Statistics – Advanced Placement G/T
Essential Curriculum

UNIT I: Exploring Data

Goal. The student will demonstrate the ability to use a problem-solving approach in employing graphical and numerical techniques to study patterns and departures from patterns. The student will interpret and summarize univariate data in graphical and numerical displays.

Objectives – The student will be able to:

a. Distinguish between categorical and quantitative variables.
b. Construct bar charts, histograms, dotplots, stemplots, and time plots.
c. Describe and interpret the overall pattern of each distribution by giving the center, spread, outliers, and shape.
d. Define the mean, variance, and the standard deviation of a distribution.
e. Compute the mean, variance, and the standard deviation of a distribution.
f. Define the median, quartiles, the five-number summary, and boxplots.
g. Compute the median and the five-number summary.
h. Construct a boxplot.
i. Compare the mean and median of normal and skewed distributions.
j. Define and compute the degrees of freedom.

UNIT II: The Normal Distribution

Goal. The student will demonstrate the ability to use a problem-solving approach in interpreting and applying properties of the standard normal distribution and *z*-scores.

Objectives – The student will be able to:

a. Describe the properties of area under a density curve.
b. Approximately locate the median and the mean of symmetric and skewed density curves.
c. State the properties of shape and symmetry of the normal distribution.
d. Calculate the *z* value of the standard normal distribution given m and s.
e. Given the proportion, calculate the *z* value.
f. Construct and interpret normal probability plots to assess normality.

UNIT III: Examining Relationships

Goal. The student will demonstrate the ability to use a problem-solving approach in exploring bivariate data. The student will use technology to compute and analyze the least squares regression line for a set of data.

Objectives – The student will be able to:

a. Construct a scatterplot to display the relation between two quantitative variables, including adding a categorical variable.
b. Interpret the direction, form, and strength of the association between variables.
c. Define and interpret the correlation coefficient *r*.
d. State the properties of the correlation coefficient in a linear relationship.
e. Define and use appropriate notation for a regression line and the least-squares regression line.
f. Calculate the least-squares regression line, using a calculator, and interpret the slope and intercept.
g. Find the slope and the intercept of the least-squares regression line from the means and standard deviations of *x* and *y* and their correlation.
h. Appropriately use the regression line to predict *y* for a given *x*.
i. Use r^2 (the coefficient of determination) to describe how much of the variation in one variable can be accounted for by a straight-line relationship with another variable.
j. Calculate and plot residuals to identify patterns of fit.

UNIT IV: Two-Variable Data

Goal. The student will demonstrate the ability to use a problem-solving approach in exploring non-linear models and categorical data.

Objectives – The student will be able to:
a. Recognize exponential growth or decay and power functions.
b. Perform a logarithmic transformation to produce a non-linear model.
c. Calculate and plot residuals to examine the fit for non-linear models.
d. State the limitations of correlation and regression by examining the effects of outliers or lurking variables.
e. Understand that a strong correlation does not imply a cause-and-effect relationship.
f. From a two-way table of counts, compute the marginal distribution in percents.
g. Describe the relationship between two categorical variables.
h. Use Simpson's Paradox to explain associations that are misleading because of lurking variables.

UNIT V: Producing Data

Goal. The student will demonstrate the ability to use a problem-solving approach to collect data according to a well-developed plan so that valid information on a conjecture can be obtained. The student will decide upon a method of data collection and analysis, and carry out experiments and observational studies.

Objectives – The student will be able to:
a. Identify the population in a sampling situation.
b. Use a table of random digits to select a simple random sample from a population and assign subjects to groups.
c. Recognize bias and sources of error in sample surveys.
d. Choose a stratified random sample from a population.

e. Recognize whether a study is an observational study or an experiment.
f. Identify the factors, treatments, response variables, and experimental units in an experiment.
g. Diagram the design of a completely randomized experiment.
h. Recognize the placebo effect and when to use the double-blind technique.
i. Explain why a randomized comparative experiment can give good evidence for cause-and-effect relationships.
j. Construct and run a simulation using a TI-83 calculator or on a computer.

UNIT VI: Probability

Goal. The student will demonstrate the ability to use a problem-solving approach to apply probability and simulation to anticipate what the distribution of data should look like under a given model. The student will recognize that the probability needed for statistical inference is primarily oriented toward describing data distributions.

