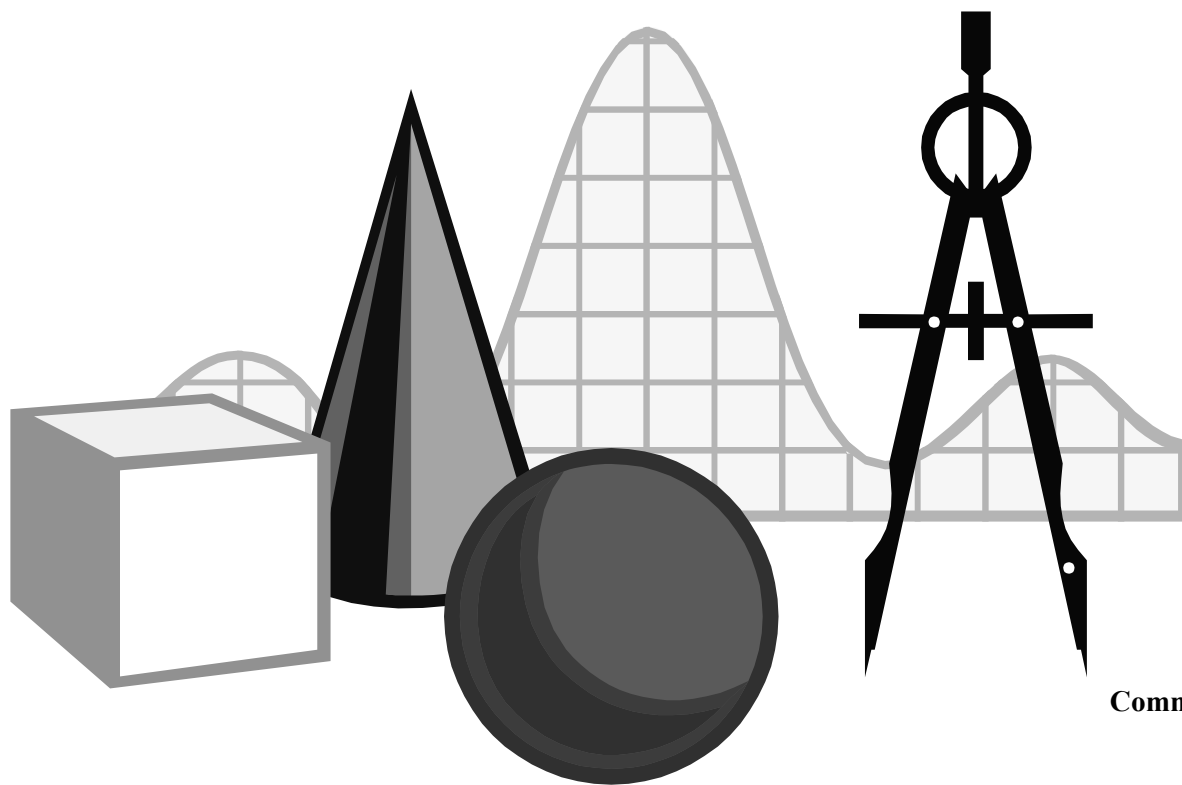


MATHEMATICS STANDARDS OF LEARNING CURRICULUM FRAMEWORK

Probability and Statistics



Commonwealth of Virginia
Board of Education
Richmond, Virginia
© 2002

Copyright © 2002

by the

Virginia Department of Education

P.O. Box 2120

Richmond, Virginia 23218-2120

<http://www.pen.k12.va.us>

All rights reserved. Reproduction of materials contained herein for instructional purposes in Virginia classrooms is permitted.

Superintendent of Public Instruction

Jo Lynne DeMary

Deputy Superintendent

M. Kenneth Magill

Assistant Superintendent for Instruction

Patricia I. Wright

Office of Secondary Instructional Services

Linda M. Wallinger, Director

Deborah Kiger Lyman, Mathematics Specialist

NOTICE TO THE READER

The Virginia Department of Education does not unlawfully discriminate on the basis of sex, race, color, religion, handicapping conditions, or national origin in employment or in its educational programs and activities.

