Probability and Statistics *Standards of Learning* - 2023 Overview of Revision

This overview includes a summary of the content embedded in four content strands.

Data in Context

* Use a statistical cycle to formulate questions, describe types of data and data sources, and constraints within the context of a problem**†**
* Compare and contrast data collection methods to plan and conduct an observational study**†**
* Utilize the principles of experimental design to plan and conduct a well-designed experiment**†**

Descriptive Statistics

* Represent and analyze data visualizations of univariate quantitative data, including dotplots, stemplots, boxplots, cumulative frequency graphs, and histograms, to identify and describe patterns and departures from patterns, using central tendency, spread, clusters, gaps, and outliers, within the context of a problem**†**
* Represent and analyze numerical characteristics of univariate quantitative data sets to describe patterns and departures from patterns within the context of a problem**†**
* Represent, compare, and analyze distributions of two or more univariate quantitative data sets, numerically and graphically**†**
* Represent and analyze categorical data, using two-way tables and other graphical displays, to describe patterns and relationships
* Represent and analyze quantitative bivariate data with scatterplots to identify and describe the relationship between two variables
* Create and interpret a linear model using the least squares regression method to assess the relationship between two quantitative variables

Probability

* Organize information and apply probability rules to compute probabilities of events within the context of a problem**†**
* Represent and interpret situations using discrete random distributions, including binomial distributions
* Represent and interpret situations using normal distributions**†**

Inferential Statistics

* Apply properties of sampling distributions and inference procedures to make decisions about population proportions
* Apply properties of sampling distributions and inference procedures to make decisions about populations

# † Content intended for a one-semester course only.

Comparison of Probability and Statistics *Standards of Learning* – 2016 to 2023

| 2016 *Standards of Learning*  Essential Knowledge and Skills (EKS) | 2023 *Standards of Learning*  Knowledge and Skills (KS)  Data in Context (DC) |
| --- | --- |
| [New Expectation] | PS.DC.1† The student will use a statistical cycle to formulate questions, describe types of data, data sources, and constraints within the context of a problem.   1. Define the stages of the statistical cycle and how each stage relates to the others. 2. Formulate questions and conclusions based on context. 3. Understand the type of data relevant to the question at hand (e.g., quantitative versus categorical). 4. Compare and contrast population and sample, and parameter and statistic. 5. Identify and explain constraints of the statistical approach. |
| Moved from PS.8† and PS.9† | PS.DC.2† The student will compare and contrast data collection methods to plan and conduct an observational study.   1. Investigate and describe sampling techniques (e.g., simple random sampling, stratified sampling, systematic sampling, cluster sampling). 2. Determine which sampling technique is best, given a particular context. 3. Investigate and explain biased influences inherent within sampling methods and various forms of response bias. 4. Use the statistical cycle to plan and conduct an observational study to answer a question or address a problem. |
| Moved from PS.10† | PS.DC.3† The student will utilize the principles of experimental design to plan and conduct a well-designed experiment.   1. Describe the principles of experimental design, including:    1. treatment/control groups;    2. blinding/placebo effects;    3. experimental units/subjects; and    4. blocking/matched pairs and completely randomized designs. 2. Evaluate the principles of experimental design to address comparison, randomization, replication, and control within the context of the problem. 3. Compare and contrast controlled experiments and observational studies and the conclusions that may be drawn from each. 4. Use the statistical cycle to plan and conduct a well-designed experiment to answer a question or address a problem. 5. Select a data collection method appropriate for a given context. |

