This is the March 2015 version of the High School Mathematics Model Curriculum for the conceptual category Statistics and Probability. (Note: The conceptual categories Number and Quantity, Algebra, Functions, Modeling, Geometry, and Statistics and Probability do not directly align with high school mathematics courses.) The current focus of this document is to provide instructional strategies and resources, and identify misconceptions and connections related to the clusters and standards. The Ohio Department of Education is working in collaboration with assessment consortia, national professional organizations and other multistate initiatives to develop common content elaborations and learning expectations.

Statistics and Probability	
Domain	Cluster
Interpreting Categorical and Quantitative Data	Summarize, represent, and interpret data on a single count or measurement variable. Summarize, represent, and interpret data on two categorical and quantitative variables. Interpret linear models.
Making Inferences and Justifying Conclusions	Understand and evaluate random processes underlying statistical experiments. Make inferences and justify conclusions from sample surveys, experiments, and observational studies.
Conditional Probability and the Rules of Probability	Understand independence and conditional probability and use them to interpret data. Use the rules of probability to compute probabilities of compound events in a uniform probability model.
Using Probability to Make Decisions	(+) Calculate expected values and use them to solve problems. (+) Use probability to evaluate outcomes of decisions.

Domain	Interpreting Categorical and Quantitative Data
Cluster	*Summarize, represent, and interpret data on a single count or measurement variable*
Standards	1. Represent data with plots on the real number line (dot plots, histograms, and box plots). 2. Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets. 3. Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers). 4. Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.

Content Elaborations

Ohio has chosen to support shared interpretation of the standards by linking the work of multistate partnerships as the Mathematics Content Elaborations. Further clarification of the standards can be found through these reliable organizations and their links:

- Achieve the Core Modules, Resources
- Hunt Institute Video examples
- Institute for Mathematics and Education Learning Progressions Narratives
- Illustrative Mathematics Sample tasks
- National Council of Supervisors of Mathematics (NCSM) Resources, Lessons, Items
- National Council of Teacher of Mathematics (NCTM) Resources, Lessons, Items
- Partnership for Assessment of Readiness for College and Careers (PARCC) Resources, Items

Expectations for Learning

Ohio has selected PARCC as the contractor for the development of the Next Generation Assessments for Mathematics. PARCC is responsible for the development of the framework, blueprints, items, rubrics, and scoring for the assessments. Further information can be found at Partnership for Assessment of Readiness for College and Careers (PARCC). Specific information is located at these links:

- Model Content Framework
- Item Specifications/Evidence Tables
- Sample Items
- Calculator Usage
- Accommodations
- Reference Sheets

Instructional Strategies and Resources

Instructional Strategies

It is helpful for students to understand that a statistical process is a problem-solving process consisting of four steps: formulating a question that can be answered by data; designing and implementing a plan that collects appropriate data; analyzing the data by graphical and/or numerical methods; and interpreting the analysis in the context of the original question. Opportunities should be provided for students to work through the statistical process. In Grades 6-8, learning has focused on parts of this process. Now is a good time to investigate a problem of interest to the students and follow it through. The richer the question formulated, the more interesting is the process. Teachers and students should make extensive use of resources to perfect this very important first step. Global web resources can inspire projects.

Although this domain addresses both categorical and quantitative data, there is no reference in the Standards 1 - 4 to categorical data. Note that Standard 5 in the next cluster (Summarize, represent, and interpret data on two categorical and quantitative variables) addresses analysis for two categorical variables on the same subject. To prepare for interpreting two categorical variables in Standard 5, this would be a good place to discuss graphs for one categorical variable (bar graph, pie graph) and measure of center (mode).

Have students practice their understanding of the different types of graphs for categorical and numerical variables by constructing statistical posters. Note that a bar graph for categorical data may have frequency on the vertical (student's pizza preferences) or measurement on the vertical (radish root growth over time - days).

Measures of center and spread for data sets without outliers are the mean and standard deviation, whereas median and interquartile ranges are better measures for data sets with outliers.

Introduce the formula of standard deviation by reviewing the previously learned MAD (mean absolute deviation). The MAD is very intuitive and gives a solid foundation for developing the more complicated standard deviation measure.

Informally observing the extent to which two boxplots or two dotplots overlap begins the discussion of drawing inferential conclusions. Don't shortcut this observation in comparing two data sets.

As histograms for various data sets are drawn, common shapes appear. To characterize the shapes, curves are sketched through the midpoints of the tops of the histogram's rectangles. Of particular importance is a symmetric unimodal curve that has specific areas within one, two, and three standard deviations of its mean. It is called the Normal distribution and students need to be able to find areas (probabilities) for various events using tables or a graphing calculator.

Instructional Resources/Tools

TI-84 and TI emulator
Quantitative Literacy Exploring Data module
NCTM Navigating through Data Analysis 9-12.

Printed media (e.g., almanacs, newspapers, professional reports)
Software such as TinkerPlots and Excel
Show World: This website offers data about the world that is up to date.
iEARN:This website offers projects that students around the world are working on simultaneously.

Common Misconceptions

Students may believe:

That a bar graph and a histogram are the same. A bar graph is appropriate when the horizontal axis has categories and the vertical axis is labeled by either frequency (e.g., book titles on the horizontal and number of students who like the respective books on the vertical) or measurement of some numerical variable (e.g., days of the week on the horizontal and median length of root growth of radish seeds on the vertical). A histogram has units of measurement of a numerical variable on the horizontal (e.g., ages with intervals of equal length).

