MOHAWK VALLEY COMMUNITY COLLEGE

 UTICA AND ROME, NEW YORK

 **COURSE OUTLINE**

 **and**

 **TEACHING GUIDE**

 **ELEMENTARY STATISTICS**

 **MA110**

Reviewed and Revised, as needed, 1998 through 2000

Reviewed and Revised, as needed, 2001 through 2007

Reviewed and Found Acceptable – 5/08

Reviewed and Revised – 10/09

Reviewed and Found Acceptable 5/10

Reviewed and Found Acceptable 5/11

Reviewed and Revised 8/12

Reviewed and Found Acceptable 5/13

Reviewed and Found Acceptable 5/14

Reviewed and Revised 5/16

Reviewed and Revised 4/18

 COURSE OUTLINE

Title: Elementary Statistics

Catalog #: MA110

Credit Hours: 3

Lecture Hours: 3 for lecture version; 2 for lab version

Lab Hours: 0 for lecture version; 2 for lab version

Prerequisite: An appropriate high school GPA, or placement test score, or MA089 Arithmetic.

Catalog Description:

This course introduces probability and statistics. Topics include graphs, tables, frequency distributions, measures of central tendency and dispersion, normal distribution, correlation and regression, probability, and inferential statistics. This course is available in two formats: lecture only, or lecture plus laboratory using technology.

**GOALS AND OUTCOMES**

The following goals and outcomes are independent of presentation format.

**COURSE TEACHING GOALS FOR ALL TOPICS:**

**GOAL A:** Use mathematical processes to acquire and convey knowledge.

**GOAL B:** Systematically solve problems and interpret information or data.

STUDENT LEARNING GOALS FOR MA110: ELEMENTARY STATISTICS

At a level appropriate for a first semester college level elementary statistics course, students successfully completing the course will be able to:

1. Demonstrate an awareness of the historical development of statistical methods and techniques and how they relate to various disciplines

2. Analyze data using a variety of statistical techniques

3. Demonstrate an ability to formulate statistical statements, reason and draw appropriate conclusions from data, and critique conclusions drawn by others

4. Demonstrate an understanding of the fundamentals of descriptive statistics, including the computation and interpretation of summary statistics, the construction and interpretation of graphs, the exploration of relationships, and the formulation of statistically-based predictions.

5. Demonstrate an understanding of the fundamentals, usefulness, and implications of random sampling and experimentation

6. Demonstrate an understanding of the fundamentals of probability-based statistical inference, including the properties of sampling distributions, the determination of normal distribution probabilities, the construction of confidence intervals and the testing of hypotheses.

7. Demonstrate the ability to use technological tools (such as a calculator and/or computer) as they relate to statistical concepts

8. Demonstrate an understanding of the course concepts by communicating through appropriate verbal, written, graphical and other means

**SUNY Learning Outcomes**

1. The student will develop well reasoned arguments.
2. The student will identify, analyze, and evaluate arguments as they occur in their own and other’s work.
3. The student will demonstrate the ability to interpret and draw inferences from mathematical models such as formulas, graphs, tables, and schematics.
4. The student will demonstrate the ability to represent mathematical information symbolically, visually, numerically, and verbally.
5. The student will demonstrate the ability to employ quantitative methods such as arithmetic, algebra, geometry, or statistics to solve problems.
6. The student will demonstrate the ability to estimate and check mathematical results for reasonableness.
7. The student will demonstrate the ability to recognize the limits of mathematical and statistical methods.

# TOPIC 1. OVERVIEW: ORIENTATION, INTRODUCTION AND TERMINOLOGY

An introduction to the statistical method, including examples of the uses and abuses of statistics, will be discussed. The historical development of statistical methods and techniques will be referenced. The definition of such terms as sample, population, inferential vs. descriptive statistics, and statistical methods will be emphasized. These terms may be scattered throughout the text (i.e., need not be in a specific chapter). Other terms, referenced in the specific Topics below, will be defined by the instructor as encountered. These include descriptive statistics, percentiles, frequency, relative frequency, mode, range, and point estimate.

**Topic Goal:** To help students develop an understanding of the fundamentals and usefulness of statistical methodology.

**Student Outcomes:** The student will:

* 1. Define and illustrate basic statistical terms (sample, population, variables, descriptive statistics, inferential statistics, etc.)
	2. Relate specific types of statistical situations to the historical development of statistical methods and techniques

**TOPIC 2. FREQUENCY DISTRIBUTIONS AND GRAPHING**

The techniques for grouping data to form frequency distributions are discussed. Types of graphs and methods of constructing graphs are discussed. Interpretation of graphs is emphasized.