| 2016 *Standards of Learning*  Essential Knowledge and Skills (EKS)  Descriptive Statistics | 2023 *Standards of Learning*  Knowledge and Skills (KS)  Descriptive Statistics (DS) |
| --- | --- |
| PS.1† The student will analyze graphical displays of univariate data, including dotplots, stemplots, boxplots, cumulative frequency graphs, and histograms, to identify and describe patterns and departures from patterns, using central tendency, spread, clusters, gaps, and outliers.  Create and interpret graphical displays of data, including dotplots, stemplots, boxplots, cumulative frequency graphs, and histograms, using appropriate technology.  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. | PS.DS.1† The student will represent and analyze data visualizations of univariate quantitative data, including dot plots, stemplots, boxplots, cumulative frequency graphs, and histograms, to identify and describe patterns and departures from patterns, using central tendency, spread, clusters, gaps, and outliers, within the context of a problem.   1. Create and interpret graphical displays of data, including dot plots, stemplots, boxplots, cumulative frequency graphs, and histograms, using appropriate technology. 2. Examine the graphs within the context of the problem by analyzing:    1. shape;    2. measures of center;    3. spread; and    4. unusual features of the data (e.g., outliers, clusters, gaps). |
| PS.2† The student will analyze numerical characteristics of univariate data sets to describe patterns and departures from patterns, using mean, median, mode, variance, standard deviation, interquartile range, range, and outliers.  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. | PS.DS.2† The student will represent and analyze numerical characteristics of univariate quantitative data sets to describe patterns and departures from patterns within the context of a problem.   1. Interpret measures of central tendency: mean, median, and mode. 2. Interpret measures of spread: range, interquartile range, variance, and standard deviation. 3. Identify possible outliers, using an algorithm. 4. Investigate and explain the influence of outliers on a univariate data set. 5. Investigate and explain ways in which standard deviation addresses variability by examining the formula for standard deviation. |
| PS.3† The student will compare distributions of two or more univariate data sets, numerically and graphically, analyzing center and spread (within group and between group variations), clusters and gaps, shapes, outliers, or other unusual features.  Compare and contrast two or more univariate data sets, numerically and graphically, 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 skewness in conjunction with measures of center and spread in a contextual framework. | PS.DS.3† The student will represent, compare, and analyze distributions of two or more univariate quantitative data sets, numerically and graphically.   1. Create graphical displays of data, including back-to-back stemplots, parallel dot plots, parallel boxplots, and histograms, using appropriate technology. 2. Compare and contrast two or more univariate data sets, numerically and graphically, within the context of a problem by analyzing:    1. shape;    2. measures of center;    3. measures of spread; and    4. unusual features of the data (e.g., clusters, gaps, outliers). |
| Moved from PS.7† | PS.DS.4 The student will represent and analyze categorical data, using two-way tables and other graphical displays, to describe patterns and relationships.   1. Create and interpret graphical displays of univariate categorical data, including bar graphs within the context of the problem, using appropriate technology. 2. Create and interpret graphical displays comparing distributions of two or more univariate categorical data sets including segmented and side-by-side bar graphs within the context of the problem, using appropriate technology. 3. Generate and interpret a two-way table as a summary of the information obtained from two categorical variables. 4. Calculate and interpret marginal, relative, and conditional frequencies to analyze data in a two-way table within the context of a problem. |
| 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.  Examine scatterplots of data, and describe skewness, 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 a great effect on a line of best fit because of extreme *x*-values) and describe the effect of the influential points. | PS.DS.5 The student will represent and analyze quantitative bivariate data with scatterplots to identify and describe the relationship between two variables.   1. Create scatterplots, using appropriate technology. 2. Examine and interpret scatterplots in the context of the problem by analyzing:    1. the form of relationship for linear and nonlinear trends;    2. the direction of the relationship for positive, negative, or no association;    3. the strength of the relationship such as strong, moderate, or weak; and    4. the presence of unusual features within the data (e.g., clusters, gaps, influential points, and/or outliers). |
| PS.5 The student will determine and interpret linear correlation, use the method of least squares regression to model the linear relationship between two variables, and use the residual plot to assess linearity.  Calculate a correlation coefficient, *r*.  Explain how the correlation coefficient, *r*, measures association by looking at its formula.  Interpret the coefficient of determination, *r*2, in a contextual framework.  Use regression lines to make predictions, and identify the limitations of the predictions.  Use residual plots to determine whether 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 determine the equation of the line of best fit for a set of data.  Interpret the slope and *y*-intercept of the least squares regression line in a contextual framework.  Explain how least squares regression generates the equation of the line of best fit by examining the formulas used in computation. | PS.DS.6 The student will create and interpret a linear model using the least squares regression method to assess the relationship between two quantitative variables.   1. Create the least squares regression model using technology to interpret the contextual meaning of the slope and *y*-intercept. 2. Using technology, calculate and interpret the correlation coefficient, *r*, within the context of a problem. 3. Using technology, calculate and interpret the coefficient of determination, *r*2, within the context of a problem. 4. Use regression lines to make predictions, and identify the limitations of the predictions, such as extrapolation. 5. Calculate and interpret a residual to understand the error of a prediction. 6. Using technology, calculate and interpret the standard deviation of the residuals, *s*. |
| PS.6 The student will make logarithmic and power transformations to achieve linearity.  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. | 1. [Included in AP Statistics] |
| PS.7† The student, using two-way tables and other graphical displays, will analyze categorical data to describe patterns and departures from patterns and to determine marginal frequency and relative frequencies, including conditional frequencies.  Produce a two-way table as a summary of the information obtained from two categorical variables.  Create and interpret graphical displays of categorical data including bar charts.  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. | 1. [Moved to PS.DS.4] |