The 2002 *Mathematics Curriculum Framework* can be found in PDF and Microsoft Word file formats on the Virginia Department of Education's website at <http://www.pen.k12.va.us>.

Introduction

Mathematics content develops sequentially in concert with a set of processes that are common to different bodies of mathematics knowledge. The content of the Mathematics Standards of Learning supports five process goals for students: becoming mathematical problem solvers, communicating mathematically, reasoning mathematically, making mathematical connections, and using mathematical representations to model and interpret practical situations. These goals provide a context within which to develop the knowledge and skills identified in the standards.

Probability and Statistics presents basic concepts and techniques for collecting and analyzing data, drawing conclusions, and making predictions. Applications may be drawn from a wide variety of disciplines ranging from the social sciences of psychology and sociology to education, allied health fields, business, economics, engineering, the humanities, the physical sciences, journalism, communications, and liberal arts. Students should be able to design an experiment, collect appropriate data, select and use statistical techniques to analyze the data, and develop and evaluate inferences based on the data.

Each topic in the Probability and Statistics Curriculum Framework is developed around the Standards of Learning. Each Standard of Learning is expanded in the Essential Knowledge and Skills column. The Essential Understandings column includes concepts, mathematical relationships, and ideas that are important to understanding and teaching the Standard of Learning effectively.

Teachers should help students make connections and build relationships among algebra, arithmetic, geometry, discrete mathematics, and probability and statistics. Connections should be made to other subject areas and fields of endeavor through applications. Using manipulatives, graphing calculators, and computer applications to develop concepts should help students develop and attach meaning to abstract ideas. Throughout the study of mathematics, students should be encouraged to talk about mathematics, use the language and symbols of mathematics, communicate, discuss problems and problem solving, and develop their competence and their confidence in themselves as mathematics students.

**PROBABILITY AND STATISTICS
STANDARD PS.1**

The student will analyze graphical displays of data, including dotplots, stemplots, and histograms, to identify and describe patterns and departures from patterns, using central tendency, spread, clusters, gaps, and outliers. Appropriate technology will be used to create graphical displays.

ESSENTIAL UNDERSTANDINGS

- Data are collected for a purpose and have meaning in a context.
- Measures of central tendency describe how the data cluster or group.
- Measures of dispersion describe how the data spread (disperse) around the center of the data.
- Graphical displays of data may be analyzed informally.
- Data analysis must take place within the context of the problem.

ESSENTIAL KNOWLEDGE AND SKILLS

- Create and interpret graphical displays of data, including dotplots, stem-and-leaf plots, and histograms.
- Examine graphs of data for clusters and gaps, and relate those phenomena to the data in context.
- Examine graphs of data for outliers, and explain the outlier(s) within the context of the data.
- Examine graphs of data, and identify the central tendency of the data as well as the spread. Explain the central tendency and the spread of the data within the context of the data.

**PROBABILITY AND STATISTICS
STANDARD PS.2**

The student will analyze numerical characteristics of univariate data sets to describe patterns and departure from patterns, using mean, median, mode, variance, standard deviation, interquartile range, range, and outliers. Appropriate technology will be used to calculate statistics.

ESSENTIAL UNDERSTANDINGS

- Data are collected for a purpose and have meaning within a context.
- Analysis of the descriptive statistical information generated by a univariate data set should include the interplay between central tendency and dispersion as well as among specific measures.
- Data points identified algorithmically as outliers should not be excluded from the data unless sufficient evidence exists to show them to be in error.

ESSENTIAL KNOWLEDGE AND SKILLS

- Interpret mean, median, mode, range, interquartile range, variance, and standard deviation of a univariate data set in terms of the problem's context.
- Identify possible outliers, using an algorithm.
- Explain the influence of outliers on a univariate data set.
- Explain ways in which standard deviation addresses dispersion by examining the formula for standard deviation.