That the lengths of the intervals of a boxplot (min,Q1), (Q1,Q2), (Q2,Q3), (Q3,max) are related to the number of subjects in each interval. Students should understand that each interval theoretically contains one-fourth of the total number of subjects. Sketching an accompanying histogram and constructing a live boxplot may help in alleviating this misconception.

That all bell-shaped curves are normal distributions. For a bell-shaped curve to be Normal, there needs to be 68% of the distribution within one standard deviation of the mean, 95% within two, and 99.7% within three standard deviations.

Diverse Learners

Strategies for meeting the needs of all learners including gifted students, English Language Learners (ELL) and students with disabilities can be found at this site. Additional strategies and resources based on the Universal Design for Learning principles can be found at www.cast.org.

Specific strategies for mathematics may include:

Help students to clearly distinguish between categorical and numerical variables by providing multiple examples of each type.

Students need to be formulating meaningful questions in the first step of the four-step process. This takes time and lots of practice, and leads to real-world contexts.

Some students may need to begin with "well-behaved" data sets. As they progress in their understanding and work with data, begin to include data sets with outliers and non-Normal shapes.

Connections:

The four-step statistical process was introduced in Grade 6, with the recognition of statistical questions. At the high school level, students need to become proficient in the first step of generating meaningful questions.

High School Conceptual Category: Statistics and Probability

Domain	**Interpreting Categorical and Quantitative Data**
Cluster	***Summarize, represent, and interpret data on two categorical and quantitative variables***
Standards	5. Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data. 6. Represent data on two quantitative variables on a scatter plot, and describe how the variables are related. a. Fit a function to the data; use functions fitted to data to solve problems in the context of the data. *Use given functions or choose a function suggested by the context. Emphasize linear and exponential models.* b. Informally assess the fit of a function by plotting and analyzing residents. c. Fit a linear function for a scatter plot that suggests a linear association.

Content Elaborations

Ohio has chosen to support shared interpretation of the standards by linking the work of multistate partnerships as the Mathematics Content Elaborations. Further clarification of the standards can be found through these reliable organizations and their links:

- Achieve the Core Modules, Resources
- Hunt Institute Video examples
- Institute for Mathematics and Education Learning Progressions Narratives
- Illustrative Mathematics Sample tasks
- National Council of Supervisors of Mathematics (NCSM) Resources, Lessons, Items
- National Council of Teacher of Mathematics (NCTM) Resources, Lessons, Items
- Partnership for Assessment of Readiness for College and Careers (PARCC) Resources, Items

Expectations for Learning

Ohio has selected PARCC as the contractor for the development of the Next Generation Assessments for Mathematics. PARCC is responsible for the development of the framework, blueprints, items, rubrics, and scoring for the assessments. Further information can be found at Partnership for Assessment of Readiness for College and Careers (PARCC). Specific information is located at these links:

- Model Content Framework
- Item Specifications/Evidence Tables
- Sample Items
- Calculator Usage
- Accommodations
- Reference Sheets

Instructional Strategies and Resources

Instructional Strategies

In this cluster, the focus is that two categorical or two quantitative variables are being measured on the **same** subject.

In the categorical case, begin with two categories for each variable and represent them in a two-way table with the two values of one variable defining the rows and the two values of the other variable defining the columns. (Extending the number of rows and columns is easily done once students become comfortable with the 2x2 case.) The table entries are the joint frequencies of how many subjects displayed the respective cross-classified values. Row totals and column totals constitute the marginal frequencies. Dividing joint or marginal frequencies by the total number of subjects define relative frequencies (and percentages), respectively. Conditional relative frequencies are determined by focusing on a specific row or column of the table. They are particularly useful in determining any associations between the two variables.

In the numerical or quantitative case, display the paired data in a scatterplot. Note that although the two variables in general will not have the same scale, e.g., total SAT versus grade-point average, it is best to begin with variables with the same scale such as SAT Verbal and SAT Math. Fitting functions to such data will avoid difficulties such as interpretation of slope in the linear case in which scales differ. Once students are comfortable with the same scale case, introducing different scales situations will be less problematic.

Instructional Resources/Tools

TI-83/84 and TI emulator
Quantitative Literacy Exploring Data module
NCTM Navigating through Data Analysis 9-12.
Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report
Software such as TinkerPlots and Excel

Common Misconceptions

Students may believe:

That a 45 degree line in the scatterplot of two numerical variables always indicates a slope of 1 which is the case only when the two variables have the same scaling.
Those residual plots in the quantitative case should show a pattern of some sort. Just the opposite is the case.

Diverse Learners

Strategies for meeting the needs of all learners including gifted students, English Language Learners (ELL) and students with disabilities can be found at this site. Additional strategies and resources based on the Universal Design for Learning principles can be found at www.cast.org.

Specific strategies for mathematics may include:

Help students clearly distinguish between categorical and numerical variables by providing multiple examples of each type.

Provide opportunities for students to formulate meaningful questions in the first step of the four-step process. This takes time and lots of practice leading to real-world contexts.

Connections:

Domain	Interpreting Categorical and Quantitative Data
Cluster	***Interpret Linear Models***
Standards	7. Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data. 8. Compute (using technology) and interpret the correlation coefficient of a linear fit. 9. Distinguish between correlation and causation.