**Topic Goal:** To help students develop an understanding of the fundamentals and usefulness of data summarization into a frequency distribution and graphical data analysis.

**Student Outcomes:** The student will:

* 1. Organize raw data into a frequency distribution
	2. Represent the data graphically using the appropriate graphical tool (bar chart or histogram)
	3. Convert frequency to relative frequency to percent
	4. Deduce characteristics of a data set by examining the graph of the distribution; these include shape, center, spread, overall pattern, and outlier identification

**TOPIC 3. MEASURES OF CENTRAL TENDENCY**

Computation of the mode, median, and mean for data is discussed. Interpretation of the measures of central tendency is emphasized. Comparisons between the measures of central tendency are made, together with some applications.

**Topic Goal**: To help students develop an understanding of the fundamentals and usefulness of the measures of central tendency: mode, median and mean.

**Student Outcomes:** The student will:

* 1. Determine a given distribution’s mode(s)
	2. Calculate a given distribution’s median
	3. Calculate a given distribution’s quartiles
	4. Calculate and interpret given percentiles for a distribution
	5. Calculate a given distribution’s arithmetic mean
	6. Perform calculations involving summation notation
	7. Deduce characteristics of a data set by examining the values of central tendency indices

# TOPIC 4. MEASURES OF DISPERSION

Computation of the range and standard deviation for data is discussed. Interpretation of the measures of dispersion is emphasized. Comparisons between the measures of dispersion are made, together with some applications.

**Topic Goals**: To help students develop an understanding of the fundamentals and usefulness of:

1. The measures of variability: range and standard deviation.

B. Measures of center and dispersion to appropriately describe a data set.

**Student Outcomes:** The student will:

## Topic 4A

* 1. Calculate a given distribution’s range
	2. Calculate a given distribution’s standard deviation.
	3. Perform calculations involving summation notation
	4. Deduce characteristics of a data set by examining the values of variability indices

# Topic 4B

* 1. Determine which measures of central tendency and dispersion best describe the data set
	2. Use location and dispersion data to make decisions for generalizations about the nature of the collected data

**TOPIC 5. CORRELATION**

Scatterplots are constructed and interpreted. The linear correlation coefficient is defined, computed, and interpreted. Correlation and causation are discussed.

**Topic Goal:** To help students develop an understanding of the fundamentals and usefulness of exploring relationships among quantitative variables.

**Student Outcomes:** The student will:

* 1. Examine paired quantitative data occurring in many areas such as science, business, economics, social sciences, and health sciences
	2. Construct a scatterplot for paired quantitative variables
	3. Recognize the response (dependent) and explanatory (independent) variable as suggested by the data
	4. Recognize outliers and identify their effects on correlation
	5. Identify positive and negative associations
	6. Recognize linear and non-linear relationships
	7. Use appropriate technology to calculate the Pearson Correlation, r
	8. Identify the range of values for r
	9. Interpret r in terms of direction and strength
	10. Recognize that correlation is not cause and effect

# TOPIC 6. RELATIONSHIPS BETWEEN VARIABLES

Relationships between paired variables are examined. The equation of the least-squares linear regression line is computed and appropriate predictions are made from it.

**Topic Goals:** To help students develop an understanding of the fundamentals and usefulness of:

1. Making predictions using the least-squares regression line.
2. Exploring relationships between paired variables in which at least one is categorical.

**Student Outcomes:** The student will:

# Topic 6A

* 1. Find the equation of the least-squares regression line
	2. Interpret the slope and the y-intercept
	3. Calculate and interpret r squared
	4. Calculate and interpret residuals
	5. Identify outliers and lurking variables and discuss their effects on the regression equation
	6. Examine the scatter plot to determine the appropriateness of the prediction

# Topic 6B

* 1. Understand the components of a two-way table
	2. Determine probability values using information given in a two-way table

# TOPIC 7. NORMAL DISTRIBUTIONS

Normal distributions are introduced and their properties discussed. Emphasis is placed on z-scores and use of the standard normal table. Applications are studied.

**Topic Goal:** To help students develop an understanding of the fundamentals and usefulness of the normal distributions.

