STATISTICS

FOR HEALTH SCIENCES

Sean Saunders
Sheridan College

Irene Lee
Humber College

Elisa Romeo
Vretta Inc

BLENDED LEARNING

CONCEPTUAL UNDERSTANDING

VOICE-ENABLED, INTERACTIVE LESSONS

DIFFERENTIATED IN-CLASS LEARNING

TEXTBOOK/ETEXTBOOK

ASSESS AND PRACTICE

ALGORITHMIC ASSESSMENTS

Vretta

COLLABORATION

MATHEMATICS
For Health Sciences

Irene Lee
Thambyrajah Kugathasan
Sean Saunders

STATISTICS
For Health Sciences

Sean Saunders
Thambyrajah Kugathasan
Irene Lee

ALONG WITH REVIEWERS FROM ONTARIO COLLEGES

Vretta

FEATURES

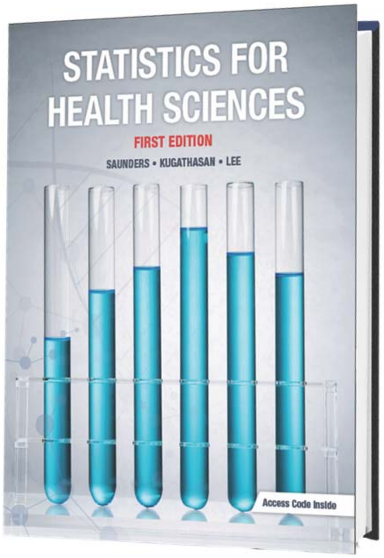
- Aligned to ONCAT's College System Exemplar Course Outlines: Standard & Advanced
- Interactive Online Resources
  - eTextbook
  - Lessons
  - Quizzes/Assignments
  - Solution Manuals
  - PPTs
- Instructor Tools
  - Dashboards
  - Administrative Features

Vretta

CONTENT

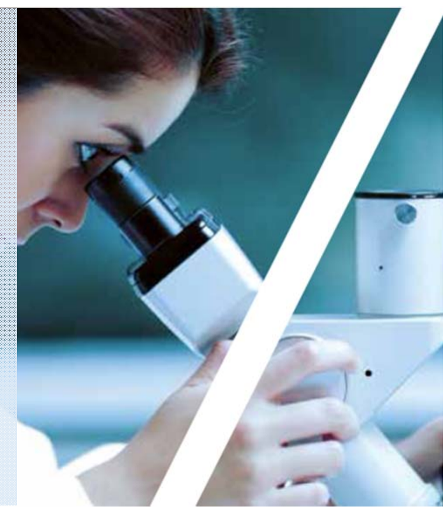
- ➢Introduction to Statistics
- ➢Measures of Central Tendency
- ➢Measures of Dispersion
- ➢Linear Correlation and Regression
- ➢Counting Techniques and Probability
- ➢Discrete Probability Distributions
- ➢Continuous Probability Distributions
- ➢Sampling Distributions & Central Limit Theorem
- ➢Estimation and Confidence Intervals
- ➢Hypothesis Testing

Vretta



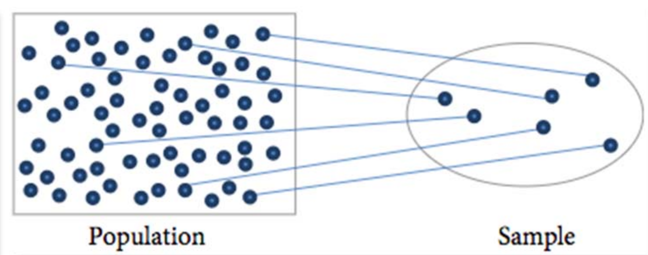
1 INTRODUCTION TO STATISTICS

- Define terms used in statistics and distinguish between descriptive and inferential statistics, populations and samples.
- Identify the benefits and drawbacks of sampling methods used in statistics, along with when to use each of them.
- Distinguish among the types of studies used in statistics, including experimental studies, observational studies, and surveys and censuses.
- Differentiate among qualitative, continuous quantitative, and discrete quantitative variables, as well as nominal, ordinal, interval, and ratio levels of measurements.
- Organize and present qualitative data using tally charts, frequency distribution tables, pie charts, and bar charts.
- Organize and present quantitative data using tools, including stem-and-leaf plots, tally charts, frequency tables, histograms, frequency polygons, ogives, and line graphs.

