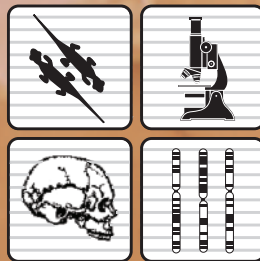


BioMath

Modeling Neuron Networks: The Neuroscience of Pain

Student Edition



COMAP





Funded by the National Science Foundation,
Proposal No. ESI-06-28091

This material was prepared with the support of the National Science Foundation. However, any opinions, findings, conclusions, and/or recommendations herein are those of the authors and do not necessarily reflect the views of the NSF.

At the time of publishing, all included URLs were checked and active. We make every effort to make sure all links stay active, but we cannot make any guaranties that they will remain so. If you find a URL that is inactive, please inform us at info@comap.com.



Published by COMAP, Inc. in conjunction with DIMACS, Rutgers University.
©2015 COMAP, Inc. Printed in the U.S.A.

COMAP, Inc.
175 Middlesex Turnpike, Suite 3B
Bedford, MA 01730
www.comap.com

ISBN: 1 933223 73 1

Modeling Neuron Networks: The Neuroscience of Pain

Overview

The human nervous system is made up of 100 to 200 billion neurons, supported by 1 to 2 trillion glial cells. Each neuron has an average of about 5 thousand synapses. Each of these synapses may release multiple neurotransmitters across a range of concentrations. It's been estimated that the number of potential neural states that these synapses can produce in a single person is more than the number of molecules in the known universe!

Every conscious experience is built from the interaction of neural activity at multiple levels within the nervous system, and the perception of pain is no exception. This summary of pain is just the tip of the iceberg. Each year we learn more and more about the nervous system and current advances are progressing at an accelerating rate. This unit explains the basics in order to show how pain from the skin (**somatosensory pain**) is processed. References will be provided that allow you to go deeper than this brief overview.

Overall, it is important to remember that pain is not a stimulus. This parallels the fact that there is no 'red' light. Red is a perceptual response to electromagnetic wavelengths within a narrow range. What we call pain is a collection of responses within the body, most often, to tissue damaging or potentially tissue damaging stimulation. Such stimuli are referred to as '**noxious**.' The receptors that have evolved to detect tissue damage are called '**nociceptors**' and they send their information through 'nociceptive' pathways from the skin, through several neural relays, to the highest levels of the brain. Pain is a hierarchy of responses organized at multiple levels of the nervous system.

Pain is a sensation. It's a complex perceptual and motivational state. It's an interpretation of an event. Our first exercise in this unit will be to try to capture some of pain's complexity by analyzing words that humans use to describe the pain experience. Some of them relate to the sensory experience--how you would describe your pain. Some relate to how pain motivates--how much you want it to stop. Some relate to the interpretation of the pain--how it affects your life.

Unit Goals and Objectives

Goal: Gain a biological and clinical appreciation of pain.

Objectives:

- Understand in general terms how somatosensory pain (pain from the skin) is processed.
- Understand how pain is an interpretation of an event.
- Be able to describe the sensory, affective and evaluative categories of pain and provide examples of each.

Goal: Gain a deeper understanding of the statistical methods used to analyze the data associated with the measurement of pain.

Objectives:

- Understand the process and methodology of data collection and develop a survey to collect data.

- Identify and understand the appropriateness of various types of data to include categorical, ordinal and numerical.
- Collect and collate data into a format conducive to analysis.
- Know the benefits and drawback of various descriptive statistic representations to include box-and-whisker plots, dot plots, histograms and pie charts.
- Understand what the results of a t-test represent.
- Understand what the results of a F-test represent.

Goal: Apply statistical methods to analyze data and make conclusions about hypotheses.

Objectives:

- Conduct an appropriate t-test to analyze and determine significance of differences in collected data.
- Conduct a F-test to complete an ANOVA (analysis of variation) in collected data.

Lesson 1 What Do You Think Pain Is?

Pain is the number one reason that people visit a physician. When a patient is experiencing pain, one of the first things a doctor must do is determine what type of pain the patient is experiencing and how intense that pain is. It is important, therefore, that both the doctor and the patient have a similar understanding of the words that are being used to describe the pain. For example, what does the patient mean when describing pain as a “stabbing” or “burning” sensation? Likewise, when a doctor asks a patient to rate the pain “on a scale of 1- 10,” how can we be sure that the patient’s scale of pain intensity is the same as what the doctor’s is? That is, does a “6” on a scale of 1- 10 mean the same thing to the doctor as well as the patient?

An Introduction to Pain

It is not surprising in human evolution that we have developed specific mechanisms to detect and respond to bodily injury. Arguably, such responses would be some of the most important for the survival of an organism. The skin contains a number of receptor types that are formed by the ends of multiple types of sensory **axons**. Pain seems to rely mostly on what are called **free nerve endings**, aptly named because they are thin branches of sensory nerves that simply branch within the skin. Other skin receptor types can have more organized endings on them and they largely play roles in other sensory qualities that can arise from the skin.

What stimulates free nerve endings includes the range of stimuli that can injure our tissues, including heat, some chemicals, and mechanical forces strong enough to threaten the integrity of cell membranes. Mechanical and chemical stimuli play roles in the activation of free nerve endings and some **analgesics** work on these chemical stimuli, aspirin being the classic example. **Capsaicin** is the chemical that puts the ‘hot’ in chili peppers. Capsaicin stimulates **vanilloid** receptors that evolved to signal the presence of very high temperatures. Mice have been created that lack these vanilloid receptors. They also lack the ability to respond appropriately to high temperatures, but do respond normally to lower temperatures and other somatosensory stimuli including non-thermal noxious stimuli. So even at the receptor level, nociception is revealing its complexity. There are different receptors for different pain qualities.

Injury causes a complex series of chemical changes in the skin. Some of them can directly activate nociceptors. Others can alter the sensitivity of nociceptors, making them more likely to be activated by subsequent stimuli. You may have noticed the red zone (flare) that sometimes develops around a slight scratch. This reddening is, in part, caused by some nociceptors releasing **substance P** and causing capillary beds to become a bit ‘leaky.’

For a time after an injury, the body increases sensitivity to further injuries after it has been damaged. Protecting injured tissue can help injuries heal. In some cases, unfortunately, these adaptations can last too long and/or be taken too far. In such cases they can contribute to the development of debilitating chronic pain states.

The nociceptive axons leading from the skin to the spinal cord are of two major types: **C fibers** and **A-delta fibers**. Both of these types are relatively thin compared to those that carry non-noxious information. C fibers have no **myelin** and are the thinnest sensory axon in the skin.

Because they are thin and because they lack myelin, action potentials in C fibers move at a 'walking pace' (approximately .5-2 meters per second or 1.1 to 4.47 miles per hour). A-delta fibers also carry nociceptive information. They are bit thicker than C fibers and they have myelin. This combination allows action potentials in A-delta fibers to move at a much more rapid pace (approximately 5 to 35 meters per second or 11 to 78 miles per hour). As we will see, these two types of nociceptive fibers tend to produce pain states with different sensory qualities.

Think back to when you perhaps accidentally burned yourself by putting your hand or foot in water that was much hotter than expected. Several things happen in fairly rapid succession. Your appendage comes out of the water quickly and reflexively. This reflexive action occurs at the spinal level, before neural information can even reach your brain. After this, you often feel a 'First Pain' that is bright, sharp, and pricking. This pain is brought to you by the faster A-delta fibers. After this, you often have time to anticipate what is about to come. The dull, throbbing, and aching 'Second Pain' soon follows and this is carried by the slower C fibers.

The complexity of pain, therefore, begins to become clear even at these basic sensory levels that are fairly well understood. This complexity grows as nociceptive information flows through ascending levels of the central nervous system. That additional complexity will be only thinly reviewed here.

In the spinal cord terminals of the A-delta and C fibers tend to synapse within a fairly short distance onto neurons in the dorsal horn of the same side of the body. These second order nociceptors may then communicate with other neurons on the same or opposite side of the spinal cord. These second order cells may also send axons that typically cross to the opposite side of the spinal cord and ascend to one of several targets in the medulla, pons, midbrain, or thalamus. Local spinal connections may trigger reflexive movements that pulled your hand or foot away from the hot water. They may also adjust your posture and center of gravity so that as the body moves away from the stimulus it does not tend to fall over in the process.

As was true in the skin, nociceptive inputs to the dorsal horn can also sensitize neural reactivity to later input. Some nociceptive neurons, the ones that respond to capsaicin, release the peptide substance P onto second order nociceptors in the spinal cord. One reason that sensitivity to pain can increase after an injury is that these second order nociceptors have 'hidden' substance P receptors that can't be stimulated under normal circumstances. After activation by injury, however, these hidden receptors can migrate to the surface of dendrites and therefore increase the number of substance P receptors available for the next burst of sensory input to stimulate. Clearly, pharmaceutical companies are very motivated to develop drugs that can affect the activity of substance P neurons and the receptors that respond to substance P. But remember that substance P is only part of the story of nociception.

At the spinal cord level, there are responses to pain that require no direct input from the brain, for example reflexive withdrawal of limbs and movements associated with jumping and pedaling movements. As you ascend through the brain, pain-related behaviors become more complex. Running and vocalizing during the delivery of a noxious stimulus involve brain stem regions up through the midbrain. If vocalization occurs after a noxious stimulus is stopped, that indicates forebrain structures such as the thalamus have been activated. The highest levels of response to

pain involve the cerebral cortex. What we would consider ‘suffering’ probably most importantly involves the prefrontal cortex. Patients who have had prefrontal lobotomies, which cut the subcortical fibers connecting the prefrontal lobes with the rest of the brain, sometimes say surprising things like: “The pain is still there but it doesn’t bother me.” or “The big pain is gone but the little pain is still there.”

One area of the brain that deserves special note in terms of pain is a little known part of the midbrain called the **periaqueductal gray**. This area surrounds the **cerebral aqueduct** (periaqueductal) and is largely made up of neurons (gray). The periaqueductal gray receives nociceptive input from the spinal cord. It also has interconnections with forebrain structures involved in several motivational states (e.g., pain and reproductive behaviors). Of perhaps greatest interest is that the periaqueductal gray has the ability to activate lower brain stem systems that send information back down to the spinal cord. They can profoundly inhibit subsequent nociceptive processing. Therefore, the brain has the ability to inhibit the perception of pain under some circumstances and inhibit it profoundly. The periaqueductal gray has access to two pain-inhibitory systems that use neurotransmitters related to well-known drugs that can inhibit pain: **opiates** and **cannabinoids**.

Reactions to pain are dynamic and organized at multiple levels of the nervous system. Plus, the nervous system has the ability to inhibit pain in certain circumstances. Our processing of pain changes with age and lifestyle. The elderly who have a sedentary lifestyle are most likely to develop chronic pain states. Pain reactivity can vary both from individual to individual and across time within an individual. Finally, there are gender differences in reactions to pain as well as in the type of pain inhibition that males versus females tend to activate when needed.

This overview should have conveyed the complexity of responses to injury. Pain is not one thing, but a hierarchy of interrelated events within the nervous system. Injury can activate multiple nociceptive pathways and produce different sensory qualities. The brain has motivational systems that drive the organism to focus on the painful stimulus in order to stop the tissue damage. At the highest level of the nervous system the cerebral cortex evaluates the sensations, motivation states, and contexts that are associated with the pain experience. Human language can tap into many of these elements of the pain experience.

The Language of Pain

Surveys and questionnaires are often the method used when gathering information from a sample of individuals on preferences, needs, or behavior. For example, surveys are used during elections to provide some insight as to which issues are important to the electorate or how voters are responding to candidates on any number of issues. Surveys can be administered in a number of ways: over the phone, in person, or in writing. The method of delivery is one of the factors taken into consideration when investigators are trying to develop a survey that is as unbiased as possible. Another factor in developing a survey that will produce good information is in selecting appropriate and clear terminology that is similarly understood by both administrators and participants in the survey. This is particularly important when a patient is being asked about pain: how severe is the pain? What type of pain is it? Is the pain consistent? Everyone experiences pain differently and may have different tolerance levels. To begin the discussion on pain, this

activity helps to build a better understanding of the words used in describing pain and to develop some uniformity in the definitions of the words that are used.

ACTIVITY 1-1 Ouch, That Hurt!

Objective: Brainstorm description words of pain and analyze types and sources of pain.

Materials:

Handout NP-H1: Ouch, That Hurt! Activity Worksheet

1. Brainstorm in your group and make a list of (single) words that are associated with the different types of pain?
2. Draft a definition of “pain.”
3. List different types of pain.
4. List different sources of pain.
5. Do all people react in the same way to pain? Explain.

Practice

1. Constructing a thorough list of words that describe different types and different degrees of pain will take more time than is available in a short class discussion. Generate additional pain descriptors (words that describe pain) by consulting resources such as the dictionary, a thesaurus, and the web.
2. Interview at least one pain sufferer; such as older adult and/or athlete, to ask what words they would use to describe their pain. These answers will be shared with the class during the next activity.

Questions for Discussion

1. How would you classify the types of pain?
2. How would you rank and describe the levels of intensity of pain?

Ranking the Words of Pain

Activity 1-2 develops the categories and classifications of pain. Three standard classifications of the pain experience will be considered: **sensory**, **affective**, and **evaluative**. These classifications are standard within the industry and are based on analysis of the dimensions of pain in the research (On the Language of Pain). Words that describe *how things feel* or *seem to a perceiving subject* are classified as sensory qualities. Sensory qualities of pain could include words that describe sensations that are temporal, spatial, pressure, and thermal. Words that describe features of an experience which render it *pleasurable* or *desirable*, or *the reverse*, or which gives it a *distinctive emotional tone* are classified as affective qualities. Affective qualities of pain could

include words that describe such things as fear, tension, or punishment. Words that describe the *subjective overall intensity of the total experience of pain* could be classified as evaluative qualities.

ACTIVITY 1-2 The Feel, Experience and Intensity of Pain

Objective: Develop categories and classifications of pain.