**Student Outcomes:** The student will:

* 1. Understand that areas under a density curve represent proportions of all observations and that the total area under the density curve is one
	2. Sketch the normal distribution with given parameter values
	3. Transform normal variable to standard normal variable for the purpose of comparing distributions and determining probabilities
	4. Determine normalized standard scores for scores of various normal distributions
	5. Algebraically convert values from one normal scale to another
	6. Identify shapes of distributions and recognize distributions that are not normal
	7. Apply the 68-95-99.7% Rule (Empirical Rule)
	8. Use the standard normal distribution tables to compute areas under normal curves
	9. Compute inverse normal density values
	10. Recognize applications of normal distributions to various fields of study

# TOPIC 8. PROBABILITY

The concept of probability, as it is used in statistical methods, is introduced. Elementary probability rules are used to compute probabilities.

**Topic Goal:** To help students develop an understanding of the fundamentals and usefulness of the basic concepts of probability.

**Student Outcomes:** The student will:

* 1. Compute empirical probabilities
	2. Apply elementary probability rules and/or Venn diagrams to compute probabilities
	3. Use probability models, such as the normal distribution, to compute probabilities
	4. Find the values bounding the shaded region of a normal curve

# TOPIC 9. RANDOM SAMPLING

The concept of a simple random sample is introduced and methods of obtaining random samples are discussed.

**Topic Goal:** To help students develop an understanding of the fundamentals and usefulness of random sampling and its implications.

**Student Outcomes:** The student will:

* 1. Define the concept of a simple random sample and apply the definition to obtain a representative sample from a given population
	2. Identify poor sampling designs
	3. Describe distinguishing characteristics of valid experiments
	4. Make and/or evaluate inferences about a population based on characteristics of a sample
	5. Relate specific types of statistical situations to the historical development of statistical methods and techniques

# TOPIC 10. SAMPLING DISTRIBUTIONS

The sampling distribution of the mean and the sampling distribution of the proportion are discussed. Applications of the Central Limit Theorem are studied.

**Topic Goal:** To help students develop an understanding of the fundamentals and usefulness of sampling distributions and the Central Limit Theorem.

**Student Outcomes:** The student will:

* 1. Define a sampling distribution of a sample statistic
	2. Understand the basic concepts of the Central Limit Theorem
	3. Identify applications of sampling distributions
	4. Provide probability statements in terms of the sample statistics
	5. Determine probabilities associated with sample statistics
	6. Understand how sampling distributions play a role in statistical inference

# TOPIC 11. ESTIMATION

Point estimates for population mean and proportion are discussed. Confidence intervals for the mean or proportion of a population are constructed. The relationship between sample size and the error in the estimate is discussed. The basic concepts of hypothesis testing are introduced. **Note**: one of Topic Goal A or B is required; the other is optional, at the discretion of the instructor.

**Topic Goals:** To help students develop an understanding of the fundamentals and usefulness of:

1. Estimating with confidence the values of the population mean.
2. Estimating with confidence the values of the population proportion when the sample size is sufficiently large.
3. Hypothesis testing and the basic terminology for testing the mean of the population or testing the proportion of the population when the sample size is sufficiently large.

**Student Outcomes:** The student will:

## Topic 11A

* 1. Define a point estimate for the population mean
	2. Construct confidence intervals for the population mean
	3. Demonstrate an understanding of how confidence intervals behave under changing conditions for sample size and confidence level (including margin of error)
	4. Determine a minimum sample size given a margin of error

# Topic 11B

* 1. Define a point estimate for the population proportion
	2. Construct confidence intervals for the population proportion for large sample size
	3. Demonstrate an understanding of how confidence intervals behave under changing conditions for sample size and confidence level (including margin of error)
	4. Determine a minimum sample size given a margin of error

# Topic 11C

* 1. Conduct a complete hypothesis test of the population mean or of the population proportion with sufficiently large sample size
	2. Distinguish the potential difference between statistical significance and practical significance
	3. Recognize several factors that can invalidate statistical inference (hypothesis tests and confidence interval estimates)

 TEACHING GUIDE

Title: Elementary Statistics

Catalog Number: MA110

Credit Hours: 3

Lecture Hours: 3 for lecture version; 2 for lab version

Lab Hours: 0 for lecture version; 2 for lab version

Prerequisites: An appropriate high school GPA, or placement test score, or MA089 Arithmetic.

