AP Statistics Unit 2 - Probability and Distributions

Content Area: **Math**
Course(s):
Time Period: **MP2**
Length: **45**
Status: **Published**

Unit Overview

Unit Summary	Unit Rationale
Unit 2 for AP Statistics guides students through a highly structured investigation of randomness, progressing from the basic rules of probability to the abstract behaviors of random variables and sampling distributions. Operating within an explicit Experience First, Formalize Later (EFFL) framework, the unit begins by demystifying chance events through applied simulations, demonstrating how steady patterns emerge from the long-run relative frequencies governed by the Law of Large Numbers. Students transition from counting raw event outcomes to applying formal rules for addition, multiplication, and conditional probability in multi-stage scenarios. Using two-way tables, Venn diagrams, and tree diagrams, students learn to test for independence mathematically and evaluate complex systems of probability, establishing a structural baseline that prepares them to move beyond simple events and explore functional models. As the curriculum advances into its second half, students formalize these probability structures as discrete and continuous random variables, learning to map out full sample spaces and assign valid probability distributions. They master calculations for finding and interpreting the parameters of center and variability, computing the expected value and variance of a variable before learning how constants transform these statistics. A significant portion of this phase challenges students to combine independent random variables, forcing them to apply the Pythagorean relationship of variance, where variability always compounds, regardless of whether variables are added or subtracted. The unit concludes by evaluating specialized probability models, including binomial settings checked via the BINS conditions and geometric distributions, before applying these models to macro-level sampling distributions. Students learn to apply the Central	This Probability and Random Variables unit is designed to serve as the primary mathematical and conceptual link between descriptive data analysis and formal statistical inference. In most high school math classes, probability is often taught as a small, separate unit about permutations, combinations, and counting problems. In AP Statistics, though, probability is the language we use to talk about uncertainty and is the foundation for all significance testing. By teaching basic probability rules alongside random variables and sampling distributions over a focused 45-day period, this unit bridges the usual gap between simple probability calculations and the construction of sampling models. It builds on the patterns students noticed during data collection and descriptive statistics in Super-Unit 1, turning those ideas into clear mathematical rules. This helps students see that while single outcomes are unpredictable, long-term patterns follow strict mathematical laws. Teaching this unit with the EFFL model in 47-minute classes helps students by starting with hands-on activities before moving to theory. Many students find it hard to understand discrete distributions or binomial formulas when they only see the algebra. Letting them first try simulations, like rolling dice, using applets, or drawing chips, helps them naturally see expected values and long-term patterns before connecting those ideas to symbols like μ or σ. Bringing the main ideas of sampling distributions into this unit also completes the study of probability distributions. Students can then see a sampling distribution as just another probability model, but one that tracks a statistic instead of a single variable. This approach makes sure that when students later learn about confidence intervals and significance tests, they see these as logical next steps based on the probability models they already know.

Limit Theorem to sample proportions and means, verify the conditions for normality and sample size, and use standard normal distributions to calculate probabilities. This sequence ensures that students develop a non-deterministic mindset and master the specific predictive methods and calculator operations required to perform in-depth statistical inference.	

NJSLS

2.1.A.1	A two-way table, also called a contingency table, can be used to summarize and compare data for two categorical variables. The entries in the cells of the table can be frequencies (i.e., counts) or relative frequencies (i.e., proportions).
2.1.A.2	Side-by-side bar charts, segmented bar charts, and mosaic plots are examples of graphs used to display the relationship between two categorical variables. In these graphs, the frequency or relative frequency of each category, or level, of one of the categorical variables is displayed for each category of the other categorical variable.
2.1.A.3	Graphical representations of two categorical variables can be used to compare the relationship of one categorical variable across the levels of the other categorical variable and determine whether the two variables are associated.
2.1.B.1	Tabular and graphical representations for the distributions of two categorical variables may reveal information that can be used to justify claims about the variable in context.
2.2.A.1	A joint relative frequency in a two-way table is a cell frequency divided by the total for the entire table.
2.2.A.2	A marginal relative frequency in a two-way table is a row total divided by the total for the entire table or a column total divided by the total for the entire table.
2.2.A.3	A conditional relative frequency is a relative frequency computed by restricting to a particular level, or category of interest. A conditional relative frequency can be a cell frequency in a row divided by the total for that row or it can be a cell frequency in a column divided by the total for that column.
2.2.B.1	Summary statistics for two categorical variables can be used to compare distributions for evidence of association between the two variables.
2.2.C.1	Summary statistics for two categorical variables may reveal information that can be used to justify claims about the variables in context.
2.3.A.1	A random process generates results that are determined by chance.
2.3.A.2	An outcome is the result of one trial of a random process.
2.3.A.3	An event is a collection of outcomes.
2.3.A.4	Simulation is a way to model random events such that simulated outcomes closely match real-world outcomes. All possible outcomes are associated with a value to be determined by chance. Record the counts of simulated outcomes and the count total.
2.3.A.5	The probability of an outcome or event is its long-run relative frequency—that is, its relative frequency over a large number of trials.
2.3.A.6	The relative frequency of an outcome or event determined from empirical data can be used to estimate the actual, or true, probability of that outcome or event.
2.3.A.7	The law of large numbers states that for independent trials, as the number of trials

