

# Analysis of Visits to the Natural History Museum of Ricardo Palma University Using the T Distribution

María Elizabeth Puelles Bulnes, Mg<sup>1</sup>, Luis Fernando Palma Yancachajlla<sup>2</sup>, Joey Gabriel Canales Huamani<sup>3</sup>, Paula Stella Sandoval Lingán<sup>4</sup>, Juan Sebastián Fleitas Marusic<sup>5</sup>

<sup>1,2,3,4,5</sup>Ricardo Palma University, Perú, [maria.puellesb@urp.edu.pe](mailto:maria.puellesb@urp.edu.pe), [202112392@urp.edu.pe](mailto:202112392@urp.edu.pe), [202111285@urp.edu.pe](mailto:202111285@urp.edu.pe), [202110436@urp.edu.pe](mailto:202110436@urp.edu.pe), [202111302@urp.edu.pe](mailto:202111302@urp.edu.pe)

**Abstract**– This case study was conducted with undergraduate students at Ricardo Palma University, Lima-Peru, in the Design of Experiments course between September and October 2023. The didactic methodology employed allowed the application of theoretical statistical knowledge to real-life situations, facilitating the development of practical and analytical skills. The study's objective is to evaluate whether there are significant differences in the number of visits by student/child audiences to the Natural History Museum of Ricardo Palma University (MHNURP) during the years 2018 and 2019. Statistical inference was performed using the t-distribution for independent samples, along with normality and variance equality tests to ensure the reliability of the results obtained. Additionally, SPSS and MINITAB were used for their respective advantages and features. It was demonstrated that there was no significant difference at the 95% confidence level in visits to MHNURP by the student/child audience between 2018 and 2019. It is important to note that this analysis was based on the available samples and the variables considered. Other relevant factors and variables could influence visits to MHNURP and should be considered in future research. It is recommended to implement various strategies to attract a greater number of visits from the studied audience, such as interactive educational programs, attractive exhibitions, special events, discounts, and collaborations with

other institutions. Furthermore, exploring other potential actions that could benefit the promotion of MHNURP programming is suggested.

**Keywords:** Natural History Museum, Ricardo Palma University, T Distribution, Student/Child Audience, Statistical Inference.

# Análisis de las visitas en el Museo de Historia Natural de la Universidad Ricardo Palma mediante la distribución T

María Elizabeth Puelles Bulnes, Mg<sup>1</sup>, Luis Fernando Palma Yancachajlla<sup>2</sup>, Joey Gabriel Canales Huamani<sup>3</sup>, Paula Stella Sandoval Lingán<sup>4</sup>, Juan Sebastián Fleitas Marusic<sup>5</sup>

<sup>1,2,3,4,5</sup>Universidad Ricardo Palma, Perú, [maria.puellesb@urp.edu.pe](mailto:maria.puellesb@urp.edu.pe), [202112392@urp.edu.pe](mailto:202112392@urp.edu.pe), [202111285@urp.edu.pe](mailto:202111285@urp.edu.pe), [202110436@urp.edu.pe](mailto:202110436@urp.edu.pe), [202111302@urp.edu.pe](mailto:202111302@urp.edu.pe)

**RESUMEN**– *El presente estudio de caso se realizó con los alumnos de graduación de la Universidad Ricardo Palma, Lima-Perú, en el curso de Diseño de Experimentos entre los meses septiembre y octubre del 2023. La metodología didáctica empleada permitió aplicar los conocimientos teóricos estadísticos a situaciones reales para desarrollar habilidades prácticas y analíticas. El objetivo del estudio es evaluar si existen diferencias significativas en la cantidad de visitas del público de estudiantes/niños al Museo de Historia Natural de la Universidad Ricardo Palma (MHNURP) durante los años 2018 y 2019. Se realizó la inferencia estadística utilizando la distribución t para muestras independientes, junto con las pruebas de normalidad e igualdad de varianzas para garantizar la confiabilidad de los resultados obtenidos. Además, se utilizaron SPSS y MINITAB por sus ventajas y facilidades de cada software. Se demostró que no hubo diferencia significativa al 95% en las visitas al MHNURP en el público de estudiantes/niños para los años 2018 y 2019. Es importante tener en cuenta que este análisis se basó en las muestras disponibles y en las variables consideradas. Otros factores y variables relevantes podrían influir en las visitas al MHNURP y deberían ser considerados en futuras investigaciones. Se recomienda implementar diversas estrategias para atraer mayor cantidad de visitas del público estudiado, como programas educativos interactivos, exposiciones atractivas, eventos especiales, descuentos, y colaboraciones con otras instituciones. Además, se sugiere explorar otras posibles acciones que pueden ser beneficiosas para promover la programación del MHNURP.*

**Palabras clave:** Inferencia estadística, distribución t, prueba de normalidad, prueba de igualdad de varianzas, grados de libertad, estudio de caso, MHNURP.

## I. INTRODUCCIÓN

El Museo de Historia Natural de la Universidad Ricardo Palma (MHNURP) también conocido como Museo Vera Alleman Haeghebaert, fue creado el 27 de noviembre de 1991 con una colección paleontológica recolectada por alumnos de la facultad de Biología. A través de convenios cooperativos

con otras instituciones, las colecciones se incrementaron en Zoología y Botánica, donde se mantienen importantes colecciones biológicas. En 1997 se abrió la sala de exposiciones permanentes, donde se muestran los distintos grupos zoológicos. Luego en 2012 se firmó un convenio con el Museo Paleontológico Meyer-Hönninger para implementar la sala de exposiciones en su sección de paleontología, con una exposición referente a la evolución de la vida en nuestro planeta, con fósiles desde los 1600 millones de años hasta los últimos 10 mil años [1]. El MHNURP cuenta actualmente con dos salas: una destinada a investigadores científicos de la biodiversidad, y otra sala de exhibición que presenta muestras paleontológicas, zoológicas, mundo marino y botánicas principalmente peruanas entre otras. Además, la diversidad de salas que ofrece el MHNURP es una oportunidad para que los visitantes puedan explorar y aprender sobre diferentes aspectos de la biodiversidad del Perú [2]. Por lo tanto, es importante que el museo tenga una estrategia clara para atraer y mantener la atención del público en cada una de sus salas.

En el marco de la Visión y Misión de la Universidad Ricardo Palma (URP), se reconoce la importancia de formar a los alumnos con un enfoque científico que les permita abordar los problemas del desarrollo nacional en los sectores productivos y sociales. Esto implica fomentar una formación integral y profesional que promueva habilidades y competencias necesarias para la resolución de problemas, la generación de soluciones innovadoras y la contribución al progreso del país [1]. Con el fin de cumplir este objetivo y como primera parte del curso de Diseño de Experimentos de la carrera de Ingeniería Industrial, se aplicó la metodología de aprendizaje basado en casos para analizar la participación de los visitantes en la MHNURP. Para ello, se contó con la participación de los alumnos que obtuvieron un rendimiento académico destacado en la primera evaluación llamada PRT1 y que expresaron su interés en formar parte de este estudio.