Materials:

Handout NP-H2: The Feel, Experience and Intensity of Pain Activity Worksheet
Consolidated list of class pain words (or use NP-H3: List of Pain Words)

1. In your small group, review the list of pain words:
 - a. Discuss which words most likely have sensory qualities.
 - b. Discuss which words most likely have affective qualities.
 - c. Discuss which words most likely have evaluative qualities.
2. Create three lists: one for sensory words, one for affective words, and one for evaluative words. Move the words back and forth until there is mutual agreement on which words represent sensory, affective, or evaluative qualities.
3. Be prepared to discuss your categories with the class.

Questions for Discussion

3. Which words most likely have sensory qualities? Affective qualities? Evaluative qualities?
4. Are there any words that contain qualities from more than one classification? How would you best classify these words?
5. How would you rank words in the same category according to the level of pain each word is perceived to describe?

The McGill Questionnaire

The McGill Questionnaire is an important clinical tool used internationally to assess pain states. The McGill pain questionnaire is a survey developed at McGill University in 1971 by Professors Melzack and Torgerson^[1]. The questionnaire asks patients to describe their pain by choosing words that describe a similar type of pain. These words are then categorized into groups. Patients are asked to circle the word in each group that best describes their pain. Patients are also asked to rate on a scale of 1 - 5 the degree of severity of their pain.

ACTIVITY 1-3 A First Look at the McGill Questionnaire

Objective: Compare McGill research with class description of pain.

Materials

Handout NP-H4: A First Look at the McGill Questionnaire Activity Worksheet

Handout NP-H5: McGill Pain Questionnaire

1. Familiarize yourself with the words and major concepts used on the McGill Pain questionnaire. Note any observations about the questionnaire.
2. Compare your class word list with the words used on the McGill questionnaire.
 - a. What are some similarities in the two lists?
 - b. What are some differences in the two lists?
3. Note that the McGill questionnaire asks patients to rank each word that describes their pain on a scale from 1 (least) to 5 (greatest). For example, if a patient is feeling a burning sensation, the patient is asked to rank the severity of that burning sensation on a scale from 1 - 5. Why do you think the questionnaire does this?

Once patients complete the McGill questionnaire, medical professionals are often interested in making comparisons between the different categories. There are many comparisons that could be studied from a statistical perspective. Future sections will consider if there is a difference in the way men and women interpret or experience pain. Another comparison might be to ask if there is a difference in the way athletes and non-athletes experience pain; or if there is a difference in the way older people and younger people experience pain. Yet another example would be comparing patients' interpretation of the different categories (sensory, affective, evaluative) of pain.

The (Your School Name Here) Survey

Your next activity will involve conducting a survey to collect data on pain. This will be analyzed in future lessons using a variety of statistical techniques and procedures. First, however, you must create the survey. Note that the McGill questionnaire is quite long and involves several subcategories of the three main categories. For the sake of simplicity, you will create a shorter survey that will be used to collect the desired information. This smaller survey will concentrate on collecting information on a subset of twelve of the McGill questionnaire words: four words from each of the three categories. To create this sub-survey, 6 words are identified below and 6 words will be selected by your class

The first six words from the McGill questionnaire that we will use, two from each category, are:

Sensory: Burning and Hurting

Affective: Sickening and Exhausting

Evaluative: Troublesome and Miserable

Review the article “On the Language of Pain” by R. Melzack and W. S. Torgerson (NP-H6)^[1]. In this paper, Drs. Melzack and Torgerson report on their survey of doctors, patients, and students to rate the classes and subclasses in the McGill Pain Questionnaire. (Your teacher may ask you to read this paper more carefully for homework.) Pages 54 and 55 of this article, list additional words. **Select six additional words, two from each category, that also have a rating somewhere around 2.5 (in the middle of the range 1 to 5). The items chosen must also be in sets of two, selected from the three pain response dimensions on the McGill. Choose items that are consistently more or less intense than the moderate items provided.**

Questions For Discussion

6. Why do you think the words above (burning, hurting, sickening, exhausting, troublesome, and miserable) were chosen for our survey?

7. Select six additional words, two from each category, that also have a rating somewhere around 2.5 (in the middle of the range 1 - 5). The items chosen must also be in sets of two, selected from the three pain response dimensions on the McGill. Choose items that are consistently more or less intense than the moderate items provided. Be prepared to discuss your choices in the class discussion.

8. You now have the 12 words, four from each category, for a survey. Create a survey that lists the 12 words in random order with instructions to rank the level of pain each word represents from 1 - 5. For example, one possible look for the survey could be:

**Ranking Words that Describe Pain
Survey**

Instructions: Listed below are twelve words that describe pain. For each word circle the value between 1 and 5 that portrays the level of pain the word represents (1 - least amount of pain, 5 - greatest amount of pain).

Sickening	1	2	3	4	5
Hurting	1	2	3	4	5
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮

Conducting a Survey

You are now ready to administer the survey to people outside of school. Each student should survey two males and two females. For each gender, one person should be another teenager and the other should be approximately the age of a parent (that is, one male teenager and one male adult; one female teenager and one female adult).

Administer the survey and collect the data based on instructions from your teacher. Your class may have an online site for uploading data or may collect data in class from paper copies of the survey.

Lesson 2 What Does Pain Look Like? Descriptive Statistics

In Lesson 1, you administered surveys to people outside of your class to obtain quantitative data on people's interpretation of words associated with pain. Throughout the remainder of this unit, we will analyze these data using descriptive and inferential statistics. The survey you developed consisted of 12 items from the McGill Pain Questionnaire; 6 items selected by the authors and 6 items selected by your class, to indicate more, or less, intense pain. The 12 McGill items come, four each, from the Sensory, Affective, and Evaluative dimensions. Early research on these items on the McGill found that people rated these items as being of about 'moderate' pain intensity (approximately 2.5 within a range from 1-5).

Psychophysics

The field of **Psychophysics** deals with the relationship between sensory stimuli and the perceptions they cause. When conducting research, painful stimuli cannot be delivered to people without their consent and certainly not for the purpose of this lesson. In your research, you did not seek out people who were actually suffering from pain and ask them to use items from the McGill to describe their pain. Rather, you asked people to voluntarily read English words describing pain, and then assign a number to each word to indicate the intensity of the imagined pain that would cause them to use that word to describe their own pain. When survey takers assign an ordinal number from 1 - 5 (or 1 - 3, 1 - 10, etc.) to each word indicating where that word falls within a given range (in this case 'pain') from lowest to greatest, this type of survey is called **cross-modality matching**. That is, the meaning of an English word is being converted into a number within a particular range (in this case 1 - 5). This type of numeric response dimension is used widely in survey research. Even though the participant can only respond at five fixed whole number points in this survey, the assumption is that these points exist on a continuous underlying sensory/perceptual continuum. This gives researchers the mathematical grounds to treat the numbers as having interval qualities, allowing the researcher to perform mathematical operations needed for the statistical analysis.

Descriptive statistics is that area of statistics that quantitatively describes the main features of a given set of data that are of interest to the researcher. That is, the researcher is summarizing the characteristics of the data set, but is making no inferences about the population from which the sample was taken. **Inferential statistics** is that area of statistics that tries to learn something about the population from which the sample data were taken. This requires the application of probability theory to well defined statistical methods. Probability theory is not needed in order to describe the main characteristics of the sample.

Before discussing the various graphs and figures associated with descriptive statistics, it is important to understand the different types of data and the role they play in research. There is a variety of ways in which data types can be described. The primary types of data are identified as **qualitative** or **quantitative**.

Quantitative data is often referred to as measurable data. Quantitative data is mathematical in nature, which allows statisticians to conduct mathematical operations on the data to determine mathematical values such as mean and standard deviation. The observations represent counts or

numerical measurements. Quantitative data can be either continuous or discrete. **Continuous data** take on all values along a continuum. **Discrete data** take on values that are separate and distinct. For example, although the length of your foot is considered continuous (your foot takes on all values along a continuum as it grows), your shoe size is discrete (sizes only come in whole and half numbers; e.g. 6, 6½, 7, 7½). Whether continuous or discrete, quantitative data have a numerical, and thus, mathematical value. We can conduct mathematical operations on these data to estimate population parameter values. Additional examples of continuous data are height or weight of individuals; length of time of an activity; and distance traveled in a given time period. Examples of discrete data include the number of children in a household; the number of employees hired by a company; or the number of computers per classroom.

Qualitative data, also referred to as **categorical** or **nominal** data, are data that place individuals or responses into groups or categories. The values of a qualitative or categorical variable are labels for the categories. A study interested in determining which ice cream flavor is a person's favorite would be an example of categorical data. Recording each response as vanilla, chocolate, strawberry, etc. places each subject into the appropriate ice cream flavor group based on a given answer.

Sometimes it is possible to order the categories. Such categorical data are referred to as **ordinal** data. For example, programs will give participants surveys to determine what the level of satisfaction is with the program. Participants may be given a series of questions in which they are asked to circle one of several responses that are ranked in order from least to greatest. This is sometimes called a **Likert scale**. Some response options are:

Strongly Disagree Disagree Neutral Agree Strongly Agree
Poor Fair Average Good Excellent

Surveys will often convert ordinal data to letter or numerical responses. For example,

a. Strongly Disagree b. Disagree c. Neutral d. Agree e. Strongly Agree
1) Poor 2) Fair 3) Average 4) Good 5) Excellent

Converting responses to a letter or numerical response often simplifies the process and allows people to respond more quickly. When talking among friends, one might say "On a scale of 1-10, how much did you like the movie?"

Questions for Discussion

1. What are some examples of quantitative data?
2. Classify the examples as continuous or discrete.
3. Provide examples of continuous data that are commonly represented as discrete data?

4. What is the difference between ordinal and discrete data?

Representations of Data and Their Descriptive Statistics

A **boxplot**, also known as a **box-and-whisker plot**, is a convenient way of depicting numerical data through their five-number summaries (minimum, first quartile, median, third quartile, and maximum). A rectangle is drawn with upper and lower limits representing the 75th and 25th Percentile Points, respectively. Therefore, the middle 50% of the data falls within this rectangle. Also, within this rectangle is a line that represents the Median (50th Percentile Point). Whiskers extend out the top and bottom of the rectangle to indicate the Upper and Lower Limits of the data being described. Box Plots, therefore, explicitly depict five points within a dataset.

No element of a Box Plot, however, involves the mean, which is the most commonly used indicator of the typical or average score within a dataset. Sometimes a line representing the mean can be added to the box plot, if the mean plays some significance when interpreting the information.

A **dot plot** is a chart consisting of data points (usually filled circles) placed along a simple scale such as a number line. There are two different versions of dot plots. One is a graph that uses columns of dots to represent the distribution of the data. Previously, this type of dot plot would have been drawn by hand. Today several software packages can reproduce such dot plots very quickly. The other type of dot plot is an alternative to bar charts where the dots are used to depict the quantitative values associated with the categorical variables. We will use this second type of dot plot.

A **bar chart** or **bar graph** is a chart that uses rectangular bars with lengths proportional to the values they represent. Bar charts can either be oriented vertically or horizontally. A vertical bar chart is sometimes called a column bar chart. A bar graph is a chart that uses either horizontal or vertical bars to show comparisons among categories. One axis of the chart shows the specific categories being compared, and the other axis represents a discrete value. Some bar graphs present bars clustered in groups of more than one (grouped bar graphs), and others show the bars divided into subparts to show cumulate effect (stacked bar graphs).

A summary of these graphs and figures is given below. Remember that even though they look similar, no statistical measure used in one is used in the other.

There are many software packages that automatically construct boxplots; others can construct boxplots with a few simple steps. Boxplots can be used to compare different data sets by placing them side-by-side or one on top of the other. The side-by-side box plots in Figure 2.1 are from the *Neuroscience Pain Data* spreadsheet. The data are Q1 - Sensory 1, Q2 - Sensory 1, and Q3 - Sensory 2. Note the lower whisker represents the minimum and the higher whisker the maximum. The box is bounded by the 25th percentile and 75th percentile values and indicated the 50th percentile within the box.

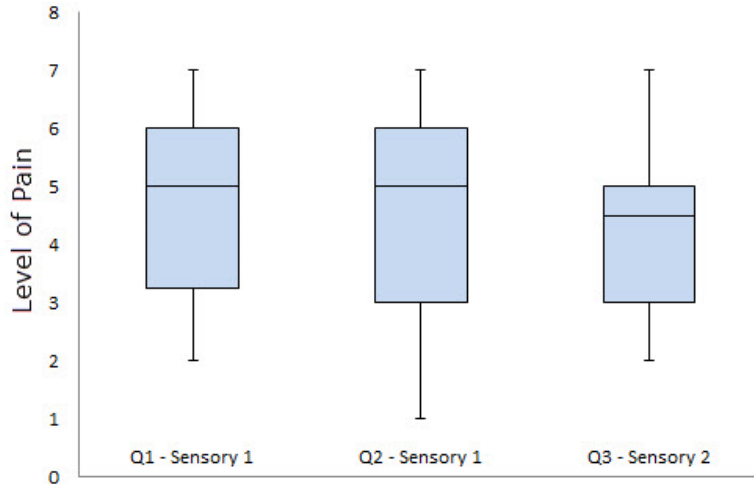


Figure 2.1: Example Box Plots of Questionnaire Data

Figure 2.2 shows examples of a single dot plot and a side-by-side dot plot. Again the data is taken from the DataEntry sheet of the *Neuroscience Pain Data* spreadsheet. The data from column Q1 - Sensory 1 is displayed in the first dot plot and a side-by-side dot plot comparison of Q1 - Sensory 1 and Q2 - Sensory 1 is displayed in the second figure. Note the number of responses for each question at each level is depicted by a dot.