| 2016 *Standards of Learning*  Essential Knowledge and Skills  Data Collection | 2023 *Standards of Learning*  Knowledge and Skills |
| --- | --- |
| 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.  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. | 1. [Moved to PS.DC.2†] |
| PS.9† The student will plan and conduct a survey. The plan will address sampling techniques and methods to reduce bias.  Distinguish between a population and a sample.  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. | 1. [Moved to PS.DC.2†] |
| PS.10† The student will plan and conduct a well-designed experiment. The plan will address control, randomization, replication, blinding, and measurement of experimental error.  Plan and conduct a well-designed experiment. The experimental design should address control, randomization, replication, blinding and minimization of experimental error.  Identify treatments, levels, factors, control groups, and experimental units in an experimental design.  Identify sources of bias and confounding, including the placebo effect.  Identify a situation when a block design, including matched pairs, would reduce the effects of confounding variables. | 1. [Moved to PS.DC.3†] |

| 2016 *Standards of Learning*  Essential Knowledge and Skills (EKS)  Probability | 2023 *Standards of Learning*  Knowledge and Skills (KS)  Probability (P) |
| --- | --- |
| PS.11† The student will identify and describe two or more events as complementary, dependent, independent, and/or mutually exclusive.  Define and give contextual examples of complementary, dependent, independent, and mutually exclusive events.  Given two or more events in a problem setting, determine whether the events are complementary, dependent, independent, and/or mutually exclusive. | PS.P.1† The student will organize information and apply probability rules to compute probabilities of events within the context of a problem.   1. Given two or more events, determine whether the events are complementary, dependent, independent, and/or mutually exclusive, and compute the probability of those events. 2. Represent and calculate probabilities using Venn diagrams, tree diagrams, and two-way tables. 3. Apply the addition rule, the multiplication rule, and complementary rule to calculate probabilities. 4. Calculate conditional probabilities to determine the association or independence of two events. |
| PS.12† The student will determine 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.  Calculate relative frequency and expected frequency.  Determine conditional probabilities for dependent, independent, and mutually exclusive events. | 1. [Moved to PS.P.1†] |
| PS.13 The student will develop, interpret, and apply the binomial and geometric probability distributions for discrete random variables, including computing the mean and standard deviation for the binomial and geometric variables.  Develop the binomial and geometric probability distributions within a practical context.  Calculate the mean and standard deviation for the binomial and geometric variables.  Use the binomial and geometric distributions to calculate probabilities associated with experiments for which there are only two possible outcomes. | PS.P.2 The student will represent and interpret situations using discrete random distributions, including binomial distributions.   1. Identify discrete random variables and create a table to represent valid discrete probability distributions within the context of a problem. 2. Calculate and interpret the mean (expected value) and standard deviation for a discrete random variable within the context of a problem. 3. Determine if a discrete random variable satisfies the conditions for a binomial distribution. 4. Design and conduct a simulation of a binomial distribution. 5. Calculate and interpret probabilities from a binomial distribution within the context of a problem. 6. Calculate the mean and standard deviation for binomial distributions. 7. Describe the center, shape, and spread of a discrete random variable within the context of a problem. |
| PS.14 The student will simulate probability distributions, including binomial and geometric.  Design and conduct a simulation of a binomial distribution.  Design and conduct a simulation of a geometric distribution.  Calculate probabilities resulting from simulations of binomial and geometric distributions. | 1. [Moved to PS.P.2] |
| PS.15 The student will identify random variables as independent or dependent and determine the mean and standard deviations for random variables and sums and differences of independent random variables.  Compare and contrast independent and dependent random variables.  Determine the mean (expected value) and standard deviation for a random variable and linear transformation of a random variable.  Determine the mean (expected value) for sums and differences of random variables.  Determine the standard deviation for sums and differences of independent random variables. | 1. [Moved to PS.P.2] |
| PS.16† The student will identify properties of a normal distribution and apply the normal distribution to determine probabilities.  Identify the properties of a normal distribution.  Describe how the standard deviation and the mean affect the graph of the normal distribution.  Calculate and interpret the *z*-score of a given data value from a normal distribution.  Determine the probability of a given event, using the normal distribution.  Use a graphing utility and a table of Standard Normal Probabilities to determine probabilities. | PS.P.3† The student will represent and interpret situations using normal distributions.   1. Compare and contrast discrete and continuous distributions. 2. Represent probability as the area under the curve of a normal distribution using the Empirical Rule and graphing technology. 3. Describe the center, shape, and spread of normal distributions within the context of a problem. 4. Compare and contrast two or more sets of normally distributed data using *z*-scores, percentiles, or probabilities within the context of a problem. 5. Standardize a data value from a normal distribution and interpret the *z*-score within the context of a problem. 6. Calculate and interpret probabilities of a normal distribution using technology within the context of a problem. |