Esta metodología permite a los estudiantes aplicar los conocimientos teóricos a situaciones reales, los cuales les brinda la oportunidad de desarrollar habilidades prácticas y analíticas, y comprender cómo se aplican los conceptos de estadística en el mundo real. Los estudiantes deben analizar la información disponible, identificar los problemas clave, evaluar diferentes opciones y tomar decisiones fundamentadas. Esto promueve el pensamiento crítico, el razonamiento lógico y reflexión sobre problemas y ayuda a

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).  
**ISSN, ISBN:** (to be inserted by LACCEI).  
**DO NOT REMOVE**

los estudiantes a desarrollar habilidades de pensamiento, análisis y síntesis de nivel superior preparándolo para desafíos futuros y encontrar soluciones innovadoras [3]-[8].

El objetivo del presente trabajo es evaluar si existen diferencias significativas en la participación de estudiantes/niños en las visitas al MHNURP durante los años 2018 y 2019. Este análisis puede ser útil para identificar si ha habido cambios significativos en la cantidad o el nivel de participación de estos grupos en diferentes períodos y evaluar el impacto de ciertas iniciativas o cambios para garantizar su sostenibilidad a largo plazo. Se realizó la Inferencia Estadística usando la Distribución t para muestras pequeñas e independientes, así como otras pruebas estadísticas necesarias para asegurar la confiabilidad de los resultados obtenidos.

El estudio fue realizado en el curso de Diseño de Experimentos, que es una herramienta clave para el ingeniero industrial y otras áreas, ya que se utiliza para la evaluación de nuevos productos, procesos o servicios, control de calidad y optimización de procesos.

## II. MATERIAL Y MÉTODOS

### 2.1 Tratamiento de datos

Los datos utilizados en este estudio se obtuvieron de [9], en lo cual se centra en la elaboración del Plan Estratégico del MHNURP para el período 2023-2025. Estos datos se basan en los informes anuales presentados por el museo sobre el registro de visitas al MHNURP en los años 2017, 2018 y 2019, considerando un escenario similar o parecido en los años 2020, 2021 y 2022 a pesar de que estuvo cerrado por la pandemia.

Estos informes proporcionan información detallada sobre la cantidad de visitantes mensualmente que acudieron al museo en cada año, permitiendo así analizar la participación del público a lo largo de ese período. En la Tabla 1, se observan las cantidades de visitas registradas en los estudiantes/niños y adultos durante los meses de enero a diciembre del 2017 al 2019. Es importante destacar que, a pesar del cierre del museo debido a la pandemia, se considera que los años 2020, 2021 y 2022 presentaron un escenario similar en cuanto a la cantidad de visitas.

En este estudio, se optó por enfocarse en el público estudiantes/niños en los años de 2018 y 2019 en lugar de los adultos. Esta elección se basó en varios factores que se tuvieron en cuenta:

1. El MHNURP cuenta con exposiciones y programas educativos dirigidos desde el 1er a 6to grado, dividido en módulos, ver Tabla N°2, donde se puede apreciar la Exposición “Los Dinosaurios: Evolución de la vida en nuestro planeta” y programación de visitas tanto para colegios o paquetes sugeridos, que se elaboraron siguiendo el plan curricular vigente del Ministerio de Educación.

Tabla 1: Datos de las visitas al MHNURP en los años 2017 al 2019 extraídos [2].

Tipo de públicos del MHNURP distribuido mensualmente 2019		
2019		
Estudiantes / niños	adultos	Número de visitantes
Enero	140	120
Febrero	166	241
Marzo	101	77
Abril	832	406
Mayo	650	148
Junio	1108	430
Julio	591	238
Agosto	726	163
Setiembre	1016	329
Octubre	704	236
Noviembre	139	60
Diciembre	107	44
Total	6280	2492

Tipo de públicos del MHNURP distribuido mensualmente 2018		
2018		
Estudiantes / niños	adultos	Número visitantes
Enero	128	39
Febrero	114	71
Marzo	25	28
Abril	532	57
Mayo	633	188
Junio	874	157
Julio	347	124
Agosto	781	413
Setiembre	780	107
Octubre	599	53
Noviembre	435	197
Diciembre	41	53
Total	5289	1487

Tipo de públicos del MHNURP distribuido mensualmente 2017		
2017		
Estudiantes / niños	adultos	Número visitantes
Enero	104	95
Febrero	130	156
Marzo	35	164
Abril	186	245
Mayo	59	332
Junio	120	456
Julio	87	339
Agosto	90	442
Setiembre	110	341
Octubre	108	307
Noviembre	0	25
Diciembre	0	12
Total	1041	2902

A través de las programaciones les permitirá comprender mejor los temas tratados en clase de los cursos de Ciencia y Ambiente y con la facilidad de optar también por un recorrido por todas las áreas del museo en una sola visita, tanto para los estudiantes de primaria o secundaria, ver Figura 1.



Figura 1. Visitas y charlas a los estudiantes y niños [2].

De esta forma, fomentando el aprendizaje y la curiosidad científica en este público de estudiantes/niños por el tipo de actividad que se realiza [2].

2. Al trabajar con este público, se pueden recopilar datos más específicos sobre su participación, como la frecuencia de visitas, el grado de interacción con las exhibiciones y el nivel de satisfacción con las actividades propuestas. Esto permitirá al MHNURP evaluar la efectividad de sus

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).  
**ISSN, ISBN:** (to be inserted by LACCEI).  
**DO NOT REMOVE**

programas educativos y realizar ajustes según los resultados obtenidos, mejorando así la experiencia de los estudiantes y niños que visitan el museo.

- Al centrarse en los años más recientes de 2018 y 2019, se puede obtener una visión más precisa y actualizada de la participación de visitas al MHNURP, una vez que hay insuficiencia de datos para los años 2020, 2021, 2022 debido al cierre del local debido a la pandemia. Luego, los años 2018 y 2019 son los datos que se tiene de las visitas presenciales delante de las estrategias y programas implementados por el museo.

Tabla 2: Sala de Exposición MHNURP para niños [2].

<p><b>Exposición:</b>          'LOS DINOSAURIOS: EVOLUCIÓN DE LA VIDA EN NUESTRO PLANETA'.          Comprende: Fósiles, flora y fauna actual.</p> <p><b>Círculo</b>          Para colegios          Paquetes Sugeridos          Se sugiere a los profesores que su visita sea realizada de la siguiente manera:</p> <p><b>Módulo I</b>          Para niños menores de 6 años: sólo el área de paleontología y vídeos</p> <p><b>Módulo II</b>          De 1° a 6° grado de primaria: Microscopía, Zoología, Botánica, Paleontología, Camélidos y Ecología*          *Se enfatizarán los temas de acuerdo al grado de enseñanza según el plan curricular emitido por el Ministerio de Educación.</p> <p><b>Módulo III</b>          De 1° a 3° grado de secundaria: Botánica, Microscopía, Ecología, Astronomía (Próximamente).</p> <p><b>Módulo IV</b>          De 4° a 5° grado de secundaria (Zoología, Paleontología, Ecología, Camélidos).</p>
--

Por lo tanto, en la Tabla N°3 se presentan las visitas de los estudiantes/niños al MHNURP en los años 2018 y 2019 en lo cual se utilizó en el presente estudio.