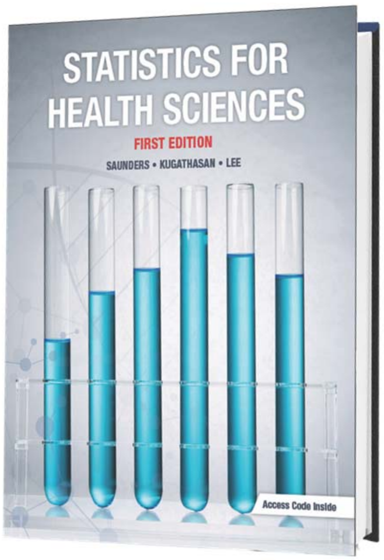


SAMPLE

Population	Sample
• refers to a set of all possible individuals, objects, or measurements of items of interest	• refers to a subset drawn from the population, meaning a portion or part of the population
• the size of a population is usually very large, or even infinite	• the size of a sample is usually much smaller and more manageable than the size of the population from which it is drawn



Vretta



2 MEASURES OF CENTRAL TENDENCY

- ■ Define and calculate various measures of central tendency for ungrouped data, including mean, median, and mode.
- ■ State the advantages and disadvantages of the various measures of central tendency and indications for the use of each.
- ■ Identify outliers and describe their effect on the arithmetic mean and compute the mean of a trimmed data set.
- ■ Calculate the weighted arithmetic mean and various non-arithmetic means, including the geometric mean and the harmonic mean.
- ■ Calculate the value of the mean, median, and mode for grouped data with unique classes or class intervals.
- ■ Identify the relationships between mean, median, and mode for symmetrical and skewed distributions

SAMPLE

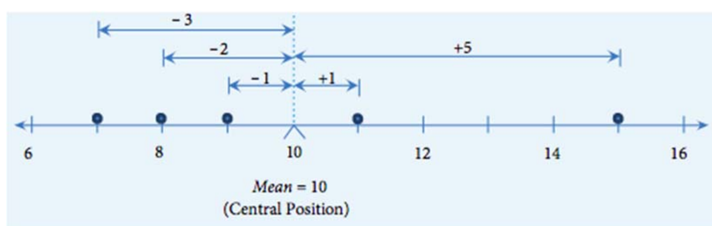
(i) {8, 15, 7, 11, 9}   There are 5 terms.

To determine the mean, add all the terms then divide by the number of terms.

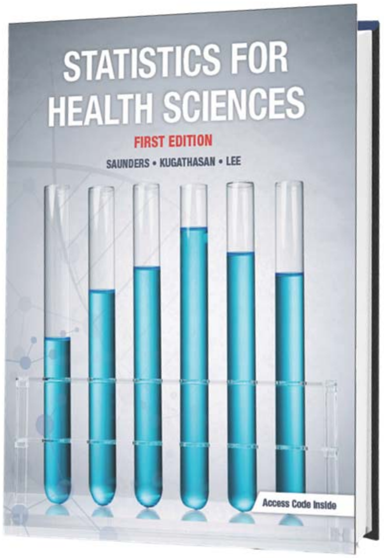
$$Mean = \frac{Sum\ of\ All\ the\ Values\ of\ the\ Terms}{Number\ of\ Terms} = \frac{(8 + 15 + 7 + 11 + 9)}{5} = \frac{50}{5} = 10$$

Therefore, the mean (or average) number of patients seen by the nurse per day this week is 10.