| 2016 *Standards of Learning*  Essential Knowledge and Skills (EKS)  Inferential Statistics | 2023 *Standards of Learning*  Knowledge and Skills (KS)  Inferential Statistics (IS) |
| --- | --- |
| PS.17 The student, given data from a large sample, will determine and interpret appropriate point estimates and confidence intervals for parameters. The parameters will include proportion and mean, difference between two proportions, difference between two means (independent and paired), and slope of a least-squares regression line.  Construct confidence intervals to estimate a population parameter, such as a proportion or the difference between two proportions; a mean or the difference between two means; or slope of a least-squares regression line.  Select a value for the confidence level of a confidence interval.  Interpret confidence intervals and confidence levels in the context of the data.  Explain the importance of random sampling for confidence intervals.  Explain how changes in confidence level and sample size effect width of the confidence interval and margin of error.  Calculate point estimates for parameters and discuss the limitations of point estimates. | PS.IS.1 The student will apply properties of sampling distributions and inference procedures to make decisions about population proportions.   1. Describe the shape, center, and spread of the sampling distribution of a proportion within the context of a problem. 2. Given a problem, construct a one sample *z* confidence interval:    1. identify the basic conditions for inference: random sample, independence, and normality;    2. calculate a confidence interval using technology; and    3. interpret the interval within the context of the problem. 3. Explain how changes in confidence level and sample size affect width of the confidence interval and margin of error. 4. Calculate and interpret a point estimate and margin of error of a confidence interval for a proportion within the context of the problem. 5. Explain how and why the hypothesis testing procedure allows one to reach a statistical decision. 6. Given a problem, apply the one sample *z* hypothesis testing procedures:    1. construct appropriate null and alternate hypotheses;    2. identify the basic conditions for inference: random sample; independence, and normality;    3. calculate and interpret the *p*-value using technology;    4. determine and justify whether to reject the null hypothesis; and    5. interpret the results within the context of the problem. 7. Use the statistical cycle to plan and conduct a statistical study about a proportion to answer a question or address a problem with inference. |
| PS.18 The student will apply and interpret the logic of an appropriate hypothesis-testing procedure. Tests will include large sample test for proportion, mean, difference between two proportions, difference between two means (independent and paired); chi-squared tests for goodness of fit, homogeneity of proportions, and independence; and slope of a least-squares regression line.  Use the chi-squared test for goodness of fit to decide whether 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, independence, and the slope of a least-squares regression line.  Compare and contrast Type I and Type II errors.  Explain how and why the hypothesis-testing procedure allows one to reach a statistical decision | PS.IS.2 The student will apply properties of sampling distributions and inference procedures to make decisions about populations.   1. Describe the shape, center, and spread of the sampling distribution of a mean within the context of a problem. 2. Calculate and interpret a point estimate and a margin of error for a confidence interval of a mean within the context of a problem. 3. Describe the use of the Central Limit Theorem in satisfying the assumptions and conditions for inference about a mean. 4. Identify the properties of a *t* distribution. 5. Given a problem, construct a one sample *t* confidence interval:    1. identify the basic conditions for inference: random sample, independence, and approximate normality;    2. calculate a confidence interval using technology; and    3. interpret the interval within the context of the problem. 6. Given a problem, apply the one sample t hypothesis testing procedures:    1. construct appropriate null and alternate hypotheses;    2. identify the basic conditions for inference: random sample, independence, and approximate normality;    3. calculate and interpret the *p* value using technology;    4. determine and justify whether to reject the null hypothesis; and    5. interpret the results within the context of the problem. |
| PS.19 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.  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. | 1. [Moved to PS.IS.1; PS.IS.2; or included in AP Statistics] |
| PS.20 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.  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. | 1. [Moved to PS.IS.2 or included in AP Statistics] |