Content Elaborations

Ohio has chosen to support shared interpretation of the standards by linking the work of multistate partnerships as the Mathematics Content Elaborations. Further clarification of the standards can be found through these reliable organizations and their links:

- Achieve the Core Modules, Resources
- Hunt Institute Video examples
- Institute for Mathematics and Education Learning Progressions Narratives
- Illustrative Mathematics Sample tasks
- National Council of Supervisors of Mathematics (NCSM) Resources, Lessons, Items
- National Council of Teacher of Mathematics (NCTM) Resources, Lessons, Items
- Partnership for Assessment of Readiness for College and Careers (PARCC) Resources, Items

Expectations for Learning

Ohio has selected PARCC as the contractor for the development of the Next Generation Assessments for Mathematics. PARCC is responsible for the development of the framework, blueprints, items, rubrics, and scoring for the assessments. Further information can be found at Partnership for Assessment of Readiness for College and Careers (PARCC). Specific information is located at these links:

- Model Content Framework
- Item Specifications/Evidence Tables
- Sample Items
- Calculator Usage
- Accommodations
- Reference Sheets

Instructional Strategies and Resources

Instructional Strategies

In this cluster, the key is that two quantitative variables are being measured on the same subject. The paired data should be listed and then displayed in a scatterplot. If time is one of the variables, it usually goes on the horizontal axis. That which is being predicted goes on the vertical; the predictor variable is on the horizontal axis.

Note that unlike a two-dimensional graph in mathematics, the scales of a scatterplot need not be the same, and even if they are similar (such as SAT Math and SAT Verbal), they still need not have the same spacing. So, visual rendering of slope makes no sense in most scatterplots, i.e., a 45 degree line on a scatterplot need not mean a slope of 1.

Often the interpretation of the intercept (constant term) is not meaningful in the context of the data. For example, this is the case when the zero point on the horizontal is of considerable distance from the values of the horizontal variable, or in some case has no meaning such as for SAT variables.

To make some sense of Pearson's r, correlation coefficient, students should recall their middle school experience with the Quadrant Count Ratio (QCR) as a measure of relationship between two quantitative variables.

Noting that a correlated relationship between two quantitative variables is not causal (unless the variables are in an experiment) is a very important topic and a substantial amount of time should be spent on it.

Instructional Resources/Tools

TI-83/84 and TI emulator
Quantitative Literacy Exploring Data module

NCTM Navigating through Data Analysis 9-12.
Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report
Software such as TinkerPlots and Excel
The Ohio Resource Center
The National Council of Teachers of Mathematics, Illuminations

Common Misconceptions

Students may believe that a 45 degree line in the scatterplot of two numerical variables always indicates a slope of 1 which is the case only when the two variables have the same scaling. Because the scaling for many real-world situation varies greatly students need to be give opportunity to compare graphs of differing scale. Asking students questions like; What would this graph look like with a different scale or using this scale? Is essential in addressing this misconception.

That when two quantitative variables are related, i.e., correlated, that one causes the other to occur. Causation is not necessarily the case. For example, at a theme park, the daily temperature and number of bottles of water sold are demonstrably correlated, but an increase in the number of bottles of water sold does not cause the day's temperature to rise or fall.

Diverse Learners

Strategies for meeting the needs of all learners including gifted students, English Language Learners (ELL) and students with disabilities can be found at this site. Additional strategies and resources based on the Universal Design for Learning principles can be found at www.cast.org.

Specific strategies for mathematics may include:

Some students may have difficulty in clearly distinguishing between correlation and causation. Provide multiple opportunities for students to see examples of each.

Provide opportunities for students to formulate meaningful questions in the first step of the four-step process. This process takes time and lots of practice and guidance. Additionally, students need practice creating these questions as it relates to real-world contexts.

Connections:

Developing a measure of relationship between two quantitative variables should have been introduced in the middle school with the Quadrant Count Ratio (QCR) as discussed in the GAISE Report. It has many shortcomings that students should identify and then try to correct as they see in the development of Pearson's r, correlation coefficient.

Domain	**Making Inferences and Justifying Conclusions**
Cluster	***Understand and evaluate random processes underlying statistical experiments***
Standards	1. Understand statistics as a process for making inferences about population parameters based on a random sample from that population. 2. Decide if a specified model is consistent with results from a given data-generating process, e.g., using simulation. *For example, a model says a spinning coin falls heads up with probability 0.5. would a result of 5 tails in a row cause you to question the model?*

Content Elaborations

Ohio has chosen to support shared interpretation of the standards by linking the work of multistate partnerships as the Mathematics Content Elaborations. Further clarification of the standards can be found through these reliable organizations and their links:

- Achieve the Core Modules, Resources
- Hunt Institute Video examples
- Institute for Mathematics and Education Learning Progressions Narratives
- Illustrative Mathematics Sample tasks
- National Council of Supervisors of Mathematics (NCSM) Resources, Lessons, Items
- National Council of Teacher of Mathematics (NCTM) Resources, Lessons, Items
- Partnership for Assessment of Readiness for College and Careers (PARCC) Resources, Items

Expectations for Learning

Ohio has selected PARCC as the contractor for the development of the Next Generation Assessments for Mathematics. PARCC is responsible for the development of the framework, blueprints, items, rubrics, and scoring for the assessments. Further information can be found at Partnership for Assessment of Readiness for College and Careers (PARCC). Specific information is located at these links:

- Model Content Framework
- Item Specifications/Evidence Tables
- Sample Items
- Calculator Usage
- Accommodations
- Reference Sheets

Instructional Strategies and Resources

Instructional Strategies

Inferential statistics based on Normal probability models is a topic for Advanced Placement Statistics (e.g., t-tests). The idea here is that all students understand that statistical decisions are made about populations (parameters in particular) based on a random sample taken from the population and the observed value of a sample statistic (note that both words start with the letter “s”). A population parameter (note that both words start with the letter “p”) is a measure of some characteristic in the population such as the population proportion of American voters who are in favor of some issue, or the population mean time it takes an Alka Seltzer tablet to dissolve.