increases, the long-run relative frequency of the outcome or event gets closer and closer to a single value.

2.4.A.1 The sample space of a random process is the set of all possible nonoverlapping outcomes. The probability of the sample space is 1.

2.4.A.2 If all outcomes in the sample space are equally likely, then the theoretical probability an event E will occur is (number of outcomes in event E)/(total number of outcomes in sample space). The probability of event E occurring is written as $P(E)$.

2.4.A.3 The probability of an event is a number between 0 and 1, inclusive.

2.4.A.4 The probability of the complement of an event E, which can be written as E', $\overline{E}$, or E^C (i.e., the probability of "not E") is equal to 1 − $P(E)$.

2.5.A.1 The probability that events A and B both will occur, sometimes called the joint probability, is the probability of the intersection of A and B. Joint probability is defined as $P(A$ intersect $B)$ or $P(A \cap B)$.

2.5.A.2 Two events are mutually exclusive, or disjoint, if they cannot occur at the same time. This means that if two events are mutually exclusive, then $P(A \cap B) = 0$.

2.6.A.1 The probability that event A will occur given that event B has occurred is called a conditional probability and is written as $P(A|B)$. Conditional probability is defined as $P(A|B) = P(A \cap B)/P(B)$

2.6.A.2 The general multiplication rule states that the probability that events A and B both will occur is equal to the probability that event A will occur multiplied by the conditional probability that event B will occur given that event A has occurred. The multiplication rule is defined as $P(A \cap B) = P(A) \cdot P(B|A)$.

2.7.A.1 Events A and B are independent if and only if knowing whether event A has occurred (or will occur) does not change the probability that event B will occur. When events A and B are independent, then $P(A|B) = PA$, $P(B|A) = P(B)$, and $P(A \cap B) = P(A) \cdot P(B)$.

2.7.A.2 The probability that event A or event B (or both) will occur is the probability of A union B. The probability of the union is defined as $P(A \cup B)$.

2.7.A.3 $P(A$ union $B) = P(A) + P(B) – P(A$ intersect $B)$, or $P(A \cup B) = P(A) + P(B) – P(A \cap B)$.

2.8.A.1 A random variable is a variable whose values have numerical outcomes that result from a random phenomenon.

2.8.A.2 A probability distribution for a discrete random variable shows the probability associated with every possible value of the random variable. The sum of the probabilities over all possible values of a discrete random variable is 1.

2.8.A.3 A discrete probability distribution can be determined using the rules of probability or estimated with a simulation.

2.8.A.4 A discrete probability distribution can be represented as a graph, table, or function showing the probabilities associated with values of a random variable.

2.8.A.5 A cumulative probability distribution can be represented as a table or function and shows the probability of being less than or equal to each value of the discrete random variable.

2.9.A.1 A numerical value measuring a characteristic of a probability distribution of a random variable, or a population, is a parameter. The value of a parameter is a single, fixed value.

2.9.A.2 The expected value (or mean) of a probability distribution is a parameter and is denoted by $E(X)$ or μ_x For a discrete random variable X, the expected value is calculated as $\mu_x = \Sigma x_i \cdot P(x_i)$, where x_i is the possible value of the random variable and $P(x_i)$ is the probability of the possible value of the random variable. The expected value can be interpreted as the long-run average outcome of the random variable. The discrete random variable can only take on values that are countable or finite.