2023 Probability and Statistics Mathematics SOL – Summary of Changes

|  |  |
| --- | --- |
| Probability and Statistics  (2016 SOL to 2023 SOL Numbering) | Parameter Changes/Clarifications (2023 SOL) |
| [New Expectation] PS.DC.1†  PS.1† PS.DS.1†  PS.2† PS.DS.2†  PS.3† PS.DS.3  PS.4† PS.DS.5  PS.5 PS.DS.6  PS.6 [Deleted]  PS.7† PS.DS.4  PS.8† PS.DC.2†  PS.9† PS.DC.2†  PS.10† PS.DC.3†  PS.11† PS.P.1†  PS.12† PS.P.1†  PS.13 PS.P.2  PS.14 PS.P.2  PS.15 PS.P.2  PS.16† PS.P.3  PS.17 PS.IS.1  PS.18 PS.IS.2  PS.19 PS.IS.1, PS.IS.2  PS.20 PS.IS.2 | PS.IS.2 [KS] - Hypothesis testing limited to one-sample tests  PS.DS.6 [KS] - Generating the equation of the line of best fit by using technology versus examining the formulas  PS.P.2 [KS] - Limited discrete random variables to binomial distributions |

|  |  |
| --- | --- |
| Deletions from Probability and Statistics (2016 SOL) | Additions to Probability and Statistics (2023 SOL) |
| PS.5 [EKS] – Use residual plots to determine whether a linear model is satisfactory for describing the relationship between two variables [Included in AP Statistics]  PS.6 – Logarithmic and power transformation to achieve linearity [Included in AP Statistics]  PS.10 [EKS] – Factors and levels of experimental design [Included in AP Statistics]  PS.13 [EKS] – Geometric distributions [Included in AP Statistics]  PS.15 [EKS] – Compare and contrast independent and dependent random variables; determine the mean (expected value) for sums and differences of random variables; determine the standard deviation for sums and differences of independent random variables [Included in AP Statistics]  PS.17 [EKS] – Given sample data, determine and interpret point estimates and confidence intervals for the parameters for the difference between two proportions, difference between two means (independent and paired), and slope of a least-squares regression line [Included in AP Statistics]  PS.18 [EKS] – Hypothesis testing for the difference between two proportions or two means, goodness of fit, homogeneity of proportions, independence, and the slope of a least-squares regression line; compare and contrast Type I and Type II errors [Included in AP Statistics]  PS.19 [EKS] – Identify and describe the characteristics of a sampling distribution of a difference between two sample proportions and a difference between two sample means [Included in AP Statistics]  PS.20 [EKS] – Use a *t*-test for two-sample data [Included in AP Statistics] | PS.DC.1†; PS.DC.2†; PS.DC.3†; and PS.IS.1 [KS] - Describe and use the statistical cycle to answer questions, solve problems, and communicate within the context of the problem  PS.DC.2 [KS] – Investigate and describe the systematic sampling technique  PS.DS.6 [KS] – Use technology to calculate and interpret the standard deviation of the residual |

**KEY:**  DC = Data in Context; DS = Descriptive Statistics; P = Probability; IS = Inferential Statistics; EKS =Essential Knowledge and Skills (2016); KS = Knowledge and Skills (2023); US = Understanding the Standard