(ii) If we plot these values on a number line, the number 10 represents the centre of these values.



Vretta



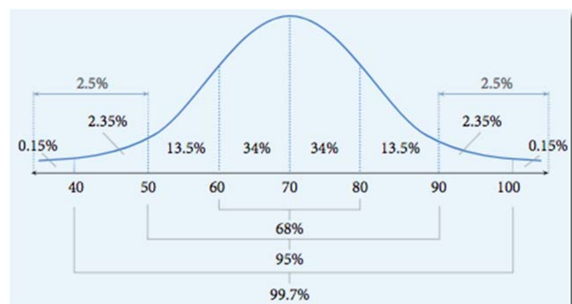
3 MEASURES OF DISPERSION

3.1 Measures of Position and Range
3.2 Measures of Deviation

- Describe data using various measures of position, including quartiles, percentiles, range, and interquartile range (IQR).
- Construct and interpret a box-and-whisker plot and use it to identify outliers.
- Calculate and interpret various measures of dispersion, including the mean absolute deviation (MAD), variance, and standard deviation.
- Compare the standard deviations of two or more data sets to identify the set with greater dispersion.
- Interpret the standard deviation of a data set using the Empirical rule and Chebyshev's theorem, and how they relate to the shape of a distribution.
- Calculate the coefficient of variation and use it to compare the variation of two or more data sets

SAMPLE

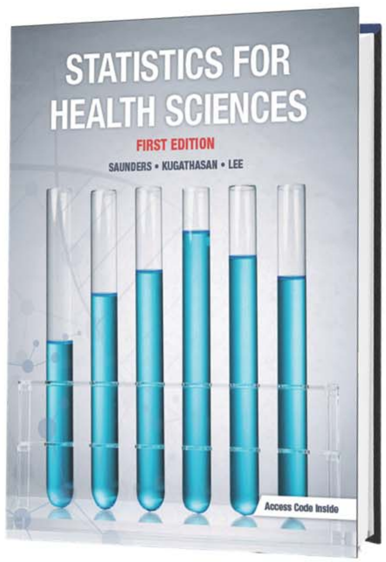
Since the we know that the mean $\bar{x} = 70$ and the standard deviation $s = 10$, we can label the diagram from Exhibit 3.2-c as follows:



(i) The empirical rule states that approximately 68% of the 200 students, or roughly 136 students, will score between 60 and 80 on the exam.

(ii) Again, we can quickly see from the above diagram that approximately the top 97.5% (100% – 2.5%) of students, or roughly 195 students, will score above 50 - i.e., they will pass the exam.

Vretta

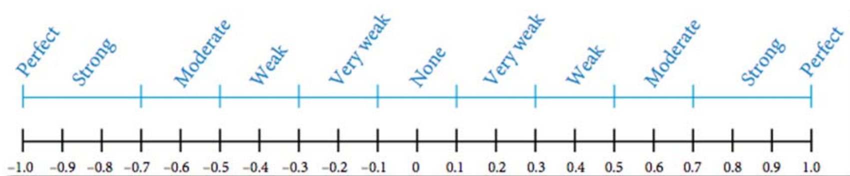


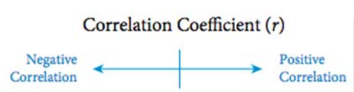
4 LINEAR CORRELATION AND REGRESSION

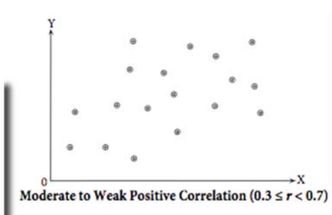
4.1 Scatter Plots and Correlation
4.2 Linear Regression

