and the standard deviation near the graph.

3. Press Enter to display the sliders. Press the down arrow (\(\downarrow\)) to move to the standard deviation slider. Use the mouse or the appropriate key to change the value of the standard deviation to 4. Press the Enter key and a normal graph will be drawn using these values.

4. Sketch the resulting graph on the same axis as the previous graphs. Include the standard deviation near the graph.

5. Press Enter to display the sliders. Move to the standard deviation slider. Change the value to 1. Leave the mean value at 10.0. Press Enter and a graph will be drawn using these values.

6. Sketch the resulting graph on the same axis as the previous graphs. Include the standard deviation near the graph.

7. Sketch on the same axis where a normal curve with a mean = 12 and a standard deviation = 2 would be positioned.

   Check your answer using Samsim.

- Make a statement about how changing this parameter effects the shape of the normal distribution.

......................................................................................................................................................
......................................................................................................................................................

......................................................................................................................................................
......................................................................................................................................................

......................................................................................................................................................
......................................................................................................................................................
• Guess the shape of the distribution if you changed the standard deviation to 10. Check your guess using Samsim.
Worksheet 2—Probability areas

**Aim**
To explore the concept that area under a normal curve represents probability for a continuous variable.

**Scenario**
You are working as the quality control officer for a company producing tea for export. This month you are particularly interested in quality aspects related to the filling and weighing of the large export packages containing tea. The actual amount of tea that the filling machine puts into a 20 kg package varies from package to package, so the actual fill may be considered a random variable, with a standard deviation of 1 kg. Use Samsim to assist you in solving the following problems.

**Problem 1**
Each tea packet holds a maximum weight of 23 kg. What proportion of package will overflow?

**Activity 1**
Sketch the area under the normal curve that will represent this proportion.
Using Samsim to check this area:

1. Select Normal Distributions from Main Menu.
2. Select Probability Areas from Normal Distributions Menu.
3. Set mean = 20.0. Use the down arrow key to move to the standard deviation slider. Set standard deviation = 1.0. Press the Enter key.
4. The Select Range menu appears. You are interested in the Greater Than range, as you want to know the area greater than 23. Press Enter to select this range.
5. Set the x value to 23 (use Page Up key). Press Enter.
6. The proportion of packages that will overflow is given as the area shaded value. This area represents the proportion greater than 23.
7. Record this value.

Problem 2

Consumer organisations can rightfully complain if they buy an export packet of tea and it weighs less than 17.5 kg. What proportion of the packages could cause this problem?

Activity 2

Sketch the area under the normal curve that will represent this proportion.

Using Samsim to check this area:

1. Press Enter to display the sliders. Leave the mean = 20 and the standard deviation = 1. Press Enter.
2. The Select Range menu appears. You are interested in the Less Than range, as you want to know the area less than 17.5. Use the down arrow key to move to this option and then press Enter to select this range.
3. Set the x value to 17.5 (use Page Down key). Press Enter.
4. The proportion of packages that you are interested in is given as the area shaded value. This area represents the proportion less than 17.5.
5. Record this value.
Problem 3

In a press release you want to be able to say x% of our export packets contain between 17.5 and 22.5 kg of tea. What is x?

Activity 3

Sketch the area under the normal curve that will represent this proportion.

Using Samsim to check this area:

1. Press Enter to display the sliders. Leave the mean = 20 and the standard deviation = 1. Press Enter.

2. The Select Range menu appears. You are interested in the Between range, as you want to know the area between 17.5 and 22.5. Use the down arrow key to move to this option and then press Enter to select this range.

3. Set the x₁ value to 17.5.
Set the x₂ value to 22.5.
Press Enter.

4. The proportion of packages that you are interested in is given as the area shaded value. This area represents the proportion between 17.5 and 22.5.

5. Record this value. ...................................................
**Problem 4**

In quality control terms, you consider that the filling machine may be malfunctioning if it is producing packets of tea weighing greater than 23.2 kg or less than 16.7 kg. Even when the machine is operating normally (i.e. not malfunctioning) what percentage of time will it produce such packages?