**PROBABILITY AND STATISTICS
STANDARD PS.3**

The student will compare distributions of two or more univariate data sets, analyzing center and spread (within group and between group variations), clusters and gaps, shapes, outliers, or other unusual features. Appropriate technology will be used to generate graphical displays.

ESSENTIAL UNDERSTANDINGS

- Data are collected for a purpose and have meaning in a context.
- Statistical tendency refers to typical cases but not necessarily to individual cases.

ESSENTIAL KNOWLEDGE AND SKILLS

- Compare and contrast two or more univariate data sets by analyzing measures of center and spread within a contextual framework.
- Describe any unusual features of the data, such as clusters, gaps, or outliers, within the context of the data.
- Analyze in context kurtosis and skewness in conjunction with other descriptive measures.

**PROBABILITY AND STATISTICS
STANDARD PS.4**

The student will analyze scatterplots to identify and describe the relationship between two variables, using shape; strength of relationship; clusters; positive, negative, or no association; outliers; and influential points. Appropriate technology will be used to generate scatterplots and identify outliers and influential points.

ESSENTIAL UNDERSTANDINGS

- A scatterplot serves two purposes:
 - to determine if there is a useful relationship between two variables, and
 - to determine the family of equations that describes the relationship.
- Data are collected for a purpose and have meaning in a context.
- Association between two variables considers both the direction and strength of the association.
- The strength of an association between two variables reflects how accurately the value of one variable can be predicted based on the value of the other variable.
- Outliers are observations with large residuals and do not follow the pattern apparent in the other data points.

ESSENTIAL KNOWLEDGE AND SKILLS

- Examine scatterplots of data, and describe skewness, kurtosis, and correlation within the context of the data.
- Describe and explain any unusual features of the data, such as clusters, gaps, or outliers, within the context of the data.
- Identify influential data points (observations that have great effect on a line of best fit because of extreme x -values) and describe the effect of the influential points.

**PROBABILITY AND STATISTICS
STANDARD PS.5**

The student will find and interpret linear correlation, use the method of least squares regression to model the linear relationship between two variables, and use the residual plots to assess linearity. Appropriate technology will be used to compute correlation coefficients and residual plots.

ESSENTIAL UNDERSTANDINGS

- Data are collected for a purpose and have meaning in a context.
- Least squares regression generates the equation of the line that minimizes the sum of the squared distances from the data points to the line.
- Each data point may be considered to be comprised of two parts: fit (the part explained by the model) and residual (the result of chance variation or of variables not measured).
- Residual = Actual – Fitted
- A correlation coefficient measures the degree of association between two variables that are related linearly.
- Two variables may be strongly associated without a cause-and-effect relationship existing between them.

ESSENTIAL KNOWLEDGE AND SKILLS

- Calculate a correlation coefficient.
- Explain how the correlation coefficient, r , measures association by looking at its formula.
- Use regression lines to make predictions, and identify the limitations of the predictions.
- Use residual plots to determine if a linear model is satisfactory for describing the relationship between two variables.
- Describe the errors inherent in extrapolation beyond the range of the data.
- Use least squares regression to find the equation of the line of best fit for a set of data.
- Explain how least squares regression generates the equation of the line of best fit by examining the formulas used in computation.

**PROBABILITY AND STATISTICS
STANDARD PS.6**

The student will make logarithmic and power transformations to achieve linearity. Appropriate technology will be used.

ESSENTIAL UNDERSTANDINGS

- A logarithmic transformation reduces positive skewness because it compresses the upper tail of the distribution while stretching the lower tail.
- Nonlinear transformations do not preserve relative spacing between data points.