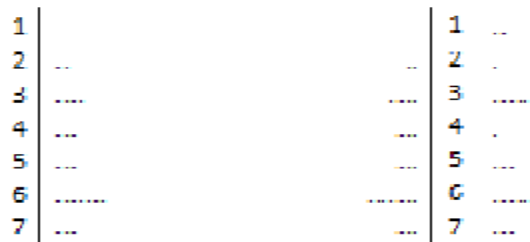


Figure 2.2: Single Dot Plot and Side-By-Side Dot Plot Examples

An example of a bar chart and bar graph appear in Figures 2.3 and 2.4. This bar chart represents the Q2 - Sensory 1 data column in the DataEntry page of the *Neuroscience Pain Data* spreadsheet. The bar graph compares the data from columns Q1 - Sensory 1 and Q2 - Sensory 1 of the DataEntry sheet. Note the number of responses at each level is depicted by a vertical bar.

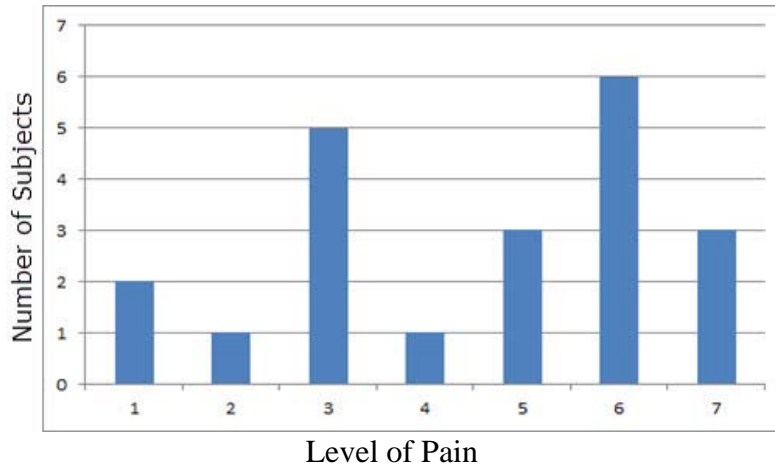


Figure 2.3: Dot Plot of Q2 - Sensory 1

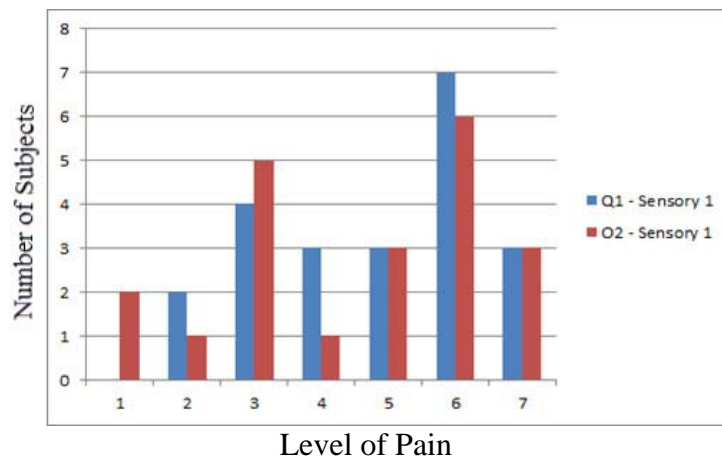


Figure 2.4: Side-By-Side Comparison of Q1 - Sensory 1 and Q2 - Sensory 1

ACTIVITY 2-1 Representing and Comparing Pain Data

Objective: Use various representations to compare and analyze data.

Materials:

Handout NP-H7: Representing and Comparing Pain Data Activity Worksheet
Neuroscience Pain Data Spreadsheet

This activity constructs box plots, dot plots, and bar charts for a subset of data that can be obtained from the *OverallSummaryData* sheet of the *Neuroscience Pain Data* spreadsheet. The first major column of this data sheet divides these data in a number of ways: 1) Overall, 2) Male, 3) Female, 4) Student-Aged, 5) Parent-Aged, 6) Male Student-Aged, 7) Female Student-Aged, 8) Male Parent-Aged, and 9) Female Parent-Aged. The rest of the columns match those of Sheet1. At the end of the data entry columns are the mean values calculated across raw data columns for a number of data subsets.

Consider each of the following comparisons of data from the *OverallSummaryData* sheet of the *Neuroscience Pain Data* spreadsheet:

- a) Provided--Overall (W7) versus Student--Overall (X7)
- b) Male (Y15) versus Female (Y23)
- c) Provided--Male (W15) versus Provided--Female (W23)
- d) Student--Male (X15) versus Student--Female (X23).

1. Graph side-by-side box plots by hand for b). What are some conclusions you can make by looking at the two box plots side-by-side?
2. Graph side-by-side box plots by hand for d). What are some conclusions you can make by looking at the two box plots side-by-side?
3. Graph side-by-side bar graphs for a). What are some conclusions you can make by looking at the side-by-side bar graphs?
4. Graph side-by-side bar graphs for c). What are some conclusions you can make by looking at the side-by-side bar graphs?
5. Construct a dot plot for Q5 - McGill Affective 1 (column I).
6. Construct a side-by-side dot plot for Q5 - McGill Affective 1 (column I) and Q6 - Student Affective 1 (column J). What conclusions can you draw by comparing these two responses?

Using Excel to Create Representations

Each of the descriptive statistics in this section could certainly be created by hand. However, it might be helpful to know how to use Excel to create each of the descriptive representations. The following is a set of instructions for creating box-and-whisker plots, dot plots, and bar graphs.

1. Creating Box-and-Whisker Plots in Excel

- a. In a vertical column of data, scroll down to the end of the data set. Click on the cell two rows below and one row to the left of the data set. In that cell and the four cells below it, create a vertical column with the labels: ***Minimum, 1st Quartile, Median, 3rd Quartile, and Maximum***
- b. Now click on the cell two rows directly below the end of the data set (it will be one column to the right of the cell labeled "Minimum." Type the Excel command ***=min(*** leaving the left parenthesis open for now. In the cell, Excel will display the set of arguments that are to be placed within the parentheses; in this case it is waiting for one argument – the data set over which the minimum is to be taken. With the cell dialogue open, click and drag from the first data entry to the last data entry, and then release the mouse (or you can type something like B2:B20, if the data set appears in the cells from B2 through B20). Press ENTER and the cell will produce the minimum value of the data set as its answer.

- c. The highlighted cell should now be the one directly under the minimum (to the right of the cell labeled *1st Quartile*). In this cell, type the command `=quartile(data array, quartile)`; that is `=quartile(B2:B20,1)` – if the data appears in cells B2 through B20 (dragging the cursor through the data set will work as well). Press ENTER and the cell will produce the first quartile value of the data set as its answer.
- d. Repeat this process for each of the remaining values typing in the respective commands for the median, third quartile, and maximum:
 - i. `=median(data array)`
 - ii. `=quartile(data array, 3)`
 - iii. `=max(data array)`
- e. Repeat this process for each data set. If the data sets are vertically side-by-side, rather than typing each command over and over again, Excel has the ability to easily copy and paste commands and scripts with each copied command carrying out the same instructions on cells proportionally situated relative to the initial cells outlined in the original command. Therefore, highlight the five descriptive values under the first data set. In the lower right hand corner will appear a small, solid square. Move the cursor over this square until it changes from a “wide” cross to a “stick” cross. Click and drag the mouse to the right (or left) of the original column. Each of the original commands will be copied in each additional column highlighted, doing the same calculations on each respective data set relative to the first data set’s position.
- f. Now that the five descriptive values are calculated, we will use the Excel graphing tools to create box-and-whisker plots. There is no direct Excel command that does this, so we will use our imaginations. Two rows below the five descriptive statistics, in the column that contains the first data set, type an equals sign (=); then click on the cell that contains the minimum value. Note that when you do this the column-row identification of that cell appears in the new cell. Press ENTER and the minimum value will again appear.
- g. The cell below it should now be highlighted. In that cell, type the command `=B23-B22`; that is, subtract the minimum value from the first quartile value, whatever the cell labels happen to be. Press ENTER.
- h. Repeat this process subsequently subtracting each of the following in each respective cell:
 - i. Median - First quartile value
 - ii. Third quartile value – Median
 - iii. Maximum – Third quartile value
- i. Copy and paste each of these values in each respective column which contains another data set.
- j. Highlight each of these newly calculated values over several (3 or 4) data sets.
- k. In the Insert menu, click the *Column* option in the *Charts* submenu. Then choose the Stacked Column option (not the *100% Stacked Column* option).
- l. Several stacked columns should appear in an Excel graphic. There should be the same number of columns as data sets. If not, the rows and columns have to be switched. This can be done easily by choosing the *Switch Row/Column* option in the *Design* submenu in the *Chart Tools* menu.

- m. Click on the top row of rectangles in the stacked columns. To make these “invisible” in the *Chart Tools* menu, select the *Format* submenu and choose the *No Fill* and *No Outline* options in the *Shape Fill* and *Shape Outline* menus.
- n. Repeat this for the bottom rectangles.
- o. Now click on the remaining top visible rectangles. In the *Chart Tools* menu, select the *Layout* submenu. In the *Layout* submenu. In the *Error Bars* menu, select the *More error bars options*.
- p. Choose the *Plus* and *Cap* options. Then select the *Custom* option and click on the *Specify Value* icon. In the *Positive Error Value* bar select the last row of values representing *Maximum – Third Quartile*. Leave the *Negative Error Value* bar alone.
- q. Now click on the remaining bottom visible rectangles. In the *Chart Tools* menu, select the *Layout* submenu. In the *Layout* submenu. In the *Error Bars* menu, select the *More error bars options*.
- r. Choose the *Negative* and *Cap* options. Then select the *Custom* option and click on the *Specify Value* icon. In the *Negative Error Value* bar select the row of values representing *First Quartile- Minimum*. This will be the second calculated row. Leave the *Positive Error Value* bar alone.
- s. All that remains is to make the inside rectangles of uniform color (usually the median is not displayed in box-and-whisker plots). Highlight each row of rectangles, choose the *No Outline* option and choose an appropriately light color. Repeat this for the remaining set of rectangles, choosing the same color.
- t. You should now have side-by-side box-and-whisker plots. If a median line is desired, simply choose the appropriate border option for one of the rows of rectangles.

2. Creating Dot Plots in Excel

This is a much easier process. Dot plots are most useful with discrete data sets. Again, in a blank area of the worksheet, vertically list each distinct value that appears in the data set (for example, 1,2,3,4...). In each cell to the right of the value use the nested command:

`=rept(“.”,countif(data array, value))`

The `rept(text criteria, number of times)` command converts a numerical value to the symbol given that number of times. The `countif(data array, value)` counts the number of times that value appears in the given data array.

For example: `=rept(“.”,countif(F2:F23,1))` will count how many values of “1” are in the cells from F2 to F23 and show that many dots.

A bar can be drawn between the values and the dots by placing the appropriate borders between the two columns. You can easily create a side-by-side dot plot by placing the nested commands on the opposite side. You may want to increase the font size of the cells with the dots so that they are clearer.

3. Creating bar charts and bar graphs in Excel

The *Charts* menu in Excel has the capacity to directly create bar charts and bar graphs. Highlighting the appropriate columns, going to the *Charts* menu within the *Insert* menu, and selecting the appropriate options is all that is needed. Then it is just a matter of selecting the appropriate format options to get your desired look. The Excel Help menu has a complete set of instructions

Box Plots are a form of ‘non-parametric’ representation of the data. Non-parametric statistics make no assumptions about the underlying distribution of the data. The scores can be ranked from low to high, but by how much one rank differs from the next is not an issue. The five scores in a Box Plot are simply 5 ranking points in a distribution: the lowest data point, the data point $\frac{1}{4}$ of the way through, the data point $\frac{1}{2}$ of the way through, the data point $\frac{3}{4}$ of the way through, and highest data point. Because non-parametric statistics make relatively few assumptions about the data, it is unlikely that researchers will violate these assumptions. They will make few mistakes with non-parametric statistics. However, there is a cost for this benefit. Non-parametric statistics do not use all of the information available in a dataset. They ‘throw away’ how tall you are in centimeters. They just say that you are taller than the person ranked below you and shorter than the person ranked above you. Throwing out information weakens statistical analyses.

Bar Graphs involve ‘parametric’ measures. Parametric statistical measures do make assumptions beyond the numbers that are recorded in the dataset. It is assumed that the data come from a continuum of data that has certain features. Here for example, if height data is available, what is known is not just that a 110 cm tall person is taller than a 100 cm person, but it is also known that these people differ by 10 cm and that this difference is twice as large as a difference in height of 5cm. With parametric data more mathematical procedures are available to be calculated. For example data values can be added up and then divided by how many elements that were added; that is, the mean of the data can be calculated. This cannot be done with non-parametric data. If the data allow for parametric analyses, researchers should use the full power of the data collected and go with parametrics. That’s what will be done with the survey data collected in Lesson 1. Even though people can only respond with the whole numbers 1-5 on this survey, it is assumed that there is an underlying pain-perception dimension, that a difference of 2 on the dimension is twice as big as a difference of 1. These survey numbers can be, and will be, used mathematically to calculate a mean.

ACTIVITY 2-2 A Look at Your Class Data

Objective: Use Excel to further compare data using various representations.

Materials:

Handout NP-H8: A Look at Your Class Data Activity Worksheet
Class survey data

1. Construct box plots, dot plots, and bar charts for the data collected from the survey constructed in class. Complete either requirement below as assigned by your teacher.

a. Construct several box plots, dot plots, and bar graphs that compare male and female responses from several of the words in the survey. What could you conclude, if anything,

about whether males and females in the survey perceived differences in the interpretation of each word?

b. Construct several box plots, dot plots, and bar graphs that compare student and adult responses from several of the words in the survey. What could you conclude, if anything, about whether students and adults in the survey perceived differences in the interpretation of each word?

