Describe The Sampling Distribution Of P. Assume The Size Of The Population Is 30,000. N=800, P=0.2 ME

Describe The Sampling Distribution Of P. Assume The Size Of The Population Is 30,000. N=800, P=0.2 ME

Introduction to Sampling Distribution of P

Understanding the sampling distribution of the sample proportion, commonly denoted as P, is fundamental in statistical inference. When researchers draw multiple samples from a population, the sample proportion ($\hat{P}$) varies from one sample to another. The collection of these sample proportions forms the sampling distribution, which provides critical insight into the behavior of $\hat{P}$, especially in estimating the true population proportion P. This article explores the characteristics of the sampling distribution of P, considering a finite population of 30,000, with a sample size of 800, and a population proportion P of 0.2. We will examine how the distribution behaves, its mean, variance, and the implications for statistical inference.

Understanding the Population and Sample Parameters

Population Size and Characteristics

- Total Population (N): 30,000
- Population Proportion (P): 0.2

The population consists of 30,000 elements, of which approximately 20% (i.e., 6,000) are classified as "successes" for the attribute of interest. The remaining 24,000 elements are "failures."

Sample Size and Its Role

- Sample Size (n): 800

Sampling involves selecting 800 individuals from the population, ideally at random, to estimate the population proportion. The sample size impacts the variability and precision of the estimate.

Sampling Distribution of the Sample Proportion $\hat{P}$

Definition and Significance

The sampling distribution of $\hat{P}$ is the probability distribution of the sample proportion calculated from all possible samples of size n drawn from the population. It helps in understanding how $\hat{P}$ varies and in constructing confidence intervals or conducting hypothesis tests about P.

Assumptions for Approximate Normality

Before applying normal approximation methods, certain conditions should be satisfied:

1. N is large enough to justify sampling distribution approximations.
2. Sample size n is sufficiently large such that $np \geq 10$ and $n(1 - p) \geq 10$.
3. The sampling is random and either with or without replacement, with adjustment if sampling is without replacement.

Given N=30,000, n=800, and P=0.2, these conditions generally hold, but the finite population correction factor must be considered due to the finite population size.

Finite Population Correction and Its Impact

Finite Population Correction (FPC)

When sampling without replacement from a finite population, the variance of $\hat{P}$ is reduced compared to infinite population assumptions. The FPC adjusts the variance accordingly:

$$
\text{FPC} = \sqrt{\frac{N - n}{N - 1}}
$$

The corrected standard error (SE) of $\hat{P}$ is:

$$
SE_{corrected} = \sqrt{\frac{p(1 - p)}{n} \times \frac{N - n}{N - 1}}
$$

Calculating the FPC for Our Scenario

- N = 30,000
- n = 800
- p = 0.2

Compute the FPC:

$$
\frac{N - n}{N - 1} = \frac{30,000 - 800}{29,999} \approx
\frac{29,200}{29,999} \approx 0.973
$$

This indicates that the variance of $\hat{P}$ is approximately 97.3% of what it would be under infinite population assumptions.

Mean and Variance of the Sampling Distribution

Expected Value (Mean) of $\hat{P}$

The mean of the sampling distribution of $\hat{P}$ equals the population proportion:

$$
\mu_{\hat{P}} = P = 0.2
$$

This means that on average, the sample proportion will be close to the true population proportion.

Variance and Standard Error

The variance of $\hat{P}$, considering the finite population correction, is:

$$
\sigma^2_{\hat{P}} = \frac{p(1 - p)}{n} \times \frac{N - n}{N - 1}
$$

Plugging in the values:

$$
\sigma^2_{\hat{P}} = \frac{0.2 \times 0.8}{800} \times 0.973 \approx \frac{0.16}{800} \times 0.973 \approx 0.000194 \times 0.973 \approx 0.000189
$$

The standard error (SE):

$$
SE = \sqrt{0.000189} \approx 0.0137
$$

This indicates that the sample proportion typically varies by about 1.37 percentage points from the true proportion.