As the statistical process is being mastered by students, it is instructive for them to investigate questions such as “If a coin spun five times produces five tails in a row, could one conclude that the coin is biased toward tails?” One way a student might answer this is by building a model of 100 trials by experimentation or simulation of the number of times a truly fair coin produces five tails in a row in five spins. If a truly fair coin produces five tails in five tosses 15 times out of 100 trials, then there is no reason to doubt the fairness of the coin. If, however, getting five tails in five spins occurred only once in 100 trials, then one could conclude that the coin is biased toward tails (if the coin in question actually landed five tails in five spins).

A powerful tool for developing statistical models is the use of simulations. This allows the students to visualize the model and apply their understanding of the statistical process.

Provide opportunities for students to clearly distinguish between a population parameter which is a constant, and a sample statistic which is a variable.

Career Connection

Students will explore the concepts of direct marketing, a marketing database, and a sales promotion as described in the High School Operations Research Modules. Use the provided case studies to lead a discussion on how this content is critical to tasks performed across various career fields (e.g., business, marketing, finance). Students will use the discussion to guide their research of related careers for developing future career goals.

Instructional Resources/Tools

TI-83/84 and TI emulator
Quantitative Literacy The Art and Techniques of Simulation module
Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report
Software such as TinkerPlots and Fathom

Common Misconceptions

Students may believe:

That population parameters and sample statistics are one in the same, e.g., that there is no difference between the population mean which is a constant and the sample mean which is a variable.

Making decisions is simply comparing the value of one observation of a sample statistic to the value of a population parameter, not realizing that a distribution of the sample statistic needs to be created.

Diverse Learners

Strategies for meeting the needs of all learners including gifted students, English Language Learners (ELL) and students with disabilities can be found at this site. Additional strategies and resources based on the Universal Design for Learning principles can be found at www.cast.org.

Connections:

The four-step statistical process was introduced in Grade 6, with the recognition of statistical questions. At the high school level, students need to become proficient in all the steps of the statistical process.

Using simulation to estimate probabilities is a part of the Grade 7 curriculum as is initial understanding of using random sampling to draw inferences about a population.

Domain	Making Inferences and Justifying Conclusions
Cluster	*Make inferences and justify conclusions from sample surveys, experiments, and observational studies*
Standards	3. Recognize the purposes of and differences among sample surveys, experiments, and observational studies; explain how randomization relates to each. 4. Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling. 5. Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant. 6. Evaluate reports based on data.

Content Elaborations

Ohio has chosen to support shared interpretation of the standards by linking the work of multistate partnerships as the Mathematics Content Elaborations. Further clarification of the standards can be found through these reliable organizations and their links:

- Achieve the Core Modules, Resources
- Hunt Institute Video examples
- Institute for Mathematics and Education Learning Progressions Narratives
- Illustrative Mathematics Sample tasks
- National Council of Supervisors of Mathematics (NCSM) Resources, Lessons, Items
- National Council of Teacher of Mathematics (NCTM) Resources, Lessons, Items
- Partnership for Assessment of Readiness for College and Careers (PARCC) Resources, Items

Expectations for Learning

Ohio has selected PARCC as the contractor for the development of the Next Generation Assessments for Mathematics. PARCC is responsible for the development of the framework, blueprints, items, rubrics, and scoring for the assessments. Further information can be found at Partnership for Assessment of Readiness for College and Careers (PARCC). Specific information is located at these links:

- Model Content Framework
- Item Specifications/Evidence Tables
- Sample Items
- Calculator Usage
- Accommodations
- Reference Sheets

Instructional Strategies and Resources

Instructional Strategies

This cluster is designed to bring the four-step statistical process (GAISE model) to life and help students understand how statistical decisions are made. The mastery of this cluster is fundamental to the goal of creating a statistically literate citizenry. Students will need to use all of the data analysis, statistics, and probability concepts covered to date to develop a deeper understanding of inferential reasoning.

Students learn to devise plans for collecting data through the three primary methods of data production: surveys, observational studies, and experiments. Randomization plays various key roles in these methods. Emphasize that randomization is not a haphazard procedure, and that it requires careful implementation to avoid biasing the analysis. In surveys, the sample selected from a population needs to be representative; taking a random sample is generally what is done to satisfy this requirement. In observational studies, the sample needs to be representative of the population as a whole to enable generalization from sample to population. The best way to satisfy this is to use random selection in choosing the sample. In comparative experiments between two groups, random assignment of the treatments to the subjects is essential to avoid damaging problems when separating the effects of the treatments from the effects of some other variable, called confounding. In many cases, it takes a lot of thought to be sure that the method of randomization correctly produces data that will reflect that which is being analyzed. For example, in a two-treatment randomized experiment in which it is desired to have the same number of subjects in each treatment group, having each subject toss a coin where Heads assigns the subject to treatment A and Tails assigned the subject to treatment B will not produce the desired random assignment of equal-size groups.

The advantage that experiments have over surveys and observational studies is that one can establish causality with experiments.