Objectives – The student will be able to:

a. Recognize random phenomena.
b. Define and compute the empirical probability of a random phenomenon.
c. Define the sample space for a random phenomenon.
d. List the sample space for a random phenomenon.
e. Use the multiplication rule to determine the number of elements in a sample space.
f. Determine the number of outcomes of an experiment both with and without replacement.
g. Define an event.
h. State the basic rules of probability (for example, the complement rule).
i. Compute the theoretical probability of an event.
j. State the definition of independent events and determine whether given events are independent.
k. Compute probabilities of disjoint or overlapping events by applying basic rules.
l. Compute joint probability.
m. Define and compute conditional probability.

UNIT VII: Random Variables

Goal. The student will demonstrate the ability to use a problem-solving approach to interpret and analyze probability distributions and apply their properties to continuous random variables.

Objectives – The student will be able to:

a. Identify and distinguish between discrete and continuous random variables.
b. Use a given probability distribution.
c. Derive a probability distribution for simple discrete random variables.
d. Interpret a probability histogram.

e. Interpret the probability of a continuous random variable as the area under a density curve.
f. Recognize the normal distribution as a probability distribution.
g. Compute the mean (expected value) of a discrete random variable from its probability distribution.
h. State the law of a large numbers.
i. Compute the mean of a linear transformation for a discrete random variable (i.e. $m_{a+bX} = a + b\, m_x$).
j. Compute the mean of a sum or difference of two discrete random variables (i.e. $m_{X+Y} = m_X + m_Y$).
k. Compute the variance and standard deviation of a discrete random variable.
l. Compute the variance and standard deviation of a linear transformation for a discrete random variable.
m. Compute the variance and standard deviation of a sum or difference for two discrete random variables.

UNIT VIII: The Binomial and Geometric Distributions

Goal. The student will demonstrate the ability to use a problem-solving approach to interpret and analyze binomial and geometric distributions and apply their properties.

Objectives - The student will be able to:

a. Identify a random variable as binomial.
b. Compute binomial probabilities using tables and using appropriate technology.
c. Construct binomial probability distribution tables and histograms.
d. Calculate cumulative distribution functions for binomial random variables and construct cumulative distribution tables and histograms.
e. Calculate means and standard deviations for binomial random variables.
f. Identify a random variable as geometric.
g. Compute geometric probabilities using tables and using appropriate technology.
h. Construct geometric probability distribution tables and histograms.
i. Calculate cumulative distribution functions for geometric random variables and construct cumulative distribution tables and histograms.
j. Calculate means for geometric random variables.

UNIT IX: Sampling Distributions

Goal. The student will demonstrate the ability to use a problem-solving approach in simulating sampling distributions.

Objectives - The student will be able to:

a. Identify parameters and statistics in a sample or experiment.
b. Recognize that the value of a statistic varies in repeated random sampling.
c. Define and interpret a sampling distribution.

d. Describe the bias and variability of a statistic in terms of the mean and spread of its sampling distribution.
e. Identify the effect of sample size on the variability of a statistic.
f. Compute the mean and standard deviation of a sampling distribution of proportions ($\hat{p}$) with population proportion *p*.
g. Recognize that the standard deviation of the sampling distribution of $\hat{p}$ gets smaller at the rate $\sqrt{n}$ as the sample size n get larger.
h. Use the normal approximation, if appropriate, to calculate probabilities that concern $\hat{p}$.
i. Compute the mean and standard deviation of a sampling distribution of means from the corresponding population parameters.
j. Recognize that the standard deviation of the sampling distribution of $\bar{x}$ gets smaller at the rate $\sqrt{n}$ as the sample size gets larger.
k. Apply the Central Limit Theorem to calculate probabilities that concern $\bar{x}$.

UNIT X: Introduction to Inference

Goal. The student will demonstrate the ability to use a problem-solving approach to report and interpret confidence intervals, test hypotheses, compute and interpret appropriate test statistics and errors. The student will write statements using appropriate language that explain how confident one can be about conclusions reached.