2.9.A.3 The standard deviation of a probability distribution is a parameter represented by $SD(X)$ or σ_x, For a discrete random variable X, the standard deviation is calculated as σ_x = √[Σ(x_i –

	$\mu_x)^2 \cdot P(x_i)]$, where x_i is the possible value of the random variable, μ_x is the mean, and $P(x_i)$ is the probability of the possible value of the random variable. The standard deviation can be interpreted as the typical deviation of the values of the random variable from the mean value (or expected value) of the random variable over the long run. The square of the standard deviation of a random variable is called the variance of the random variable and is denoted as $V(X)$ or σ_x^2.
2.9.B.1	The mean and standard deviation for the probability distribution of a discrete random variable should be interpreted in the context of a specific population.
2.10.A.1	A binomial random variable, X, is a discrete random variable that counts the number of successes in repeated independent trials, n, that have only two possible outcomes (success or failure), with the probability of success p and the probability of failure $1 - p$.
2.10.B.1	If a random variable is binomial, its mean, μ_x, is np and its standard deviation, σ_x, is $\surd[np(1 - p)]$.
2.10.C.1	The mean, standard deviation, and probabilities for a binomial distribution should be interpreted in context.
2.10.D.1	A probability distribution can be constructed using the rules of probability or estimated with a simulation.
2.10.E.1	The probability that a binomial random variable, X, has exactly x successes for n independent trials, when the probability of success is p, is calculated as $P(X = x)$ = (n over x) $p^x (1 - p)^{n-x}$, x = 0, 1, 2, . . . , n. This is called the binomial probability function.
2.11.A.1	A continuous random variable is a variable that can take on any value within a specified domain. Every interval within the domain has a probability associated with it.
2.11.A.2	Many continuous random variables are well-modeled by a normal distribution.
2.11.A.3	A normal distribution can be described as a continuous, unimodal, bell-shaped, and symmetric curve.
2.11.A.4	A normal curve can be used to model a distribution of data and a continuous random variable.
2.11.A.5	The normal distribution, or the normal curve, is identified by two parameters, the mean, μ, and the standard deviation, σ. The smaller the standard deviation, the taller and more concentrated the normal curve is around its mean. The larger the standard deviation, the shorter and less concentrated the normal curve is around its mean.
2.11.B.1	A standard normal distribution is a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$.
2.11.C.1	The empirical rule can be used to estimate the area of a region under the graph of the normal distribution curve. For a normal distribution, approximately 68% of observations are within 1 standard deviation of the mean, approximately 95% of observations are within 2 standard deviations of the mean, and approximately 99.7% of observations are within 3 standard deviations of the mean. This is called the empirical rule, or the 68–95–99.7 rule.
2.11.D.1	If the distribution of a random variable is approximately normal, the probability that the random variable takes on values within a particular interval of the random variable is determined by the area under the normal curve within that interval. The total probability or area under the normal curve is 1.
2.11.E.1	The boundaries of an interval associated with a given area in a normal distribution can be determined using technology or using z-scores and a standard normal table.
2.11.E.2	Intervals associated with a given area in a normal distribution can be determined by assigning appropriate inequalities to the boundaries of the intervals. To determine the intervals, p is defined as a number between 0 and 100, x_a is the lower bound, and x_b is the upper bound on a normal distribution.
2.11.E.2.i	$P(X < x_a) = p/100$ means that the lowest p% of the values lie to the left of x_a.

2.11.E.2.ii	$P(x_a < X < x\_b) = p/100$ means that the lowest p% of the values lie between x_a and x_b.
2.11.E.2.iii	$P(X > x\_b) = p/100$ means that the highest p% of the values lie to the right of x_b.
2.11.E.2.iv	To determine the most extreme p% of values on both sides requires dividing the area associated with p% into two equal areas on either extreme of the distribution: $P(X < x_a) = 1/2(p/100)$ and $P(X > x\_b) = 1/2(p/100)$ mean that half of the p% most extreme values lie to the left of x_a and half of the p% most extreme values lie to the right of x_b.
2.11.F.1	Percentiles and proportions may be used to compare relative positions of individual values within a normal distribution or between normal distributions.
2.12.A.1	A sampling distribution of a statistic is the distribution of values of the statistic for all possible samples of a given size from a given population.
2.12.A.2	The sampling distribution of a statistic can be simulated by repeatedly generating a large number of random samples from the population assuming known value(s) for the parameter(s). The value of the statistic is determined and recorded for each sample. The resulting distribution of the sample statistic values approximates the sampling distribution of the statistic.
2.12.A.3	A randomization distribution is the distribution of a statistic generated by simulation from repeatedly randomly reallocating, or reassigning, the response values to treatment groups. The value of the statistic is determined and recorded for each reallocation, or reassignment. The resulting distribution of the statistic values approximates the sampling distribution of the statistic.
2.12.A.4	The central limit theorem (CLT) states that the sampling distribution of a mean of a random sample has a shape that can be approximated by a normal distribution. The larger the sample is, the better the approximation will be.