Standard 4 addresses estimation of the population proportion parameter and the population mean parameter. Data need not come from just a survey to cover this topic. A margin-of-error formula cannot be developed through simulation, but students can discover that as the sample size is increased, the empirical distribution of the sample proportion and the sample mean tend toward a certain shape (the Normal distribution), and the standard error of the statistics decreases (i.e. the variation) in the models becomes smaller. The actual formulas will need to be stated.

Standard 5 addresses testing whether some characteristic of two paired or independent groups is the same or different by the use of resampling techniques. Conclusions are based on the concept of p-value. Resampling procedures can begin by hand but typically will require technology to gather enough observations for which a p-value calculation will be meaningful.

Use a variety of devices as appropriate to carry out simulations: number cubes, cards, random digit tables, graphing calculators, computer programs.

Instructional Resources/Tools

TI-83/84 and TI emulator
Quantitative Literacy *The Art and Techniques of Simulation* module
Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report
Making Sense of Statistical Studies
Focus in High School mathematics: Statistics and Probability
Navigating through Data Analysis in Grades 9-12
"A Sequence of Activities for Developing Statistical Concepts," Franklin and Kader, The Statistics Teacher Network Number 68 Winder 2006,
Software such as TinkerPlots and Fathom

Common Misconceptions

Students may believe:

That collecting data is easy; asking friends for their opinions is fine in determining what everyone thinks.

That causal effect can be drawn in surveys and observational studies, instead of understanding that causality is in fact a property of experiments.

That inference from sample to population can be done only in experiments. They should see thatinference can be done in sampling and observational studies if data are collected through a random process.

Diverse Learners

Strategies for meeting the needs of all learners including gifted students, English Language Learners (ELL) and students with disabilities can be found at this site. Additional strategies and resources based on the Universal Design for Learning principles can be found at www.cast.org.

Specific strategies for mathematics may include:
Students need to understand the differences between sample surveys, observational studies, and experiments and the role that randomness plays in each.

Students like to ask each other questions, but constructing meaningful, unbiased survey questions is not easy. Begin by critiquing published surveys before having students design their own.

Connections:

The four-step statistical process was introduced in middle school, with the first step likely more often generated by teachers than students. At the high school level, students need to become proficient in the first step of generating meaningful questions, as well as designing a plan to collect their data using the three primary methods: surveys, observational studies, and experiments.

Using simulation to estimate probabilities is a part of the Grade 7 curriculum, as is introductory understanding of using random sampling to draw inferences about a population.

Domain	**Conditional Probability and the Rules of Probability**
Cluster	***Understand independence and conditional probability and use them to interpret data***
Standards	1. Describe events as subsets of a sample space (the set of outcomes) using characteristics (or categories) of the outcomes, or as unions, intersections, or complements of other events (“or,”“and,” “not”). 2. Understand that two events **A** and **B** are independent if the probability of **A** and **B** occurring together is the product of their probabilities, and use this characterization to determine if they are independent. 3. Understand the conditional probability of **A** given **B** as **P (A and B)/P (B)**, and interpret independence of **A** and **B** as saying that the conditional probability **A** given **B** is the same as the probability of **A**, and the conditional probability of **B** given **A** is the same as the probability of **B.** 4. Construct and interpret two-way frequency tables of data when two categories are associated with each object being classified. Use the two-way table as a sample space to decide if events are independent and to approximate conditional probabilities. *For example, collect: data from a random sample of students in your school on their favorite subject among math, science, and English. Estimate the probability that a randomly selected student from your school will favor science given that the student is in tenth grade. Do the same for other subjects and compare the results.* 5. Recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations. *For example, compare the chance of having lung cancer if you are a smoker with the chance of being a smoker if you have lung cancer.*

Content Elaborations

Ohio has chosen to support shared interpretation of the standards by linking the work of multistate partnerships as the Mathematics Content Elaborations. Further clarification of the standards can be found through these reliable organizations and their links:

- Achieve the Core Modules, Resources
- Hunt Institute Video examples
- Institute for Mathematics and Education Learning Progressions Narratives
- Illustrative Mathematics Sample tasks
- National Council of Supervisors of Mathematics (NCSM) Resources, Lessons, Items
- National Council of Teacher of Mathematics (NCTM) Resources, Lessons, Items
- Partnership for Assessment of Readiness for College and Careers (PARCC) Resources, Items

Expectations for Learning

Ohio has selected PARCC as the contractor for the development of the Next Generation Assessments for Mathematics. PARCC is responsible for the development of the framework, blueprints, items, rubrics, and scoring for the assessments. Further information can be found at Partnership for Assessment of Readiness for College and Careers (PARCC). Specific information is located at these links:

- Model Content Framework
- Item Specifications/Evidence Tables
- Sample Items
- Calculator Usage
- Accommodations
- Reference Sheets

Instructional Strategies and Resources

Instructional Strategies

The **Standard for Mathematical Practice**, *precision* is important for working with conditional probability. Attention to the definition of an event along with the writing and use of probability function notation are important requisites for communication of that precision. For example: Let A: Female and B: Survivor, then P(A|B) =. The use of a vertical line for the conditional “given” is not intuitive for students and they often confuse the events B|A and A|B. Moreover, they often find identifying a conditional difficult when the problem is expressed in words in which the word “given” is omitted. For example, find the probability that a female is a survivor. The standard ***Make sense of problems and persevere in solving them*** also should be employed so students can look for ways to construct conditional probability by formulating their own questions and working through them such as is suggested in standard 4 above. Students should learn to employ the use of Venn diagrams as a means of finding an entry into a solution to a conditional probability problem.

