LABORATORY 1

Practice with Microsoft Excel: Collecting and Analyzing Data on Risk Perception and the Availability Heuristic

Purpose

- to introduce the concepts of the **availability heuristic** and **risk perception**;
- to give you experience in tabulating, summarizing and graphing data using Microsoft Excel;
- to give you practice with calculating and interpreting the **Pearson product-moment correlation coefficient**, ***r***.

Introduction

One of the determinants of what we choose to do in some situations is our perception of the risk associated with taking a particular course of action, being exposed to a particular agent, or being placed in a particular situation. For example, some people refuse to fly in airplanes because they feel that doing so puts them in an unacceptably risky situation. Of course, our perception of risk does not always correspond to actual risk. Transportation experts, for example, calculate that we are at much greater risk of dying when we drive than when we fly.

Of course, the calculations of experts are based on as much objective data as possible. A key issue for understanding people's reactions in situations involving risk is knowing what information they have about the risk or hazard associated with the situation. Psychologists who study risk perception try to address this issue in various ways, and one way is to ask for judgments of relative risk or relative hazard. In such judgments, people are presented with pairs of alternatives that differ in objective risk and choose which one of the alternatives is perceived as the more risky or more hazardous.

The following page contains a questionnaire that asks for judgments of the relative *lethality* of pairs of events. Each judgments presents the names of two potential causes of death, and the task for the respondent is to select which of the two is a more frequent cause of death. To provide a little more information about the relative danger associated with each lethal event, the respondents are also asked to give a "risk ratio" indicating how many times more deaths are caused by one event than the other. For example, if one event is thought to cause twice as many deaths as the other, the number for the risk ratio would be 2, for a 2 : 1 ratio.

Getting Started and Collecting Data

When you are instructed to do so, turn to the Lethal Events Questionnaire and read the instructions for completing it. Then, as indicated by your instructor, select in each pair the event that is the most frequent cause of death in the United States and, after making your selection, indicate how many times more deaths are caused by your selection than by the other alternative. When you are finished with the entire survey, turn to the section in this manual on "Data Analysis".

LETHAL EVENTS QUESTIONNAIRE

Instructions: Please write your answers on this sheet. In each question, you are presented with a pair of events which are potentially lethal. Your task is, first, to decide which of the two events causes more deaths per year in the US, and, second, to indicate how many times more deaths are caused by that event than by the other alternative.

Example. You are presented with wildfires and hurricanes - decide which event causes more deaths per year in the US and circle that event. After you circle the more frequent "lethal event", estimate how many times more deaths it causes annually than the other, less frequent event. If you think more people die of from wildfires than from hurricanes, circle wildfires. Then, if you think that wildfires cause 3 more deaths per year than hurricanes, write "3" in the blank.

Which of the following events causes more deaths per year?

1. Lightning or Tornado ?
   How many times more deaths ? _____

2. Suicide or Diabetes ?
   How many times more deaths ? _____

3. Stomach Cancer or Drowning ?
   How many times more deaths ? _____

4. Motor Vehicle Accidents or Lung Cancer ?
   How many times more deaths ? _____

5. Leukemia or Fire ?
   How many times more deaths ? _____

6. Heart Disease or Accidents of any kind ?
   How many times more deaths ? _____

7. Accidental shooting or Skin cancer?
   How many times more deaths ? _____

8. Syphilis or Malaria ?
   How many times more deaths ? _____

9. Influenza or Alzheimer's Disease ?
   How many times more deaths ? _____

Data Analysis

The purpose of the data analysis is to determine the role of *availability* in judging how likely an event is as a cause of death. *Availability* refers to how easily an event comes to mind; in a famous paper, Amos Tversky and Daniel Kahnemann (1973) argued that people tend to rely on availability when judging the frequencies of events, giving higher judgments to events that more easily come to mind. Using availability as an index of frequency is an example of a **judgment heuristic**, a "rule-of-thumb" procedure that is easy to use and generally works. That is, in general, if something does not come to mind readily, it is probably not very frequent.

In the Lethal Events Questionnaire, each pair consists of one event that is presumed to have a high availability and another that has low availability, based on how likely the event is to be noted in a news item, but is objectively more frequent. The sequence of pairs in the questionnaire has also been ordered to exhibit increasing risk ratios between the event that is objectively more often a cause of death and the one that is less often a cause of death. (Data on frequencies of causes of death come from the National Vital Statistics Survey maintained by National Center for Health Statistics of the Centers for Disease Control. Links to the data can be found online at http://www.cdc.gov/nchs/about/major/dvs/mortdata.htm, except for the data on lightning strikes and tornadoes, which do not appear on the current tables. These estimates come from individual reports on the lethality of these events).

The data analysis consists of first determining how many choices were of the more available event and then of determining the risk ratio at which availability is overcome by objective frequency.

Table 1: Frequencies[a] and risk ratios for events in Lethal Events Questionnaire

More frequent event (F)		More available event (A)		Risk ratio
Cause of death	# deaths	Cause of death	# deaths	F / A
Lightning[b]	100	Tornado[b]	70	1.4
Diabetes	69,301	Suicide	29,350	2.4
Stomach Cancer	12,645	Drowning	3,482	3.3
Lung Cancer	155,521	Motor Vehicle	43,364	3.6
Leukemia	21,339	Fire	3,377	6.3
Heart Disease	710,760	Accidents, any kind	97,900	7.3
Skin cancer	7,420	Accidental shooting	776	9.6
Syphilis	58	Malaria	3	19.3
Alzheimer's	49,558	Influenza	1,765	28.0

[a] Data from CDC Final Report on Deaths for 2000

[b] Data on lightning strikes and tornadoes are not in the CDC Final Report for 2000 but are annual estimates based on CDC data from other papers.