Shape of the Sampling Distribution

Approximate Normality

Because the sample size $n=800$ is large, and both np and $n(1-p)$ exceed 10, the Central Limit Theorem applies, and the sampling distribution of $\hat{P}$ is approximately normal.

- Shape: Symmetric, bell-shaped.
- Center: Mean at $P=0.2$.
- Spread: Determined by the standard error (~0.0137).

Implications of Normal Approximation

- Facilitates constructing confidence intervals.
- Simplifies hypothesis testing about P.
- Allows the use of z-scores for inference.

Practical Applications and Confidence Intervals

Constructing a Confidence Interval

Suppose we observe a sample proportion $\hat{P}$ from our sample. To estimate the true population proportion P with a confidence level (e.g., 95%), we would use:

$$
\text{CI} = \hat{P} \pm z_{\alpha/2} \times SE_{corrected}
$$

Where $z_{\alpha/2}$ is the z-value corresponding to the desired confidence level (for 95%, $z \approx 1.96$).

Example Calculation

Assuming the sample proportion $\hat{P} = 0.21$ (just as an example):

$$
\text{Margin of Error} = 1.96 \times 0.0137 \approx 0.0269
$$

Confidence interval:

$$
0.21 \pm 0.0269 \Rightarrow (0.1831, 0.2369)
$$

This interval suggests we are 95% confident that the true population proportion lies between approximately 18.3% and 23.7%.

Implications of Sample Size and Population Size

Effect of Increasing or Decreasing N and n

- Larger N: Reduces the effect of finite population correction, making the variance closer to that of an infinite population.
- Larger n: Decreases the standard error, leading to more precise estimates.
- Small N relative to n: Not applicable here, as the sample size is small compared to the population, but the FPC still applies.

Trade-offs in Sampling

Choosing the right sample size involves balancing cost, time, and the desired precision. Larger samples reduce variability but increase resource expenditure.

Conclusion

The sampling distribution of the sample proportion $\hat{P}$, given a population of 30,000 with a true proportion P of 0.2 and a sample size of 800, is approximately normal due to the large sample size. Its mean equals the population proportion (0.2), while its standard error is approximately 0.0137, accounting for the finite population correction factor. This distribution enables statisticians to make inferences about the population proportion with quantifiable confidence. Proper understanding of its shape, mean, variance, and the influence of the finite population is essential for accurate statistical analysis and decision-making.

Summary:

- The mean of the sampling distribution of $\hat{P}$ is 0.2.
- The standard error, considering the finite population correction, is about 0.0137.
- The distribution is approximately normal due to the large sample size.
- Finite population correction slightly reduces the variance.
- Confidence intervals can be constructed using the normal approximation, facilitating inference about the population proportion P.

This comprehensive understanding of the sampling distribution of P allows researchers and statisticians to draw reliable conclusions from sample data, ensuring effective and accurate statistical decision-making.

Frequently Asked Questions

What is the sampling distribution of p when sampling size N=800 from a population of 30,000 with P=0.2?

The sampling distribution of p is approximately normal with mean equal to the population proportion P=0.2 and standard deviation (standard error) calculated as sqrt[P(1-P)/N] = sqrt[0.2 0.8 / 800] ≈ 0.0141.

How do you calculate the standard deviation of the sampling distribution of p in this scenario?

The standard deviation (standard error) is calculated as sqrt[P(1-P)/N] = sqrt[0.2 0.8 / 800] ≈ 0.0141, which measures the variability of sample proportions around the true proportion P.

Is the sampling distribution of p approximately normal, and why?

Yes, because with a large sample size (N=800), the Central Limit Theorem ensures that the sampling distribution of p is approximately normal, especially since both np and n(1-p) are greater than 5.

What is the probability that a sample proportion p is greater than 0.25 in this sampling distribution?

First, calculate the z-score: z = (0.25 - 0.2) / 0.0141 ≈ 3.55. Using standard normal tables, the probability P(p > 0.25) ≈ 1 - 0.9998 = 0.0002.

How does the population size of 30,000 impact the sampling distribution of p for N=800?