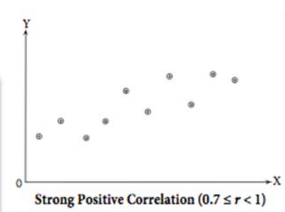
- Differentiate between independent variables and dependent variables.
- Create a scatter plot and identify possible relationships between the two variables.
- Compute the linear regression equation (line of best fit) using the least squares method and plot it against a scatter plot of the data.
- Interpret the slope and y-intercept of the line and use the regression equation to predict the value of the dependent variable for a selected value of the independent variable.
- Calculate the Pearson linear coefficient of correlation (r) to explain the type of relationship and calculate the coefficient of determination (r2) to explain the strength of the relationship.
- Construct a residual plot and identify the outliers from a scatter plot.
- Define the terms 'explained deviation', 'unexplained deviation', and 'total deviation', calculate their values for a given data set, and use them to compute the coefficient of determination (r 2).

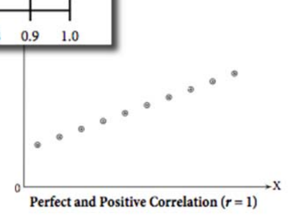
SAMPLE

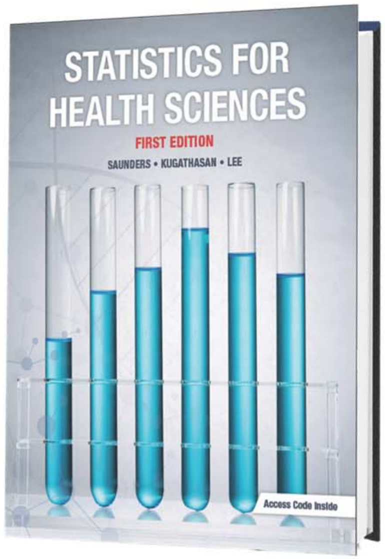












5 ELEMENTARY PROBABILITY THEORY

- Distinguish between theoretical, empirical, and subjective approaches to probability.
- Determine the various outcomes and events for a given experiment.
- Use outcome tables, tree diagrams, and Venn diagrams to construct the sample space for an experiment.
- Distinguish between mutually exclusive and non-mutually exclusive events, as well as independent and dependent events.
- Calculate probabilities using a variety of approaches, including the general sample space approach, addition rule, multiplication rule, complementary event rule, and Bayes' Theorem.
- Calculate the odds against and the odds in favour of an event.
- Solve problems and calculate probabilities using a variety of counting rules, including the fundamental counting rules, factorial rules, permutations rule, and combinations rule.

SAMPLE

Since the students may only write down one choice of program, this tells us the two groups (those who are planning to enrol in a Nursing program and those who are planning to enrol in a Pharmacy program) are mutually exclusive. Hence,

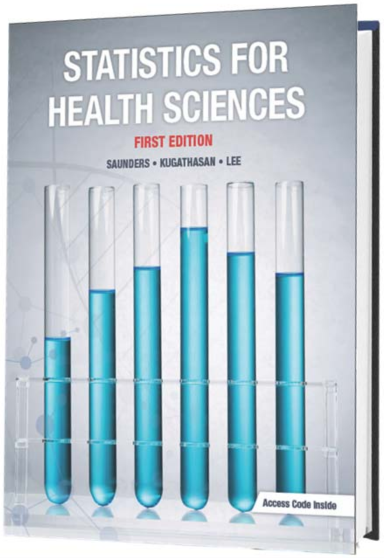
$$P(\textit{Nursing} \text{ or } \textit{Pharmacy}) = P(\textit{Nursing}) + P(\textit{Pharmacy})$$

Moreover, we know that $P(\textit{Nursing}) = \frac{17}{40} = 0.425$ and $P(\textit{Pharmacy}) = \frac{13}{40} = 0.325$. Therefore,

$$P(\textit{Nursing} \text{ or } \textit{Pharmacy}) = \frac{17}{40} + \frac{13}{40} = \frac{30}{40} = \frac{3}{4} = 0.75$$

This tells us that the probability that a randomly-selected student is interested in enrolling in a Nursing program or a Pharmacy program is 0.75.