2. Construct box plots, dot plots, and bar charts for the data collected from the survey constructed in class. Complete the requirements below as assigned by your teacher.

a. Construct side-by-side boxplots for each of the following:

- i. The McGill sensory words “Burning” and “Hurting.”
- ii. The McGill affective words “Sickening” and “Exhausting.”
- iii. All four sensory words, the McGill sensory words and the sensory words developed in class.
- iv. All four evaluative words, the McGill sensory words and the sensory words developed in class.

b. Construct dot plots for each of the following:

- i. The McGill sensory word “Burning.”
- ii. The McGill affective word “Sickening.”
- iii. The McGill affective word “Exhausting.”
- iv. The McGill evaluative word “Troublesome.”
- v. One of the sensory words developed in class.
- vi. One of the evaluative words developed in class.

c. Construct side-by-side dot plots for each of the following:

- i. The McGill sensory words.
- ii. The sensory words developed in class.
- iii. The McGill affective words.
- iv. The affective words developed in class.
- v. The McGill affective word “Exhausting” and one of the affective words developed in class.
- vi. The McGill evaluative word “Miserable” and one of the evaluative words developed in class.

d. Construct bar charts or bar graphs for each of the following.

- i. The McGill sensory word “Hurting.”
- ii. The McGill affective word “Sickening.”
- iii. One of the affective words developed in class.
- iv. One of the evaluative words developed in class.
- v. The McGill affective word “Sickening” and one of the affective words developed in class.
- vi. Both of the sensory words developed in class.
- vii. Both of the McGill evaluative words.

Practice

Use the data from the *Neuroscience Pain Data* spreadsheet to construct representations below as assigned by your teacher.

1. Construct side-by-side box plots for Q1 - McGill Sensory 1 (OverallSummaryData column E) and Q7 - Student Sensory 1 (OverallSummaryData column K). Would you conclude there is a significant difference between the McGill responses and the Student responses? Explain.
2. Construct side-by-side box plots for Q1 - McGill Sensory 1 (OverallSummaryData column E) and Q7 - Student Sensory 1 (OverallSummaryData column K). Would you conclude there is a significant difference between the McGill responses and the Student responses? Explain.
3. Construct side-by-side box plots for McGill-Overall (OverallSummaryData column W) and Student Overall (OverallSummaryData column X). Would you conclude there is a significant difference between the responses? Explain.
4. Construct side-by-side dot plots for Q9 - McGill Evaluative 1 (DataEntry column M) and Q10 - Student Evaluative 1 (DataEntry column N). Would you conclude there is a significant difference between the responses? Explain.
5. Construct side-by-side dot plots for Q2 - Student Sensory 1 (DataEntry column F) and Q4 - Student Sensory 2 (DataEntry column H). Would you conclude there is a significant difference between the responses? Explain.
6. Construct a bar graph comparing male and female responses for Q7 - McGill Affective 2 (DataEntry column K). Is there a difference between male and female responses? Explain.
7. Construct a bar graph comparing student and adult responses for Q7 - McGill Affective 2 (DataEntry column K). Is there a difference between student and adult responses? Explain.

LESSON 3 What Does Pain Look Like? Inferential Statistics

In this lesson, we will use parametric statistical analyses to make statistical inferences. Researchers use inferential statistics to infer something about the population from which data are drawn. They are also interested in whether data collected differ due to simple chance or whether there is something other than chance that causes differences researchers observe in the data.

Pain and Sample Differences

In the survey developed in class and administered outside of class, we chose words intending to represent pain levels that were consistently higher, or lower, than the middle-level pain reflected in the items provided (they had mean values in published research of around 2.5). Hopefully, the graphs generated from these data showed a difference in provided words versus student words in the direction that was predicted. Do the observed differences reflect that the words really tapped into different pain levels, or were the differences obtained simply due to chance differences? Is our set of questions really tapping into the same pain level?

The survey actually only represents a **sample** of people from a much larger **population**, the population of males/females, as well as student/parent-aged people. If the survey were administered again, collecting data from a different sample of people, would the results be exactly the same? Almost certainly they would not. This is a problem faced by scientists all the time when they collect data to investigate any number of research questions. Inferential statistics assist in analyzing data, like the survey results distributed here, and predict, or infer, statistical features of the population from which the sample data were drawn. If two samples differ, do they differ because they were drawn from different populations with different statistical properties or do they differ because of chance differences that are expected to occur when samples are randomly taken from a single population? We never know for sure which possibility is true. Inferential statistics attempt to determine the probability of those differences that are due to random variance.

Many times researchers want to know whether or not there is a difference between two populations. For example, is there a difference between males and females in pain perception? If there is, then males and females are drawing their behaviors from different potential behaviors, and scientists expect to consistently observe differences in pain behavior related to gender. Words for this survey were chosen by the class with the hope those words would be different from those words provided by the authors and would tap into a population of pain rankings that were either higher or lower than those provided. If successful in that selection, then there should be a difference between the provided words and the student-chosen words and this difference should occur with regularity whenever this survey is administered over different samples. In reality, inferential statistics never tells us what is really true. Researchers will never know with certainty if the sampling data is from one, or more than one, population. At best, inferential statistics calculates the probability that something makes a difference, and that the difference is significant. Unfortunately, scientists have to live with varying degrees of uncertainty. Additionally, results need to be able to be replicated across various samples, regardless of what any single inferential test indicates, in order to be considered valid.

Student's t-test

One standard inferential statistical procedure is called the *Student's t-test*, or simply the t-test. The Student's t-test is used when there is one nominal variable (say, gender: male/female) and one measurement variable (say, height) and the aim is to compare the mean values of the measurement variable (for example, determining if there is a difference in the mean heights of males and females). The nominal variable can only have two values, such as male or female, student or adult, or treated or untreated. A chemist working for the Guinness Brewing Company developed the Student's t-test to test the quality of different stout brewing procedures. He invented a model of how samples would vary by chance if the samples were drawn by chance from the same population. His student-t distribution describes the probabilities that various results could happen by chance.

The Student's t-test attempts to determine if the means are the same or different by first making a statement that assumes the means are the same. This statement is referred to as the **null hypothesis** and symbolized by H_0 . The null hypothesis in inferential statistics refers to the general or default position, which in this case assumes the means are the same. The null hypothesis is usually written as some variation of $H_0: \mu_1 - \mu_2 = 0$.

A Student t test statistic, t_s , is then generated through a formula that uses the available data. Different scenarios call for different versions (calculations) of the t-test statistic. One version of the Student's t-test statistic is used, for example, if the two samples are of different sizes. Other versions are used with samples that are paired or matched (for example, pre/post-tests), samples that are overlapping (paired samples missing some data), or samples that have different variations/standard deviations. A one-sampled t-test is used when the mean of the sample is compared to a fixed value μ_0 . Additional versions include what is referred to as one-tailed or two-tailed t-tests. The primary difference between a one-tailed and two-tailed test is the null hypothesis that is being tested:

$$\begin{array}{ll} \text{Two-tailed} & H_0: \mu_1 - \mu_2 = 0 \\ \text{One-tailed} & H_0: \mu_1 - \mu_2 < 0 \text{ or } H_0: \mu_1 - \mu_2 > 0 \end{array}$$

When the **standard deviation** of a statistic is estimated from the data, the result is called the **standard error** of the statistic. In some texts the term standard error is used for the actual standard deviation of a statistic. In that case, the estimated value is called the estimated standard error. When the standard error is substituted for the standard distribution of the sample mean, the statistic no longer has a normal distribution. It has a Student t-distribution. Figure 3.1 shows a comparison of a normal distribution graph and a t-distribution graph. Note that the graphs are very similar; the most obvious difference is the t-distribution has more area (that is, more of the population) in the tails than the normal distribution.

Population Distribution and Probability Density Function

When statisticians talk about a **population distribution**, they are referring to the area under a given curve. Mathematically, the area under a probability curve should be equal to 1 unit. This represents 100% of a population, since 100% is numerically equivalent to 1 or the whole. The curve defines how the population is distributed across some well-defined domain. Although any curve that encloses an area of 1 unit is a population distribution, statisticians are usually interested in a few special distributions that include the normal distribution, the t distribution, the F distribution, the Poisson distribution, and the geometric distribution. Most population measures fall into one of these special distributions. The t distribution and the normal distribution are two of the most common distributions. Both are symmetric; that is, there are as many representatives below the mean as there are above the mean, and the curves are reflective about the mean.

The curve itself (not the area underneath) is called a **probability density function** and is commonly abbreviated **pdf** by statisticians. The area of under the entire pdf for any distribution is 1. The pdf for the t-distribution and the normal distribution are displayed below, side by side.

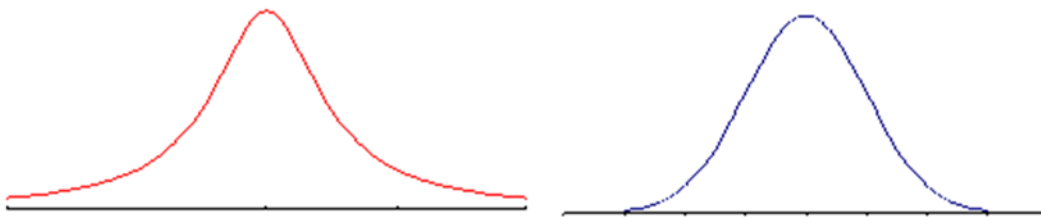


Figure 3.1: pdf of t-distribution (population) and pdf of normal distribution (population)

The t-tests

Activity 3-1 concentrates on two different versions of two-tailed t-tests: one in which two samples are independently drawn, like males versus females; and one based upon related or paired samples, for example, how the same people reacted to the pre-selected words and how they reacted to words that were selected to be more, or less, indicative of severe pain states. The first version of the t-test is sometimes referred to as **between groups** and the second version of the t-test is sometimes referred to as **within groups**.

The general formula for the two independent samples t-test (between groups) is:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \text{ with degrees of freedom } df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1} \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2-1} \left(\frac{s_2^2}{n_2}\right)^2}$$

The null hypothesis for this scenario is $H_0: \mu_1 - \mu_2 = 0$ (meaning the two measures are the same; their difference is 0).

The **degrees of freedom** are usually approximated by software and there are several formulas depending on the calculations and situation. The formula given above is a common one.

The goal of the two-sample t-test is to compare the mean of the responses in the two groups. Each group is considered to be a sample from a distinct population and the responses in each group are independent of those in the other group. Small samples can create problems with the assumptions that underlie the analysis. For example, the data could contain outliers or the sample could not be approximately normal. Careful scrutiny of the data to determine if there are outliers will help.

The general formula for the paired samples t-test (within group) is:

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{s^2}{n}}}, \text{ with degrees of freedom } df = n - 1$$

The null hypothesis for this scenario is $H_0: \mu_1 - \mu_2 = 0$.

For a paired or matched t-test to be valid, the data samples must come from the same entity or unit for each sample. A good example of this scenario is an experiment that determines if there has been improvement in a pre-test/post-test scenario among the same student population. Another example would be comparing the grades in math and history of the same group of students. In this case, the same students are taking both math and history and their grades in each course would be compared.

The Student's t-tests were originally designed as a one-sample test. However, it is most useful when the t-test procedure is adapted to other situations that involve more than one sample. Both the paired samples and independent samples are examples of this extension. The paired t-test is a natural extension of this. The **difference variable**, ($Y = X_1 - X_2$), is the 't-variable' of interest.

Since this lesson provides only a thin overview of parametric statistics, it will not discuss the assumptions made when performing a Student's t-test. One of them is that, since the data are being tested against whether they came from the same population, it is assumed that the variability of the two are comparable and that the means of the two samples, on average, do not show a difference. When **variance** is not equal, there are statistical corrections that should be made, but we will not undertake a discussion of this here.

To broadly overview the concept of the between groups Student's t-test, there are two (independent) samples that may be drawn from a common population with a single mean and a single variance. That is, the null hypothesis assumes the means of the two samples is the same. If this is true (that any differences between the two samples are simply reflective of the effects of random variation), then the Student's t-test provides a method to determine the probability that the mean difference observed is due to chance.

Researchers often use the **.05 level of significance** to evaluate inferential statistics. If the statistic being used tells the investigator that the observed results could be due to chance, but the chance probability is equal to or less than 1 chance in 20, that is 5 chances in 100, then the investigator

rejects the assumption that the two samples come from one population. The investigator states that there is a ‘significant’ difference between the sample means and assumes that the difference observed is NOT due to chance, but rather to something that has caused there to be two populations of results. For example, males use the same pain words to reflect less severe pain than females. That is, there is a ‘significant’ difference in the way males and females use the words of the McGill pain Questionnaire.

Two-Sample t-Test: Is There a Gender Difference in Pain Perception?

Do men and women have a different perception of pain? This question cannot be completely answered by examining this dataset. Perhaps it can provide some insight; however, as to whether or not this is a reasonable query. This is an example of a two-sample t-test.

The data provided in the *Neuroscience Pain Data* spreadsheet file will provide some insight into whether men and women perceive pain differently. The *DataEntry* page contains “raw” data collected from subjects at a regional university. The *OverallSummaryData* page contains various calculations of means, standard deviations, degrees of freedom, and other critical descriptive measures. There is one score for each male and female in the dataset.

Recall that to calculate the student-t the variance must be calculated in a slightly different way to reflect that it is estimating an unknown population based upon this sample. Going into the necessary depth to explain why estimates of variance are different than variance that is not an estimate would take too much time and is not important to demonstrating the use of the t-test in this instance. If you wish to read further about the t-test, consult the Research Methods Knowledge Base website at http://www.socialresearchmethods.net/kb/stat_t.php^[2].

Column Y in the *Neuroscience Pain Data* spreadsheet summarizes the “combined” overall responses of the subjects by taking the average of all the scores of each subject. This column can be used to consider whether men and women experience pain differently. Let μ_M be the population mean response for males and μ_F be the population mean response for females. Let X_M be the sample mean response for males and X_F be the sample mean response for females. The procedure for the two-sample t-test is as follows:

Step 1. State the null hypothesis:

$$H_0: \mu_M - \mu_F = 0$$

Step 2. Calculate the standard deviation:

$$\sigma = \sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}} = \sqrt{\frac{(0.53)^2}{7} + \frac{(0.27)^2}{14}} = 0.21$$

These numbers were obtained from column Y in the *OverallSummaryData* page.

Step 3. Calculate the t-statistic (numbers found in column Y):

$$t_{calc} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{(5.22 - 4.65) - 0}{0.21} = 2.68$$

Step 4. Calculate the degrees of freedom (taken from column Y): $df = [7.62] = 7$.

