Orange Unified School District

# STATISTICS AP

Year Course

**GRADE LEVEL:** 11-12

**PREREQUISITES:** Completion of Pre-Calculus with Trigonometry with a grade of C or better or concurrent enrollment in Pre-Calculus with Trigonometry

or

Completion of Integrated Mathematics III with a grade of B or better and teacher recommendation

## INTRODUCTION TO THE SUBJECT:

Advanced Placement Statistics is a year course introducing students to the major concepts and tools for collecting, analyzing, and drawing conclusions from data. Students are exposed to four broad conceptual themes:

1. Exploring Data: Observing patterns and departures from patterns
2. Planning a Study: Deciding what and how to measure
3. Anticipating Patterns: Producing models using a probability theory and simulation
4. Statistical Inference: Confirming models

Students who successfully complete the course and examination may receive credit and/or advanced placement for one semester introductory college statistics course.

## COURSE OBJECTIVES:

## BY THE END OF THE COURSE THE STUDENT WILL BE ABLE TO:

Appreciate the usefulness of obtaining and analyzing data for making decisions and advancing knowledge.

Understand the importance of data collection and be able to critique the quality of studies based upon issues of data collection.

Apply basic data analytical techniques to uncover patterns and truths with data sets and understand the primary importance of graphing the data.

Apply basic techniques of statistical inference to data, interpret the results of a statistical analysis using the concepts of confidence interval or tests of significance, and assess when particular inferential procedures are appropriate.

Communicate the results of statisical analyses or quantitative findings in writing and speaking.

**COURSE OVERVIEW AND APPROXIMATE UNIT TIME ALLOTMENTS:**

**WEEKS**

I. Exploring Data: Observing Patterns and Departures from Patterns 11

*Exploratory analysis of data makes use of graphical and numerical techniques to study patterns and departures from patterns. Emphasis should be placed on interpreting information from graphical and numerical displays and summaries.*

A. Interpreting graphical displays of distributions of univariate data (dotplot, stemplot, histogram, cumulative frequency plot)
   1. Center and spread
   2. Clusters and gaps
   3. Outliers and other unusual features
   4. Shape

B. Summarizing distributions of univariate data
   1. Measuring center: median, mean
   2. Measuring spread: range, interquartile range, standard deviation
   3. Measuring position: quartiles, percentiles, standardized scores (z-scores)
   4. Using boxplots
   5. The effect of changing units on summary measures

C. Comparing distributions of univariate data (dotplots, back-to-back stemplots, parallel boxplots)
   1. Comparing center and spread: within group, between group variation
   2. Comparing clusters and gap
   3. Comparing outliers and other unusual features
   4. Comparing shapes

D. Exploring bivariate data
   1. Analyzing patterns in scatterplots
   2. Correlation and linearity
   3. Least-squares regression line
   4. Residual plots, outliers, and influential points
   5. Transformations to achieve linearity: logarithmic and power transformations

**WEEKS**

E. Exploring categorical data: frequency tables
   1. Marginal and joint frequencies for two-way tables
   2. Conditional relative frequencies and association

II. Planning a Study: Deciding What and How to Measure 7

*Data must be collected according to a well-developed plan if valid information on a conjecture is to be obtained. This plan includes clarifying the question and deciding upon a method of data collection and analysis.*

A. Overview of methods of data collection
   1. Census
   2. Sample survey
   3. Experiment
   4. Observational study

B. Planning and conducting surveys
   1. Characteristics of a well-designed and well-conducted survey
   2. Populations, samples, and random selection
   3. Sources of bias in surveys
   4. Simple random sampling
   5. Stratified random sampling

C. Planning and conducting experiments
   1. Characteristics of a well-designed and well-conducted experiment
   2. Treatments, control groups, experimental units, random assignments, and replication
   3. Sources of bias and confounding, including placebo effect and blinding
   4. Completely randomized design
   5. Randomized block design, including matched pairs design

D. Generalizability of results from observational studies, experimental studies, and surveys

**WEEKS**

III. Anticipating Patterns: Producing Models Using Probability Theory and Stimulation — 7

*Probability is the tool used for anticipating what the distribuiton of data should look like under a given model.*

A. Probability as relative frequency
   1. "Law of large numbers" concept
   2. Addition rule, multiplication rule, conditional probability, and independence
   3. Discrete random variables and their probability distributions, including binomial
   4. Simulation of probability distributions, including binomial and geometric
   5. Mean (expected value) and standard deviation of a random variable, and linear transformation of a random variable

B. Combining independent random variables
   1. Notion of independence versus dependence
   2. Mean and standard deviation for sums and differences of independent random variable

C. The normal distribution
   1. Properties of the normal distribution
   2. Using tables of the normal distribution
   3. The normal distribution as a model for measurements

