Show That The Variance Of A Random Variable Following A Geometric Distribution Is P 21p.

Show That The Variance Of A Random Variable Following A Geometric Distribution Is P 21p.

Understanding the variance of a random variable following a geometric distribution is fundamental in probability theory and statistics. This concept plays a vital role in modeling real-world phenomena such as the number of trials needed to achieve the first success in a sequence of Bernoulli trials. In this comprehensive guide, we will walk through the derivation of the variance for a geometric distribution, clarify the underlying assumptions, and explore its implications, all while employing a structured and SEO-optimized approach to facilitate better understanding and accessibility.

Introduction to the Geometric Distribution

Definition and Basic Concepts

The geometric distribution models the number of Bernoulli trials needed for the first success. It is one of the simplest discrete probability distributions and is widely used in fields like reliability testing, queuing theory, and survival analysis.

- Bernoulli Trials: Independent experiments with only two possible outcomes: success (with probability p) and failure (with probability 1 - p).
- Random Variable (X): Represents the trial count at which the first success occurs.

Probability Mass Function (PMF):

The probability that the first success occurs on the k-th trial is given by:

$$P(X = k) = (1 - p)^{k-1} p, \quad \text{for } k = 1, 2, 3, \ldots$$

where:

- p: probability of success on each trial.
- $(1 - p)^{k-1}$: probability of $k-1$ consecutive failures before the first success.

Mathematical Derivation of Variance for the Geometric Distribution

Expected Value $E[X]$

Before computing the variance, it's essential to understand the expected value of the geometric random variable.

The expectation for $X \sim \text{Geometric}(p)$ is:

$$E[X] = \frac{1}{p}$$

Derivation of $E[X]$:

1. Use the definition of expectation:

$$E[X] = \sum_{k=1}^{\infty} k \cdot P(X=k)$$

2. Substitute the PMF:

$$E[X] = \sum_{k=1}^{\infty} k (1 - p)^{k-1} p$$

3. Recognize that this sum is a standard form, where:

$$\sum_{k=1}^{\infty} k r^{k-1} = \frac{1}{(1 - r)^2} \quad \text{for } |r| < 1$$

4. Applying this with $r = 1 - p$:

$$E[X] = p \cdot \frac{1}{(1 - (1 - p))^2} = p \cdot \frac{1}{p^2} = \frac{1}{p}$$

Second Moment $E[X^2]$

The variance is calculated using the formula:

$$\operatorname{Var}(X) = E[X^2] - (E[X])^2$$

Therefore, we need to find $E[X^2]$.

Calculating $E[X^2]$:

1. Express as a sum:

$$E[X^2] = \sum_{k=1}^{\infty} k^2 P(X=k)$$

2. Substitute the PMF:

$$ E[X^2] = \sum_{k=1}^{\infty} k^2 (1 - p)^{k-1} p $$

3. Recognize a standard summation involving the second moment:

$$ \sum_{k=1}^{\infty} k^2 r^{k-1} = \frac{1 + r}{(1 - r)^3} \quad \text{for } |r| < 1 $$

4. Applying with $r = 1 - p$:

$$ E[X^2] = p \cdot \frac{1 + (1 - p)}{(p)^3} = p \cdot \frac{2 - p}{p^3} $$

5. Simplify:

$$ E[X^2] = \frac{2 - p}{p^2} $$

Note: The derivation of this sum involves generating functions or differentiation of the sum for $E[X]$, but for brevity, these standard results are employed.

Calculating Variance $\operatorname{Var}(X)$

Using the formulas:

$$ \operatorname{Var}(X) = E[X^2] - (E[X])^2 $$

and substituting the derived expectations:

1. $E[X] = \frac{1}{p}$

2. $E[X^2] = \frac{2 - p}{p^2}$

we get:

$$ \operatorname{Var}(X) = \frac{2 - p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{2 - p}{p^2} - \frac{1}{p^2} = \frac{(2 - p) - 1}{p^2} = \frac{1 - p}{p^2} $$

Final Result:

$$
\boxed{
\operatorname{Var}(X) = \frac{1 - p}{p^2}
}
$$

This formula indicates the variance of a geometric random variable with success probability p.