Since the population is much larger than the sample size (N=800) and the sampling is assumed to be without replacement, the finite population correction factor is negligible, so the standard error remains approximately the same as in an infinite population.

Additional Resources

Sampling Distribution of p: An Expert Analysis

Understanding the sampling distribution of the sample proportion, denoted as p, is fundamental in statistics—particularly in inferential statistics where we aim to make predictions or generalizations about a population based on sample data. This article delves into the intricacies of the sampling distribution of p, especially in the context of a large population, and explores how the parameters, such as population size, sample size, and the true proportion, influence its behavior. Using a concrete example where the population size is 30,000, the sample size is 800, and the true population proportion p is 0.2, we will examine the theoretical foundations, assumptions, calculations, and practical implications of the sampling distribution of p.

Understanding the Sampling Distribution of p

At the heart of statistical inference lies the concept of a sampling distribution. It describes the probability distribution of a statistic—here, the sample proportion $\hat{p}$—obtained by repeatedly drawing samples of a fixed size from a population.

What Is the Sample Proportion $\hat{p}$?

The sample proportion, $\hat{p}$, is a statistic that estimates the true population proportion p. It is calculated as:

$\hat{p}$ = (Number of successes in the sample) / (Sample size n)

For example, if we survey 800 individuals and find that 160 of them exhibit a certain characteristic, then:

$\hat{p}$ = 160 / 800 = 0.2

This estimate varies from sample to sample, and understanding how it varies across samples is the core of the sampling distribution.

Why Is the Sampling Distribution of p Important?

Knowing the shape, center, and spread of the sampling distribution of p allows statisticians to:

- Construct confidence intervals for p.
- Conduct hypothesis tests about p.
- Assess the variability of $\hat{p}$ across different samples.

In essence, it provides the probabilistic foundation for making inferences about the population based on sample data.

Parameters and Assumptions in Our Scenario

Our specific scenario involves the following parameters:

- Population size (N): 30,000
- Sample size (n): 800
- Population proportion (p): 0.2

Before analyzing the sampling distribution, it’s crucial to understand the impact of these parameters and the assumptions underlying the distribution.

Population Size and Sampling Method

Given that the population size N=30,000 is considerably large relative to the sample size n=800, the sample can be considered a sampling without replacement but with negligible dependency effects. Typically, when N is large relative to n (specifically, when the sample is less than 5% of the population), the binomial distribution approximates the sampling distribution well, and the finite population correction (FPC) factor can be applied if needed.

Sample Size and Its Effect

A sample size of 800 is substantial, ensuring that the sampling distribution of p is approximately normal, thanks to the Central Limit Theorem, provided that certain conditions are met.

Population Proportion p=0.2

This indicates that, in the entire population, roughly 20% possess the characteristic of interest. Our goal is to understand how the sample proportion $\hat{p}$ behaves around this true proportion.

Theoretical Foundations of the Sampling Distribution of p

The behavior of $\hat{p}$ is governed by the properties of the binomial distribution, which models the number of successes in n independent Bernoulli trials with probability p.

Binomial Distribution and Approximation

- Binomial distribution: B(n, p)
- Sample proportion $\hat{p}$: $\hat{p} = X/n$, where X ~ Binomial(n, p)

The probability mass function (pmf) of X is:

$$P(X = k) = C(n, k)\, p^k (1 - p)^{n - k}$$

where C(n, k) is the binomial coefficient.

Since $\hat{p} = X/n$, the distribution of $\hat{p}$ is derived from the binomial.

Normal Approximation to the Distribution of $\hat{p}$

When n is large, the binomial distribution can be approximated by a normal distribution:

- Mean of $\hat{p}$: $\mu_{\hat{p}} = p$
- Standard deviation (standard error): $\sigma_{\hat{p}} = \sqrt{p(1 - p) / n}$

Provided that the following conditions hold, the normal approximation is valid:

- $np \geq 10$
- $n(1 - p) \geq 10$

In our case:

- $np = 800\ 0.2 = 160 \geq 10$
- $n(1 - p) = 800\ 0.8 = 640 \geq 10$

Thus, the distribution of $\hat{p}$ can be approximated as normal with mean 0.2 and standard deviation calculated below.