**Activity 4**

1. Press Enter to display the sliders. Leave the mean = 20 and the standard deviation = 1. Press Enter.

2. The Select Range menu appears. You are interested in the Outside range, as you want to know the area outside 16.7 and 23.2. Use the down arrow key to move to this option and then press Enter to select this range.

3. Set the $x_1$ value to 16.7.
   
   Set the $x_2$ value to 23.2.
   
   Press Enter.

4. The proportion of packages that you are interested in is given as the area shaded value. This area represents the proportion outside 16.7 and 23.2.

5. Record this value. ..........................

6. Press F6 twice to return to The Main Menu, or F1 to exit.
Worksheet 3—Sampling

Aims

- To demonstrate how the elements of a sampling distribution for the means are established.
- To explore the effects of changing the sample size on the shape of the sampling distribution.

Activity 1: Sampling simulation

1. Select Sampling Distributions from Main Menu.
2. Select Sampling Simulation from Sampling Distributions Menu.
3. Three sliders will be displayed, Population Size, Sample Size, and Number of Samples Drawn. Using the arrow keys, Home, End, Page Up, Page Down or the mouse:
   - set Population Size to 60,
   - set Sample Size to 10,
   - set Number of Samples Drawn to 30.

   Press the Enter key to start the simulation. The F5 function key turns the animation process off and thereby speeds up the sampling process.
4. Observe the sampling process and how the mean of the sample of 10 values is calculated and becomes an element in the new sampling distribution. Observe how the value of the sample means compare with the population mean.
5. Observe the distribution of the x bar values.
6. Make a comment on the shape of the distribution of these values.

......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
Activity 2: Varying the sample size

1. Press the Enter key to display the sliders.

2. Use the mouse or appropriate keys to change the sliders to display the following values:
   - set Population Size to 60
   - set Sample Size to 5
   - set Number of Samples Drawn to 100.

   Press the Enter key to start the simulation. The F5 function key turns the animation process off and thereby speeds up the sampling process.

3. Sketch the shape of the distribution of x bar on the axis below. Mark on it the smallest and largest values of x.

4. Press the Enter key to display the sliders once again.

5. Use the mouse or appropriate keys to change the sliders to display the following values:
   - set Population Size to 60
   - set Sample Size to 30
   - set Number of Samples Drawn to 100.

   Press the Enter key to start the simulation. The F5 function key turns the animation process off and thereby speeds up the sampling process. It will take a while to complete this sampling process. While the distribution is building up try to guess how its shape will differ from that of the distribution of sample size 10. Write this down.

   ......................................................................................................................................................
   ......................................................................................................................................................

6. Sketch on the same axis as before the shape of this sampling distribution. Mark on it the smallest and largest values of x bar.

7. Comment on how these distributions differ.

8. Make a statement as to why this is so.
Worksheet 4—Sampling from a normal population

Aims

• To demonstrate the shape of the sampling distribution of the mean when sampling from a normal population.
• To demonstrate the effect of changing the sample size on the spread of the sampling distribution.

Activity 1: Sampling distribution shape

1. Select Sampling Distributions from Main Menu.
2. Select Normal Distribution from Sampling Distributions Menu.
3. Two sliders will be displayed; Sample Size, and Number of Samples Drawn. Using the arrow keys, Home, End, Page Up, Page Down or the mouse:
   • set Sample Size to 30,
   • set Number of Samples Drawn to 100.

   Press the Enter key to start the sampling process.
4. Observe the values of the sample means and compare this to population mean.

   When the sampling is complete, use the histogram and the graphs of the distribution to answer the following questions.
5. Write down a description of the differences between the general shape of the histogram of the population means and the population distribution shape.

   ......................................................................................................................................................
   ......................................................................................................................................................
6. Explain why these differences occur.

   ......................................................................................................................................................
   ......................................................................................................................................................
7. What would you expect the mean and standard deviation of the sampling distribution to be?