Understanding the Variance Formula $\frac{1 - p}{p^2}$

Implications of the Variance Expression

- Dependence on p: As p increases, the variance decreases, indicating less variability in the number of trials needed for success.
- Limit Behavior:

- When $p \to 1$:

$$\operatorname{Var}(X) \to 0$$

Because success is almost certain on the first trial, so variability diminishes.

- When $p \to 0$:

$$\operatorname{Var}(X) \to \infty$$

Since the number of trials can be very large with high variability.

- Practical Interpretation:

- The formula reflects the uncertainty or spread in the number of trials before the first success.

Applications in Real-world Scenarios

- Quality Control: Estimating the variability in the number of items produced before detecting a defect.
- Reliability Engineering: Assessing the variability in the number of cycles before failure.
- Marketing Campaigns: Modeling the number of attempts needed before achieving a successful conversion.

Additional Insights and Related Distribution Properties

Relation to Other Distributions

The geometric distribution is closely related to the negative binomial distribution, which generalizes the number of trials until the r-th success, with the geometric being the special case where $r=1$.

Key properties:

- The mean and variance change with the number of required successes.
- The geometric distribution is memoryless, which means:

$$P(X > s + t \mid X > s) = P(X > t)$$

for all $s, t \geq 0$.

Memoryless property simplifies the analysis of stochastic processes and is unique among discrete distributions.

Alternative Definitions of the Geometric Distribution

Note that some literature defines the geometric distribution as the number of failures before the first success, leading to different formulas for mean and variance:

- Number of failures before first success: $Y = X - 1$

- Mean: $E[Y] = \frac{1 - p}{p}$

- Variance: $\operatorname{Var}(Y) = \frac{1 - p}{p^2}$

The relationship between these variants highlights the importance of understanding the specific definition used.

Conclusion: Summary of Variance Derivation and Significance

The derivation of the variance of a geometric distribution, resulting in:

$$
\operatorname{Var}(X) = \frac{1 - p}{p^2}
$$

is a fundamental aspect of understanding the distribution's behavior. It emphasizes how the success probability p influences the spread or variability in the number of trials until the first success.

This variance formula is pivotal in statistical modeling and decision-making processes across numerous fields, including manufacturing, healthcare, and social sciences. Recognizing the variance's dependence on p allows practitioners to estimate risks, plan resources, and optimize systems effectively.

By mastering the derivation and interpretation of this variance, statisticians and analysts can better leverage the geometric distribution for practical applications, ensuring more accurate predictions and informed decisions.

Meta-Description:
Learn how to derive and understand the variance of a random variable following a geometric distribution, with detailed explanations and applications. Discover why the variance is $\frac{1 - p}{p^2}$ and its significance in probability and statistics.

Frequently Asked Questions

What is the variance of a geometric random variable with success probability p?

The variance of a geometric random variable with success probability p is given by $Var(X) = (1 - p) / p^2$.

How do you derive the variance of a geometric distribution?

The variance is derived using the properties of the geometric distribution, starting from its mean and variance formulas, often involving the expected value of the square of the random variable and the mean.

Why is the variance of a geometric distribution expressed as $(1 - p) / p^2$?

Because the geometric distribution models the number of trials until the first success, and its variance reflects the spread of this count, which mathematically simplifies to $(1 - p) / p^2$ based on its probability mass function and moments.

What does the formula $Var(X) = (1 - p) / p^2$ tell us about the variability of the distribution as p approaches 0?

As p approaches 0, the variance increases dramatically, indicating higher variability and greater dispersion in the number of trials needed for the first success.

Is the variance of a geometric distribution always larger than its mean?

