**Definition 1.1**

The mean of a sample of n

**Q:** An analyst is studying the scoring trends in NBA games from 1947 to 2015. They want to calculate the **average points scored by a single team in games played during the 1947 season**. Using the dataset, determine the average points scored by one team for all games in the 1947 season.

**A:** The average points scored by a single team in games played during the 1947 season is **67.92 points per game**.

**Definition 1.2**

The variance of a sample of measurements

**Q:** An NBA statistician is analyzing the scoring consistency of teams during the 1950 season. To measure how much the points scored by teams in individual games varied from the average points scored, the statistician wants to calculate the **sample variance** of the points scored for that season. Using the dataset, determine the sample variance of points scored by a team during the 1947 season.

**A:** The sample variance (s2s^2s2) of the points scored by teams during the 1947 NBA season is approximately **146.40**.

**Definition 1.3**

The standard deviation of a sample of measurements is the positive square root of that variance

**Q:** A sports historian is analyzing the scoring performance of NBA teams from the league's inception in 1947 through the 1955 season. The historian wants to measure the consistency of scoring during this period by calculating the **standard deviation** of points scored per game. Calculate the **standard deviation**, which represents the typical deviation of a team's score from the average score during the 1947–1955.

**A:** The typical deviation of a team's points scored in a game from the average is approximately **12.04 points**.

**Definition 2.6**

Suppose S is a sample space associated with an experiment. To every event A in S (A is a subset of S), we assign a number, P(A), called the probability of A, so that the following axioms hold:

Axiom 1:

Axiom 2:

Axiom 3: If

**Q:** An NBA analyst is studying games from the 1947–1955 seasons and is interested in the likelihood that a team scores within specific point ranges in a single game. The analyst defines the following mutually exclusive events:

1. **: The probability that a team scores between 40 and 59 points.**
2. **The probability that a team scores between 60 and 79 points.**
3. **​: The probability that a team scores between 80 and 99 points.**

The analyst wants to calculate the probability that a team scores **within any of these three ranges** in a randomly selected game.

**A:**

* (): The probability that a team scores between 40 and 59 points is approximately **0.0028** (0.28%).
* (): The probability that a team scores between 60 and 79 points is approximately **0.0539** (5.39%).
* (): The probability that a team scores between 80 and 99 points is approximately **0.3555** (35.55%).
* The total probability (: The probability that a team scores within any of these ranges is approximately **0.4122** (41.22%).

**Definition 2.7/Theorem 2.2**

An ordered arrangement of r distinct objects is called a permutation. The number of ways of ordering n distinct objects taken r at a time will be designated by the symbol

**Q:** An NBA analyst is studying the performance of teams from the **1960–1970 seasons**. The analyst wants to analyze the top **5 teams** based on total points scored during this period and create different ordered rankings of these teams for a report.

 If there are teams being considered from the 1960–1970 period, how many different ways can the analyst rank the top teams?

 What would be the total number of rankings if the analyst decides to only rank the top teams instead?

**A:**

* If the analyst ranks the top teams from , there are **30,240 unique rankings** possible.
* If the analyst ranks the teams from , there are **720 unique rankings** possible.

**Definition 2.9**

The conditional probability of an event A, given that an event B has occurred

Q: An NBA analyst is studying data from games played up to and including the 2015 season. The analyst wants to calculate the conditional probability that:

1. A team wins a game given that they scored more than 110 points.

Let:

* Event AAA: The team wins the game.
* Event BBB: The team scores more than 110 points.

**A:**

* : The probability that a team scores more than 110 points is approximately 29.24%.
* : The probability that a team wins and scores more than 110 points is approximately 21.18%.
* : The conditional probability that a team wins, given that they scored more than 110 points, is approximately 72.44%.

**Definition 2.10**

Two events A and B are said to be independent if any one of the following holds:

**Q:** An NBA analyst is studying games from the **1990–2000 seasons** to determine if two events are independent:

1. **Event A**: A team scores more than 100 points in a game.
2. **Event B**: The same team wins the game.

**A:**

* Since scoring more than 100 points and winning are **not independent**.
* Similarly, and , confirming that the two events are **dependent**.

**Definition 2.11**

For some positive integer k, let the sets be such that

**Q:** An NBA analyst is studying the scoring patterns of teams in the **2000–2010 seasons**. The analyst decides to divide all games into **three mutually exclusive categories** based on the total points scored by the team:

1. : Games where the team scores less than 80 points.
2. ​: Games where the team scores between 80 and 100 points (inclusive).
3. ​: Games where the team scores more than 100 points.