Tabla 3: Muestras de Estudiantes/niños 2018-2019

mes	Estud_niños_2018	Estud_niños_2019
1	128	140
2	114	166
3	25	101
4	532	832
5	633	650
6	874	1108
7	347	591
8	781	726
9	780	1016
10	599	704
11	435	139
12	41	107

## 2.2 Inferencia Estadística y Distribución t

La Estadística se divide en dos partes principales: la estadística descriptiva y la Inferencia Estadística [10], [11]. En la estadística descriptiva se utilizan técnicas y métodos estadísticos para organizar, resumir y visualizar los datos.

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).  
**ISSN, ISBN:** (to be inserted by LACCEI).  
**DO NOT REMOVE**

Incluyen medidas de tendencia central, como la media y la mediana, así como medidas de dispersión, como la desviación estándar.

Además, se pueden utilizar gráficas y tablas para representar los datos de manera más comprensible. El objetivo es obtener una descripción clara y concisa de los datos recopilados. Por otro lado, la Inferencia Estadística, requiere herramientas que permiten validar que las conclusiones obtenidas a partir de las muestras también son válidas para toda la población [12]. La inferencia estadística según [10], [11] y [12] es el proceso de utilizar la información de una muestra para hacer inferencias acerca de la población de la cual se extrajo la muestra. La inferencia estadística, es de gran importancia en la toma de decisiones porque nos permite obtener conclusiones sobre una población utilizando información de una muestra representativa de esa población. A través de la inferencia estadística, podemos hacer generalizaciones y estimaciones sobre características o parámetros de interés en una población, como la media, la proporción o la desviación estándar. Entonces, la inferencia estadística nos ayuda a tomar decisiones basadas en evidencia y nos permite generalizar los resultados obtenidos en una muestra a toda una población, lo que resulta fundamental para la toma de decisiones acertadas.

La distribución t es un método utilizado en inferencia estadística para estimar la media de una población a partir de una muestra pequeña. También conocida como la distribución de t-Student, fue descubierta por el estadístico inglés William Sealy Gosset en 1908. Gosset trabajaba en la cervecería Guinness en Dublín, Irlanda, y su tarea era analizar los datos de calidad de la cerveza para mejorar su producción. Durante su análisis, Gosset se encontró con un desafío: los datos que estaba estudiando no seguían una distribución normal y la desviación estándar poblacional era desconocida. Esto significaba que no podía utilizar las herramientas estadísticas existentes para realizar inferencias precisas sobre los datos, como solución, Gosset desarrolló la distribución t, que le permitió analizar los datos de calidad de la cerveza y realizar inferencias precisas sobre la media poblacional, incluso con muestras pequeñas. La distribución t se convirtió en una herramienta fundamental en la estadística para casos en los que no se conocen las características de la población y se trabaja con muestras limitadas [13], [14].

Gracias a la distribución t, Gosset pudo proporcionar a la cervecería Guinness valiosos conocimientos estadísticos que ayudaron a mejorar la calidad de su cerveza. Desde entonces, la distribución t, ha sido ampliamente utilizada en diferentes campos de investigación y análisis estadístico.

En el campo de la probabilidad y estadística, la distribución t, es una función de densidad de probabilidad que se utiliza para estimar la media de una población normalmente distribuida cuando el tamaño de la muestra es pequeño. Además, esta distribución también se utiliza en la teoría de pequeñas muestras o teoría exacta del muestreo, que se aplica tanto a muestras pequeñas como grandes. Para

comprender la distribución t, es necesario familiarizarse con el concepto de “grados de libertad” y que son una medida estadística que refleja el número de valores que pueden variar libremente en una muestra. En otras palabras, los grados de libertad indican cuántos valores pueden ser elegidos arbitrariamente sin afectar la validez de la muestra. Para definir los grados de libertad, es necesario hacer referencia a la varianza muestral que es una medida estadística que se utiliza para estimar la varianza de una población a partir de una muestra.

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (1)$$

La fórmula (1) es utilizada para calcular la varianza y los grados de libertad se encuentran en el denominador y equivale al número de observaciones menos 1. Esta terminología se debe a que, la varianza  $s^2$  está basada en  $n$  cantidades y  $x_1 - \bar{x}, x_2 - \bar{x}, x_3 - \bar{x}, \dots, x_n - \bar{x}$ , estas suman cero. Por lo tanto, especificar los valores de cualquier  $n - 1$  de estas cantidades puede determinar el valor restante [10] y [13].

La distribución t, se utiliza para estimar la media de una población cuando la desviación estándar poblacional es desconocida y se utiliza la desviación estándar muestral en su lugar y es especialmente útil cuando el tamaño de la muestra es pequeño ( $n < 30$ ). La distribución t, se puede aplicar para una muestra y dos muestras. En el caso de dos muestras, pueden ser muestras dependientes donde el grado de libertad es  $n - 1$  y muestras independientes donde el grado de libertad es  $n_1 + n_2 - 1$ ,  $n_i$  tamaño de la muestra  $i$ . El tratamiento para ambas muestras es diferente. En el caso de muestras independientes, es necesario probar la normalidad de los datos y la igualdad de varianza para que los resultados de la inferencia estadística sean confiables [12], [14] y [15].

El intervalo de confianza también se utiliza como una medida de la precisión de una estimación estadística. Es utilizada para estimar el rango dentro del cual se espera que se encuentre un parámetro poblacional desconocido, basándose en los datos de una muestra. Esta se interpreta como un rango de valores que contiene el verdadero valor del parámetro con un cierto nivel de confianza. Por ejemplo, un intervalo de confianza del 95% significa que, si se repitiera el experimento muchas veces, el 95% de los intervalos de confianza calculados contendrían el verdadero valor del parámetro.

La función de densidad [15] y [16] de una distribución t se define de la siguiente manera:

$$h_k(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}} \left(1 + \frac{t^2}{k}\right)^{-\frac{(k+1)}{2}}, \infty < t < \infty, k \in Z^+$$

y

$$\Gamma(p) = \int_0^{\infty} e^{-x} x^{p-1} dx \quad (2)$$

Donde:

$\Gamma$  representa la función gamma.

$k$  es el número de grados de libertad.

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).

**ISSN, ISBN:** (to be inserted by LACCEI).

**DO NOT REMOVE**

$\mu$  es la media de la distribución.

$\sigma$  es la desviación estándar de la distribución.

La función de densidad (2), describe la probabilidad de obtener un determinado valor para la variable aleatoria en una distribución t con  $k$  grados de libertad y la gráfica de esta función de densidad es simétrica, respecto del eje de ordenadas, con  $k$  grados de libertad. Se observa que la Figura 2, algo semejante a la distribución normal.

Se puede observar también que el valor máximo de la ordenada alcanza en  $\mu = 0$ . Esto significa que los valores de la variable aleatoria se distribuyen de manera equitativa a ambos lados del eje central.

