**Anexo 4 – Determinación de tamaño de la muestra**

La determinación del tamaño de muestra es un elemento que permite asegurar que los resultados obtenidos sean representativos y que puedan inferirse con un alto grado de precisión. Se pretende que con la muestra obtenida minimizar el error de muestreo y garantizar la validez estadística de las visitas a ser realizadas.

El cálculo del tamaño de muestra necesario para las visitas a ser realizadas en el marco del convenio con el Fondo Todos Somos Pazcifico, se empleará una fórmula específica que contempla el caso de una población finita con reemplazo.

La utilización de la siguiente metodología busca asegurar la fiabilidad y validez de los datos recopilados en las visitas a ser realizadas, así como para garantizar que las conclusiones derivadas sean correctas.

Para la determinación de esta muestra se usa la formula asumiendo que la muestra se comporta como una normal estándar. El empleo de la distribución normal estándar en la determinación del tamaño de muestra debido a que permite la estimación de intervalos de confianza y la realización de pruebas de hipótesis de una manera más efectiva debido a lo siguiente:

**1. Propiedades de la Distribución Normal**

La distribución normal estándar, caracterizada por una media de cero y una desviación estándar de uno, se usa debido a sus propiedades matemáticas precisas y simétricas. Su forma de campana facilita el cálculo de probabilidades y la determinación de áreas bajo la curva, lo anterior permitiendo el establecimiento de intervalos de confianza y la ejecución de pruebas de hipótesis.

**2. Aplicación del Teorema Central del Límite**

Según el Teorema Central del Límite, la distribución de las medias muestrales se aproxima a una distribución normal a medida que el tamaño de la muestra aumenta, independientemente de la distribución original de la población. Esta propiedad permite la utilización de la distribución normal estándar para realizar inferencias estadísticas a partir de muestras de tamaño considerable, lo que facilita la estimación de parámetros poblacionales.

**3. Estandarización y Facilidad de Cálculo**

La normalización de datos mediante su conversión a la distribución normal estándar simplifica los cálculos estadísticos y permite la comparabilidad entre distintos conjuntos de datos. Los valores críticos asociados a la distribución normal estándar, como aquellos correspondientes a niveles de confianza habituales son de conocimiento general, lo cual estandariza el proceso de determinación de intervalos de confianza y facilita la comunicación de resultados de manera clara y consistente.

**4. Precisión en la Estimación de Parámetros**

La adopción de la distribución normal estándar posibilita una estimación precisa de los parámetros poblacionales, especialmente en el caso de muestras grandes. Esto se debe a que la estructura de la distribución normal permite cuantificar de manera precisa la incertidumbre asociada a las estimaciones, facilitando así la determinación de márgenes de error exactos y confiables.

La fórmula de la normal para la determinación del tamaño de la muestra a ser usada es la siguiente:

$$n=\frac{Nσ^{2}Z\_{α}^{2}}{e^{2}\left(N-1\right)+σ^{2}Z\_{α}^{2}}$$

Donde:

$N$**:** Representa el tamaño de la población total, es decir, el número total de elementos en la población de interés.

$σ^{2}$**:** Corresponde a la varianza de la población, que mide la dispersión de los datos en la población.

$Z\_{α}$**:** Es el valor crítico de la distribución normal estándar asociado con el nivel de confianza deseado.

$e$**:** Representa el margen de error aceptable, que es la tolerancia máxima para la diferencia entre la media estimada de la muestra y la media verdadera de la población.

El cociente entre el numerador y el denominador ajusta el tamaño de muestra requerido para asegurar que la muestra sea representativa de la población con un nivel de precisión y confianza definidos.

Ahora bien, en virtud del convenio de cooperación 80905-026-2019 con el Fondo Todos Somos Pazcífico, y con el objetivo de determinar la muestra que me permita definir que las instalaciones derivadas del convenio se encuentran correctas se determinan las variables de la siguiente manera:

$N$**:** La población es **TRES MIL SEISCIENTOS DIEZ** **(3.610)** instalaciones.

$σ^{2}$**:** Debido a que se asume normal estándar, esta será de 1

$Z\_{α}$**:** Este corresponderá a un nivel de confianza de 95% siendo entonces $Z\_{α}$=1,64

$e$**:** Este margen será del 10%.

Para lo cual el valor de la muestra es de **DOSCIENTOS CINCUENTA Y DOS (252)** **de instalaciones** a ser visitados, este número considera reemplazo. Con lo anterior, el muestreo debe ser al menos 252 instalaciones.

# Bibliografía

Triola, M. F. (2009). *Estadistica.* Mexico: Pearson, Educación.