With degrees of freedom always truncate to the next lowest integer; that is, the use the greatest integer function.

The t-statistic and degrees of freedom will be used to determine if there is a **statistically significant** difference in the answers given by males and females in the survey. To make that determination, the t-statistic is compared to t scores that would be expected to be observed given a specified probability. That is, what is the probability that a t-statistic of that value would be observed with the given degrees of freedom?

Step 5. Determine the probability that the t-statistic would be observed given the degrees of freedom.

$$Pr(|T| \geq t_{calc} = 2.68) = Pr(T \leq -2.68 \text{ or } T \geq 2.68)$$

This probability is referred to as the **p-value** for a test of the null hypothesis, $H_0: \mu_M - \mu_F = 0$ against the “alternative” hypothesis, $H_a: \mu_M - \mu_F \neq 0$. Graphically, this is represented by the diagram below:

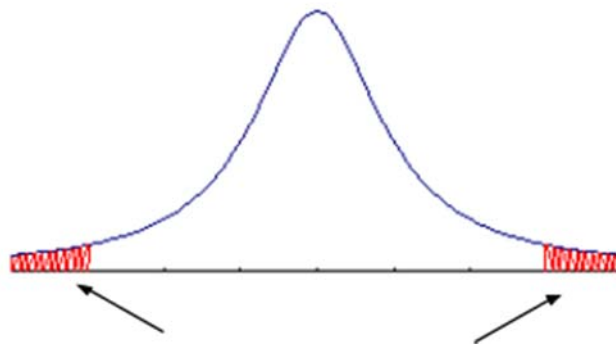


Figure 3.2: Comparing t-statistic Value to pdf Curve

The probability that $H_a: \mu_M - \mu_F \neq 0$ is $Pr(|T| \geq t_{calc})$ (area shown in red).

Using excel, we can calculate the value: “=tdist(x, deg_freedom, tails).”

In this example, the Excel function looks like: = TDIST(2.68,7,2), and returns the value 0.032.

Step 6. Rather than calculating the red area by hand, these values are recorded in a **t-Table** for specific probabilities, usually $\alpha = 0.1, 0.05,$ and 0.01 . A t-Table is given on the last page of the *Neuroscience Pain Data* spreadsheet. For a 2-tailed test, the probability is divided in half, evenly

along both tails. Thus, at the $\alpha = 0.05$ level, according to the t-Table in the *Neuroscience Pain Data* spreadsheet,

$$t_{0.975}(7) = 2.3646$$

The value $t_{0.975}(7) = 2.3646$ can be obtained using the Excel function = TINV(α , deg_freedom) = TDIST(0.05,7) = 2.3646.

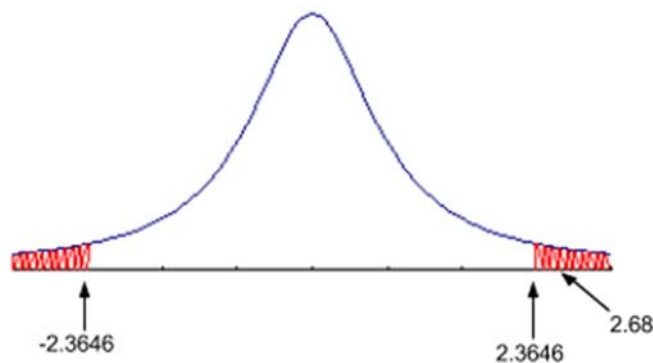


Figure 3.3: $Pr(|T| \geq t_{calc}) = Pr(|T| \geq 2.68) = 0.032$

Since $Pr(|T| \geq 2.68) = 0.032 < 0.05 = \alpha$, this suggests that we reject the null hypothesis, $H_0: \mu_M - \mu_F = 0$.

Another way to describe this is by referring to the diagram above. The “red” area is referred to as the “reject” region. The t-Table score, $t_{\frac{\alpha}{2}}(df)$, is the boundary of the reject region. If the calculated t-statistics, t_{calc} , falls in the reject region, that is, $|t_{calc}| > t_{\frac{\alpha}{2}}(df)$, then the data does not support the null hypothesis.

Summary: It is important to understand the procedure involved with the Student’s t-test and the following guidelines:

- Assume the null hypothesis H_0 , which is the status quo, is true.
- The p-value is the probability that the gathered data will actually happen assuming the null hypothesis is true.
- A small p-value suggests there is a very small probability that the observed data would have happened, if the null hypothesis were true. In this case, what is the logical conclusion? The logical conclusion with a small p-value is to **reject** the null hypothesis. (The calculated t statistic falls in the reject region.)
- A large p-value suggests there is a reasonable probability that the observed data would have happened, if the null hypothesis were true. In this case, what is the logical conclusion? The logical conclusion with a larger p-value is to **not reject** the null hypothesis. (The calculated t statistic does not fall in the reject region.)
- So what makes rejection of the null hypotheses statistically significant? This is a result of comparing the p-value of the t-statistic to the $\alpha = 0.1, 0.05, \text{ or } 0.01$ levels as determined by the conditions of the study.

Note: The α -level is often set in advance of the analysis and is usually determined by the researcher's aims and goals for the analysis.

ACTIVITY 3-1 Explore Two-Sample t-Test and Perception of Pain

Objective: Conduct and analyze results of a two-sample t-test.

Materials:

Handout NP-H9: Explore Two-Sample t-Test and Perception of Pain Activity Worksheet
Neuroscience Pain Data Spreadsheet and t-table

1. Using the data in *Neuroscience Pain Data* spreadsheet, conduct a two-sample t-test to determine if there is a difference in the perception of pain between teenagers (students) and adults.
2. Using the data collected from the survey constructed in class, conduct a two-sample t-test to determine if there is a difference in the perception of pain between males and females. Does your conclusion agree or disagree with the in-class example? What do you think this means?

Paired t-test: Differences in an Individual's Response

This example will again rely on the data in the *Neuroscience Pain Data* spreadsheet. This time the question will be: Is there a difference in responses to the McGill-Sensory words and the McGill-Affective words. The cumulative data for this comparison is found in columns Q and R of the *Neuroscience Pain Data* spreadsheet.

This is an example of a “paired” or “matched” t-test since the responses of each individual in one category are being compared to that individual's responses in another category. That is, the responses are “paired” since they belong to the same individual. The procedure is the same as the previous example (the two independent sample t-test). The difference will be in how the statistics and critical values are calculated.

Step 1. State the null hypothesis and set the α -level:

$$H_0: \mu_S - \mu_A = 0; \alpha = 0.01$$

Step 2. Calculate the standard deviation:

$$\sigma = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{(0.2)^2}{21}} = 0.04$$

These numbers were obtained from column Z in the *OverallSummaryData* page.

Step 3. Calculate the t-statistic (numbers found in column Z):

$$t_{calc} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{s^2}{n}}} = \frac{0.29}{0.04} = 6.64$$

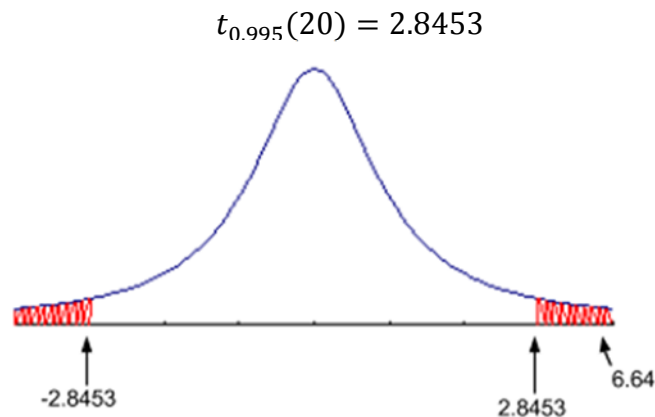
Step 4. Calculate the degrees of freedom: $df = n - 1 = 20$

Again, the t-statistic and degrees of freedom will be used to determine if there is a statistically significant difference in the answers given for McGill-Sensory words and McGill-Affective words.

Step 5. To make that determination, the t-statistic is again compared to t scores that would be expected to be observed given a specified probability. That is, the p-value is:

$$Pr(|T| \geq t_{calc} = 6.64) = Pr(T \leq -6.64 \text{ or } T \geq 6.64) = 1.8 \times 10^{-6}$$

Step 6. Using the t-Table given on the last page of the *Neuroscience Pain Data* spreadsheet, at the $\alpha = 0.01$ level,



Since $Pr(|T| \geq 6.64) = 1.8 \times 10^{-6} < 0.01 = \alpha$, this suggests that we reject the null hypothesis, $H_0: \mu_S - \mu_A = 0$.

ACTIVITY 3-2 Explore Paired t-Test and Difference in Data Items

Objective: Conduct and analyze results of a paired t-test.

Materials:

Handout NP-H10: Explore Paired t-Test and Difference Data Items Activity Worksheet
Neuroscience Pain Data Spreadsheet

1. Repeat the paired t-test process to determine if there is a difference between McGill-Sensory and McGill-Evaluative data found in the *Neuroscience Pain Data* spreadsheet.

2. Conduct a paired t-test using the data collected from the survey developed in class to determine if there is a difference between the McGill words and the words developed in class.

Do We Ever Not Reject H_0 ?

This example demonstrates the case when H_0 is not rejected. Consider the data in column E from the *Neuroscience Pain Data* spreadsheet. This column contains the responses by subjects to the first word given: McGill-Sensory 1. Do you think the data suggest a difference in male/female perception of pain with this word? To gain some insight into a potential answer, a two-sample t-test will be conducted on the data. Once again, note that many of the required initial calculations can be found in the *Neuroscience Pain Data* spreadsheet on the *OverallSummaryData* page.

Step 1. State the null hypothesis and α -level:

$$H_0: \mu_M - \mu_F = 0; \alpha = 0.05$$

Step 2. Calculate the standard deviation:

$$\sigma = \sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}} = \sqrt{\frac{(1.03)^2}{7} + \frac{(0.34)^2}{15}} = 0.40$$

The standard error can be calculated in Excel using the “SQRT” Excel function:

$$=SQRT(((1.03^2)/7)+((0.34^2)/15))$$

Step 3. Calculate the t-statistic (numbers found in column E):

$$t_{calc} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(4.86 - 4.80) - 0}{0.21} = 0.15$$

Step 4. Calculate the degrees of freedom (taken from column Y): $df = [6.63] = 6$

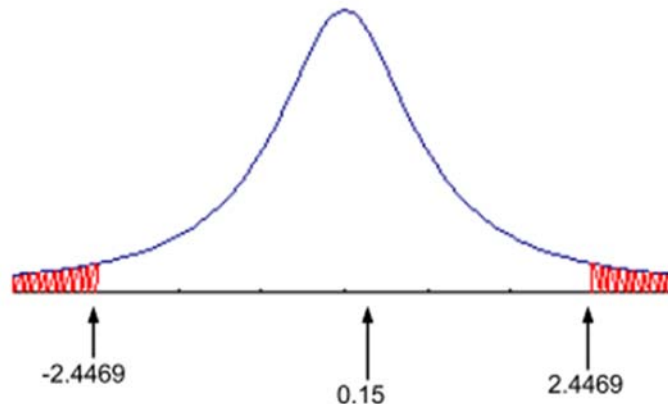
This value can be calculated in Excel using the formula given on page 26. For convenience, it also appears at the bottom of column E on the *OverallSummaryData* page.

Step 5. Next, calculate the probability that a t-statistic of that value would be observed with the given degrees of freedom: Excel: “=TDIST(0.15, 6, 2)”

$$Pr(|T| \geq t_{calc} = 0.15) = Pr(T \leq -2.68 \text{ or } T \geq 0.15) = 0.89$$

Step 6. This is the *p-value* for a test of the null hypothesis, $H_0: \mu_M - \mu_F = 0$ against the “alternative” hypothesis, $H_a: \mu_M - \mu_F \neq 0$. Again, this is represented graphically by the diagram below: Excel: “=TINV(0.05, 6)”

$$t_{0.975}(6) = 2.4469$$



The probability that $H_a: \mu_M - \mu_F \neq 0$ is $Pr(|T| \geq t_{calc})$. At the $\alpha = 0.05$ level, since $t_{0.975}(6) = 2.4469$ and $Pr(|T| \geq 0.15) = 0.89 > 0.05 = \alpha$, this suggests that we do **not** reject the null hypothesis, $H_0: \mu_M - \mu_F = 0$.

Note in the diagram above, the calculated t-score, $t_{calc} = 0.15$, falls inside the “acceptable” region bounded by the t-Table score, $t_{\frac{\alpha}{2}}(df) = 2.4469$. This suggests the null hypothesis is not rejected.

Both the data in the *Neuroscience Pain Data* spreadsheet and the data collected from the survey developed in class present ample opportunities to practice both types of t-tests. For additional practice, complete the exercises below.

Practice

1. Determine if there is a difference in perceived pain between males and females for each of the following from the *Neuroscience Pain Data* spreadsheet. Explain why you are using a paired t-test or a two-sample t-test in each case.

- a. Q3-McGill Affective 1 (column G)
- b. Q6-McGill Evaluative 2 (column J)
- c. Q8-Student Sensory 2 (column L)
- d. Q10-Student Affective 2 (column N)
- e. McGill Average Affective (column R)
- f. Student Average Evaluative (column V)

2. Determine if there is a difference in perceived pain between adults and students for each of the following from the *Neuroscience Pain Data* spreadsheet. Explain why you are using a paired t-test or a two-sample t-test in each case.

- a. Q3-McGill Affective 1 (column G)
- b. Q6-McGill Evaluative 2 (column J)
- c. Q8-Student Sensory 2 (column L)
- d. Q10-Student Affective 2 (column N)
- e. McGill Average Affective (column R)
- f. Student Average Evaluative (column V)

3. Determine if there is a difference in perceived pain between males and females for each of the following from the data collected from the survey constructed in class. Explain why you are using a paired t-test or a two-sample t-test in each case.