Su valor medio y varianza son:

$$E(T) = \mu = \int_{-\infty}^{\infty} t \cdot h_k(t) dt = \int_{-\infty}^{\infty} t \cdot \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}} \left(1 + \frac{t^2}{k}\right)^{-\frac{(k+1)}{2}} dt = 0 \quad (3)$$

Si  $k > 3$

$$Var(T) = \sigma^2 = E \quad (4)$$

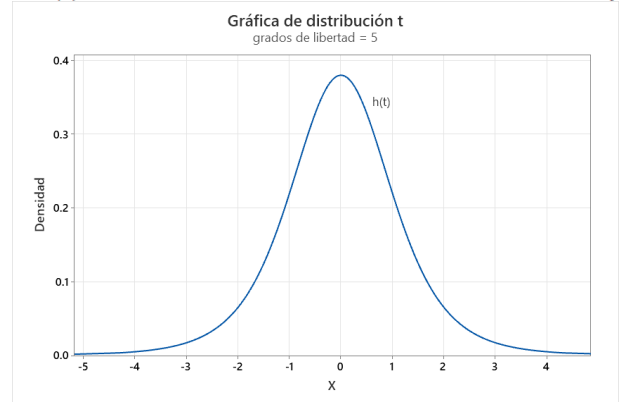


Figura 2. Distribución T, grados libertad = 5

En la Figura 3, presenta la gráfica de varias distribuciones t con diferentes grados de libertad. Se observa, cuando el número de datos es menor que 30, la distribución presenta una mayor dispersión y es más gruesa en las colas. Además, la distribución es más corta en el centro, lo que indica que los valores centrales son menos probables que la distribución normal estándar, pero a medida que el número de datos aumenta siendo  $n > 30$ , la distribución t se aproxima a una distribución normal estándar [16].

Propiedades de la distribución t

Es importante considerar las siguientes propiedades de la distribución t [13], [15] y [16]:

1. Cada curva tiene una forma de campana con centro en cero.
2. Las curvas t son más dispersas que la curva normal estándar.
3. A medida que aumenta el valor de  $k$ , la dispersión de la curva t disminuye.
4. A medida que  $k \rightarrow \infty$ , las curvas t se acercan a la curva normal estándar.

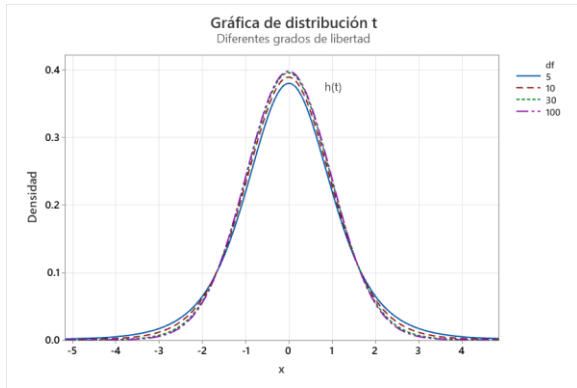


Figura 3. Distribución t con diferentes grados de libertad

**Paso 1: Formulación de hipótesis**

En las muestras independientes se refiere a la comparación de dos grupos o poblaciones diferentes y consiste en establecer una hipótesis nula  $H_0$  y una hipótesis alternativa  $H_1$  que se someterán a prueba. La  $H_0$  se formula como una afirmación de igualdad o de no diferencia entre los dos grupos o poblaciones, mientras que la  $H_1$  se formula como una afirmación de diferencia o de desigualdad entre los dos grupos o poblaciones.

**Paso 2: Calcular el valor crítico  $t_c$**

Se halla de la Tabla T student con una cola o dos colas dependiendo del objetivo de la investigación. Considerando un nivel de significancia  $\alpha$  y los grados de libertad específicos. Luego el valor obtenido divide la distribución t en región de aceptación y rechazo  $H_0$ .

**Paso 3: Calcular T estadístico ( $t_0$ )**

Se calcula el  $t_0$  utilizando la media y la desviación estándar. El valor del estadístico  $t_0$  se compara con  $t_c$  para determinar si se rechaza o no la hipótesis nula.

$$t_0 = \frac{(x_1 - x_2)}{\sqrt{\sigma_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (5)$$

y la varianza ponderada:

$$\sigma_p^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 + \sum_{j=1}^{n_2} (x_j - \bar{x}_2)^2}{n_1 + n_2 - 2} \quad (6)$$

**Paso 4: Regla de decisión**

Si el problema es de una cola del lado izquierdo: Si  $t_0 < t_c$  hay evidencia para rechazar  $H_0$  y si  $t_0 > t_c$  hay evidencia para aceptar  $H_0$ , ver Figura 4.

Si el problema es de una cola del lado derecho: Si  $t_0 < t_c$  hay evidencia para aceptar  $H_0$  y si  $t_0 > t_c$  hay evidencia para rechazar  $H_0$ , ver Figura 5.

Si el problema es de dos colas: Si  $-t_c < t_0 < t_c$  hay evidencia para aceptar  $H_0$ , caso contrario, hay evidencia para rechazar  $H_0$ , ver Figura 6.

Aplicando la regla de decisión, se determina si hay evidencia para aceptar o rechazar la hipótesis nula (de una cola o dos colas).

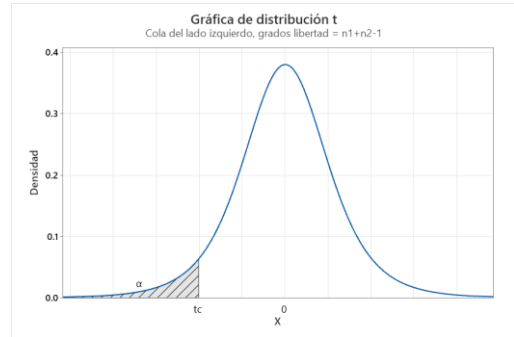


Figura 4. Distribución T con cola del lado izquierdo

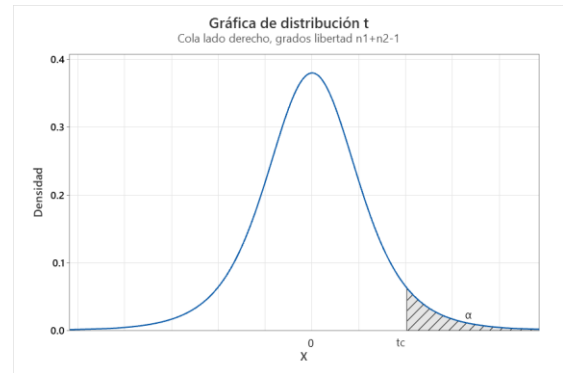


Figura 5. Distribución T con cola del lado derecho

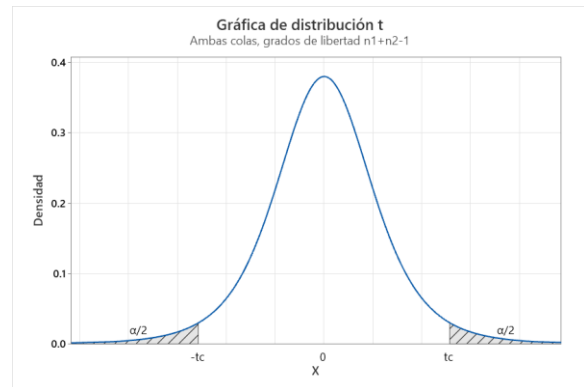


Figura 6. Distribución T con ambas colas

**2.3 Prueba de Normalidad**

Es una herramienta estadística que se utiliza para determinar si una muestra de datos sigue una distribución normal. La prueba de normalidad es importante en el contexto de la distribución t, ya que esta distribución asume que los datos se distribuyen normalmente.