Calculating the Sampling Distribution

Parameters

Using the provided data, we now compute the key parameters of the sampling distribution of p̂.

Mean of the Sampling Distribution

The mean of p̂, denoted as μ_{p̂}, is equal to the true population proportion p:

$\mu_{\hat{p}} = p = 0.2$

This indicates that, on average, the sample proportion p̂ will be close to 0.2 across repeated samples.

Standard Error of p̂

The standard deviation of p̂, also known as the standard error (SE), measures the typical deviation of p̂ from p:

SE = sqrt[p(1 - p) / n]

Plugging in the values:

SE = sqrt[0.2 0.8 / 800] = sqrt[0.16 / 800] = sqrt[0.0002] ≈ 0.01414

This small standard error reflects the expected precision of p̂ as an estimator of p.

Finite Population Correction (FPC)

Since our population is finite, and the sample size constitutes a significant portion of the population (n/N ≈ 0.0267), the FPC can be applied to adjust the standard error:

FPC factor = sqrt((N - n) / (N - 1))

Calculating:

FPC = sqrt((30,000 - 800) / (30,000 - 1)) = sqrt(29,200 / 29,999) ≈ sqrt(0.9733) ≈ 0.9866

Adjusted standard error:

$SE_{adj} = SE\ FPC \approx 0.01414\ 0.9866 \approx 0.01394$

Given the small size of n relative to N, the correction slightly reduces the standard error, making the distribution more concentrated around the mean.

Shape and Properties of the Sampling Distribution

Based on the calculations above, the sampling distribution of $\hat{p}$ can be summarized as follows:

- Shape: Approximately normal due to the large n, per the Central Limit Theorem.
- Center: At $p = 0.2$.
- Spread: Standard error ≈ 0.01394 with the finite population correction.

This distribution is symmetric and bell-shaped, centered at 0.2, with about 95% of the sample proportions lying within roughly ±2 standard errors:

- 95% confidence interval (approximate):

$\hat{p} \in [\ p - 2\ SE_{adj},\ p + 2\ SE_{adj}\ ] = [\ 0.2 - 2\ 0.01394,\ 0.2 + 2\ 0.01394\ ] \approx [0.1721, 0.2279]$

This interval indicates the expected range of $\hat{p}$ values obtained from repeated samples of size 800, assuming the true p is 0.2.

Implications for Statistical Inference

Understanding the sampling distribution allows practitioners to perform various inference tasks with confidence.

Constructing Confidence Intervals

Using the normal approximation, a 95% confidence interval for the true proportion p can be constructed as:

$\hat{p} \pm z_{0.025} SE_{adj} = \hat{p} \pm 1.96 \; 0.01394$

If in a particular sample, $\hat{p}$ is observed to be, say, 0.195, then:

CI: $0.195 \pm 1.96 \; 0.01394 \approx 0.195 \pm 0.0273 = [0.1677, 0.2223]$

This interval suggests that, with 95% confidence, the true p lies within this range.

Hypothesis Testing

Suppose we want to test:

- Null hypothesis: $p = 0.2$
- Alternative hypothesis: $p \neq 0.2$

Given a sample $\hat{p}$, we calculate the test statistic:

$Z = (\hat{p} - p_0) / SE_{adj}$

where $p_0 = 0.2$

If the computed Z exceeds critical values (±1.96 for a 5% significance level), we reject the null hypothesis.

Practical Consider

[Describe The Sampling Distribution Of P Assume The Size Of The Population Is 30 000 N 800 P 0 2 Me]

Describe The Sampling Distribution Of P Assume The Size Of The Population Is 30 000 N 800 P 0 2 Me

Related Articles

- [California Sea Lions Can Swim As Fast As 40.0 Km/h. Suppose A Sea Lion Begins To Chase A Fish At This]
- [Below Are Descriptions Of Internal Control Problems. In The Space To The Left Of Each Item, Enter The]
- [Boundary Work Is Work Done To Maintain The Border Between Legitimate And Nonlegitimate Science Within]

[Back to Home]