Objectives – The student will be able to:

a. Define and interpret the meaning of "95% confidence" and other statements of confidence.
b. Calculate a confidence interval for the mean m.
c. Determine when a sample is large enough and normal enough to use a confidence level safely.
d. Understand the factors that affect the margin or error of a confidence interval.
e. Determine the sample size necessary for a specified error.
f. State the null and alternative hypotheses in a testing situation.
g. Explain the meaning of the *P*-value when given the numerical value of *P*.
h. Calculate the *z* statistic and the *P*-value for one-sided and two-sided tests about the mean m of a normal population.
i. Use decision analysis to determine Type I and Type II errors.
j. Define power of a significance test and determine what factors increase and decrease power.
k. Assess the statistical significance at standard levels.
l. Recognize that significance testing does not measure the size or importance of an effect.

UNIT XI: Inference for Distributions

Goal. The student will demonstrate the ability to use a problem-solving approach to use statistical inference to select appropriate models for testing hypotheses about a

population mean. The student will write statements using appropriate language that explain how confident one can be about conclusions reached.

Objectives – The student will be able to:

a. Recognize when a problem requires inference about a mean or comparing two means.
b. Recognize from the design of a study whether one-sample, matched pairs or two-sample procedure are needed.
c. Use *t* procedures to obtain a confidence interval at a stated level of confidence of the population mean for one-sample and two-sample sets of data.
d. Using the *P*-value, test the hypothesis that a population mean has a specified value against either a one-sided or a two-sided alternative.
e. Using the *P*-value, test the hypothesis that two populations have equal means against either a one-sided or a two-sided alternative.
f. Recognize when one-sample and two-sample procedures are appropriate in practice.
g. Recognize the influence of the design of a study, outliers, or a small sample from a skewed distribution on *t* procedures.
h. Recognize matched pairs data and use *t* procedures to obtain confidence intervals and to perform tests of significance.
i. Compute the confidence interval for the difference between the means of two populations.
j. Test the hypothesis that two populations have equal means against a one-sided or two-sided alternative.
k. Compute and interpret appropriate degrees of freedom.

UNIT XII: Inference for Proportions

Goal. The student will demonstrate the ability to use a problem-solving approach to use statistical inference to select appropriate models for testing hypotheses about a population proportion. The student will write statements using appropriate language that explain how confident one can be about conclusions reached.

Objectives – The student will be able to:

a. Recognize when a problem requires inference about a proportion or comparing two proportions.
b. Calculate from sample counts the sample proportion or proportions that estimate the parameters of interest.
c. Compute the confidence interval for a population proportion *p*.
d. Carry out a test of significance for the hypotheses that proportions are equal against either a one-sided or a two-sided alternative.
e. State the conditions of sample size necessary to use these *z* procedures.
f. Use two-sample *z* procedures to give a confidence interval for the difference between two proportions based on independent samples from the populations.
g. Test the hypothesis that proportions in two distinct populations are equal.
h. Determine when a pooled sample proportion is appropriate.
i. State the conditions when these *z* procedures can be used.

UNIT XIII: Chi-Square Procedures

Goal. The student will demonstrate the ability to use a problem-solving approach in applying various tests of significance, and in stating and checking types of regression.

Objectives - The student will be able to:

a. Calculate the chi-square statistic and the *P*-value.
b. State the null and alternative hypothesis for chi-square goodness of fit.
c. Draw appropriate conclusions from the chi-square statistic using deviations between observed and expected counts.
d. Arrange data in a two-way table (matrix)
e. Calculate the expected cell counts to determine the chi-square statistic.
f. Compute the chi-square statistic and its *P*-value using appropriate technology.
g. State the null and alternative hypotheses for a chi-square test of independence.
h. Determine what deviations from the null hypothesis are most important if the chi-square test is significant.
i. Calculate the component of the chi-square statistic for any cell.
j. Identify the degrees of freedom of a chi-square statistic.
k. Use chi-square critical values to approximate the *P*-value from a table.

UNIT XIV. Inference for Regression

Goal. The student will demonstrate the ability to use a problem-solving approach in computing and analyzing properties of regression slope.

Objectives – The student will be able to:

a. Interpret the coefficients ? and ? in the linear regression model ($\hat{y} = a + bx$) and the standard deviation ? of the observed y values.
b. Interpret the coefficients a and b in the TRUE regression model ($m_y = a + bx$) and compute the standard error about this line.
c. Determine a confidence interval for regression slope.
d. Determine confidence and prediction intervals for regression slope.
e. State and check regression assumptions.