It will take a lot of practice to master the vocabulary of “or,” “and,” “not” with the mathematical notation of union (∪), intersection (∩), and whatever notation is used for complement.

The independence of two events is defined in Standard 2 using the intersection. It is far more intuitive to introduce the independence of two events in terms of conditional probability (stated in Standard 3), especially where calculations can be performed in two-way tables.

The Standards in this cluster deliberately do not mention the use of tree diagrams, the traditional way to treat conditional probabilities. Instead, probabilities of conditional events are to be found using a two-way table wherever possible. However, tree diagrams may be a helpful tool for some students. The difficulty is realizing that the second set of branches are conditional probabilities.

Instructional Resources/Tools

The *Titanic Problem* is a well-known problem that revolves around conditional probability.

Focus in High School Mathematics: Reasoning and Sense Making in Statistics and Probability, NCTM.

Navigating through Probability in Grades 9-12, NCTM.

Common Misconceptions

Students may believe:

That multiplying across branches of a tree diagram has nothing to do with conditional probability.

That independence of events and mutually exclusive events are the same thing.

Diverse Learners

Strategies for meeting the needs of all learners including gifted students, English Language Learners (ELL) and students with disabilities can be found at this site. Additional strategies and resources based on the Universal Design for Learning principles can be found at www.cast.org.

Specific strategies for mathematics may include:

There are many good problems that can appeal to students’ sensitivities of fairness and justice in society. Students can formulate their questions that concern how certain characteristics of their own identity groups are viewed by society and understand how conditional probability is often misunderstood by society as whole.

Using a two-way table begins with calculation of marginal probabilities. Conditional probabilities and determination of independent events follow. Even a Bayes problem that in a tree-diagram context asks for the problem of a particular prior event having happened in the first set of branches when the given information is about what specifically happened in the second set of branches, becomes very straightforward when using a two-way table, while also avoiding a lot of confusing tree-diagram notation.

Connections:

Beginning work with categorical variables and two-way tables occurs in Grade 8. It is likely that these standards will need to be revisited on a deeper level in order to be successful with these standards.

Domain	Conditional Probability and the Rules of Probability
Cluster	***Use the rules of probability to compute probabilities of compound events in a uniform probability model***
Standards	6. Find the conditional probability of A given B as the fraction of B's outcomes that also belong to A, and interpret the answer in terms of the model. 7. Apply the Addition Rule, P(A or B) = P(A) + P(B) – P A and B), and interpret the answer in terms of the model. 8. (+) Apply the general Multiplication Rule in a uniform probability model, P(A and B) = P(A)P(B|A) = P(B)P(A|B), and interpret the answer in terms of the model. 9. (+) Use permutations and combinations to compute probabilities of compound events and solve problems.

Content Elaborations

Ohio has chosen to support shared interpretation of the standards by linking the work of multistate partnerships as the Mathematics Content Elaborations. Further clarification of the standards can be found through these reliable organizations and their links:

- Achieve the Core Modules, Resources
- Hunt Institute Video examples
- Institute for Mathematics and Education Learning Progressions Narratives
- Illustrative Mathematics Sample tasks
- National Council of Supervisors of Mathematics (NCSM) Resources, Lessons, Items
- National Council of Teacher of Mathematics (NCTM) Resources, Lessons, Items
- Partnership for Assessment of Readiness for College and Careers (PARCC) Resources, Items

Expectations for Learning

Ohio has selected PARCC as the contractor for the development of the Next Generation Assessments for Mathematics. PARCC is responsible for the development of the framework, blueprints, items, rubrics, and scoring for the assessments. Further information can be found at Partnership for Assessment of Readiness for College and Careers (PARCC). Specific information is located at these links:

- Model Content Framework
- Item Specifications/Evidence Tables
- Sample Items
- Calculator Usage
- Accommodations
- Reference Sheets

Instructional Strategies and Resources

Instructional Strategies

Identifying that a probability is conditional when the word "given" is not stated can be very difficult for students. For example, if a balanced tetrahedron with faces 1, 2, 3, 4 is rolled twice, what is the probability that the sum is prime (A) of those that show a 3 on at least one roll (B)? Whether what is asked for is P(A and B), P(A or B), or P(A|B) can be problematic for students. Showing the outcomes in a Venn Diagram may be useful. The calculation to find the probability that the sum is prime (A) given at least one roll shows 3 (B) is to count the elements of B by listing them if possible, namely in this example, there are 7 paired outcomes (31, 32, 33, 34, 13, 23, 43). Of those 7 there are 4 whose sum is prime (32, 34, 23, 43). Hence in the long run, 4 out of 7 times of rolling a fair tetrahedron twice, the sum of the two rolls will be a prime number under the condition that at least one of its rolls shows the digit 3.

Note that if listing outcomes is not possible, then counting the outcomes may require a computation technique involving permutations or combinations, which is a STEM topic.

In the above example, if the question asked were what is the probability that the sum of two rolls of a fair tetrahedron is prime (A) or at least one of the rolls is a 3 (B), then what is being asked for is P(A or B) which is denoted as $P(A \cup B)$ in set notation. Again, it is often useful to appeal to a Venn Diagram in which A consists of the pairs: 11, 12, 14, 21, 23, 32, 34, 41, 43; and B consist of 13, 23, 33, 43, 31, 32, 34. Adding P(A) and P(B) is a problem as there are duplicates in the two events, namely 23, 32, 34, and 43. So P(A or B) is 9/16 + 7/16 – 4/16 = 12/16 or 3/4, so 3/4th of the time, the result of rolling a fair tetrahedron twice will result in the sum being prime, or at least one of the rolls showing a 3, or perhaps both will occur.