Vretta



6 DISCRETE PROBABILITY DISTRIBUTIONS

- Differentiate between discrete random variables and continuous random variables.
- Construct probability distributions for discrete random variables.
- Calculate the expected value (mean), variance, and standard deviation of a discrete random variable using its probability distribution.
- State the properties of the binomial probability distribution.
- Interpret and apply the binomial formula to calculate probabilities for a binomial random variable.
- Calculate the expected value (mean), variance, and standard deviation of a binomial random variable.
- Calculate probabilities using the hypergeometric, geometric, and Poisson probability distributions.
- Calculate the expected value (mean), variance, and standard deviation of hypergeometric, geometric, and Poisson random variables.

SAMPLE

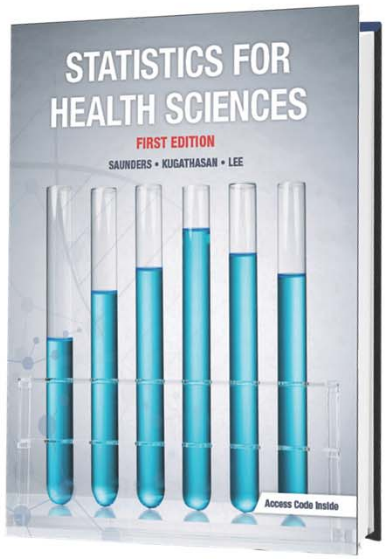
Binomial Formula

$$P(x) = \binom{n}{x} p^x \cdot q^{(n-x)}$$

23. According to the Canada Obesity Network, one in four adult Canadians has clinical obesity. A sample of 32 adult Canadians is taken. What is the probability that exactly half of the 32 adults are obese?

$$P(x) = \binom{32}{16}(0.25)^{16}(0.75)^{16} \approx 0.0014$$

Vretta



7 CONTINUOUS PROBABILITY DISTRIBUTIONS

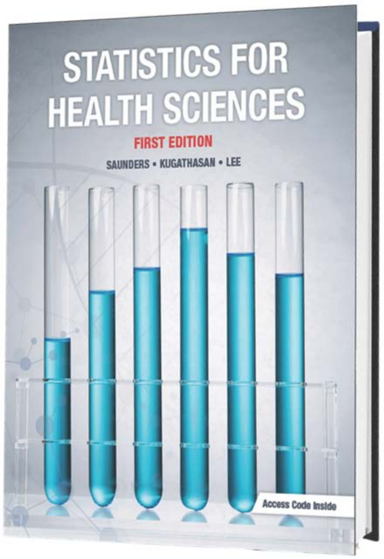
- State the characteristics and the properties of a continuous probability distribution.
- Distinguish between the uniform and exponential continuous probability distributions.
- Calculate probabilities using the continuous uniform and exponential probability distributions, and solve application problems involving both distributions.
- Calculate the expected value (mean), variance, and standard deviation for uniform and exponential continuous random variables.
- Compare the properties of the normal distribution and standard normal distribution.
- Calculate the probabilities for the normal distribution by calculating the z-score and using the table of areas under the standard normal distribution curve and the Z-scores.
- Solve normal distribution problems corresponding to intervals above, below, or between given values.
- Convert raw X-scores to standardized Z-scores, and vice versa.
- Use the normal distribution to approximate the binomial probability distribution.

SAMPLE

32. Human pregnancies of healthy singleton fetuses (i.e., excluding pregnancies with early-term complications or miscarriages, and multiple births) are normally distributed with a mean of 266 days and a standard deviation of 13 days.

c. If we stipulate that a baby is overdue when born more than 2 weeks after the expected date, what is the probability a randomly selected baby is overdue?