   ......................................................................................................................................................
   ......................................................................................................................................................
8. Sketch the histogram of the sample means below.
Activity 2: Change sample size

1. Press the Enter key to display the sliders.

2. Use the mouse or appropriate keys to change the sliders to display the following values:
   - set Sample Size to 10
   - set Number of Samples Drawn to 100.

   Press the Enter key to start the sampling process.

3. What would you expect the mean and standard deviation of the sampling distribution to be?

   ......................................................................................................................................................
   ......................................................................................................................................................

4. Sketch the histogram of the sample means below.

   ......................................................................................................................................................
   ......................................................................................................................................................

5. Compare the shape of the histogram above for samples of size 10 to the shape of the histogram for samples of size 30. Describe the differences in shape.

   ......................................................................................................................................................
   ......................................................................................................................................................

6. Explain why these differences exist.

   ......................................................................................................................................................
   ......................................................................................................................................................
Activity 3: Change number of samples drawn

1. Press the Enter key to display the sliders.

2. Use the mouse or appropriate keys to change the sliders to display the following values:
   - set Sample Size to 30
   - set Number of Samples Drawn to 50.
   Press the Enter key to start the sampling process.

3. What would you expect the mean and standard deviation of the sampling distribution to be?

   ......................................................................................................................................................
   ......................................................................................................................................................

4. Sketch the histogram of the sample means below.

5. The histogram provides a discrete approximation to the probability distribution of the continuous random variable. Comment on how close the histogram shape is to the ideal sampling distribution.

   ......................................................................................................................................................
   ......................................................................................................................................................

6. Press the Enter key to display the sliders.

7. Use the mouse or appropriate keys to change the sliders to display the following values:
   - set Sample Size to 30
   - set Number of Samples Drawn to 300.
   Press the Enter key to start the sampling process.

8. What would you expect the mean and standard deviation of the sampling distribution to be?

   ......................................................................................................................................................
   ......................................................................................................................................................

9. Sketch the histogram of the sample means below.
10. How close does the histogram fit the ideal sampling distribution shape?

......................................................................................................................................................

......................................................................................................................................................

11. Is it a better fit for 50 or 300 samples drawn? .........................
    Comment why this may be so?

......................................................................................................................................................

......................................................................................................................................................

12. What shape would the histogram tend to if you could simulate a very large number of samples?

......................................................................................................................................................

......................................................................................................................................................
Worksheet 5—Sampling from non-normal populations

Aim

To demonstrate the central limit theorem, that is, if the sample size is large the sampling distribution of the mean is normal even if the parent population is not normal.

Activity 1: Sampling from negative exponential

1. Select Sampling Distributions from Main Menu.
2. Select Negative Exponential Distribution from Sampling Distributions Menu.
3. Two sliders will be displayed; Sample Size, and Number of Samples Drawn. Using the arrow keys, Home, End, Page Up, Page Down or the mouse:
   • set Sample Size to 5,
   • set Number of Samples Drawn to 100.
   Press the Enter key to start the sampling process.
4. Sketch the shape of the histogram on the first set of axes on the following page. Note the largest and smallest x bar value.
   **Repeat the experiment with a sample size of 10.**
5. Press the Enter key to display the sliders.
6. Use the mouse or appropriate keys to change the sliders to display the following values:
   • set Sample Size to 10
   • set Number of Samples Drawn to 100.
   Press the Enter key to start the sampling process.
7. Sketch the shape of the histogram on the second set of axes on the attached page. Note the largest and smallest x bar value.
Repeat the experiment with a sample size of 30.

8. Press the Enter key to display the sliders.

9. Use the mouse or appropriate keys to change the sliders to display the following values:
   • set Sample Size to 30
   • set Number of Samples Drawn to 100.

Press the Enter key to start the sampling process.

10. Sketch the shape of the histogram on the third set of axes on the attached page. Note the largest and smallest \( \bar{x} \) value.