D. Sampling distributions
   1. Sampling distribution of a sample proportion
   2. Sampling distribution of a sample mean
   3. Central Limit Theorem
   4. Sampling distribution of a difference between two independent sample proportions
   5. Sampling distribution of a difference between two independent sample means
   6. Simulation of sampling distributions

**WEEKS**

IV. Statistical Inference: Confirming Models 11

*Statistical inference guides the selection of appropriate models.*

A. Confidence intervals
   1. The meaning of a confidence interval
   2. Large sample confidence interval for a proportion
   3. Large sample confidence interval for a mean
   4. Large sample confidence interval for a difference between two proportions
   5. Large sample confidence interval for a difference between two means (unpaired and paired)

B. Tests of significance
   1. Logic of significance testing, null and alternative hypotheses: p-values; one- and two-sided test; concepts of Type I and Type II errors; concept of power
   2. Large sample test for a proportion
   3. Large sample test for a mean
   4. Large sample test for a difference between two proportions
   5. Large sample test for a difference between two means (unpaired and paired)
   6. Chi-square test for goodness of fit, homogeneity of proportions, and independence (one- and two-way tables)

C. Special case of normally distributed data
   1. T-distribution
   2. Single sample t procedures
   3. Two sample (independent and matched pairs) t procedures
   4. Inference for the slope of least-squares regression line

V. Review and Preparation for AP Examination 2

**DATE OF CONTENT REVISION:** NEW

**DATE OF BOARD APPROVAL:** April 15, 1999

# APPENDIX

## Probability and Statistics

This discipline is an introduction to the study of probability, interpretation of data, and fundamental statistical problem solving. Mastery of this academic content will provide students with a solid foundation in probability and facility with processing statistical information.

1. Students know the definition of the notion of independent events, and can use the addition, multiplication, and complementation rules to solve for probabilities of particular events in finite sample spaces.
2. Students know the definition of conditional probability, and use it to solve for probabilities in finite sample spaces.
3. Students demonstrate understanding of the notion of discrete random variable by using them to solve for the probabilities of outcomes, such as the probability of the occurrence of five heads in fourteen coin tosses.
4. Students are familiar with the standard distributions (normal, binomial, and exponential), and can use them to solve for events in problems where the distribution belongs to these families.
5. Students determine the mean and standard deviation of a normally distributed random variable.
6. Students know the definitions of the mean, median, and mode of distribution of real valued data, and can compute them in particular situations.
7. Students compute the variance and standard deviation of a distribution of data.
8. Students organize and describe distributions of data using a number of different methods, including frequency tables, histograms, standard line and bar graphs, stem and leaf displays, scatter plots, and box and whisker plots.
9. Students find the line of best fit to a given distribution of data using least squares regression.

## Probability and Statistics - Advanced

This discipline is a technical and in depth extension of probability and statistics. In particular, mastery of advanced placement academic content gives students the background of success on the Advanced Placement exam in the subject.

1. Students solve probability problems with finite sample spaces using the addition, multiplication, and complementation rules for probability distributions, and understand the simplifications which arise with independent events.
2. Students know the definition of conditional probability, and use it to solve for probabilities in finite sample spaces.
3. Students demonstrate understanding of the notion of discrete random variables by using them to solve for the probabilities of outcomes, such as the probability of the

occurrence of five or fewer heads in fourteen coin tosses.

4. Students understand the notion of a continuous random variable, and can interpret the probability of an outcome as the area of a region under the graph of the probability density function associated with the random variable.
5. Students know the definition of the mean of a discrete random variable, and can determine it for a particular discrete random variable.
6. Students know the definition of the variance of a discrete random variable, and can determine it for a particular discrete random variable.
7. Students demonstrate understanding of the standard distributions (normal, binomial, and exponential), and can use them to solve for events in problems where the distribution belongs to these families.
8. Students determine the mean and standard deviation of a normally distributed random variable.
9. Students know the Central Limit Theorem, and can use it to obtain approximations for probabilities in finite sample spaces problems whose probabilities are distributed binomially.
10. Students know the definitions of the mean, median, and mode of distribution of real valued data, and can compute them in particular situations.
11. Students compute the variance and standard deviation of a distribution of data.
12. Students find the line of best fit to a given distribution of data using least squares regression.
13. Students know the definition of the correlation coefficient of two variables, and are familiar with its properties.
14. Students organize and describe distributions of data using a number of different methods, including frequency tables, histograms, standard line and bar graphs, stem and leaf displays, scatter plots, and box and whisker plots.
15. Students are familiar with the notions of a statistic of a distribution of values, of the sampling distribution of a statistic, and of the variability of a statistic.
16. Students know basic facts concerning the relation between the mean and standard deviation of a sampling distribution of the mean and standard deviation of the population distribution.
17. Students determine confidence intervals of a simple random sample from a normal distribution of data, and determine the sample size required for a desired margin of error.
18. Students determine the P-value for a statistic for a simple random sample from a normal distribution.
19. Students are familiar with the chi-square distribution and test, and understand its uses.