Yes, since the mean is $1/p$ and the variance is $(1 - p) / p^2$, the variance is generally larger than the mean for p in (0,1), especially as p becomes small.

Can the variance formula for the geometric distribution be written as $P(1 - p) / p^2$?

No, the correct variance formula is $(1 - p) / p^2$; P is not used in this context unless referring to probability, but in the variance formula, it is simply p, the success probability.

How does the variance change when the success probability p increases?

As p increases, the variance decreases because $(1 - p)$ diminishes and the entire expression $(1 - p) / p^2$ becomes smaller, indicating less variability in the number of trials until the first success.

Additional Resources

Variance of a Geometric Distribution: An In-Depth Analysis

Understanding the behavior of probability distributions is fundamental to statistics and probability theory, especially when it comes to modeling real-world phenomena. One such distribution that is both foundational and widely applicable is the geometric distribution. Its variance, a measure of the spread or dispersion of the distribution, is a critical parameter for statisticians and data scientists alike. This article delves into the proof that the variance of a random variable following a geometric distribution is $\frac{1-p}{p^2}$, often expressed as $\frac{(1-p)}{p^2}$, or in some contexts, as $P 2 1 p$ (more precisely, $\frac{1-p}{p^2}$). We will unpack the mathematical derivation, provide contextual understanding, and explore implications for statistical modeling.

Introduction to the Geometric Distribution

The geometric distribution models the number of trials needed to achieve the first success in a sequence of independent Bernoulli trials, each with success probability p. It is one of the simplest discrete probability distributions, characterized by its "memoryless" property, meaning that the probability of success remains constant regardless of prior failures.

Key characteristics:

- Random variable X: Number of trials until the first success.
- Probability mass function (pmf):

$$
P(X = k) = (1 - p)^{k-1} p, \quad k = 1, 2, 3, \ldots
$$

- Parameters:
- p: Probability of success on each trial.
- $1 - p$: Probability of failure on each trial.

Expected Value of a Geometric Random Variable

Before tackling the variance, it is essential to understand the expected value or the mean, $E[X]$, of a geometric distribution.

Derivation of $E[X]$:

$$
E[X] = \sum_{k=1}^{\infty} k \cdot P(X = k) = \sum_{k=1}^{\infty} k (1 - p)^{k-1} p
$$

Recognizing the sum as a weighted geometric series, we can factor out p:

$$
E[X] = p \sum_{k=1}^{\infty} k (1 - p)^{k-1}
$$

Recall the sum:

$$
\sum_{k=1}^{\infty} k r^{k-1} = \frac{1}{(1 - r)^2}, \quad \text{for } |r| < 1
$$

Substituting $r = 1 - p$:

$$
E[X] = p \times \frac{1}{(1 - (1 - p))^2} = p \times \frac{1}{p^2} = \frac{1}{p}
$$

Result:

$$
\boxed{
E[X] = \frac{1}{p}
}
$$

This indicates that, on average, about $1/p$ trials are needed to achieve the first success.

Variance of a Geometric Distribution: The Goal

The variance $\operatorname{Var}(X)$, measures the spread of the distribution around its mean:

$$
\operatorname{Var}(X) = E[X^2] - (E[X])^2
$$

Thus, to compute the variance, we need to determine $E[X^2]$, the second moment of X.

Key steps:

1. Find $E[X^2]$.
2. Use the relation $\operatorname{Var}(X) = E[X^2] - (E[X])^2$.

Calculating $E[X^2]$: The Second Moment

The second moment is:

$$
E[X^2] = \sum_{k=1}^{\infty} k^2 \cdot P(X = k) = p \sum_{k=1}^{\infty} k^2 (1 - p)^{k-1}
$$

The sum:

$$
S = \sum_{k=1}^{\infty} k^2 r^{k-1}
$$

for $|r| < 1$, is a standard sum in series theory, and can be derived or looked up as:

$$
\sum_{k=1}^{\infty} k^2 r^{k-1} = \frac{1 + r}{(1 - r)^3}
$$

Proof Sketch:

- The sum of $k r^{k}$ is $\frac{r}{(1 - r)^2}$.
- Differentiating sum formulas and algebraically manipulating leads to the sum involving k^2.