Answer the following:

1. Verify that ​ and that ​ are mutually exclusive.
2. If the total number of games in the dataset is NNN, calculate the proportion of games in each category .
3. If a random game is selected, what is the probability that the team scored more than 100 points

**A:**

1. **Verification of Partition**: The subsets form a valid partition of the sample space because their union equals 100% of the games, and they are mutually exclusive.
2. **Proportion of Games**:

* Scoring less than 80 points : **7.22%**
* Scoring between 80 and 100 points: **54.42%**
* Scoring more than 100 points: **38.36%**

1. **Probability of Scoring Over 100 Points**: The probability that a team scores more than 100 points in a random game is **38.36%**.

**Definition 3.4**

Let Y be a discrete random variable with the probability function p(y)

**Q:** An NBA analyst is evaluating the performance of teams during the **2010–2015 seasons**. Let represent the number of points scored by a team in a game. The analyst assigns probabilities to different point totals based on their frequency in the dataset.

Answer the following:

1. Calculate the expected number of points scored by a team during this time period.
2. Interpret the expected value in the context of average team performance.

**A:**

1. **Expected Value**  The expected number of points scored by a team in a game during the **2010–2015 seasons** is approximately **99.16 points**.
2. **Interpretation**: On average, teams scored about **99 points per game** during this time-period. This represents the typical offensive performance for teams during the 2010–2015 seasons based on the dataset.

**Definition 3.5**

If Y is a random variable with mean E(Y) = , the variance of a random variable Y is defined to be the expected value of (Y-

**Q:** An NBA analyst is studying the variability in team scoring during the **2005–2010 seasons**. Let YYY represent the points scored by a team in a single game.

1. Compute the variance of points scored by teams during this period.

**A:**

1. The **mean score** during the **2005–2010 seasons** is 98.72 points, and the variance is 153.48 points.

**Definition 3.7**

A random variable Y is said to have a binomial distribution based on n trials with success probability p if and only if

**Q:** An NBA analyst is examining team scoring patterns in the **2010–2015 seasons**. The analyst defines:

* n: The total number of games played by a team during this period.
* p: The proportion of games in which the team scored more than 100 points.
* y: The number of games in which the team scores more than 100 points.

1. What is the probability that a team scores more than 100 points in **exactly 10 games** during this period?

**A:** The probability is effectively **0**. This result occurs because the total number of games across all teams is very large, making the binomial probability for such a small y negligible.

**Definition 3.8**

A random variable Y is said to have a geometric probability distribution if and only if

Q: An NBA analyst is studying the distribution of games where a team scores over 120 points during the **1990–1995 seasons**. The analyst models this as a geometric distribution where:

* y: The number of games it takes for a team to score over 120 points for the first time.
* p: The probability of a team scoring over 120 points in a single game.
* : The probability of a team not scoring over 120 points in a single game.

1. Calculate the probability that a team scores over 120 points for the first time in their **5th game**.

**A:** There is a **6.78% chance** that a team will score over 120 points for the first time in their 5th game of the **1990–1995 seasons**

**Definition 3.9**

A random variable Y is said to have a negative binomial probability distribution if and only if

**Q:** An NBA analyst is studying the scoring trends of teams during the **2000–2005 seasons**. The analyst focuses on games where teams score **at least 110 points**.

* y: The number of games played until a team scores **at least 110 points** for the **3rd time**.
* : The number of successes (scoring at least 110 points).
* p: The probability of scoring **at least 110 points** in a single game (calculated from the dataset).
* : The probability of not scoring **at least 110 points** in a single game.

1. What is the probability that a team scores at least 110 points for the **3rd time in their 8th game** during the **2000–2005 seasons**?

**A:** There is a **1.95% chance** that a team will score at least 110 points for the 3rd time in their 8th game during the **2000–2005 seasons**

**Theorem 2.2/Definition 2.7**

An ordered arrangement of r distinct objects is called a permutation. The number of ways of ordering n distinct objects taken r at a time will be designated by the symbol

**Q:** During the **1995–2000 seasons**, an NBA analyst is evaluating the performance of the top-scoring teams. The analyst wants to rank the **top 5 scoring teams** in all possible orders using the dataset of all teams that scored **at least 100 points** in a game during this period.

* n: The total number of teams that scored at least 100 points.
* r=5: The number of teams to rank.