Vretta



8 SAMPLING DISTRIBUTIONS AND THE CENTRAL LIMIT THEOREM

- Identify the characteristics of the sampling distributions of the sample means, sample proportions, and sample correlation coefficients for a random variable.
- Distinguish between the individual data values in a population distribution of a random variable and the grouped data values in a sampling distribution for that random variable.
- State the central limit theorem and use it to calculate the mean and standard error of the sampling distribution of sample means for a random variable.
- Define sampling error and standard errors and compute the standard error of the mean.
- Compute probabilities and predict ranges of values for the X-distribution.
- Apply the finite population correction factor to correct the standard error of the mean.
- Compute the mean and standard error for the sampling distribution of sample proportions.
- Use the normal distribution to approximate the p-distribution and predict a range of values for the p-distribution.

SAMPLE

	Population Distribution	Sampling Distribution
What is being sampled?	Individual elements drawn randomly from the population	Samples of a fixed size *n*, drawn randomly from the population
What data is being recorded?	Individual values of each of the elements drawn	Statistics from each of the samples drawn
What is the shape of the distribution (i.e., what does the distribution look like)?	Could be anything, based on all the possible individual outcomes of the population	As the sample size increases, the sampling distribution tends to approach a "nice" shape (e.g., a normal distribution curve)

Population vs Sampling Distributions

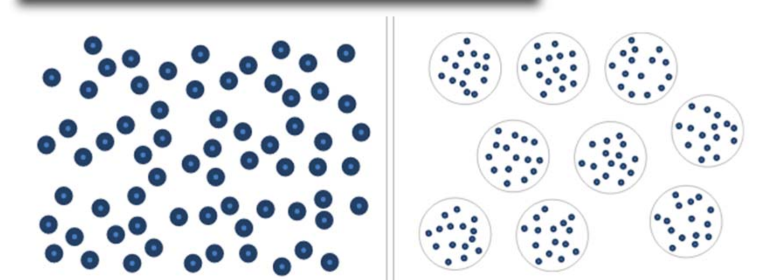
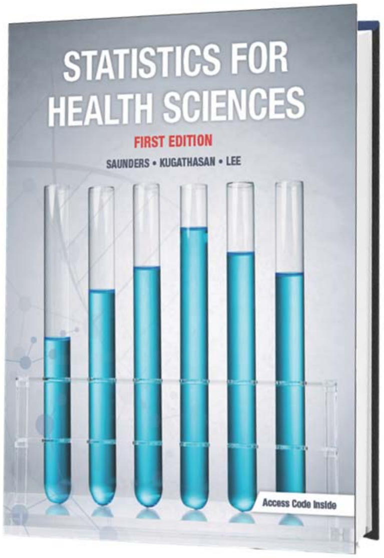


Exhibit 8.1-a Population Distribution

Exhibit 8.1-b Sampling Distribution

Vretta



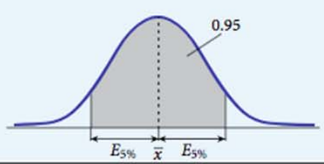
9 ESTIMATION AND CONFIDENCE INTERVALS

- Define inferential statistics, population parameters, and biased and unbiased estimators.
- Differentiate between point estimates and range (or interval) estimates.
- Identify the appropriate population parameter based on the sample data and the estimator for that parameter.
- Define and interpret confidence level, level of significance, and margin of error.
- Construct confidence intervals for the population mean when the population standard deviation is known, using the normal Z-distribution.
- Compare accuracy and precision of confidence intervals and determine the sample size required for a given accuracy and precision.
- State the characteristics and properties of the Student's t-distribution.
- Construct confidence intervals for the population mean when the population standard deviation is unknown, using the Student's t-distribution.
- Construct confidence intervals for population proportions.