- a. Responses to the McGill sensory word “Burning.”
- b. Responses to the average of the McGill sensory words “Burning” and “Hurting.”
- c. Responses to the McGill affective word “Exhausting.”
- d. Responses to the average of the McGill affective words “Sickening” and “Exhausting.”
- e. Responses to the McGill evaluative word “Miserable.”
- f. Responses to the overall average of all the McGill words.
- g. Responses to the average of the affective words developed in class.
- h. Responses to the average of the evaluative words developed in class.
- i. Responses to the overall average of all the words developed in class.
- j. Responses to the overall average of all the words, McGill and those developed in class.

4. Determine if there is a difference in perceived pain between adults and students for each of the following from the data collected from the survey constructed in class. Explain why you are using a paired t-test or a two-sample t-test in each case.

- a. Responses to the McGill sensory word “Burning.”
- b. Responses to the average of the McGill sensory words “Burning” and “Hurting.”
- c. Responses to the McGill affective word “Exhausting.”
- d. Responses to the average of the McGill affective words “Sickening” and “Exhausting.”
- e. Responses to the McGill evaluative word “Miserable.”
- f. Responses to the overall average of all the McGill words.
- g. Responses to the average of the affective words developed in class.
- h. Responses to the average of the evaluative words developed in class.
- i. Responses to the overall average of all the words developed in class.
- j. Responses to the overall average of all the words, McGill and those developed in class.

5. Using the data collected in the survey developed in class, determine if there is a difference in perceived pain between the McGill words and the words developed in class.

Lesson 4 What Does Pain Look Like? - ANOVA

The previous lesson introduced two applications of the Student's t-test. The Student's t-test can be an effective test when comparing the means of two samples, independent or paired, or when comparing the mean of a sample against a given mean value. This was demonstrated with several examples that included comparing sensory responses to evaluative responses and comparing male responses to female responses. However, there are many instances where more complex comparisons are required. In these situations, the Student's t-test does not work. These comparisons require different types of inferential tests.

Neuroscientists might be interested in comparing responses across all three classifications of pain: sensory, affective, and evaluative. Are the means of the responses of all three classifications the same or do they differ? What if neuroscientists were interested in comparing responses between multiple age groups, for example: children, teenagers, young adults, and elderly adults? Would the mean responses across these different age groups be the same or would they differ? How do statisticians compare the means of multiple groups?

Another type of comparison that is heavily used in survey research is a *contingency table* or a *cross tabulation*. This is a comparison of different groups in terms of their response. For example, is it possible to compare the “high” or “low” pain responses between a males and females? This type of comparison is called a “two-by-two” comparison and can be represented by the following table.

Two-by-Two Comparison	Low (≤ 4)	High (>5)
Male	a	b
Female	c	d

Comparing the values a, b, c, and d in the two-by-two table might provide the neuroscientist with some insight on the type of responses produced by males and females.

This lesson will explore some of the other statistical methods that could be used to compare data among different groups and categories. Groups in statistical studies are often called *blocks*. It would be helpful to remember that a concise and accurate description of statistics is simply “the study of variation.” In the end, statisticians analyze variation in data to make inferences about the population(s) in question. What, if anything, is causing the variation? Where is the majority of variation occurring? Is the variation occurring across blocks or is the variation occurring within a group or blocks? The type of statistical method used depends on the question asked in the null hypothesis. Different questions require different methods. Using the wrong method or making the wrong assumptions about the methodology could result in an incorrect interpretation. The various statistical methods that study variation are collectively referred to as *Analysis of Variance*. Analysis of Variance is often abbreviated ANOVA.

One-Way ANOVA: The F-test

The first type of ANOVA to be explored in this lesson is *One-Way ANOVA*, also referred to as the *F-test*. It was mentioned in the previous lesson that the pdf's for the t-distribution and the normal distribution are symmetric (they are reflective images of each other over the mean). The F-distribution is an example of a non-symmetric distribution. A graph of the pdf for the F-distribution (using 6 degrees of freedom) appears in figure 4.1.

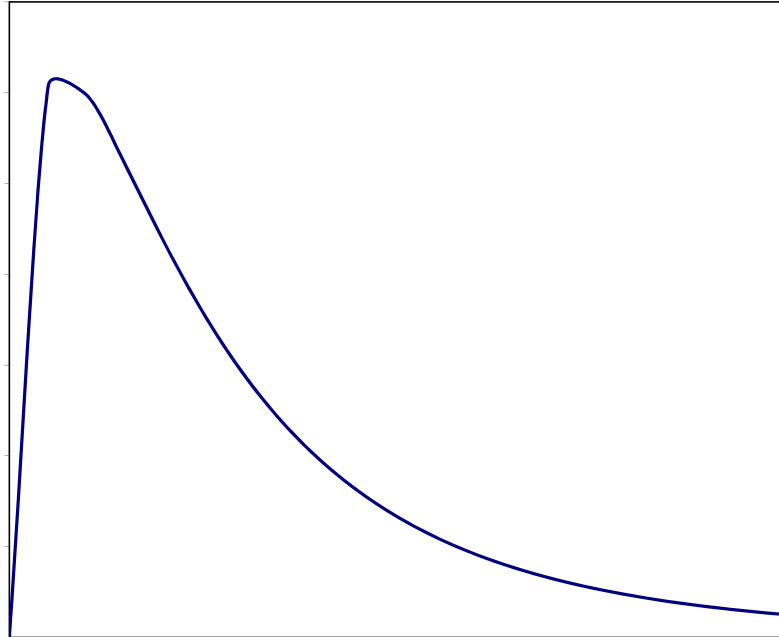


Figure 4.1: Pdf for F-distribution (note most of the population is to the left)

Although the curve is very different and distinct from the Student's t-distribution, the process by which the F-statistic is evaluated is the same. In fact, the process for each ANOVA method is basically the same:

Step 1. State the null hypothesis H_0 and α -level.

Step 2. Appropriate calculations: standard error, test statistic (in this case, the F-statistic), degrees of freedom, etc. (All calculations are done using technology, not by hand.)

Step 3. Calculate the p-value, the probability that a test statistic of that value would be observed with the given degrees of freedom.

Step 4. At the stated α -level, determine if the null hypothesis is rejected or not rejected based on the p-value or, equivalently, where the test statistic falls along the graph of the population distribution in comparison to the critical score obtained from a table of values or the appropriate statistical software command.

The diagram below describes this process in terms of a typical probability density function for the F-distribution.

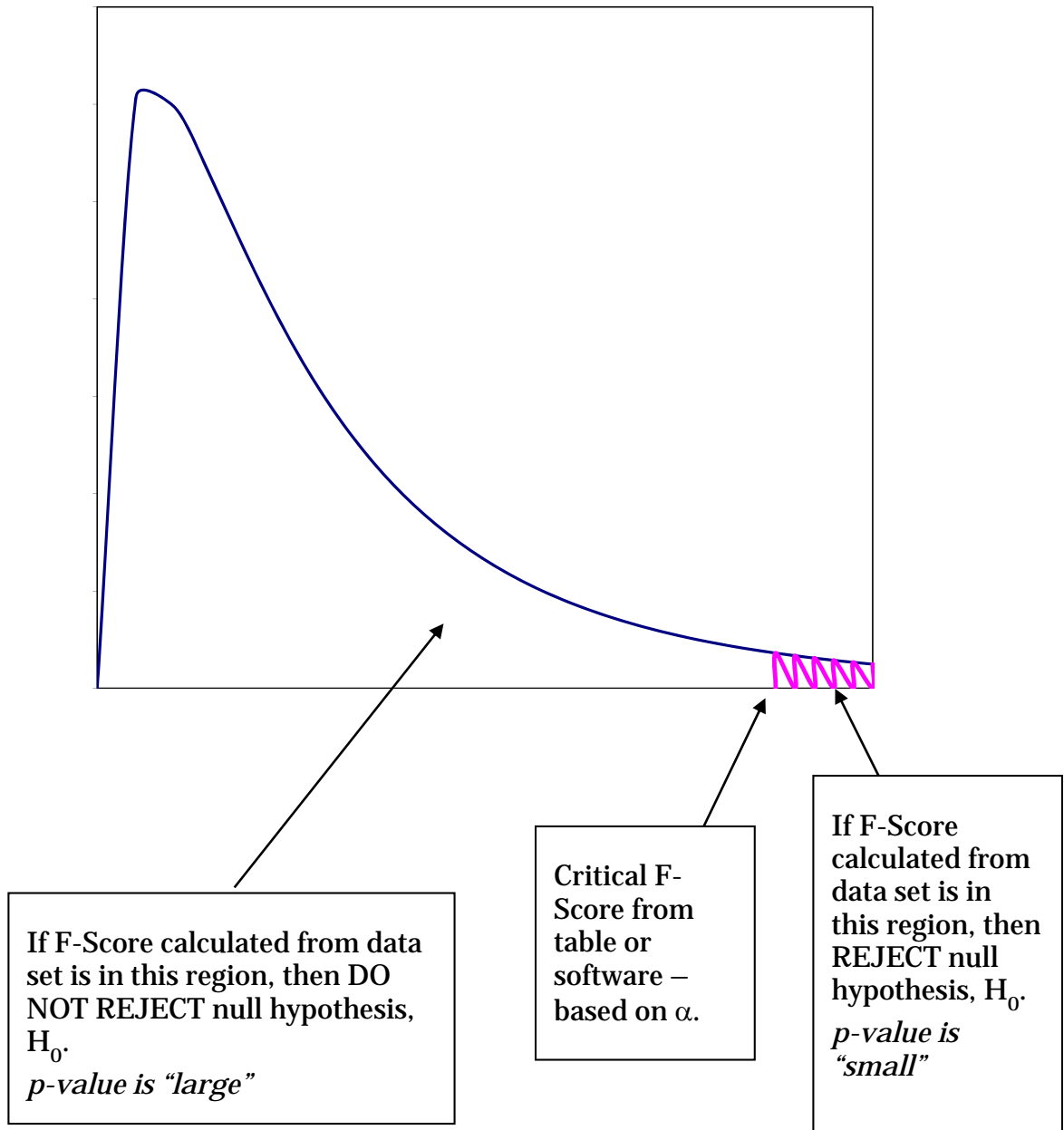


Figure 4.2: Using F-Distribution and Score

To infer if the means of several blocks are the same or different, the F-test analyzes the variation in the data by comparing how much of the variation occurs *between* the means of the blocks and how much of the variation occurs *within* the blocks. Figure 4.3 shows data from three blocks. The three blocks have relatively equal means. This implies there is little variation between the three means; in this case we say there is little variation *between* the blocks. The variation in the data *within* each group accounts for most of the variation in the data. In this case, we say there is

variation *within* the blocks. Most of the variation within blocks implies that the null hypothesis, which assumes the means are equal, should not be rejected.

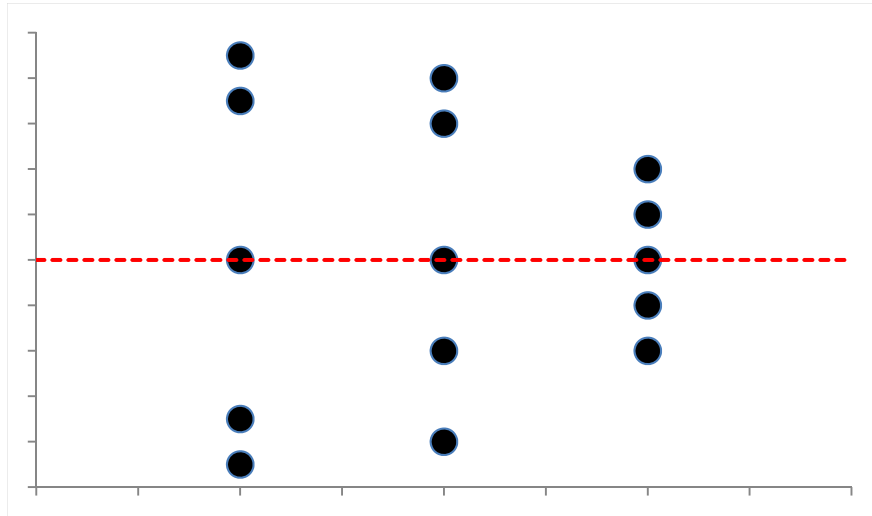


Figure 4.3: Three Blocks with Equal Means

Figure 4.4 shows a case in which data from three blocks has less variation within the blocks and shows variation *between* the three group means. In this case, it would make sense to reject the null hypothesis. The question is how to develop a test that reasonably measures the variation between blocks and within blocks, and then compares those measurements to determine where the majority of that variation is in order to come to a conclusion as to whether or not to reject the null hypothesis.

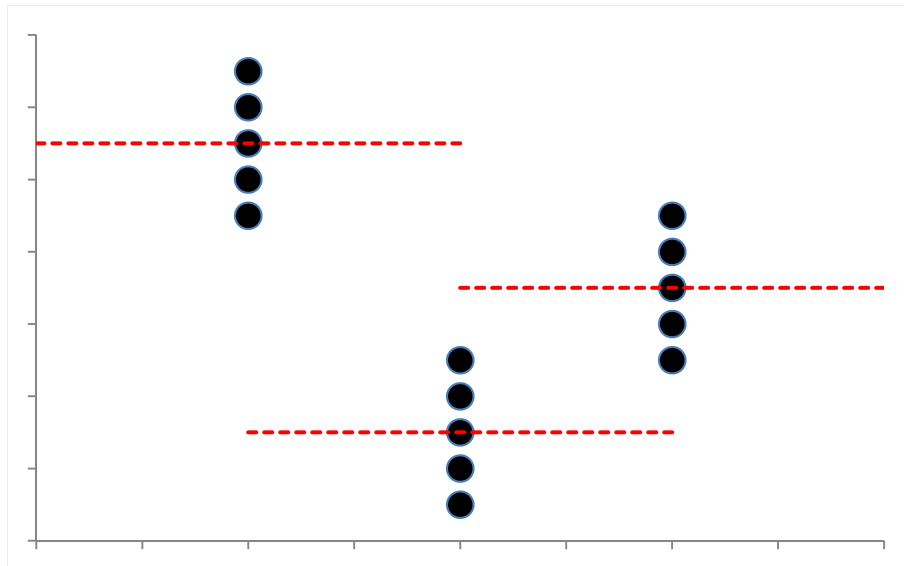


Figure 4.4: Three Blocks with Unequal Means

The F-test is one ANVOA method that makes this comparison. Software and technology are used to calculate the F-statistic. Nevertheless, before looking at an example, it would be beneficial to look at how the F-statistic is calculated. Suppose there are mn data points in a data set; m blocks and n data points within each group. The data could be represented in the m -by- n matrix as shown in table 4.2.

x_{11}	x_{21}	x_{31}	\cdots	x_{m1}
x_{12}	x_{22}	x_{32}	\cdots	x_{m2}
\vdots	\vdots	\vdots	\vdots	\vdots
x_{1n}	x_{2n}	x_{3n}	\cdots	x_{mn}

Table 4.2: m -by- n Data Matrix

Define

x_{ij} to be the j^{th} data entry in the i^{th} block;

\bar{x} to be the overall average of all the data points;

\bar{x}_i to be the average of the data points in the i^{th} block;

$x_{ij} - \bar{x}_i$ to be the “variation” the ij^{th} entry is from the i^{th} block mean;

$\bar{x}_i - \bar{x}$ to be the “variation” the i^{th} block mean is from the overall mean.