Step 1. Score questionnaire.

Use Table 1 to identify which event in each pair was the more available event, then score the selection in each pair as a "1" if the more available event was chosen and as a "0" if the more frequent event was chosen.

If the selection was of the more frequent event, then write the reported risk ratio as the perceived risk ratio. However, if the selection was of the more available event, then calculate the inverse of the reported risk ratio (divide 1 by the number given by the respondent) to get the perceived risk ratio. For example, if the respondent selected "tornado" (the more available) over "lightning" (the more frequent) and reported that 2.5 times as many people are killed by tornadoes as by lightning, the perceived risk ratio would be 1 / 2.5, or .40

Step 2. Put results from 7 questionnaires in spreadsheets.

Set up a spreadsheet in Microsoft Excel with 10 columns and 9 rows. Starting in the second column of the first row, label the columns with the numbers 1 through 9, and starting in the second row of the first column, label the rows with the subject numbers S1, S2, and so on to S7. In the last row of the first column, enter the expression "P(Availability)".

Set up a second spreadsheet with 10 columns and 9 rows, organized like the first one, except label the last row of the first column with the expression "Mean Risk Ratio".

Get data from 6 other people's questionnaires to combine with your data, so the results for 7 respondents are entered in each of the two spreadsheets.

In the first spreadsheet, enter for each subject the "1"'s and "0"'s for all the test pair in the order the pairs were tested. In the second spreadsheet, enter the perceived risk ratios for each pair in the order the pairs were tested.

Step 3. Summarize results using the mean.

In the last row of each column of the spreadsheet, enter the formula that instructs Excel to calculate a mean value for the data of the 7 subjects in the table. This formula will look like the following:

=average(b2:b9)

The first cell address ("b2") is the cell with the data for the first subject, and the second cell

address ("b9") is the cell with data for the last subject. The colon sign (":") indicates that all values between the first and last cell will be included in the calculation.

Adjust the decimal points of your answer to 2 significant places to the right of the decimal point. Then copy the formula and paste it into the remaining 8 cells for the average value.

Repeat these steps for the second spreadsheet.

Step 4. Graph P(Availability) as a function of test pair.

Create a small table with 9 columns and 2 rows. In the top row, enter the numbers 1 to 9 for the 9 test pairs. Then, in the first column of the second row, enter the formula:

=b10

This formula instructs Excel to put the results of your calculation of the mean in that cell. Copy the formula and paste it into each of the remaining cells of this table.

Select all 18 cells (two rows of 9 columns) in this table, then click on the *Chart Wizard*, indicated by the icon with 3 small towers in the Excel Menu bar. Using the Chart Wizard involves 4 steps.

Step 1: Choose, as the type of Chart, "XY Scatter", with the option of "data points connected by lines", and click "Next".

Step 2: Inspect the sample graph to make sure it looks about right, and click "Next".

Step 3: Add a title, a label for the x-axis, and a label for the y-axis by clicking in each box and entering the text. Click on the tab marked "Gridlines", and uncheck the box labeled "Major gridlines", then click "Next".

Step 4: Click on "As new sheet" and enter "Risk Ratio" in the box, then click "Finish".

With your graph displayed, double-click anywhere in the gray interior of the graph to open a "Format Plot Area" window, and select the white square under "Area", then click "Ok".

Step 5. Graph Mean Risk Ratio as a function of Actual Risk Ratio.

Repeat the operations described in Step 4 to create another table with 2 rows of 9 columns for the spreadsheet with risk ratios, and make a graph showing the relation between

these two mean perceived risk ratio and actual risk ratio. Print the graphs from Step 4 and Step 5.

Step 6. Find the correlation of Mean Risk Ratio with Actual Risk Ratio.

Add to the table created in Step 5 a third row, in which the actual risk ratio is below the perceived risk ratio. In a cell just below the table, enter "Pearson r =" then press "Enter" and, with the next cell below highlighted, select the *Function Wizard*. In the Function Wizard window that opens, select "Statistical" from the "Function Category" column on the left and select "Pearson" from the "Function name" on the right, then click "OK". A window will open that has two open boxes, labeled "Array 1" and "Array 2", each with a small red arrow on the right hand side of the box. Click the red arrow for Array 1, then select all the values in your table for Mean Risk Ratio, and click the small red arrow in the box at the top of the spreadsheet. Next, click the red arrow for Array 2 and select all the values in your table for Actual Risk Ratio, and click the small red arrow in the box at the top of the spreadsheet. Your value of the correlation coefficient should appear in the selected cell of your spreadsheet.

Step 7. Say what it all means.

Answer the following 3 questions:

1. How does the judgment of relative risk depend on availability in your data?
2. How does it depend on actual risk ratio?
3. What is the relationship between perceived risk ratio and actual risk ratio?

Lab Report

Your report of this lab will consist of:

1) your answers to the 3 questions of Step 7;
2) your two graphs; and
3) your spreadsheet showing the correlation coefficient data and result.

References

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207-232.