ESSENTIAL KNOWLEDGE AND SKILLS

- Apply a logarithmic transformation to data.
- Explain how a logarithmic transformation works to achieve a linear relationship between variables.
- Apply a power transformation to data.
- Explain how a power transformation works to achieve a linear relationship between variables.

**PROBABILITY AND STATISTICS
STANDARD PS.7**

The student, using two-way tables, will analyze categorical data to describe patterns and departure from patterns and to find marginal frequency and relative frequencies, including conditional frequencies.

ESSENTIAL UNDERSTANDINGS

- Data are collected for a purpose and have meaning in a context.
- Simpson’s paradox refers to the fact that aggregate proportions can reverse the direction of the relationship seen in the individual parts.
- Two categorical variables are independent if the conditional frequencies of one variable are the same for every category of the other variable.

ESSENTIAL KNOWLEDGE AND SKILLS

- Produce a two-way table as a summary of the information obtained from two categorical variables.
- Calculate marginal, relative, and conditional frequencies in a two-way table.
- Use marginal, relative, and conditional frequencies to analyze data in two-way tables within the context of the data.

**PROBABILITY AND STATISTICS
STANDARD PS.8**

The student will describe the methods of data collection in a census, sample survey, experiment, and observational study and identify an appropriate method of solution for a given problem setting.

ESSENTIAL UNDERSTANDINGS

- The value of a sample statistic varies from sample to sample if the simple random samples are taken repeatedly from the population of interest.
- Poor data collection can lead to misleading and meaningless conclusions.

ESSENTIAL KNOWLEDGE AND SKILLS

- Compare and contrast controlled experiments and observational studies and the conclusions one can draw from each.
- Compare and contrast population and sample, and parameter and statistic.
- Identify biased sampling methods.
- Describe simple random sampling.
- Select a data collection method appropriate for a given context.

**PROBABILITY AND STATISTICS
STANDARD PS.9**

The student will plan and conduct a survey. The plan will address sampling techniques (e.g., simple random and stratified) and methods to reduce bias.

ESSENTIAL UNDERSTANDINGS

- The purpose of sampling is to provide sufficient information so that population characteristics may be inferred.
- Inherent bias diminishes as sample size increases.

ESSENTIAL KNOWLEDGE AND SKILLS

- Investigate and describe sampling techniques, such as simple random sampling, stratified sampling, and cluster sampling.
- Determine which sampling technique is best, given a particular context.
- Plan a survey to answer a question or address an issue.
- Given a plan for a survey, identify possible sources of bias, and describe ways to reduce bias.
- Design a survey instrument.
- Conduct a survey.

**PROBABILITY AND STATISTICS
STANDARD PS.10**

The student will plan and conduct an experiment. The plan will address control, randomization, and measurement of experimental error.

ESSENTIAL UNDERSTANDINGS

- Experiments must be carefully designed in order to detect a cause-and-effect relationship between variables.
- Principles of experimental design include comparison with a control group, randomization, and blindness.

ESSENTIAL KNOWLEDGE AND SKILLS

- Plan and conduct an experiment. The experimental design should address control, randomization, and minimization of experimental error.

**PROBABILITY AND STATISTICS
STANDARD PS.11**

The student will compute and distinguish between permutations and combinations and use technology for applications.

ESSENTIAL UNDERSTANDINGS

- The fundamental counting principle states that if one decision can be made n ways and another can be made m ways, then the two decisions can be made nm ways.
- The number of subsets of a set with n elements is 2^n .

ESSENTIAL KNOWLEDGE AND SKILLS

- Compare and contrast permutations and combinations.
- Calculate the number of permutations of n objects taken r at a time.
- Calculate the number of combinations of n objects taken r at a time.

**PROBABILITY AND STATISTICS
STANDARD PS.12**

The student will identify and describe two or more events as complementary, dependent, independent, and/or mutually exclusive.