The calculations for the F-statistic are recorded in a table in the following way:

<i>Source</i>	<i>d.f.</i>	<i>Sum of Squares</i>	<i>Mean Squares</i>	<i>F</i>
<i>Between</i>	$m - 1$	$SSB = \sum_{i=1}^m n(\bar{x}_i - \bar{x})^2$	$MSB = \frac{SSB}{m - 1}$	$\frac{MSB}{MSW}$
<i>Within</i>	$mn - m$	$SSW = \sum_{j=1}^n \sum_{i=1}^m (x_{ij} - \bar{x}_i)^2$	$MSW = \frac{SSW}{mn - m}$	
<i>Total</i>	$mn - 1$	$SS_{Total} = SSB + SSW$		

Table 4.3: Calculations for F-Statistic

One way to interpret the F-statistic is as a ratio between the “average” variation between groups and the “average” variation within groups. A “large” ratio of these values would suggest most of the variation is in the numerator; that is, between the groups. That would imply that a “large” F-statistic should reject the null hypothesis. A “small” ratio of these values would suggest most of the variation is in the denominator; that is, within groups. That would imply that a “small” F-statistic should not reject the null hypothesis. The p-value and a table of F-values determine what is a “small” or “large” F-statistic.

F-statistic ANOVA: Is There a Difference in Responses among McGill Words?

This example uses the data from the *Neuroscience Pain Data* spreadsheet to determine if there is a difference in the mean responses among the McGill words (columns E, G, I, K, M, and O).

Step 1. State the null hypothesis and α -level:

$$H_0: \mu_E = \mu_G = \mu_I = \mu_K = \mu_M = \mu_O; \alpha = 0.05$$

Step 2. Construct the F statistic table

<i>Source</i>	<i>d.f.</i>	<i>Sum of Squares</i>	<i>Mean Squares</i>	<i>F</i>
<i>Between</i>	3	31.3075	10.4358	4.6687
<i>Within</i>	$128 - 3 = 125$	279.4113	2.2353	
<i>Total</i>	127	310.7188		

Step 3. Calculate the p-value

$$Pr(|F| \geq 4.6687) = fdist(4.6687, df \text{ numerator}, df \text{ denominator}) = 0.0040$$

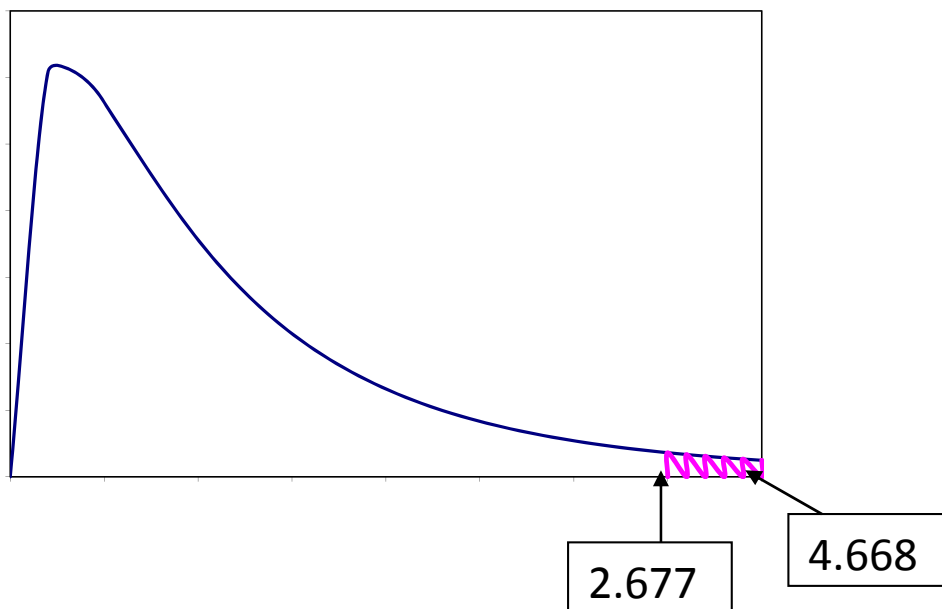
Step 4. Compare the results:

$$Pr(|F| \geq 4.6687) = 0.0040 < 0.05 = \alpha$$

or

$$F - \text{Table Value} = finv(0.05, 3, 125) = 2.6771 < 4.6687$$

This would suggest we reject the null hypothesis (see diagram below). That is, the analysis suggests that mean responses for the McGill words are different.



Since the calculated F-statistic, 4.6687 is in the reject region, the appropriate conclusion is to reject H_0 .

ACTIVITY 4-1 Another Look at McGill Pain Words using F-Statistic ANOVA

Objective: Conduct and analyze results of an F-statistic ANOVA

Materials:

Handout NP-H11: Another Look at McGill Pain Words Activity Worksheet
Neuroscience Pain Data Spreadsheet

1. Use the *Neuroscience Pain Data* spreadsheet to determine if the mean responses for the Student columns (F, H, J, L, N, P) are the same.

2. Describe specifically using the groups and pain words the results of your analysis.

Contingency Tables: The Chi-squared test

The introduction to this lesson described another type of comparison that is heavily used in survey research: the contingency table or cross tabulation. This simplest form of a contingency table is a two-by-two comparison of two groups in terms of their “binary” response to a question. In terms of our pain survey, we might be interested in determining if males and females have different perceptions of “high” versus “low” pain responses to some of the questions. The survey asks respondents to rank pain levels on a scale of 1 to 7. “High” pain could be categorized as a 5, 6, or 7 response; whereas “low” pain could be categorized as a 1, 2, 3, or 4 response. With this somewhat arbitrary scaling of “high” and “low,” it could then be possible to compare the “high” or “low” pain responses between a males and females. This “two-by-two” comparison can be represented by the table that appears in the beginning of this section, repeated here with column and row totals added.

Two-by-Two Comparison	Low (< 4)	High (>5)	Total
Male	a	b	a+b
Female	c	d	c+d
Total	a+c	b+d	a+b+c+d

There are several types of statistical tests (or measures of association) used to compare the values of a contingency table. Some more common ones include Pearson’s chi-squared test (χ^2), the G-test, Fisher’s exact test, and Barnard’s test. The most commonly used test in comparing values a, b, c, and d in a two-by-two table is the chi-squared test ($\chi^2(k)$, with k degrees of freedom). The Chi-squared distribution is another example of a non-symmetric distribution.

The χ^2 -squared test is used with categorical data to evaluate the likelihood that an observed difference between the groups occurred by chance. Contingency tables are one such application of the χ^2 -squared test. Calculating the χ^2 -squared statistic (X^2) for a two-by-two table is done using the following formula:

$$X^2 = \frac{(ad - bc)^2(a + b + c + d)}{(a + b)(c + d)(b + d)(a + c)}$$

A graph of the pdf for the Chi-squared distribution appears below.

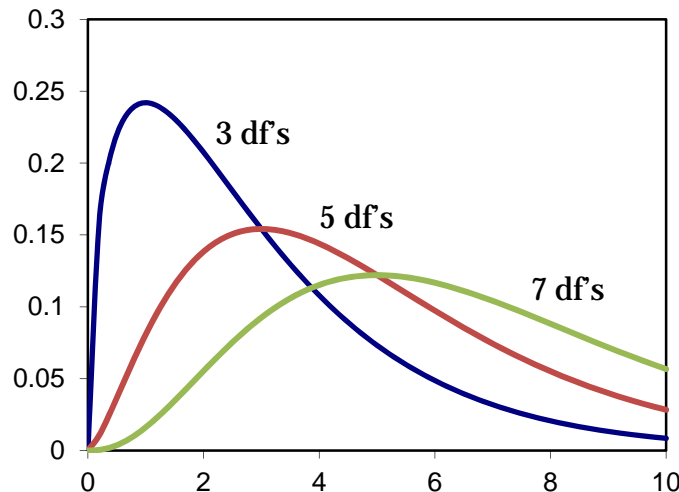


Figure 4.5: Pdf for $\chi^2(k)$ -distribution ($k = 3, 5, \& 7$ degrees of freedom)

The process by which the chi-squared (χ^2) distribution is calculated is once again similar to the Student's t-distribution and the F-statistic:

Step 1. State the null hypothesis H_0 and α -level.

Step 2. Appropriate calculations: standard error, test statistic (in this case, the χ^2 statistic), degrees of freedom, etc. (Again, all calculations are done using technology, not by hand.)

Step 3. Calculate the p-value, the probability that a test statistic of that value would be observed with the given degrees of freedom.

Step 4. At the stated α -level, determine if the null hypothesis is rejected or not rejected based on the p-value or, equivalently, where the test statistic falls along the graph of the population distribution in comparison to the critical score obtained from a table of values or the appropriate statistical software command.

Is There a Difference in Responses to McGill Question 5 between teenage students and adults?

This example uses the data from the Neuroscience Pain Data spreadsheet to determine if there is a difference in responses to the McGill question 5 (column I) between teenage students and adults. The first step is to represent the data in a “binary” fashion. The binary representation will compare “high” responses to “low” responses. Any response of 5, 6, or 7 in Q3 will be labeled as “high.” Any response of 1, 2, 3, or 4 will be labeled as “low.” Two-by-two contingency tables

compare binary categorical data. With this labeling, the contingency table can now be constructed for Q3 from the Neuroscience Pain Data spreadsheet:

Two-by-Two Comparison	Low (< 4)	High (>5)	Total
Teenager	10	3	13
Adult	3	5	8
Total	13	8	21

Step 1. State the null hypothesis and α -level: The chi-squared test helps decide if there is a statistically significant difference in the response between the groups.

H_0 : *There is no significant difference between the category of Teenager – Adult and the frequency of high – low response;*
 $\alpha = 0.05$

Step 2. Calculate the X^2 statistic:

The first step in calculating the X^2 statistic is to determine the degrees of freedom. With contingency tables, the degrees of freedom are one less than the number of categories. Since there are two categories, the degrees of freedom in this example is 1: $df = 1$. Next, calculate X^2 :

$$X^2 = \frac{(ad - bc)^2(a + b + c + d)}{(a + b)(c + d)(b + d)(a + c)} = \frac{(10 \cdot 5 - 3 \cdot 3)^2(10 + 5 + 3 + 3)}{(10 + 5)(3 + 3)(5 + 3)(10 + 3)} = 3.77$$

Step 3. Calculate the p-value: Using excel and/or a chi-squared distribution table for α level 0.05,

$$Pr(|\chi^2| \geq 3.77) = \chi^2 dist(3.77, df = 1) = 0.0522$$

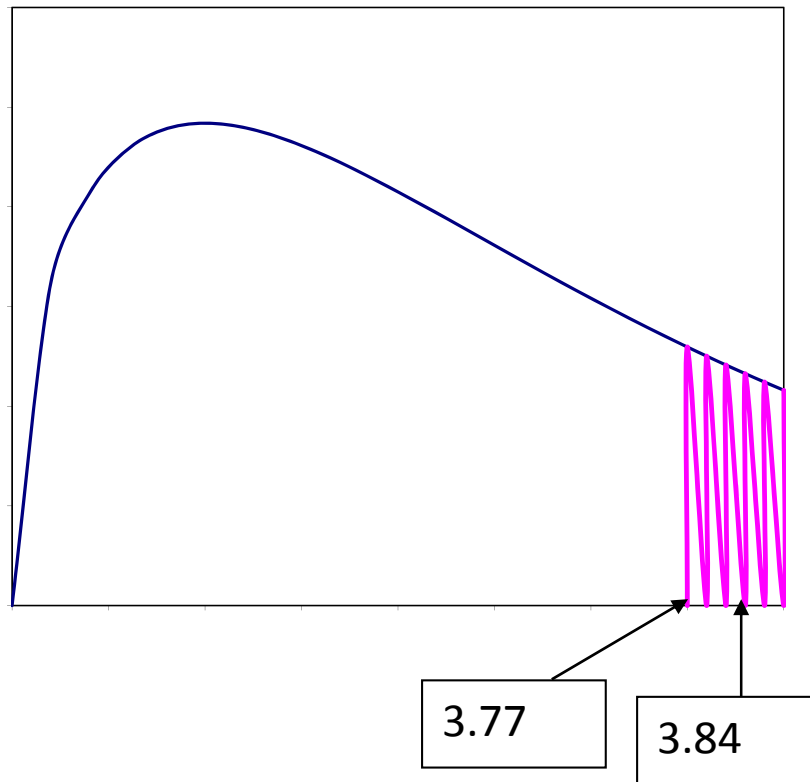
Step 4. Compare the results: Using excel and a chi-squared table of values,

$$Pr(|\chi^2| \geq 3.77) = 0.0522 > 0.05 = \alpha$$

or

$$\chi^2 - Table Value = 3.84 > 3.77$$

This would suggest we do not reject the null hypothesis (see diagram below). That is, the analysis suggests that there is no significant difference in the high-low response between teenagers and adults for Q3 on the McGill questionnaire.