11. Observe the three histograms. Describe what happens to the general shape of the histogram as the sample size increases.
Activity 2: Sampling from uniform distribution

1. Press F6 to return to the Sampling Distributions Menu.
2. Select Uniform Distribution from Sampling Distributions Menu.
3. Two sliders will be displayed: Sample Size, and Number of Samples Drawn. Using the arrow keys, Home, End, Page Up, Page Down or the mouse:
   - set Sample Size to 5,
   - set Number of Samples Drawn to 100.
   Press the Enter key to start the sampling process.
4. Sketch the shape of the histogram on the first set of axes on the following page. Note the largest and smallest x bar value.
   **Repeat the experiment with a sample size of 10.**
5. Press the Enter key to display the sliders.
6. Use the mouse or appropriate keys to change the sliders to display the following values:
   - set Sample Size to 10
   - set Number of Samples Drawn to 100.
   Press the Enter key to start the sampling process.
7. Sketch the shape of the histogram on the second set of axes on the following page. Note the largest and smallest x bar value.
   **Repeat the experiment with a sample size of 30.**
8. Press the Enter key to display the sliders.
9. Use the mouse or appropriate keys to change the sliders to display the following values:
   - set Sample Size to 30
   - set Number of Samples Drawn to 100.
   Press the Enter key to start the sampling process.
10. Sketch the shape of the histogram on the third set of axes on the following page. Note the largest and smallest x bar value.
    8. Observe the three histograms. Describe what happens to the general shape of the histogram as the sample size increases.

......................................................................................................................................................
......................................................................................................................................................
9. From the results of activities 1 and 2 make a statement about the sampling distribution of the mean when sample size is large and parent populations are non normal.

......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................
......................................................................................................................................................

10. These activities have demonstrated the Central Limit Theorem. Do some research and write down reasons why the Central Limit Theorem is so important.
Worksheet 6: Hypothesis on a normal mean, \( s \) known

Aim

To guide the user through a set of conditions that lead to a

“Z test and estimator of \( \mu \)”

being the appropriate hypothesis procedure for a particular problem.

Problem

The president of a union claims that her health care workers (whose average income is $45,000) are underpaid, since she believes the average annual health care income for the country exceeds $45,000.

Management claims the workers are well paid as they believe the average annual health care workers’ income for the country is less than $45,000. An independent arbitrator believes the average is quite close to $45,000 and that the standard deviation of the population of health care workers incomes is $5,000. The population of incomes is normally distributed.

The disputants and the arbitrator decide to base their decision on a sample of size 100 drawn from the population of all health care workers in the country. Taking a simple random sample of 100 such health care workers annual incomes gives a mean annual income of $43,500.

Management claims the average annual income for health care workers is less than $45,000 while the union believes that the average annual income for health care workers exceeds $45,000. Test the hypothesis that the average annual income of health care workers in this country is $45,000 against the alternative that it is not $45,000. Use \( \alpha = 0.05 \).

Use StatTest to assist you to determine the appropriate hypothesis testing procedure for this problem.
Activity

• Move to the Data Type screen.
• The problem asks whether the average income is $45,000. This is an interval data type.
• Use the down arrow key to highlight the interval option. Press F8 to display information about interval data. Press F6 to move back to the Data Type screen. Press F5 to see examples of different type of interval data. Press F6 to move back to the Data Type screen.
• With ‘Interval’ highlighted press the Enter key to select ‘Interval’ data type.
• The problem is looking at only one population—annual average income of health care workers. Press F8 to display information on single population, press F6 to return to Type of Problem screen. Press F5 to see an example, press F6 to return to Type of Problem screen.
• Press Enter to select ‘single population’.
• The problem stated you could assume the population was normally distributed. Press F8 for information. Press F6 to return to Type of Population screen.
• Press Enter to select for a ‘normally distributed population’.
• At this stage you might like to press the F4 function button. You will have displayed on your screen a tree like structure that shows you all the decisions you have made so far. Press F6 to return back to the Decision Screen. You will now continue to make decisions concerning your data.
• The problem requires the testing of whether the average annual income is $45,000 against the hypothesis that it is not. This is a measure relating to the mean. Press F8 to display information and press F6 to return to Type of Information screen.
• Press Enter to select ‘mean’.
• The problem stated that the standard deviation of the population’s income was $5,000. Press F8 to see information and then F6 to return.
• Press Enter to select ‘yes’ the standard deviation of the population is known.
• The Advice given is to use:

  **Z-test and estimator of μ**

• Now apply this test to your data. Press F5 to see an example of how to use this test.
Worksheet 7: Hypothesis on a normal mean, \( \sigma \) unknown

Aim

To guide the user through a set of conditions that lead to a

“t test and estimator of \( \mu \) (x bar)”

being the appropriate hypothesis procedure for a particular problem.

Problem

A dietitian claims that the average Australian is at least 4 kg overweight. To test
his claim a random sample of 50 Australians are weighed, and the difference
between their actual and ideal weight was calculated. The mean and standard
deviation was 3.87 and 1.5 kg respectively.

Assuming that the population of weight differences is normally distributed, can
we conclude with \( \alpha = 0.01 \), that there is enough evidence to show his claim is true?

Use StatTest to assist you to determine the appropriate hypothesis testing
procedure for this problem.

Activity

• Move to Data Type screen.
• The problem is looking at whether Australians are at least 4 kg overweight. This
  is an interval data type. Use the down arrow key to move to this option
• Press F8 to display information relating to interval data or/and press F5 to view
  examples of interval data. Press F6 to return to the Data Type screen.
• The problem involves a single population - Australians. Press F8 to see
  information about a single population, press F5 to view an example of a Single
  Population with Interval Data. Press F6 to return to Type of Problem screen.
  Select ‘single population’ by pressing Enter.
• The F4 (Decisions Made) function key allows you to view the Decisions Made so
  far, while the F7 (Tree) function key displays all the options available for the
  current Data Type ( in this instance Interval) and the final test used.
  Use the function keys when you wish to view extra information.
• The problem stated that the population of weight differences was normally
distributed. Select ‘normally distributed’ from the menu options.
• The problem is looking at the mean difference between actual and ideal weight.
  Therefore, the ‘Type of Information’ to select is ‘mean’.
• The problem states the standard deviation of the sample is 1.5 kg. However, the standard deviation of the population is not known. Therefore, select 'no', the standard deviation of the population is not known.

• The Advice given is to use:
  
  t–test and estimator of μ

• Now apply this test to your data. Press F5 to see an example of how to use this test.
Worksheet 8: Hypothesis on a binomial p

Aim

To guide the user through a set of conditions that lead to a

“Z test and estimator of p”

being the appropriate hypothesis procedure for a particular problem.

Problem

A package designer for a biscuit company wishes to determine whether there is any difference in the way customers react to two new packet designs. She places equal numbers of packets on the shelf in a supermarket and observes how many of each packet are purchased each day.

She rotates the position of the two packet types from one day to the next to eliminate the effect of shelf position and makes sure that there are always equal numbers of the two packet types on the shelf.

After 500 purchases have been made, 200 of the first packet design and 300 of the second design have been purchased. At the $\alpha = 0.05$ level test the hypothesis that there is no difference in the packet designs.

Use StatTest to assist you to determine the appropriate hypothesis testing procedure for this problem.

Advice

• Move to the Data Type screen.
• The problem is testing the hypothesis that there is no difference in packet design. This is nominal data—packets are either one design or the other (A or B).
• Press F8 to display information relating to nominal data, press F5 to view examples of nominal data. Press F6 to return to the Data Type screen.
• Press Enter to select ‘Nominal’ data.
• There is only one population—packets of biscuits. Press F8 to see information about a single population, press F5 to view an example of a Single Population with Nominal Data. Press F6 to return to Type of Problem screen.
• Press Enter to select ‘single population’.
• The data can be allocated to two categories—packets of type A or type B. Press F8 to view information on types of categories. Press F6 to return to Number of Categories screen. Press Enter to select ‘two categories’.
• The Advice given is to use:

  **Z-test and estimator of p**

• Now apply this test to your data. Press F5 to see an example of how to use this test.
Worksheet 9: Hypothesis on the difference between two normal means

Aim
To guide the user through a set of conditions that lead to a
“Z test and estimator of $\mu_1$ and $\mu_2$”
being the appropriate hypothesis procedure for a particular problem.