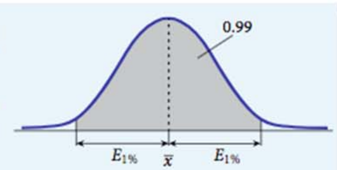
SAMPLE

The sample mean is $\bar{x} = 5.2$ hours and the margin of error at the 5% level of significance is $E = 0.35$ hours. Hence, we are 100% − 5% = 95% confident that the population mean will fall within the following interval:

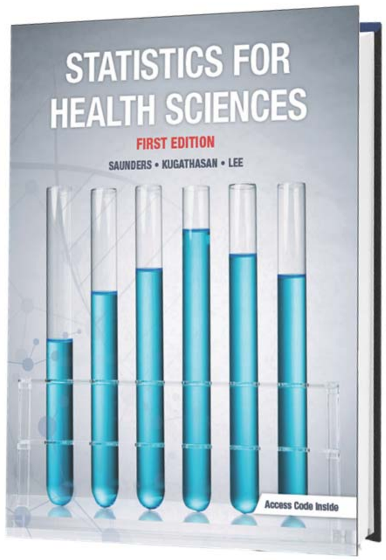
$$\bar{x} - E < \mu < \bar{x} + E$$
$$5.2 - 0.35 < \mu < 5.2 + 0.35$$
$$4.85 < \mu < 5.55 \text{ hours}$$



Since the probability that the population mean will not be contained within this range is only 1% (99% confidence interval), which is less than the original probability of 5% (95% confidence interval), the margin of error at the 1% level of significance must be **greater than** 0.35 hours.



Vretta



10 HYPOTHESIS TESTING

- Define the terms used in hypothesis testing, including null and alternative hypotheses, one- and two-tailed tests, test statistic, p-value, level of significance, and type I and type II errors.
- State the procedure for evaluating hypotheses using the 5 steps for formal hypothesis testing.
- Conduct and evaluate hypothesis tests of a population mean using the normal Z-distribution when the population standard deviation is known (or a large sample size is used).
- Conduct and evaluate hypothesis tests of a population mean using the Student's t-distribution when the population standard deviation is unknown.
- Conduct and evaluate hypothesis tests of a population proportion using the normal Z-distribution.
- Conduct and evaluate hypothesis tests for the linear correlation coefficient using the test statistic approach and the critical r-value approach.

SAMPLE

A researcher believes that people living in the GTA are at a higher risk of diabetes than the general population in Canada. A random sample of 400 residents of the GTA is collected, of which 43 have diabetes. Test the researcher's claim at the 10% level of significance if the proportion of all Canadians who have been diagnosed with diabetes is 0.093.

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.1075 - 0.093}{\sqrt{\frac{0.093(0.907)}{400}}} \approx 1.00 \; and \; p = 0.1587 > \alpha = 0.10$$

Since p > α, there is not sufficient evidence to reject the null hypothesis that the proportion of residents of the GTA is the same as the national proportion. Hence, the researcher cannot conclude that the prevalence of diabetes in the GTA is any higher than it is in Canada.

Vretta

Colleges Currently Using Statistics For Health Sciences

Seneca

HUMBER

Lambton College

Sheridan

ALGONQUIN COLLEGE

St. Lawrence College

nc Niagara College Canada

SAULT COLLEGE

Vretta

STUDENT SURVEY

STATISTICS FOR HEALTH SCIENCES

94%
Agreed or **Strongly Agreed**

"The Vretta Online Quizzes Helped Me Prepare For The Tests"