ESSENTIAL UNDERSTANDINGS

- The complement of event A consists of all outcomes in which event A does not occur.
- Two events, A and B, are independent if the occurrence of one does not affect the probability of the occurrence of the other. If A and B are not independent, then they are said to be dependent.
- Events A and B are mutually exclusive if they cannot occur simultaneously.

ESSENTIAL KNOWLEDGE AND SKILLS

- Define and give contextual examples of complementary, dependent, independent, and mutually exclusive events.
- Given two or more events in a problem setting, determine if the events are complementary, dependent, independent, and/or mutually exclusive.

**PROBABILITY AND STATISTICS
STANDARD PS.13**

The student will find probabilities (relative frequency and theoretical), including conditional probabilities for events that are either dependent or independent, by applying the “law of large numbers” concept, the addition rule, and the multiplication rule.

ESSENTIAL UNDERSTANDINGS

- Data are collected for a purpose and have meaning in a context.
- Venn diagrams may be used to find conditional probabilities.
- The “law of large numbers” states that as a procedure is repeated again and again, the relative frequency probability of an event tends to approach the actual probability.

ESSENTIAL KNOWLEDGE AND SKILLS

- Calculate relative frequency and expected frequency.
- Find conditional probabilities for dependent, independent, and mutually exclusive events.

**PROBABILITY AND STATISTICS
STANDARD PS.14**

The student will develop, interpret, and apply the binomial probability distribution for discrete random variables, including computing the mean and standard deviation for the binomial variable.

ESSENTIAL UNDERSTANDINGS

- A probability distribution is a complete listing of all possible outcomes of an experiment together with their probabilities. The procedure has a fixed number of independent trials.
- A random variable assumes different values depending on the event outcome.
- A probability distribution combines descriptive statistical techniques and probabilities to form a theoretical model of behavior.

ESSENTIAL KNOWLEDGE AND SKILLS

- Develop the binomial probability distribution within a real-world context.
- Calculate the mean and standard deviation for the binomial variable.
- Use the binomial distribution to calculate probabilities associated with experiments for which there are only two possible outcomes.

**PROBABILITY AND STATISTICS
STANDARD PS.15**

The student will simulate probability distributions, including binomial and geometric.

ESSENTIAL UNDERSTANDINGS

- A probability distribution combines descriptive methods and probabilities to form a theoretical model of behavior.
- A probability distribution gives the probability for each value of the random variable.

ESSENTIAL KNOWLEDGE AND SKILLS

- Design and conduct an experiment that simulates a binomial distribution.
- Design and conduct an experiment that simulates a geometric distribution.

**PROBABILITY AND STATISTICS
STANDARD PS.16**

The student will identify random variables as independent or dependent and find the mean and standard deviations for sums and differences of independent random variables.

ESSENTIAL UNDERSTANDINGS

ESSENTIAL KNOWLEDGE AND SKILLS

- A random variable is a variable that has a single numerical value, determined by chance, for each outcome of a procedure.

- Compare and contrast independent and dependent random variables.
- Find the standard deviation for sums and differences of independent random variables.

**PROBABILITY AND STATISTICS
STANDARD PS.17**

The student will identify properties of a normal distribution and apply the normal distribution to determine probabilities, using a table or graphing calculator.

ESSENTIAL UNDERSTANDINGS

- The normal distribution curve is a family of symmetrical curves defined by the mean and the standard deviation.
- Areas under the curve represent probabilities associated with continuous distributions.
- The normal curve is a probability distribution and the total area under the curve is 1.

ESSENTIAL KNOWLEDGE AND SKILLS

- Identify the properties of a normal probability distribution.
- Describe how the standard deviation and the mean affect the graph of the normal distribution.
- Determine the probability of a given event, using the normal distribution.

**PROBABILITY AND STATISTICS
STANDARD PS.18**

The student, given data from a large sample, will find and interpret point estimates and confidence intervals for parameters. The parameters will include proportion and mean, difference between two proportions, and difference between two means (independent and paired).