It should be noted that the Multiplication Rule in Standard 8 is designated as STEM when it is connected to the discussion of independence in Standard 2 of the previous S-CP cluster. The formula P(A and B) = P(A)P(B|A) is best illustrated in a two-stage setting in which A denotes the outcome of the first stage, and B, the second. For example, suppose a jar contains 7 red and 3 green chips. If one draws two chips without replacement from the jar, the probability of getting a red followed by a green is P(red on first and green on second) = P(red on first)P(green on second given a red on first) = (7/10)(3/9) = 21/90. Demonstrated on a tree diagram indicates that the conditional probabilities are on the second set of branches.

Instructional Resources/Tools

NCTM Navigating through Probability 9-12.
The National Council of Teachers of Mathematics, Illuminations

TI-83/84 and TI emulator (for permutations and combinations calculations)

Common Misconceptions

Students may believe:

That the probability of A or B is always the sum of the two events individually.
That the probability of A and B is the product of the two events individually, not realizing that one of the probabilities may be conditional.

Diverse Learners

Strategies for meeting the needs of all learners including gifted students, English Language Learners (ELL) and students with disabilities can be found at this site. Additional strategies and resources based on the Universal Design for Learning principles can be found at www.cast.org.

Connections:

Probability is introduced in Grade 7. The concepts of independence and conditional probabilty are the extended topics in domain S-CP.

Domain	**Using Probability to Make Decisions**
Cluster	***Calculate expected values and use them to solve problems***
Standards	1. (+) Define a random variable for a quantity of interest by assigning a numerical value to each event in a sample space; graph the corresponding probability distribution using the same graphical displays as for data distributions. 2. (+) Calculate the expected value of a random variable; interpret it as the mean of the probability distribution. 3. (+) Develop a probability distribution for a random variable defined for a sample space in which theoretical probabilities can be calculated; find the expected value. *For example, find the theoretical probability distribution for the number of correct answers obtained by guessing on all five questions of a multiple-choice test where each question has four choice, and find the expected grade under various grading schemes.* 4. (+) Develop a probability distribution for a random variable defined for a sample space in which probabilities are assigned empirically; find the expected value. *For example, find a current data distribution on the number of TV sets per household in the United States, and calculate the expected number of sets per household. How many TV sets would you expect to find in 100 randomly selected households?*

Content Elaborations

Ohio has chosen to support shared interpretation of the standards by linking the work of multistate partnerships as the Mathematics Content Elaborations. Further clarification of the standards can be found through these reliable organizations and their links:

- Achieve the Core Modules, Resources
- Hunt Institute Video examples
- Institute for Mathematics and Education Learning Progressions Narratives
- Illustrative Mathematics Sample tasks
- National Council of Supervisors of Mathematics (NCSM) Resources, Lessons, Items
- National Council of Teacher of Mathematics (NCTM) Resources, Lessons, Items
- Partnership for Assessment of Readiness for College and Careers (PARCC) Resources, Items

Expectations for Learning

Ohio has selected PARCC as the contractor for the development of the Next Generation Assessments for Mathematics. PARCC is responsible for the development of the framework, blueprints, items, rubrics, and scoring for the assessments. Further information can be found at Partnership for Assessment of Readiness for College and Careers (PARCC). Specific information is located at these links:

- Model Content Framework
- Item Specifications/Evidence Tables
- Sample Items
- Calculator Usage
- Accommodations
- Reference Sheets

Instructional Strategies and Resources

Instructional Strategies

This domain and cluster belong to STEM, and hence need not be for all students.

Probability assignments to the simple events (i.e., outcomes) listed in a sample space must be positive numbers that sum to one over all the simple events. Note that the simple events are mutually exclusive.

Probability assignments may be determined theoretically such as assigning 1/8 to each of the outcomes {1,2,3,4,5,6,7,8} in one roll of a fair (balanced) octahedron. They may also be assigned empirically in experiments for which the theoretical probabilities are too difficult to find, such as how often the possible hands of poker occur, or cannot be found such as how often a thumbtack lands up or on its side, or require a condition to be specified to make sense of the problem such as the number of children in an American household. To make sense of the latter, a random sample of 150 households needs to be taken and the number of children per household determined. For example, suppose that the values are (0,1,2,3,4,5,6,7,8} with frequencies {50,35,70,27,7,2,4,0,5}. The empirical assignment is by relative frequencies given by f/150 for each value, respectively.

For data sets, the arithmetic average (fair share or balance point of deviations) was calculated as a measure of center for the data. Analogously, for probability distributions, the weighted average of the values with their probabilities or values with their relative frequencies $\sum_i x_i p(x_i)$ or $\sum_i x_i \left(f_i / n \right)$ give a measure of center for the probability distribution. This is called the expected value of the random variable.

Instructional Resources/Tools

NCTM Navigating through Probability 9-12.
The National Council of Teachers of Mathematics, Illuminations

Common Misconceptions

Students may believe:

That as long as a fair device is being used, the values in the sample space are equally likely. For example, if a fair die is rolled twice but the sample space consists of the sum of the two rolls, i.e., value {2,3,4,5,6,7,8,9,10,11,12}, it is not the case that the probability assignment to the sums is 1/11.

Diverse Learners

Strategies for meeting the needs of all learners including gifted students, English Language Learners (ELL) and students with disabilities can be found at this site. Additional strategies and resources based on the Universal Design for Learning principles can be found at www.cast.org.