Problem
The selection of a new store location depends on many factors, one of which is the level of household incomes in the area around the proposed site. Suppose that a large department store chain is trying to decide whether to build a new store in Montlene or in the nearby city of Glades. Building costs are lower in Montlene and the company decides it will build there unless the average household income is higher in Glades than in Montlene.

In a survey of 100 residences in each of the cities the mean annual household income was $29,980 in Montlene and $28,650 in Glades. From other sources it is known that the population standard deviation for annual household incomes are $4,740 in Montlene and $5,365 in Glades. Assume that incomes are normally distributed.

At the 5% significance level can it be concluded that the mean household income in Montlene exceeds that of Glades.

Use StatTest to assist you to determine the appropriate hypothesis testing procedure for this problem.

Activity
• Move to Data Type screen.
• The problem is looking at incomes—this in interval data. Select ‘Interval’ on the menu.
• There are two populations—Montlene and Glades. Use the down arrow key to highlight ‘comparison of two populations’. Press F8 to see information about comparing two population. Press F5 to see an example on Comparing Two Population using Interval Data.
• Select the option ‘comparison of two populations’.
• The problem has assumed that the incomes are normally distributed.
• Select ‘normally distributed’.
• The problem is concerned about mean incomes. The Type of Information to be tested is ‘mean’. Select this option.
• The samples are independent as they are two distinct populations, one in Montlene and one in Glades. Press F8 to see information relating to independent samples. Press F6 to return to the Experimental Design screen.

• Select ‘independent samples’.

• The population standard deviations are given:
  Montlene = $4,740
  Glades = $5,365

• Select ‘yes’, the standard deviations of the two populations are known.

• The Advice given is to use:

  **$Z$-test and estimator of $\mu_1$ and $\mu_2$**

• Now apply this test to your data. Press F5 to see an example of how to use this test.
Worksheet 10: Hypothesis for independence of categorical variables

Aim
To guide the user through a set of conditions that lead to a “χ² test of a contingency table” being the appropriate hypothesis procedure for a particular problem.

Problem
Alcohol and nicotine consumption during pregnancy may harm children. Because drinking and smoking behaviours may be related, it is important to understand the nature of this relationship when assessing the possible effects on children.

One study classified 452 mothers according to their alcohol intake prior to pregnancy recognition and their nicotine intake during pregnancy. The data are summarised in the following table.

<table>
<thead>
<tr>
<th>Alcohol (ounces/day)</th>
<th>Nicotine (milligrams/day)</th>
<th>None</th>
<th>1-15</th>
<th>16 or more</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td></td>
<td>105</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>.01-.10</td>
<td></td>
<td>58</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>.11-.99</td>
<td></td>
<td>84</td>
<td>37</td>
<td>42</td>
</tr>
<tr>
<td>1.00 or more</td>
<td></td>
<td>57</td>
<td>16</td>
<td>17</td>
</tr>
</tbody>
</table>

(Data from Ann P. Streissguth et al. 1984, “Intrauterine alcohol and nicotine exposure: Attention and reaction time in 4-year-old children”, Developmental Psychology, 20, pp. 533-541.)

Using this data determine if the variables alcohol intake prior to pregnancy recognition and nicotine intake during pregnancy are independent. Use α = .05.

Use StatTest to assist you to determine the appropriate hypothesis testing procedure for this problem.