Since the calculated X^2 -statistic, 3.77 is not in the reject region, the appropriate conclusion is to not reject H_0 .

ACTIVITY 4-2 Looking at High and Low Pain Using Chi Squared Test

Objective: Conduct and analyze results of a Chi Squared test to determine if there is a difference in high (5, 6, or 7) – low (1, 2, 3, or 4) responses to the Student survey question 6 (column J) between males and females.

Materials:

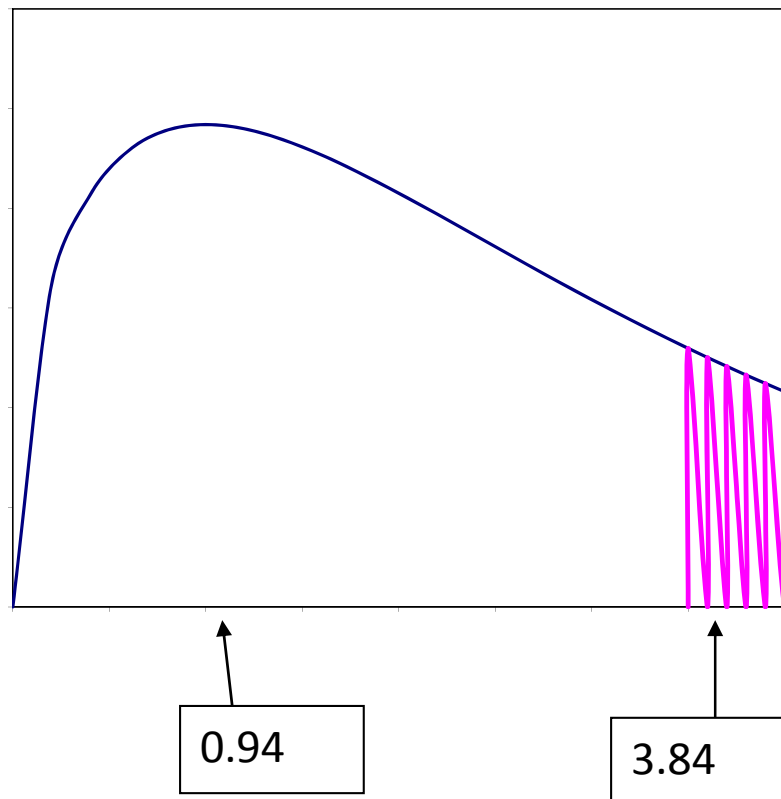
Handout NP-H12: Looking at High and Low Pain Using Chi Squared Test Activity Worksheet

Neuroscience Pain Data Spreadsheet

Use data from the Neuroscience Pain Data spreadsheet to determine if there is a difference in high (5, 6, or 7) – low (1, 2, 3, or 4) responses to the Student survey question 6 (column J) between males and females.

Two-by-Two Comparison	Low (< 4)	High (>5)	Total
Males	3	4	7
Females	8	6	14
Total	11	10	21

1. State the null hypothesis and α -level: The chi-squared test helps decide if there is a statistically significant difference in the response between the groups.
2. Calculate the X^2 statistic:
3. Calculate the p-value: Using a chi-squared distribution table for α level 0.05,
4. Compare the results to determine whether or not to reject the null hypothesis.



Practice

1. Using the *Neuroscience Pain Data* spreadsheet, determine if there is a difference in the mean responses among...
 - a. McGill words in columns G, K, and M.
 - b. McGill words in columns E, I, and K.
 - c. McGill words in columns E, G, K, and M.
 - d. Student survey words in columns F, H, and J.
 - e. McGill words in columns H, J, and L.

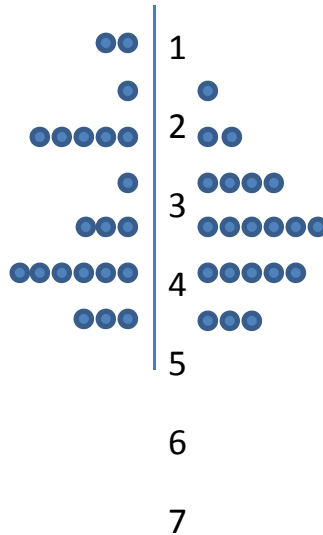
2. Using the *Neuroscience Pain Data* spreadsheet, determine if there is a difference in high (5, 6, or 7) – low (1, 2, 3, or 4) responses to...
 - a. McGill question 3 (column G) between males and females.
 - b. Student survey question 2 (column F) between adults and teenagers.
 - c. McGill question 9 (column M) between adults and teenagers.
 - d. Student survey question 4 (column H) between males and females.

ASSESSMENT - Modeling Neuron Networks: The Neuroscience of Pain

Part I. Fill in the Blank

1. _____ fibers are a type of nociceptive axon that have no myelin and are the thinnest sensory axon in the skin.
2. _____ fibers are a type of nociceptive axon that have myelin and move at a much more rapid pace.
3. Words that describe how things feel or seem to a perceiving subject are classified as _____ qualities.
4. Words that describe features of an experience which render it pleasurable, or desirable, or the reverse, or which gives it a distinctive emotional tone are classified as _____ qualities.
5. Words that describe the subjective overall intensity of the total experience of pain are classified as _____ qualities.
6. When it is possible to order the categories, such categorical data is referred to as _____ data.
7. A survey that assigns an ordinal number to each choice indicating where that choice falls within a given range from lowest to greatest is called _____.
8. _____ data is mathematical in nature allowing statisticians to conduct mathematical operations on the data.
9. A _____ is a chart consisting of data points (such as filled circles) placed along a simple scale such as a number line.
10. A convenient way of depicting numerical data through their five-number summaries (minimum, first quartile, median, third quartile, maximum) is called a _____.

Part II. The side-by-side dot plot contains the responses from a survey on pain. The dots on the left reflect subjects' responses to a sensory question. The dots on the right reflect subjects' responses to an affective question.



1. Calculate the first quartile for both Q1 and Q2.
2. The other 4 numbers for the 5-number summary for Q1 are minimum=1, median=5, 3rd quartile=6, max=7. Using the first quartile found in 1 above, construct a boxplot for the Q1 data.
3. The other 4 numbers for the 5-number summary for Q2 are minimum=2, median=5, 3rd quartile=6, max=7. Using the first quartile found in 1 above, construct a boxplot for the Q2 data.
4. Draw the boxplots side by side and interpret your answer.

Part III. The dataset below consists of two groups to a question on pain: Q1 represents the responses by males to a pain Sensory question, Q2 represents the responses by females to the same Sensory question. Use this dataset to answer each of the following questions:

Q1 4 5 4 4 2 6 4 6 7 5 7 6 3 5 6 6 5 5 3 7 5
Q2 4 5 5 5 6 4 3 5 4 4 5 2 6 6 6 4 5 2 2 2 4

1. Explain what t-test is used to compare male and female responses (Q1 and Q2) to this question.
2. State the null hypothesis.

3. Calculate the degrees of freedom for the t-test.
4. Calculate the t-statistics for the t-test.
5. Using $\alpha = 0.05$, and the t-table provided, what conclusions can be drawn regarding male and female responses to this sensory question?

Part IV: The dataset below consists of two responses to two questions on pain by the same group of people: Q1 represents the responses by the group to a pain Sensory question, Q2 represents the responses by the same group to an Affective question. Use this dataset to answer each of the following questions:

Q1 2 3 6 5 4 4 4 2 4 6 5 2 3 7 5 2 5 3 2 1 4
Q2 5 3 6 4 3 5 3 3 6 6 5 2 2 6 5 3 4 3 4 1 7

1. Explain what t-test is used to compare male and female responses (Q1 and Q2) to this question.
2. State the null hypothesis.
3. Calculate the degrees of freedom for the t-test.
4. Calculate the t-statistics for the t-test.
5. Using $\alpha = 0.01$, and the t-table provided, what conclusions can be drawn regarding male and female responses to this sensory question?

Glossary

A-delta fibers – a type of sensory fiber that transmits information to the brain in response to “fast” pain stimuli such as cold or pressure.

Affective – relating to, arising from, or influencing feelings or emotions; refers to the unpleasantness and urge to escape the unpleasantness.

Analgesics – members of the group of drugs used to relieve pain.

Analysis of Variance (ANOVA) – a collection of statistical models used to analyze the differences between group means and their associated procedures.

Axons – sensory fibers leading from the skin to the spinal cord.

Bar chart – a chart that uses horizontal or vertical rectangular boxes with lengths proportional to the value they represent. Also called a bar graph.

Bar graph – a chart that uses horizontal or vertical rectangular boxes with lengths proportional to the value they represent. Also called a bar chart.

Between groups – analysis that is done to compare two independent samples from the population. For example, to compare males vs. females.

Blocks – identified groups of data in statistical studies.

Box-and-whisker plot – a representation of numerical data using its five-number summary statistics (minimum, first quartile, median, third quartile, maximum). Also known as a boxplot.

Cannabinoids – a type of drug that inhibits pain.

Capsaicin – a chemical sensed as “hot” in our taste sense.

Categorical Data – data from observations whose value is one of a fixed number of nominal categories.

Cerebral aqueduct – a channel in the brain connecting the third and fourth ventricles and surrounded by the periaqueductal gray.

C-fibers – a type of sensory fiber that transmits information to the brain in response to lingering pain stimuli such as burning or aching.

Continuous data – data that can take on any value within a range.

Cross-modality matching – a type of survey using assignment of ordinal numbers within a given range as responses to show lowest to greatest.

Degrees of freedom – the number of values in the final calculation of a statistic that are free to vary.

Descriptive Statistics – mathematical quantities (such as mean, median, mode, etc.) that summarize and interpret some of the properties of a set of data (sample) but do not infer the properties of the population from which the sample was drawn.

Difference variable – the variable of interest in finding whether or not there is a significant difference between or within data sets.

Discrete data – data that can take on only a finite number of values within a range.

Dot plot – a graphical display of data using dots usually depicting the frequency of possible groups within categorical data.

Evaluative – to determine the significance, worth, or condition of, usually by careful appraisal or study; refers to cognitions such as appraisal, cultural values, distraction, and hypnotic suggestion.

Free nerve endings – thin branches of sensory nerves that simply branch with the skin.

F-test – an ANOVA (analysis of variance) test done to calculate a F-statistic to determine whether or not there is variance in sample means. Also known as a one-way ANOVA.

Inferential Statistics – mathematical methods that employ probability theory for deducing (inferring) the properties of a population from the analysis of the properties of a data sample drawn from that population.

Level of significance – the probability of rejecting the null hypothesis of a statistical test when it is actually true.

Likert scale – a method of assigning a quantitative value to qualitative data usually within the creation of a survey or other data collection.

Myelin – a mixture of proteins and phospholipids that form a sheath around nerve fibers.

Nociceptor – a receptor (as in the skin) for nociceptive stimuli; a pain sense organ.

Nominal Data – data that establishes an arbitrary numerical scale where the exact numerical quantity of a particular value has no significance beyond its ability to establish a ranking over the set of data points.

Noxious – constituting a harmful influence on mind or behavior.

Null Hypothesis – a proposition that undergoes verification to determine if it should be rejected or not rejected in favor of an alternative hypothesis.

One-way ANOVA – an analysis of variance test to determine whether three or more samples from a population have the same means. Also known as a F-test.

Opiates – a type of drug that inhibits pain.

Ordinal Data – a data type consisting of numerical scores that exist on an ordinal scale; that is, an arbitrary numerical scale where the exact numerical quantity of a particular value has no significance beyond its ability to establish a ranking over the set of data points.

p-Value – the probability of observing a test statistic that is as extreme or more extreme than currently observed assuming that the null hypothesis is true.

Pain – an unpleasant feeling often caused by intense or damaging stimuli, such as stubbing a toe, burning a finger, putting alcohol on a cut, and mental, emotional, or behavioral factors.

Periaqueductal gray – the area of the brain made up of neurons that receive nociceptive input from the spinal cord.

Population – any entire group from which we collect data.

Population distribution – the arrangement or distribution of data of a population across a well-defined domain as the area under a probability curve. The total area under the probability curve equals 1 unit.

Probability Density Function (pdf) – of a random variable is a function that describes the relative likelihood for that random variable to take on a given value.

Psychophysics – the field of study dealing with the relationship between sensory stimuli and the perceptions they cause.

Qualitative Data – data that approximates or characterizes, but does not measure the attributes, characteristics, properties, etc. of a thing or phenomenon.

Quantitative Data – any data that is in numerical form.

Sample – a small part or quantity examined as an indicator of what the whole is like.

Sensory – relating to sensation or to the senses; refers to the intensity, location, and duration of the pain.

Somatosensory pain – pain associated with the interaction of our bodies with our external environment through touch.

Standard deviation – a statistical measure indicating the amount of separation or dispersion from the average of a set of data as a whole.

Standard error – a statistical measure quantifying the accuracy with which a sample represents a population.

Statistically significant – the case when the result of a statistical test indicates that a difference is caused by something other than mere random chance.

Student's t-test – a statistical hypothesis test to calculate a t-statistic to determine differences in samples based on a null hypothesis. Also known simply as the t-test.

Substance P – a neuropeptide that functions as a neurotransmitter of pain.

t-table – a table showing the areas (probabilities) under the probability density function of the t-distribution for different degrees of freedom.

Vanilloid – a type of receptor that contributes to the detection of pain and stimuli.

Variance – a measure of the amount of difference or spread in a set of data.

Within group – analysis that is done to compare two random samples from a population or from an independent group in the population. For example, two samples of women.

References

[1] Melzack, R. and Torgerson, W.S. (1971, January). On the language of pain. *Anesthesiology*. 34(1), 50-9.

[2] Trochim, W.M.K. (2006). The t-test. Research Methods Knowledge Base. Found at http://www.socialresearchmethods.net/kb/stat_t.php.