Connections:

Probability is introduced in Grade 7.

Domain	Using Probability to Make Decisions
Cluster	***Use probability to evaluate outcomes of decisions***
Standards	5. (+) Weigh the possible outcomes of a decision by assigning probabilities to payoff values and finding expected values. a. Find the expected payoff for a game of chance. *For example, find the expected winnings from a state lottery ticket or a game at a fast-food restaurant.* b. Evaluate and compare strategies on the basis of expected values. *For example, compare high-deductible versus a low-deductible automobile insurance policy using various, but reasonable, chances of having a minor or a major accident.* 6. (+) Use probabilities to make fair decisions (e.g., drawing by lots, using a random number generator). 7. (+) Analyze decisions and strategies, using probability concepts (e.g., product testing, medical testing, pulling a hockey goalie at the end of a game.)

Content Elaborations

Ohio has chosen to support shared interpretation of the standards by linking the work of multistate partnerships as the Mathematics Content Elaborations. Further clarification of the standards can be found through these reliable organizations and their links:

- Achieve the Core Modules, Resources
- Hunt Institute Video examples
- Institute for Mathematics and Education Learning Progressions Narratives
- Illustrative Mathematics Sample tasks
- National Council of Supervisors of Mathematics (NCSM) Resources, Lessons, Items
- National Council of Teacher of Mathematics (NCTM) Resources, Lessons, Items
- Partnership for Assessment of Readiness for College and Careers (PARCC) Resources, Items

Expectations for Learning

Ohio has selected PARCC as the contractor for the development of the Next Generation Assessments for Mathematics. PARCC is responsible for the development of the framework, blueprints, items, rubrics, and scoring for the assessments. Further information can be found at Partnership for Assessment of Readiness for College and Careers (PARCC). Specific information is located at these links:

- Model Content Framework
- Item Specifications/Evidence Tables
- Sample Items
- Calculator Usage
- Accommodations
- Reference Sheets

Instructional Strategies and Resources

Instructional Strategies

This domain and cluster belong to STEM, and hence need not be for all students.

A game of chance is said to be fair if the expected net winnings are 0. If the expected net winnings is negative, then the player needs to decide if the game is worth playing. For example, a spinner has 18 red, 18 black and 2 green sections. Suppose, players gain a one score point if the spinner lands on red, otherwise the players loose a one score point. The probability the spinner lands on red is $\frac{18}{38}$. The probability it lands elsewhere is $\frac{20}{38}$. So, the expected probability is $1 \times \frac{18}{38} + (-1) \times \frac{20}{38} = -.053$ score points. This means that players should expect to lose a little over .05 of a score point every time they play the game. Calculating an expected value enables players to decide whether or not the game is worth playing

Expected values may be used to decide between two strategies. For example, suppose shop owner needs to decide whether to stock product A or product B and can only stock one of them. Profit margins for A follow the distribution (in thousands of dollars): 5,4,3,2,1 with probabilities .1,.45,.3,.1,.05, respectfully. Those for B follow: 8,7,6,5,4,3,2,1,0 with probabilities: .1,.15,.15,.1,.1,0,0,0,.4. The expected profit by stocking A is 5(.1)+4(.45)+3(.3)+2(.1)+1(.05) = 3.45 thousands of dollars. The expected profit by stocking B is 8(.1)+7(.15)+6(.15)+5(.1)+4(.1)+0(.4) = 3.65 thousands of dollars. So, based on expected values of profit margins, the better choice would be to stock product B.

Conditional probabilities are situations where the interpretation of an observation is dependent upon or “conditioned on” some other factor. For example, a blood test has been shown to indicate the presence of a particular disease 95% of the time when the disease is actually present. The same blood test gives a false positive result 0.5% of the time. A false positive result suggests that even though the blood test indicates that the person has the disease (the positive part) but subsequent, additional testing indicates the person does not have that disease (hence positive but false or a false positive).

Suppose that one percent of the population actually has the disease. If a person’s blood test is positive, how likely is it that the person has the disease?

This scenario can be restated as the following conditional probability problem:

“What is the probability that a person actually has the disease given that (or conditioned on) the blood test indicates the person has the disease?”
There are two possibilities for a person to produce a positive blood test result: the person has the disease or the person does not have the disease.

The probability that a person has the disease given a positive blood test result is $0.01 \times 0.95 = 0.0095$. This represents the part of the population that is expected to both provide a positive blood test result and have the disease. The probability that a person gives a positive blood test result but does not have the disease is $0.99 \times 0.005 = 0.00495$. This represents the part of the population that is expected to give a positive blood test result but does not have the disease. The total, or $0.01445 = 0.0095 + 0.00495$, represents a part of the population that has a positive blood test, whether it has a disease or not. The conditional probability that a person actually has the disease when the blood test is positive is $\frac{.0095}{.01445} = 0.657$, i.e., the probability that a person actually has the disease is about 66% given that the blood test results are positive.

Instructional Resources/Tools

NCTM Navigating through Probability 9-12.
Data Driven Mathematics module, Probability Models

Common Misconceptions

Students may believe:

Probabilities and expected values aren’t useful in making decisions that affect one’s life. Students need to see that these are not merely textbook exercises.

Diverse Learners

Strategies for meeting the needs of all learners including gifted students, English Language Learners (ELL) and students with disabilities can be found at this site. Additional strategies and resources based on the Universal Design for Learning principles can be found at www.cast.org.

Connections:

Probability is introduced in Grade 7.