Activity
• Move to Data Type screen.
• The problem is looking at whether there is a relationship between alcohol and nicotine intake. There are several categories or “levels” of alcohol and nicotine intake and the data represents frequency counts in these categories. This is nominal data.
• Press F8 to display information about nominal data. Press F6 to move back to the Data Type screen. Press F5 to see examples of different type of nominal data. Press F6 to move back to the Data Type screen.
• With ‘Nominal’ highlighted press the Enter key to select ‘Nominal’ data type.
• The problem is determining if there is a relationship between the variables nicotine intake and alcohol intake. In other words whether they are independent or not. Move to the option ‘analysis of relationship 2 variables’. Press F8 to display information, press F6 to return to ‘Type of Problem’ screen. Press F5 to see an example, press F6 to return to Type of Problem screen.
• Ensure that the option ‘analysis of relationship 2 variables’ is highlighted and press Enter to select.
• The Advice given is to use:
  \[ \chi^2 \text{ test of a contingency table} \]
• Now apply this test to your data. Press F5 to see an example of how to use this test.
Worksheet 11: Hypothesis for relationship between two variables, interval data

Aim

To guide the user through a set of conditions that lead to
“Regression and Correlation”
being the appropriate testing procedure for a particular problem.

Problem

The human body takes in more oxygen when exercising than when it is at rest. To
deliver the oxygen to the muscles, the heart must beat faster. Heart rate is easy to
measure, but measuring oxygen uptake requires elaborate equipment. If oxygen
uptake (VO2) can be accurately predicted from heart rate (HR), the predicted
values can replace actually measured values for various research purposes.

Unfortunately, not all human bodies are the same, so no single prediction
equation works for all people. Researchers can, however, measure both HR and
VO2 for one person under varying sets of exercise conditions and calculate a
regression equation for predicting that person’s oxygen uptake from heart rate.
They can then use predicted oxygen uptakes in place of measured uptakes for this
individual in later experiments.

Here are data for one individual.

<table>
<thead>
<tr>
<th>HR</th>
<th>94</th>
<th>96</th>
<th>95</th>
<th>95</th>
<th>94</th>
<th>95</th>
<th>94</th>
<th>104</th>
<th>104</th>
<th>106</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO2</td>
<td>.473</td>
<td>.753</td>
<td>.929</td>
<td>.939</td>
<td>.832</td>
<td>.983</td>
<td>1.049</td>
<td>1.176</td>
<td>1.292</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HR</th>
<th>108</th>
<th>110</th>
<th>113</th>
<th>113</th>
<th>118</th>
<th>115</th>
<th>121</th>
<th>127</th>
<th>131</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO2</td>
<td>1.403</td>
<td>1.499</td>
<td>1.529</td>
<td>1.599</td>
<td>1.749</td>
<td>1.746</td>
<td>1.897</td>
<td>2.040</td>
<td>2.231</td>
</tr>
</tbody>
</table>

(Data provided by Paul Waldsmith from experiments conducted in Don Corrigan’s laboratory at Purdue University.)

Using this data test to see if there is correlation between the variables heart rate
and oxygen uptake. Use $\alpha = .01$.

Use StatTest to assist you to determine the appropriate hypothesis testing
procedure for this problem.
Activity

- Move to Data Type screen.
- The problem is looking at interval data.
- Move to the interval option. Press F8 to display information about interval data. Press F6 to move back to the Data Type screen. Press F5 to see examples of different type of interval data. Press F6 to move back to the Data Type screen.
- With ‘Interval’ highlighted press the Enter key to select ‘Interval’ data type.
- The problem is looking at whether there is a relationship between oxygen uptake (VO2) and heart rate (HR). This is an ‘analysis of the relationship between two variables’. Move to this option and press F8 to display information, press F6 to return to Type of Problem screen. Press F5 to see an example, press F6 to return to Type of Problem screen.
- Ensure that ‘analysis of the relationship between two variables’ is highlighted and press Enter to select this option.

The Advice given is to use:

Regression and Correlation

- Now apply this test to your data. Press F5 to see an example of how to use this test.