Existen varias pruebas de normalidad disponibles, como la prueba de Shapiro-Wilk, la prueba de Kolmogorov-Smirnov y la prueba de Anderson-Darling. Estas pruebas evalúan si la distribución de los datos es similar a una distribución normal

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).  
**ISSN, ISBN:** (to be inserted by LACCEI).  
**DO NOT REMOVE**

teórica [12] y [17]. En la Figura 7, se muestran diversos métodos para realizar la prueba de normalidad.

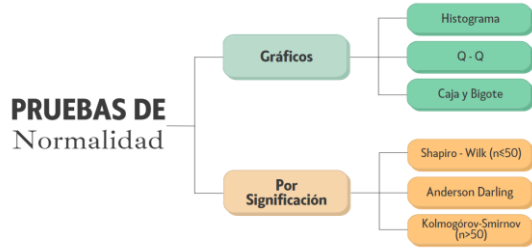


Figura 7. Métodos de pruebas de normalidad. Elaboración propia

### 2.3.1 El gráfico de Q-Q y la Prueba Anderson-Darling

Los métodos gráficos, pueden ser una forma sencilla de interpretar si existe normalidad en los datos, sin embargo, pueden ser subjetivos. Por ejemplo, si hacemos un histograma tendríamos que complementarlo con un gráfico Q-Q para poder tener suficiente evidencia para aceptar o rechazar si existe normalidad en los datos. El gráfico de Q-Q, compara los cuantiles teóricos de una distribución con los cuantiles de los datos observados. Esto nos permite evaluar si los datos se ajustan a una distribución específica, como una distribución normal [15], [16] y [17].

En base a ello, tenemos las pruebas de normalidad por significación como Shapiro Wilk el cual se usa cuando el número de muestras es menor a 50 y por el contrario Kolmogorov-Smirnov cuando es mayor a 50. La prueba de Anderson-Darling (AD) se puede utilizar para cualquier tamaño de muestra y actualmente, se considera que es la más robusta e ideal para estudiar la normalidad. El estadístico AD, evalúa la bondad de ajuste midiendo la discrepancia entre la línea ajustada según la distribución normal y la función de distribución empírica basada en los datos. Se trata de una distancia elevada al cuadrado que da más importancia a las colas de la distribución [17]. A diferencia de otras pruebas de normalidad, la prueba de Anderson-Darling da mayor peso a las colas de la distribución, lo que permite detectar con mayor facilidad las diferencias. Al ingresar los datos en algún software estadístico, se genera un gráfico de normalidad y un p\_valor correspondiente. Si consideramos que el p\_valor como una medida de la evidencia que debemos comparar con el nivel de significancia  $\alpha$  para aceptar o rechazar la hipótesis nula, un p\_valor pequeño comparado con  $\alpha$  indica una evidencia fuerte en contra de la hipótesis nula, luego los datos no siguen una distribución normal. Caso contrario, hay evidencia a favor de aceptar la hipótesis nula y los datos siguen una distribución normal [16] y [17].

### 2.3.2 Prueba de Igualdad de Varianza

Estadístico de prueba de Bartlett

El software estadístico Minitab proporciona una prueba de Bartlett y un p\_valor para evaluar la igualdad de varianzas. En estas pruebas, la hipótesis nula es que las varianzas son

iguales, mientras que la hipótesis alternativa es que las varianzas no son iguales.

La prueba de Bartlett se utiliza cuando los datos provienen de distribuciones normales. Sin embargo, esta prueba no es robusta cuando los datos se desvían de la normalidad. El estadístico de prueba de Bartlett calcula promedios ponderados de las varianzas de muestra, teniendo en cuenta los grados de libertad. Si hay una diferencia significativa en los promedios, es más probable que las varianzas de las muestras no sean iguales. El estadístico B sigue una distribución  $\chi^2$  con  $k - 1$  grados de libertad. Si el p\_valor es menor que el nivel de significancia  $\alpha$  establecido, se rechaza la hipótesis nula de igualdad de varianzas. Esto indica que hay evidencia suficiente para concluir que las varianzas no son iguales entre las muestras analizadas [16].

$$B = \frac{(\sum V_i) \ln \left( \frac{\sum V_i S_i^2}{\sum V_i} \right) - \sum V_i \ln S_i^2}{1 + \frac{\left\{ \sum \left( \frac{1}{V_i} \right) - \frac{1}{\sum V_i} \right\}}{3(k-1)}} \quad (7)$$

Donde:

$$S_i^2 \quad \sum_{j=1}^N \frac{(x_{ij} - \bar{x})^2}{(n_i - 1)}$$

$k$  número de muestras

$V_i$   $n_i - 1$

$n_i$  número de observaciones al i-ésimo nivel del factor

Estadístico de prueba F

Cuando se tienen únicamente dos niveles llamados muestras, Minitab utiliza una prueba F en lugar de la prueba de Bartlett. Esta prueba F se utiliza cuando los datos provienen de distribuciones normales. En la interpretación de los resultados, si el p\_valor obtenido es menor que el nivel de significancia  $\alpha$  establecido, se rechaza la hipótesis nula de igualdad de varianzas. Esto indica que existen diferencias significativas entre las varianzas de las muestras analizadas. Por lo tanto, se puede concluir que las varianzas no son iguales.

Es importante destacar que, si el valor p es mayor o igual que  $\alpha$ , no se cuenta con suficiente evidencia para rechazar la hipótesis nula y se concluye que las varianzas son iguales [16].

$$F = \frac{S_1^2}{S_2^2} \quad (8)$$

donde:

$S_1^2$  varianza de la muestra 1

$S_2^2$  varianza de la muestra 2

$n_1 - 1$  grados de libertad para el numerador

$n_2 - 1$  grados de libertad para el denominador

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).  
**ISSN, ISBN:** (to be inserted by LACCEI).  
**DO NOT REMOVE**

### III. RESULTADOS

Se utilizaron los programas estadísticos SPSS y Minitab para generar los gráficos y realizar pruebas estadísticas específicas. Cada uno de estos programas tiene sus propias ventajas, los cuales se emplearon para mejorar la calidad de los resultados obtenidos.

#### Procesamiento de Datos

El Resumen de procesamiento de datos es mostrada en la siguiente Tabla 4.

Tabla 4. Procesamiento de los datos

Resumen de procesamiento de los datos						
grupos	N	Válido		Casos Perdidos		Total
		N	Porcentaje	N	Porcentaje	
visitas estud_niños 2018	12	12	100,0%	0	0,0%	12 100,0%
estud_niños 2019	12	12	100,0%	0	0,0%	12 100,0%

En la Tabla 4, muestra los grupos de estudiantes/niños de los años 2018 y 2019. Ambos grupos tienen el porcentaje del 100%, indicando que se consideraron los datos de los 12 meses y que no hay casos perdidos y todos los datos fueron utilizables para el análisis.