1. How many ways can the analyst rank the top 5 scoring teams during the **1995–2000 seasons**?

**A:** The number of ways to rank the top 5 scoring teams in all possible orders is approximately **6.69 × 10¹⁸**

**Theorem 2.3**

The number of ways of partitioning n distinct objects into k distinct groups

**Q:** During the **1995–1996 NBA season**, an analyst is interested in games where teams scored exactly **90, 95, or 100 points**. The analyst wants to calculate the number of unique ways to organize the scores for these games, considering:

1. : Total games where teams scored exactly 90, 95, or 100 points.
2. ​: Games where teams scored exactly 90 points.
3. ​: Games where teams scored exactly 95 points.
4. : Games where teams scored exactly 100 points.
5. How many unique ways can the scores be arranged for these games during the **1995–1996 season**?

**A:** There were 230 games where teams scored exactly:

* 65 games with 90 points.
* 67 games with 95 points.
* ​98 games with 100 points.

1. The total number of unique arrangements for these scores is approximately

**Theorem 2.4**

The number of unordered subsets of size r chosen (without replacement) from n available objects

**Q:** **Question:**

During the **2010–2015 NBA seasons**, an analyst is studying the distribution of games where teams scored **exactly 100 points**. Out of these games, the analyst wants to choose **5 games** to feature in an analysis, regardless of the order.

* n: Total games where teams scored exactly 100 points.
* r=5: The number of games to choose.

1. How many ways can the analyst select **5 games** from all games where teams scored exactly 100 points during the **2010–2015 NBA seasons**?

**A:** In the 2010–2015 seasons, there were 496 games where teams scored exactly 100 points. The number of ways to select and rank 5 of these games in combinations is **245,157,170,544**

**Theorem 2.5**

The Multiplicative Law of Probability: The probability of the intersection of two events A and B

**A:** Using the NBA dataset, consider the following scenario from the 2010–2015 seasons:

* **Event A**: A game where the home team scores more than 110 points.
* **Event B**: Among the games satisfying Event A, the away team also scores more than 100 points.

Using the **Multiplicative Law of Probability**, compute the probability that both Event A and Event B occur, .

**Q:**

* .**:** 16.69% (0.1669).
* **:** 69.74% (0.6974).
* **:**  11.64% (0.1164). ​​

**Theorem 2.6**

The Additive Law of Probability: The probability of the union of two events A and B

**Q:** An analyst is studying the relationship between home and away team outcomes during the seasons covered. Let:

* **Event A**: The home team wins a game.
* **Event B**: The away team wins a game.

1. Calculate the probability that either the home team or the away team wins a game (not considering ties).
2. Determine Probability of the home team winning.
3. Determine Probability of the away team winning.

**A:**

* P(A): **0.3010**.
* P(B): **0.1817**.

**Theorem 2.7**

If a is an event

**Q:** An analyst is interested in the probability of the home team **not winning** during the 2010 season.

* Event A: The home team wins a game.
* Event : The home team does not win (this includes ties).

1. Calculate the probability of the home team winning during the 2010 season.

**A:**

P (A): 0.5

P (): 0.5

These probabilities indicate an equal chance of winning or not winning for home teams in that season. ​​

**Theorem 2.8**

Assume that { is a partition of S (see definition 2.11) such that . Then for any event A

**Q:** An analyst is studying the relationship between team wins and their scoring performance during the 2015 season. The dataset contains games played by various NBA teams. Using the total dataset, let:

* A: A team scores more than 100 points in a game.
* ​: The home team wins the game.
* ​: The away team wins the game.

1. Determine the probability that a team scores more than 100 points in a game.

**A:** The total probability of scoring more than 100 points during the 2015 season is approximately **47.8%.**

**Theorem 2.9**

Assume that { if a partition of S (see definition 2.11) such that . Then

**Q:** An analyst is interested in understanding the likelihood of a team winning (home or away) given that they scored more than 100 points during the 2015 season.

Let:

* Event A: A team scores more than 100 points in a game.
* Event: The home team wins the game.
* Event: The away team wins the game.

What is the probability that the **home team wins** given that more than 100 points are scored in a game during the 2015 season?

**A:** This means the probabilities are evenly distributed between home and away wins when more than 100 points are scored.

**Theorem 3.2**

Let Y be a discrete random variable with probability function p(y) and g(Y) be a real-valued function of Y

**Q:** An analyst is studying the expected weighted scores of teams during the 2015 NBA season. Let:

* Y: The points scored by a team in a game.
* : The square of the points scored.

1. Calculate the expected value where , for all points scored during the 2015 season.

**A:** The expected value is approximately **10,147.84**

**Theorem 3.4**

Let Y be a discrete random variable with probability function p(y), g(Y) be a function of Y, and c be a constant

**Q:** An NBA analyst is analyzing the scoring patterns of teams in the 2010 season. They want to explore the effect of scaling the average points scored in games by a constant factor .