Standards for Mathematical Practice

1	Make sense of problems and persevere in solving them
2	Reason abstractly and quantitatively
3	Construct viable arguments and critique the reasoning of others
4	Model with mathematics
5	Use appropriate tools strategically
6	Attend to precision
7	Look for and make use of structure
8	Look for and express regularity in repeated reasoning

Unit Focus

Enduring Understandings	Essential Questions
1. While individual outcomes in a random phenomenon are entirely unpredictable in the short run, a clear, deterministic mathematical pattern emerges over a large number of repeated trials (the Law of Large Numbers). 2. Probability rules and mathematical sample spaces allow us to quantify risk, calculate the likelihood of complex, multistage events, and	1. How can short-term chaos and long-term predictability coexist in a random process? 2. How do the rules of probability help us calculate the likelihood of multiple real-world events occurring simultaneously or sequentially? 3. How do individual random variables combine mathematically, and how do sampling

evaluate whether two variables are independent or influence one another. 3. Random variables and sampling distributions translate raw chance events into functional probability models, mapping out the expected long-run center, shape, and compounding variability of sample statistics.	distributions serve as the ultimate bridge between probability models and formal decision-making?

Instructional Focus

Learning Targets

Part 1: Probability

- I can interpret probability as a long-run relative frequency using the Law of Large Numbers.
- I can design and implement a simulation using a random number generator or a random number table to estimate probabilities.
- I can construct a sample space and use basic probability rules, including the complement rule and the Addition Rule for Mutually Exclusive events.
- I can calculate and interpret conditional probabilities using formulas, two-way tables, and Venn diagrams.
- I can determine if two events are independent mathematically by checking if $P(A|B) = P(A)$.
- I can utilize the General Multiplication Rule and construct tree diagrams to solve complex, multi-stage probability problems.

Part 2: Random Variables and Probability Distributions

- I can construct a discrete probability distribution, verify its validity, and calculate/interpret its Expected Value (Mean) and Standard Deviation in context.
- I can determine the structural effects of linear transformations (adding, subtracting, multiplying, and dividing by a constant) on the mean and standard deviation of a random variable.
- I can calculate the mean and standard deviation of the sum or difference of two independent random variables using the Pythagorean variance relationship.
- I can identify a Binomial setting using the BINS conditions and compute exact and cumulative probabilities using formulas and calculator functions (binompdf / binomcdf).
- I can compute and interpret the mean and standard deviation of a Binomial random variable.
- I can identify a Geometric setting using the BITS conditions, compute the probability of a success occurring on a specific trial, and calculate its expected value.
- I can describe the sampling distributions of a sample proportion (p-hat) and a sample mean (x-bar), verify their normality using safety checks ($np \geq 10$ and $n \geq 30$), and compute standard normal probabilities using the Central Limit Theorem.

Prerequisite Skills

Part 1: Probability • Students need to have fluency in multiplying probabilities (fractions, decimals, and percentages) across multi-step events. • Familiarity with algebraic notations of intersections and unions ("and" ∩, "or" ∪) • Navigating basic fraction structures within a table to find parts of a restricted whole. *Part 2: Random Variables and Probability Distributions* • Understanding that Σ represents the sum of the elements. • Confidently executing values into complex multi-step radical formulas, such as the standard deviation of combined variables. • Baseline conceptual awareness of combinations to determine the number of paths to a specific success structure.

Common Misconceptions
 	
Part 1: Probability
 	
• Students believe a random process must "even out" in the short run
• Students confuse independent with mutually exclusive and believe they mean the same thing.
    ○ Reinforce that if two events are mutually exclusive (cannot happen at the same time), they are automatically dependent because knowing one happened tells you the probability of the other happening is exactly 0.
 	
Part 2: Random Variables and Probability Distributions
 	
• When combining random variables (X + Y or X - Y), students attempt to add or subtract the standard deviations ($\sigma_x \pm \sigma_y$).
    ○ Enforce that variances add, standard deviations never do. Students must convert the variances to standard deviations, add them, and take the square root.
• Students assume that when calculating the standard deviation of a difference (X - Y), they must subtract the variances.
    ○ Emphasize that combining two uncertain variables always increases total variability, regardless of whether you add or subtract the values. You always add the variances.
• **On free-response questions, students write out calculator speak without labels such as binomcdf(10, 0.5, 3). Under the College Board guidelines, this would receive no credit. Students must explicitly define the parameters: binomcdf(n = 10, p = 0.5, x ≤ 3).