#### Análisis de la Estadística Descriptiva

Se obtuvieron las siguientes informaciones, ver Tabla 5:

Tabla 5: Estadística descriptiva

#### Estadísticas

Variable	Media	Desv.Est.	Varianza	CoefVar
estud_niños 2018	440.8	306.7	94085.8	69.59
estud_niños 2019	523	375	140746	71.69

Se observa, que el incremento en la media de visitas de 440.8 en 2018 a 523 en 2019 indica que, en promedio, más estudiantes/niños visitaron el museo en 2019. Este aumento puede ser resultado de nuevas exhibiciones, o mayores esfuerzos de marketing y publicidad por parte del museo.

El aumento en la desviación estándar y la varianza de un año a otro sugiere que hubo una mayor variabilidad en el número de visitas diarias o mensuales en 2019. Esto podría significar que algunos meses tuvieron picos significativamente alto de visitas, mientras que otros podrían haber tenido menos, debido a eventos especiales, condiciones climáticas o vacaciones escolares.

El aumento en el coeficiente de variación indica que la variabilidad relativa de las visitas fue mayor en 2019 en 3.01% en comparación con 2018.

#### Análisis de gráfico de valores individuales

Se presenta el gráfico de valores individuales de los estudiantes/niños que visitaron el MHNURP en los años 2018 y 2019, ver la Figura 8.

Se observa que los datos están bastante dispersos en ambos

grupos. Esto se puede inferir a partir de la alta desviación estándar en ambos años, así como del coeficiente de variación relativamente alto 70% aproximadamente. Esto sugiere que hay variabilidad en la cantidad de estudiantes/niños que visitaron el museo en cada año, especialmente en alta en los meses julio, agosto y septiembre en el año 2018 y abril, julio y septiembre en 2019. Además, es posible que existan factores externos que influyan en la cantidad de visitas, como días feriados nacionales o condiciones climáticas.

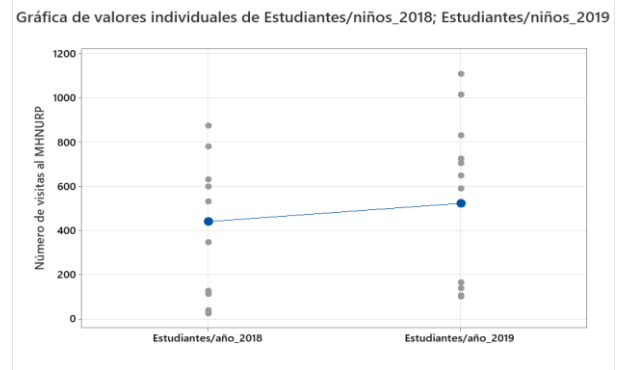


Figura 8. Gráfico de valores individuales

En conclusión, se recomienda una mayor investigación para determinar las causas subyacentes de la variabilidad en las visitas al museo. Esto podría ayudar a desarrollar estrategias para mejorar la cantidad y la consistencia de las visitas en el futuro.

#### Análisis de los histogramas

En la Figura 9, se muestra el histograma de las visitas al MHNURP de los estudiantes/niños del 2018. Se observa que la mayoría de las visitas se agrupan en el extremo superior de la escala y disminuyen conforme nos movemos hacia el extremo inferior.

Por otro lado, en la Figura 10, muestra el histograma de las visitas al MHNURP de los estudiantes/niños 2019. Se observa, una distribución más uniforme, sugiere que hubo un mayor número de visitas al inicio del año y un aumento en la cantidad de visitas durante la mitad del año. Esto podría indicar que se implementaron estrategias exitosas para atraer a más visitantes al museo en esos períodos de tiempo específicos.

Es alentador ver que el esfuerzo por promover el museo ha dado resultados positivos y ha generado un mayor interés en los estudiantes/niños. Sin embargo, es importante analizar más a fondo los factores que contribuyeron a este aumento y evaluar su efectividad para continuar atrayendo visitantes en el futuro.

En resumen, los histogramas pueden proporcionar una visión detallada de la distribución de los datos y pueden ayudar a identificar patrones o tendencias en la cantidad de visitas al museo a lo largo del tiempo.

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).  
**ISSN, ISBN:** (to be inserted by LACCEI).  
**DO NOT REMOVE**



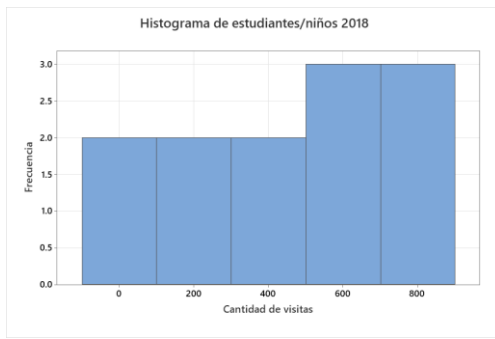


Figura 9: Histograma de estudiantes/niños 2018

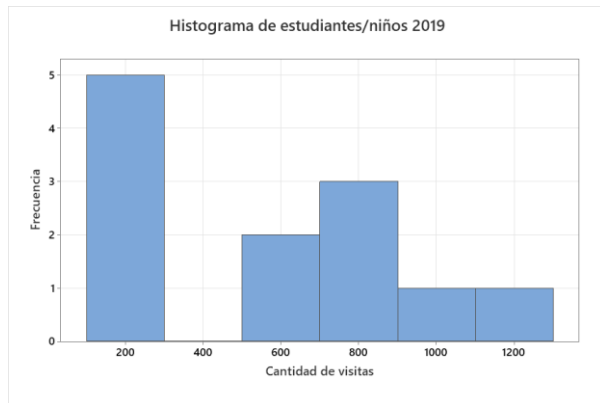


Figura 10: Histograma de estudiantes/niños 2018

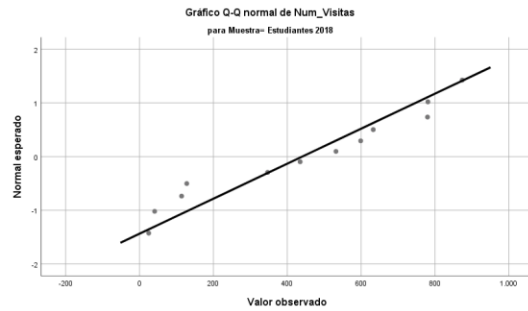


Figura 11. Gráfico Q-Q normal – Estudiantes/niños 2018

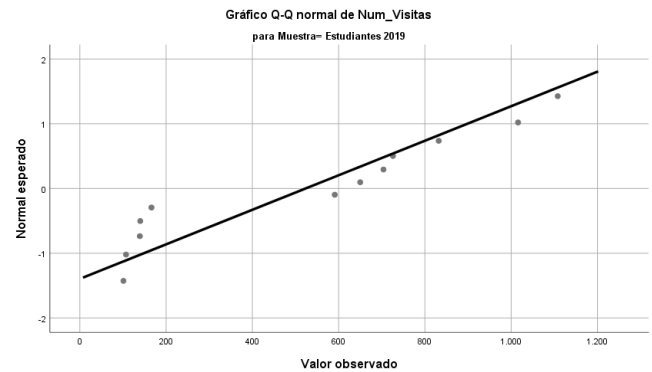


Figura 12. Gráfico Q-Q normal – Estudiantes/niños 2019

### Sobre la prueba de Normalidad

Planteamos las siguientes hipótesis:

H0: Los datos siguen una distribución normal.