* g(Y) is the number of points scored by a team,
* represents the scaling factor,
* E[g(Y)] is the expected value of points scored, which can be calculated using the dataset.

1. Using the dataset, calculate E[cg(Y)] for c=1.5 and interpret the scaled expected value of points scored by teams in the 2010 season.

**A:** The expected value of is approximately 150.40150.40150.40

**Definition 3.10**

A random variable Y is said to have a hypergeometric probability distribution if and only if

**Q:** An NBA analyst is reviewing the dataset to study games where the home team scored more than 100 points during the **1995–2000 seasons**.

* Let N represent the total number of games played in this range.
* Let r represent the total number of games where the home team scored more than 100 points.
* The analyst randomly selects n games from the entire dataset for this time period.

1. Calculate the probability that exactly y of the selected n games involves home teams scoring more than 100 points.

**A:** The probability that exactly 3 out of 10 randomly selected games from the 1995–2000 NBA seasons involve the home team scoring more than 100 points is approximately **22.82%**.

**Definition 3.11**

A random variable Y is said to have a Poisson probability distribution if and only if

**Q:** An analyst is investigating the number of games where the home team scored exactly 120 points during the 2010 NBA season.

* y: The number of games in which the home team scored 120 points.
* λ: The average number of games per team in the 2010 season where the home team scored exactly 120 points.

1. Calculate λ based on the total number of games played and the number of times the home team scored 120 points.

2. Using λ, compute the probability p(y) that the home team scored 120 points in exactly y=5 games during the season.

What is the probability?

**A:**

* Number of games where the home team scored exactly 120 points in the 2010 season: 19 games.
* Total number of games in the 2010 season: 2,624 games.
* Average number of games with 120 points per game (λ): 0.0072

This means the likelihood of observing exactly 5 games where the home team scored 120 points in the 2010 season is

**Definition 3.15**

Let Y be an integer-valued random variable for which where The probability-generating function P (t) for Y is defined to be

**Q:** An NBA statistician is analyzing the scoring patterns of home teams in the 2005 season.

* Y represent the number of games in which the home team scored exactly t points, where t ranges over the different possible scores in that season.
* The probabilities correspond to the probabilities of the home team scoring 0, 1, 2, ..., points respectively.

1. Calculate the probabilities for scores t = 100,105, and 110.

2. Use the generating function to compute P(t) for a specific value of t = 2.

**A:**

* Probability of scoring 100 points: **0.0312**
* Probability of scoring 105 points: **0.0282**
* Probability of scoring 110 points: **0.0164**
* The result when t = 2: **2.24**

**Definition 3.16**

The kth factorial moment for a random variable Y is defined to be

**Q:** An NBA analyst is studying the scoring performance of teams during the 2010 season. Let Y represent the points scored by a team in a game. The analyst wants to compute the **k-th factorial moment** of Y, to understand the variability in scoring performance. Calculate ​ for the points scored by teams during the 2010 season.

**A:** 10,094.5

**Definition 4.3**

Let F(y) be the distribution function for a continuous random variable Y. Then f (y), given by

Q: In the 1970 NBA season, an analyst is studying the distribution of points scored by teams in a game.

* F(y): represents the probability that a team's score is less than or equal to y.
* f(y): The probability density

1. Calculate f (100), the probability density of scoring exactly 100 points

**A:** The probability density function f(100), which approximates the likelihood of a team scoring exactly 100 points, is approximately **1.75%**

**Definition 4.5**

The expected value of a continuous random variable Y is

**Q:** An NBA analyst is studying the scoring patterns of teams during the 1995 season. Let Y represent the points scored by teams in a single game.

* Estimate the probability density function f(y)for the points scored by teams.
* Calculate the expected value E(Y), representing the average points scored by a team in the 1995 season

**A:** The expected value of points scored by teams during the 1995 season is approximately **101.28 points**. ​​

**Definition 4.6**

If < , a random variable Y is said to have a continuous uniform probability distribution on the interval ( ) if and only if the density function of Y is

**Q:** An analyst wants to analyze the uniform distribution of home team scores during the 1995 season. Assume the scores are uniformly distributed between the minimum home team score () and the maximum home team score (for that season.

* Calculate the probability density function f(y) for any score y within the range [].
* Determine f(y) for a specific score y=100.

A: The uniform distribution probability density at y=100 is 0.0108, considering the scores range from 63 to 156.

**Definition 4.8**

A random variable Y is said to have a normal probability distribution if and only if, for the density unction of Y is

**Q:** An NBA analyst is modeling the distribution of points scored by home teams during the 1995 season as a normal distribution. Let:

* μ: The mean points scored by home teams during the 1995 season.
* σ: The standard deviation of points scored by home teams during the 1995 season.
* y: The points scored by a home team in a specific game.