ESSENTIAL UNDERSTANDINGS

- A primary goal of sampling is to estimate the value of a parameter based on a statistic.
- Confidence intervals use the sample statistic to construct an interval of values that one can be reasonably certain contains the true (unknown) parameter.
- Confidence intervals and tests of significance are complementary procedures.
- Paired comparisons experimental design allows control for possible effects of extraneous variables.

ESSENTIAL KNOWLEDGE AND SKILLS

- Construct confidence intervals to estimate a population parameter, such as a proportion or the difference between two proportions; or a mean or the difference between two means.
- Select a value for alpha (Type I error) for a confidence interval.
- Interpret confidence intervals in the context of the data.
- Explain the importance of random sampling for confidence intervals.
- Calculate point estimates for parameters, and discuss the limitations of point estimates.

**PROBABILITY AND STATISTICS
STANDARD PS.19**

The student will apply and interpret the logic of a hypothesis-testing procedure. Tests will include large sample test for proportion, mean, difference between two proportions, and difference between two means (independent and paired) and Chi-squared test for goodness of fit, homogeneity of proportions, and independence.

ESSENTIAL UNDERSTANDINGS

- Confidence intervals and tests of significance are complementary procedures.
- Paired comparisons experimental design allows control for possible effects of extraneous variables.
- Tests of significance assess the extent to which sample data support a hypothesis about a population parameter.
- The purpose of a goodness of fit test is to decide if the sample results are consistent with results that would have been obtained if a random sample had been selected from a population with a known distribution.
- Practical significance and statistical significance are not necessarily congruent.

ESSENTIAL KNOWLEDGE AND SKILLS

- Use the Chi-squared test for goodness of fit to decide if the population being analyzed fits a particular distribution pattern.
- Use hypothesis-testing procedures to determine whether or not to reject the null hypothesis. The null hypothesis may address proportion, mean, difference between two proportions or two means, goodness of fit, homogeneity of proportions, and independence.
- Compare and contrast Type I and Type II errors.
- Explain how and why the hypothesis-testing procedure allows one to reach a statistical decision.

**PROBABILITY AND STATISTICS
STANDARD PS.20**

The student will identify the meaning of sampling distribution with reference to random variable, sampling statistic, and parameter and explain the Central Limit Theorem. This will include sampling distribution of a sample proportion, a sample mean, a difference between two sample proportions, and a difference between two sample means.

ESSENTIAL UNDERSTANDINGS	ESSENTIAL KNOWLEDGE AND SKILLS
<ul style="list-style-type: none"> • The Central Limit Theorem states: <ul style="list-style-type: none"> – The mean of the sampling distribution of means is equal to the population mean. – If the sample size is sufficiently large, the sampling distribution approximates the normal probability distribution. – If the population is normally distributed, the sampling distribution is normal regardless of sample size. • Sampling distributions have less variability with larger sample sizes. 	<ul style="list-style-type: none"> • Describe the use of the Central Limit Theorem for drawing inferences about a population parameter based on a sample statistic. • Describe the effect of sample size on the sampling distribution and on related probabilities. • Use the normal approximation to calculate probabilities of sample statistics falling within a given interval. • Identify and describe the characteristics of a sampling distribution of a sample proportion, mean, difference between two sample proportions, or difference between two sample means.

**PROBABILITY AND STATISTICS
STANDARD PS.21**

The student will identify properties of a t-distribution and apply t-distributions to single-sample and two-sample (independent and matched pairs) t-procedures, using tables or graphing calculators.

ESSENTIAL UNDERSTANDINGS

- Paired comparisons experimental design allows control for possible effects of extraneous variables.
- The sampling distribution of means with a small sample size follows a t-distribution.

ESSENTIAL KNOWLEDGE AND SKILLS

- Identify the properties of a t-distribution.
- Compare and contrast a t-distribution and a normal distribution.
- Use a t-test for single-sample and two-sample data.