Applying this to our case with $r = 1 - p$:

$$
E[X^2] = p \times \frac{1 + (1 - p)}{(1 - (1 - p))^3} = p \times \frac{2 - p}{p^3}
$$

Simplify numerator and denominator:

$$
E[X^2] = p \times \frac{2 - p}{p^3} = \frac{2 - p}{p^2}
$$

Deriving the Variance

Now, with $E[X] = \frac{1}{p}$ and $E[X^2] = \frac{2 - p}{p^2}$, the variance becomes:

$$
\operatorname{Var}(X) = E[X^2] - (E[X])^2 = \frac{2 - p}{p^2} - \left(\frac{1}{p}\right)^2
$$

Simplify:

$$
\operatorname{Var}(X) = \frac{2 - p}{p^2} - \frac{1}{p^2} = \frac{2 - p - 1}{p^2} = \frac{1 - p}{p^2}
$$

Final Result:

$$
\boxed{
\operatorname{Var}(X) = \frac{1 - p}{p^2}
}
$$

This formula confirms that the variance of a geometric distribution depends on the success probability p. As p approaches 0, the variance increases dramatically, reflecting greater uncertainty and spread in the number of trials needed for success.

Understanding the Result: Intuitive Insights

The derived formula, $\operatorname{Var}(X) = \frac{1 - p}{p^2}$, has several important implications:

- Inversely proportional to p^2: As the probability of success p increases, the variance decreases. When successes are more likely, the number of trials needed becomes more predictable.

- Dependence on $1 - p$: The numerator $(1 - p)$ reflects the failure probability. Higher failure rates lead to a larger spread in the number of trials until success.

- Real-world applications: For example, in quality control, if the probability of a defective item (failure) is high, the variance in the number of inspections until finding a non-defective item is also high.

Practical Examples and Applications

Example 1: Quality Control

Suppose a factory produces items with a 10% defect rate ($p = 0.9$). The variance in the number of items inspected until the first defect appears:

$$
\operatorname{Var}(X) = \frac{1 - 0.9}{(0.9)^2} = \frac{0.1}{0.81} \approx 0.123
$$

This low variance indicates that the number of inspections until the first defect is fairly predictable.

Example 2: Medical Testing

In a scenario where the probability of detecting a disease in a test is 0.2 ($p=0.2$), the variance in the number of tests until a positive detection:

$$
\operatorname{Var}(X) = \frac{1 - 0.2}{(0.2)^2} = \frac{0.8}{0.04} = 20
$$

A high variance suggests significant fluctuation in the number of tests needed for the first positive result.

Conclusion: Summary and Significance

The variance of a geometric distribution, $\operatorname{Var}(X) = \frac{1 - p}{p^2}$, is a fundamental result that encapsulates the distribution's variability. Its derivation hinges on the calculation of the second moment $E[X^2]$, which involves summing a weighted series. The result underscores the intuitive understanding that as the success probability p increases, the process becomes more predictable, and the variability diminishes. Conversely, low success probabilities lead to high variance, indicating greater uncertainty.

Understanding this variance is crucial for designing experiments, quality assurance, risk assessment, and many other fields where modeling the number of trials until an event occurs is essential. This derivation not only reinforces core principles of

[Show That The Variance Of A Random Variable Following A Geometric Distribution Is P 21p]()

Show That The Variance Of A Random Variable Following A Geometric Distribution Is P 21p

Related Articles

- A Typical Cell Phone Uses Wavelength Of 0.36meters. How Much Energy Does This Wave Have?
- Is The Income You Could Earn After Going To College Part Of The Cost Of Going To College?
- What Do The New Wave Films 400 Blows & Breathless Have In Common With Happy New Year?

Back to Home