84%
Agreed or **Strongly Agreed**

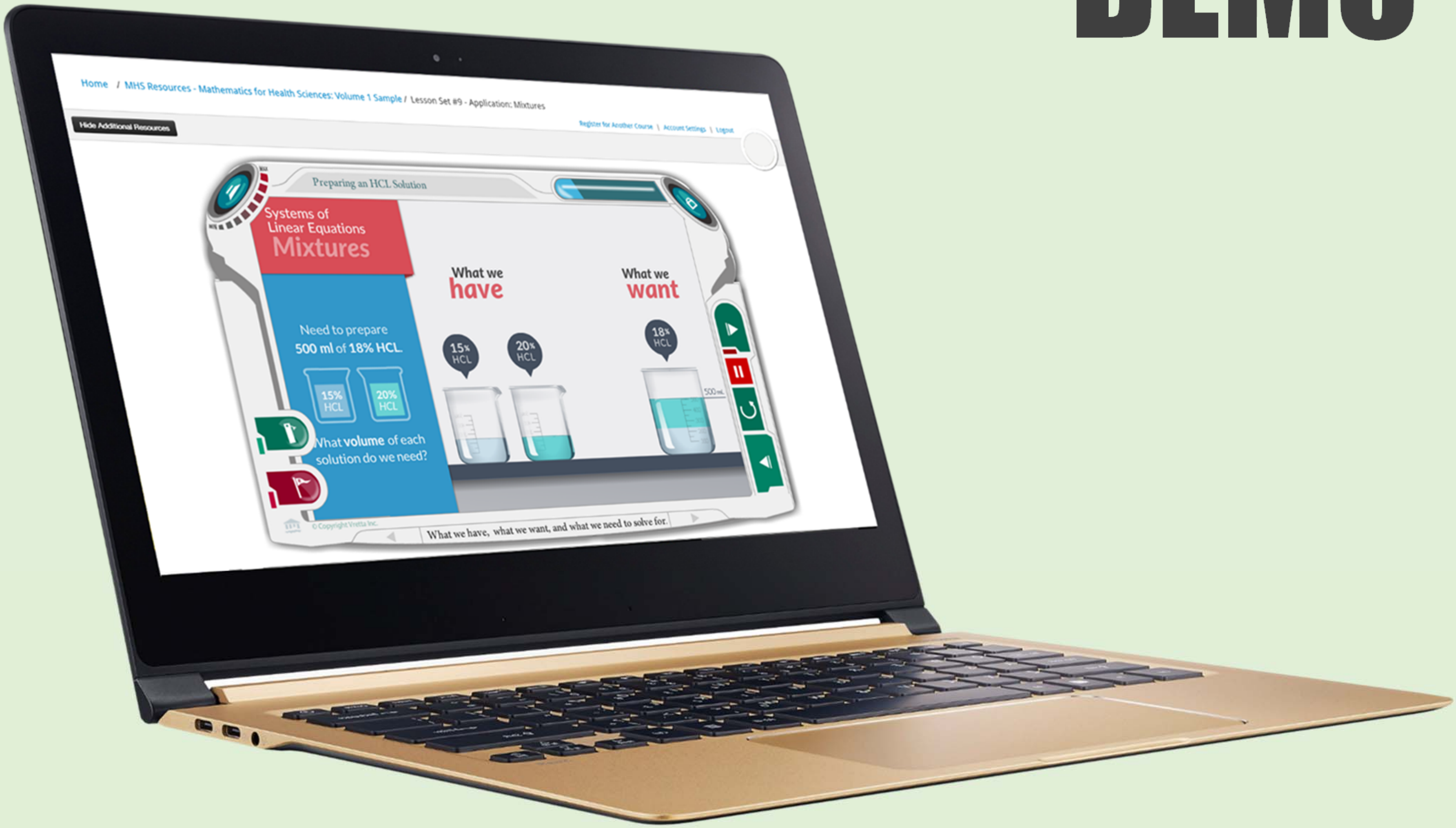
"The Textbook is Easy to Understand with Lots of Examples."

83%
Agreed or **Strongly Agreed**

"The Vretta Resources (Textbook and Online Lessons) Helped Me Understand the Learned Material"

Vretta

DEMO


Home / MHS Resources - Mathematics for Health Sciences: Volume 1 Sample / Lesson Set #9 - Application: Mixtures
Hide Additional Resources
Preparing an HCL Solution
Systems of Linear Equations
Mixtures
Need to prepare 500 ml of 18% HCL.
15% HCL
20% HCL
What volume of each solution do we need?
What we have
What we want
15% HCL
20% HCL
18% HCL
500 mL
What we have, what we want, and what we need to solve for.

Vretta