1. Calculate μ and σ from the dataset.
2. Using the calculated values, find the probability density f(y) for a game where the home team scored **120 points**.

**A:**

* The μ of the home team points (PTS) is approximately **101.28**.
* The σ of the home team points (PTS) is approximately **12.52**.
* The probability density function for a score of 120 points is approximately **0.01041**.

**Definition 4.9**

A random variable Y is said to have a gamma distribution with parameters α > 0 and β > 0 if and only if the density function of Y is

**Q:** An NBA analyst is studying the distribution of points scored by home teams in the 1998 season. The analyst models the points scored by home teams (Y)

1. Compute the mean μ and variance of the home team points scored in the 1998 season.
2. Using the gamma distribution properties, estimate α and β from the mean and variance.
3. Using the estimated α and β, calculate the probability density f(110), representing the likelihood of a home team scoring exactly 110 points.

**A:**

* **(μ):** 95.33
* **(σ²):** 140.46
* **Parameters for the Gamma distribution:**
  + **Alpha:** 64.70
  + **Beta:** 1.47
* f(110) = 0.0145

**Theorem 3.10**

If Y is a random variable with a hypergeometric distribution

**Q:** An NBA analyst is studying the scoring performance of home teams during the 1995–2000 seasons. Let:

* N: Total number of games played in this period.
* r: Total number of games where the home team scored at least 100 points.
* n: Total number of games sampled randomly without replacement.

1. The expected number of games in the sample where the home team scored at least 100 points.

2. The variance of this count.

**A:**

1. Mean: 20.76 (approximately)

2. Variance: 12.10 (approximately)

**Theorem 3.11**

If Y is a random variable possessing a Poisson distribution with parameter λ, then

**Q:** An NBA analyst is interested in studying the number of games where the home team scored over 120 points during the 1985 season. Let the random variable Y represent the number of such games.

Calculate:

1. The expected number of games.
2. The variance.

**A:**

1. The expected value (μ) is 462.

2. The variance (σ²) is also 462.

**Theorem 3.14**

Tchebysheff’s Theorem Let Y be a random variable with mean μ and finite variance σ2. Then, for any constant k > 0

**Q:** An analyst is examining the consistency of home team scores during the 1995 NBA season using Chebyshev's inequality.

* Y: The points scored by the home team in a game.
* μ: The mean points scored by home teams during the 1995 season.
* σ: The standard deviation of points scored by home teams during the 1995 season.
* k: A positive integer indicating the number of standard deviations from the mean.

1. Compute the mean and standard deviation of home team scores for the 1995 season from the dataset.

2. Calculate the probability.

**A:**

1. The mean of home team scores: 101.28 points.

2. The standard deviation of home team scores: 12.52 points.

3. Using Chebyshev's inequality with k=2, the probability that the home team scores fall within + or - 2σ of the mean is at least 75%.

**Theorem 4.3**

If the random variable Y has density function f (y) and a < b, then the probability that Y falls in the interval [a, b] is

**Q:** An NBA analyst is studying the distribution of points scored by home teams during the 1995 season. Let Y represent the points scored by a home team in a game. Using the probability density function f(y) derived from the dataset, calculate the probability that a home team scores between 90 and 110 points in a single game during the 1995 season.

**A:**

1. Total number of games: 2360
2. Number of games where the home team scored between 90 and 110 points: 1409
3. Probability that the home team scored between 90 and 110 points: 0.5970

**Theorem 4.4**

Let g(Y ) be a function of Y; then the expected value of g(Y ) is given by

**Q:** During the 1970 season, an NBA analyst wants to analyze the scoring patterns of home teams. Let Y represent the points scored by a home team in a game. The analyst defines g(Y)=, where g(Y) represents the squared score of the home team. Calculate the expected value of g(Y) for home teams during the 1970 season.

**A:** The expected value of the squared scores for the home teams during the 1970 season is approximately **13,387.92**

**Theorem 4.8**

If Y has a gamma distribution with parameters α and β, then

**Q:** An analyst is studying the scoring trends of NBA games during the 1980 season. Let Y represent the points scored by the home teams in this season. Assume that Y follows a Gamma distribution with shape parameter α and scale parameter β. Use the dataset to estimate α as the average number of wins per team in the 1980 season and β as the average points scored by a home team divided by the number of games played.

**A:**

1. (α): 43.18 (estimated as the average number of wins per team in the 1980 season).
2. (β): 0.0574 (calculated as the average points scored by home teams divided by the total number of games).