Spiraling For Mastery

Current Unit Content/Skills	Spiral Focus	Activity

Part 1: Probability
 	
• Applying Addition and Multiplication Rules
• Calculating conditional probability
• interpreting solutions
 	
Part 2: Random Variables and Distributions
 	
• Calculating expected value and standard deviation
• Combining independent variables.
	Part 1: Probability
	
	• Categorical Data displays from Unit 1
	• Transitioning from visual charts to mathematical counts
	
	Part 2: Random Variables and Distributions
	
	• Shapes, Centers, and Transformations from Unit 1
	• Recognizing the effect of linear transformations on data sets.
		Part 1: Probability
		
		• **The Mosaic Table Pivot**
		    o Present a mosaic plot or two-way table from Unit 1. Give students 3 minutes to convert it into a probability sample space and calculate P(A|B) to evaluate if the variables are independent.
		
		Part 2: Random Variables and Distributions
		
		• **The Better Game Shift**
		    o When teaching shifts in random variables (X + c vs. cX), display side-by-side boxplots from Super-Unit 1 showing a dataset multiplied by c. Ask students to justify why the standard deviation scales by the constant but the variance squares.

Assessment

Formative Assessment	Summative Assessment
• Homework • Lesson Checks • Quizzes • Exit Tickets • Lesson Reflections • Performance Tasks • AP Classroom Progress Checks	• Topic Tests • Unit 2 Benchmark (AP Classroom)

Resources

Key Resources	Supplemental Resources
AP Classroom AP Statistics CED [Pacing Guide]	iXL Delta Math Desmos Khan Academy Math Medic Skew the Script Teacher Made worksheets

Career Readiness, Life Literacies, and Key Skills

9.2.12.CAP.4	Evaluate different careers and develop various plans (e.g., costs of public, private, training schools) and timetables for achieving them, including educational/training requirements, costs, loans, and debt repayment.
9.2.12.CAP.6	Identify transferable skills in career choices and design alternative career plans based on those skills.
9.4.12.CT.2	Explain the potential benefits of collaborating to enhance critical thinking and problem solving (e.g., 1.3E.12profCR3.a).
9.4.12.DC.6	Select information to post online that positively impacts personal image and future college and career opportunities.
9.4.12.IML.3	Analyze data using tools and models to make valid and reliable claims, or to determine optimal design solutions (e.g., S-ID.B.6a., 8.1.12.DA.5, 7.1.IH.IPRET.8).
9.4.12.IML.7	Develop an argument to support a claim regarding a current workplace or societal/ethical issue such as climate change (e.g., NJSLSA.W1, 7.1.AL.PRSNT.4).
9.4.12.TL.3	Analyze the effectiveness of the process and quality of collaborative environments.
9.1.12.FI.3	Develop a plan that uses the services of various financial institutions to prepare for long term personal and family goals (e.g., college, retirement).
9.1.12.RM.1	Describe the importance of various sources of income in retirement, including Social Security, employer-sponsored retirement savings plans, and personal investments.

Interdisciplinary Connections

S.ID.B.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.
6.1.12.GeoGI.1.a	Explain how geographic variations impacted economic development in the New World, and its role in promoting trade with global markets (e.g., climate, soil conditions, other natural resources).
S.IC.A.2	Decide if a specified model is consistent with results from a given data-generating process, e.g., using simulation.
RL.CR.11–12.1	Accurately cite strong and thorough textual evidence and make relevant connections to strongly support a comprehensive analysis of multiple aspects of what a literary text says explicitly and inferentially, as well as interpretations of the text; this may include determining where the text leaves matters uncertain.
S.IC.B.3	Recognize the purposes of and differences among sample surveys, experiments, and observational studies; explain how randomization relates to each.
8.1.12.DA.5	Create data visualizations from large data sets to summarize, communicate, and support different interpretations of real-world phenomena.
RI.MF.9–10.6	Analyze, integrate, and evaluate multiple interpretations (e.g., charts, graphs, diagrams, videos) of a single text or text/s presented in different formats (visually, quantitatively) as well as in words in order to address a question or solve a problem.
RI.AA.9–10.7	Describe and evaluate the argument and specific claims in an informational text, assessing whether the reasoning is valid and the evidence is relevant and sufficient; identify false statements and reasoning.
W.AW.11–12.1	Write arguments to support claims in an analysis of substantive topics or texts, using valid reasoning and relevant and sufficient evidence.
8.2.12.ED.3	Evaluate several models of the same type of product and make recommendations for a new design based on a cost benefit analysis.
HS-LS2-6	Evaluate the claims, evidence, and reasoning that the complex interactions in ecosystems maintain relatively consistent numbers and types of organisms in stable conditions, but changing conditions may result in a new ecosystem.
	Global economic activities involve decisions based on national interests, the exchange of different units of exchange, decisions of public and private institutions, and the ability to distribute goods and services safely.
HS-ETS1-4	Use a computer simulation to model the impact of proposed solutions to a complex real-world problem with numerous criteria and constraints on interactions within and between systems relevant to the problem.