H1: Los datos no siguen una distribución normal.

$\alpha = 0.05$

En la Figura 11 y Figura 12, se muestran los gráficos Q-Q, de los estudiantes/niños 2018 y 2019 respectivamente. Se observan que los puntos se ajustan a la línea diagonal, podemos decir que hay evidencia para aceptar H0. Por tanto, la muestra de estudiantes/niños 2018 y 2019 siguen una distribución normal.

La prueba de normalidad aplicando Anderson Darling con un  $\alpha = 0.05$ , se observa en la Figura 13, los p\_valores de los años 2018 es 0.345 y 2019 es 0.064 respectivamente y ambos son mayores que el nivel de significancia  $\alpha$ .

Por tanto, hay evidencia para no rechazar H0, los datos siguen una distribución normal al 95% de confianza.

### Sobre la prueba de Igualdad de Varianza

H0: Los datos tienen igualdad de varianza.

H1: Los datos no tienen igualdad de varianza.

$\alpha = 0.05$

Se utilizó la prueba F, ver Tabla 6.

Se observa, que el p\_valor igual a  $0.515 > \alpha$ , hay evidencia para aceptar H0.

Entonces, no hay diferencia significativa en las varianzas de los datos sobre la cantidad de vistas de los estudiantes/niños

al MHNURP en los años 2018 y 2019. Esto sugiere que la cantidad de visitas al museo en ambos años tienen las mismas varianzas, vienen de la misma población.

### Inferencia Estadística

Cumpliendo las pruebas de normalidad e igualdad de varianza, procederemos a realizar la prueba estadística que corresponde a la distribución t.

En el presente estudio, se eligieron los años específicos de 2018 y 2019 para comparar las visitas al MHNURP y no hubo un seguimiento de los mismos sujetos en ambos años. Por tal motivo, se considera muestras independientes.

Las hipótesis están dadas por:

H0:  $\mu_1 = \mu_2$ , No hay diferencia significativa en las visitas al MHNURP en los estudiantes/niños en el año 2018 y en los estudiantes/niños en el año 2019.

H1:  $\mu_1 \neq \mu_2$ , Hay diferencia significativa en las visitas al MHNURP en los estudiantes/niños en el año 2018 y en los estudiantes/niños en el año 2019. Considerar  $\alpha = 0.05$ .

Aplicando la distribución t para muestras pequeñas e independientes, se obtuvieron los siguientes resultados, ver la Tabla 7.

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).

**ISSN, ISBN:** (to be inserted by LACCEI).

**DO NOT REMOVE**

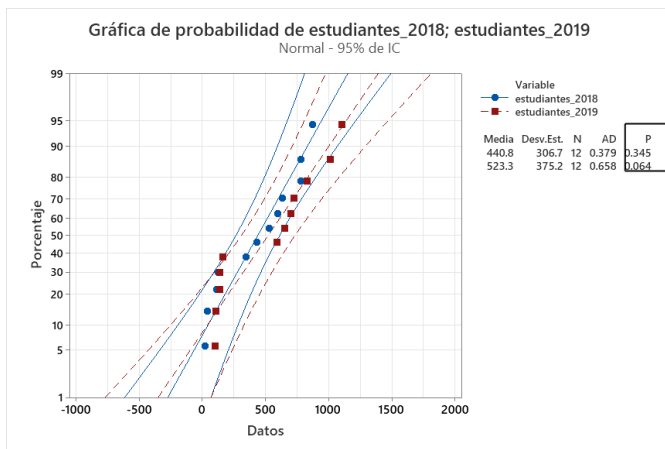


Figura 13. Prueba de normalidad de Anderson Darling.

Tabla 6. Resultado de la Prueba F

Método	Estadística de prueba	GL1	GL2	Valor p
F	0.67	11	11	0.515

Tabla 7. Resultado de la distribución t

### Prueba

Hipótesis nula  $H_0: \mu_1 - \mu_2 = 0$   
 Hipótesis alterna  $H_1: \mu_1 - \mu_2 \neq 0$

Valor T	GL	Valor p
-0.59	22	0.561

Basándonos en los resultados obtenidos, se observa que  $p\_valor = 0.561 > \alpha = 0.05$ , podemos concluir que no hay suficiente evidencia para rechazar  $H_0$ . Esto significa que no podemos afirmar con 95% de confianza que exista una diferencia significativa en las medias de las visitas al MHNURP en los estudiantes/niños en el año 2018 y en las medias de las visitas de los estudiantes/niños en el año 2019.

## IV. DISCUSIONES

El MHNURP juega un papel importante en difundir información y conocimiento acerca de la biodiversidad contribuyendo en crear conciencia sobre la necesidad de valorar, proteger y conservar la biodiversidad. Además, el museo puede ser un espacio de reflexión y discusión acerca de las acciones que debemos tomar como sociedad para lograr un manejo sustentable de los recursos y asegurar la vida en nuestro planeta. Se trabajó con los programas estadísticos SPSS y Minitab con el fin de obtener gráficos y pruebas estadísticas específicas. Cada uno de estos programas tiene sus propias ventajas y beneficios que se han utilizado para mejorar la calidad de los resultados obtenidos.

La inferencia estadística es una herramienta fundamental en la investigación científica y en la toma de decisiones basadas en datos. Permite hacer generalizaciones y conclusiones sobre una población a partir de una muestra o dos de datos, lo que es esencial cuando no es posible o práctico medir todos los

individuos de una población. Cabe resaltar que la Inferencia Estadística es un proceso riguroso que debe realizarse siguiendo los pasos adecuados para obtener resultados confiables y precisos. Actualmente es un campo ampliamente discutidos y debatidos en el ámbito de la estadística y la investigación científica, tales como el tamaño de muestra, unos argumentan que una muestra grande es necesaria para obtener datos confiables, el cumplimiento de la normalidad solo en una sola muestra es suficiente y no en las dos muestras, aunque Gosset desarrolló la distribución t, propuso que las muestras deben cumplir la normalidad pero todo dependerá del estudio a realizar y todas estas discusiones reflejan el constante desarrollo y refinamiento de los métodos estadísticos a emplear.

La distribución t se utiliza para estimar y probar la media de una población cuando el tamaño de la muestra es pequeño y/o la desviación estándar poblacional no se conoce. Al utilizar la distribución t en estos casos, se tiene en cuenta la mayor incertidumbre asociada con muestras pequeñas, lo que ayuda a obtener resultados más precisos y confiables en la inferencia estadística. También se deberá tener en cuenta el nivel de significancia establecido previamente, así como las pruebas de normalidad e igualdad de varianza. Para comprobar la normalidad de los datos se utilizaron dos métodos: el gráfico y la prueba de Anderson-Darling. Estos métodos permitieron obtener resultados más objetivos y evitar la subjetividad asociada al gráfico Q-Q o Histogramas. Por otro lado, la prueba de Bartlett es una herramienta útil para evaluar la igualdad de varianzas en datos provenientes de distribuciones normales. Sin embargo, es importante tener en cuenta que esta prueba no es robusta cuando los datos se desvían de la normalidad. Es fundamental considerar otros métodos y herramientas estadísticas para evaluar la igualdad de varianzas en estos casos. En el caso, la prueba F es una herramienta estadística utilizada para evaluar si hay diferencias significativas en las varianzas entre los grupos y asume que los datos siguen una distribución normal. Además, es necesario considerar la validez de los supuestos subyacentes, como la normalidad de los datos, e igualdad de varianza antes de realizar conclusiones basadas en la distribución t para muestras pequeñas e independientes. Si los datos no cumplen los supuestos de normalidad, se puede considerar utilizar pruebas no paramétricas, como la prueba de Mann-Whitney U o la prueba de Wilcoxon, que no asumen una distribución específica.