### Catalog

Description: This course introduces probability and statistics. Topics include graphs, tables, frequency distributions, measures of central tendency and dispersion, normal distribution, correlation and regression, probability, and inferential statistics. This course is available in two formats: lecture only, or lecture plus laboratory using technology

**Required Topics: (35 hours)**

Module 1: Overview (3 hours)

 Note: The emphasis should be on the discussion of the simple random sample but the coverage of other sampling techniques is encouraged.

Module 2: Visualizing Data (3 hours)

 Note: The following graphs are optional: Dotplot ; Pareto Chart ; Stem-and-Leaf Plot ; Pie Chart ; Time Series

 Note: The discussion of ways graphs can mislead is at the instructor’s discretion.

Module 3: Center, Spread, & Position (7 hours)

 Note: Approximating the mean and standard deviation of grouped data are optional topics.

 Note: Students are required to be able to determine the standard deviation using appropriate technology. Computation details are at the instructor’s discretion.

 Note: Using the IQR method to identify outliers is an optional topic

Module 4: Bivariate Data (7 hours)

 Note: Students are required to be able to determine the correlation coefficient, the regression line coefficients, and the coefficient of determination using appropriate technology. The computational formula for r and the procedure for the equation coefficients illustrated on are at the instructor’s discretion.

Module 5: Discrete Probability (2 hours)

 Note: The calculation of probabilities, including conditional probabilities, from a two-way table are required. The use of probability formulas are optional.

 Note: Calculating compound probabilities for independent events are optional.

Module 6: The Normal Curve (7 hours)

Note: The Central Limit Theorem for Proportions is an optional topic.

Module 7: Estimation (6 hours)

**Optional Topics: (6 hours)**

Module 1: Overview

 Additional sampling techniques, including cluster & stratified samples

Module 2: Visualizing Data

 Additional graphs, including: Dotplot ; Pareto Chart ; Stem-and-Leaf Plot ; Pie Chart ; Time Series

 The ways that graphs can mislead

Module 3: Center, Spread, & Position

 Approximating the mean & std. deviation of grouped data

 Using the IQR method to identify outliers is an optional topic

 The Coefficient of Variation and Chebychev’s Inequality may be discussed, but as of 4/6/18, they are not included in the Guided Notes of the Toolkit.

Module 5: Discrete Probability

 The use of probability formulas

 Calculating compound probabilities for independent events

 The Binomial & Poisson Distributions may also be discussed, but as of 4/6/18, they are not included in the Guided Notes of the Toolkit.

Module 6: The Normal Curve

The Central Limit Theorem for Proportions

The Normal Approximation to the Binomial Distribution may also be discussed, but as of 4/6/18, it is not included in the Guided Notes of the Toolkit.

Module 7: Estimation

Estimation of the other parameters, in addition to the one selected by the instructor as required, may be discussed. As of 4/6/18, the Toolkit does not contain guided notes on estimation for the proportion.

Hypothesis Testing of other parameters, in addition to the one selected by the instructor as required may be discussed. As of 4/6/18, the Toolkit only contains guided notes for hypothesis tests for the mean with known sigma.

**Calculator Usage Statement**

Note: The following statement was approved by the Mathematics Department on May 5, 1995.

1. All students in MA110 will be required to purchase a calculator capable of two-variable statistics or to take an alternative lab/lecture format section.
2. Instructors may recommend a particular model of calculator but should not expect 100% of the students to actually obtain that particular model.
3. It is left up to the instructor’s discretion how much class time is devoted to instruction in the use of the calculator.
4. Students should be able to calculate the mean and standard deviation of a set of raw data by carrying out all necessary calculations and substitutions into the defining formula. Means and standard deviations of grouped data, linear correlation coefficients, and coefficients of the regression equation should be handled solely using the statistical functions on the calculator, and emphasis given in class to interpretation of these results.
5. Test questions on these topics should stress concepts and interpretations rather than results of calculations.
6. Although it would possible, using a programmable graphics calculator, to “automate” the process of calculating binomial probabilities, confidence intervals, and (on some calculator models) normal probabilities, the added complexity of using the programmable graphics calculator does , and (on some calculator models) normal probabilities, not justify the time saved in computing these quantities by the usual process.
7. The use of calculators in this way does not preclude or substitute for the use of other technological tools in the MA110 classroom, but rather is an adjunct to other tools as may be effective, or even as a potential lead-in to the introduction of such other tools.