El estudio fue realizado en el curso de Diseño de Experimentos, que es una herramienta clave para el ingeniero industrial y otras áreas, ya que se utiliza en el desarrollo de nuevos productos, procesos o servicios, control de calidad y optimización de procesos. Es esencial que los alumnos estén bien capacitados en esta área para aprovechar al máximo sus beneficios. Además, es importante que los alumnos lleven lo aprendido en el curso al campo experimental, ya que esto les permitirá aplicar los conocimientos adquiridos en situaciones

reales y mejorar su capacidad para resolver problemas de manera creativa y efectiva.

## VII. CONCLUSION

El análisis de Inferencia Estadística realizado utilizando dos muestras pequeñas e independientes reveló que no hay evidencia suficiente para rechazar la hipótesis nula. Esto indica que no se encontró una diferencia significativa al 95% en las visitas al MHNURP en los estudiantes/niños en el año 2018 y en los estudiantes/niños en el año 2019. Es importante tener en cuenta que este análisis se basó en las muestras disponibles y en las variables consideradas. Otros factores y variables relevantes podrían influir en las visitas al MHNURP y deberían ser considerados en futuras investigaciones. El MHNURP tiene un papel fundamental en la conservación y difusión del patrimonio cultural de nuestro país. Por tanto, es fundamental que se promueva el desarrollo y fortalecimiento de los museos para que puedan cumplir con su misión de manera efectiva y alcanzar un mayor impacto en la sociedad.

El estudio permitió a los estudiantes comprender el papel fundamental que juega la Inferencia Estadística en la toma de decisiones, al aplicar la distribución T en un escenario de dos muestras independientes. Además, les permitió realizar el tratamiento y análisis de los datos, formular las hipótesis, realizar diversas pruebas estadísticas, y finalmente, interpretar los resultados obtenidos con datos reales. Al aplicar adecuadamente herramientas estadísticas como la distribución T, los estudiantes pudieron comprender la importancia de utilizar herramientas estadísticas precisas y confiables para realizar análisis precisos en la toma de decisiones, permitiendo desarrollar las habilidades prácticas y analíticas.

## VIII. RECOMENDACIÓN

Implementar diversas estrategias para atraer mayor cantidad de visitas del público estudiado, como programas educativos interactivos, exposiciones atractivas, eventos especiales, descuentos, y colaboraciones con otras instituciones.

En base a la información proporcionada, se podría ampliar el estudio con más muestras ahora que el museo está abierto al público. Si se obtienen más de tres muestras, se podría aplicar el diseño completamente al azar, diseño de bloques o diseño factorial. Sin embargo, la elección del enfoque estadístico dependerá de los objetivos de investigación y de las características específicas de los datos.

## AGRADECIMIENTOS

Se agradece a la Escuela de Ingeniería Industrial de la Universidad Ricardo Palma, a los estudiantes que cursaron la asignatura de Diseño de Experimentos, entre los semestres académicos 2023-II. Sin ellos no hubiera sido posible la ejecución de este artículo.

## REFERENCIAS

[1] Universidad Ricardo Palma. Museo de historia natural, 2023. <https://www.urp.edu.pe/centros-institutos-y-museo/museo-de-historia-natural/>

[2] Museo de Historia Natural. Vera Alleman H., 2023. [Museo de Historia Natural URP - Información \(mhnurp.com\)](https://www.mhnurp.com/)

[3] G. Panikos, B. Kayleigh. "Sustainability teaching using case-based debates". Journal of International Education in Business. Vol15. Issue 1. Page 147-163. April 2022.

[4] E. Gunilla. B. Staffan. "Reflections in Case-Based Learning: Experiences of Computer-Based Simulation in Social Work Education". Department of Social Work, Jönköping University, Jönköping, Sweden. DOI 10.1080/10437797.2022.2033655. 2022.

[5] Z. Xinhong, Z. Boyan, Z. Fan. "Student-centered case-based teaching and online-offline case discussion in postgraduate courses of computer science". International Journal of Educational Technology in Higher Education. Vol. 20. Issue 1. Article number 06. 10.1186/s41239-022-00374-2. 2023.

[6] F. Cárcel. "El aprendizaje mediante estudio de caso en la empresa industrial". 3C Empresa. Área de Innovación y Desarrollo. Ed. 24. Vol 4. N°2. DOI: <http://dx.doi.org/10.17993/3cemp.2015.040424.248-260.2015>

[7] Aggregate production planning considering organizational learning with case-based analysis. Ainal. Shams Engineering Journal. Vol 13. Issue 2. N° 101575. DOI 10.1016/j.asej.2021.09.002. 2022.

[8] Ch. Chih-Chan, C. Ya-Ting. "Impact of smart classrooms combined with student-centered pedagogies on rural students' learning outcomes: Pedagogy and duration as moderator variables". Institute of Education & Center for Teacher Education, National Cheng Kung University, Taiwan. 2023.

[9] V. Palomino. Plan Estratégico del Museo de la Universidad Ricardo Palma 2023-2025 - Trabajo de Suficiencia Profesional para optar al Título Profesional de Licenciado en Administración. Universidad San Ignacio de Loyola, 2023. <https://repositorio.usil.edu.pe/server/api/core/bitstreams/3ec02282-20fb-442f-85fd-4ef8b5ee5055/content>

[10] R. Millones, E. Barreno, F. Vásquez, C. Castillo. "Estadística aplicada a la ingeniería y los negocios". Universidad de Lima. ISBN: 978-9972-45-297-0. 2015.

[11] J. García, J. Bachero, A. Ivars, I. López, C. Rojo, F. Ruiz "Inferencia Estadística". Garceta. 2ed. ISBN: 978-84-9281-232-5. 2018.

[12] DC. Montgomery. Design and analysis of experiments. 9th ed. USA: Wiley; 2017.

[13] By Student. The probable error of a mean. Biometrika. 1908; 6: 1-25.

[14] R. Sánchez. "t-Student- Usos y abusos". Revista Mexicana Cardiología. Vol. 26. N°1. Pág. 59-61. <https://www.scielo.org.mx/pdf/rmc/v26n1/v26n1a9.pdf>

[15] R. Fisher «Aplicaciones de la distribución 'de Student'». Metron 5: 90-104. Archivado desde [pdf el original](#) el 5 de marzo de 2016.

[16] R. Walpole, R. Myers, K. Ye. Probability & Statistics for Engineers & Scientists (7th edition). New Delhi: Pearson. p. 237. 2006.

[17] C. Flores, K. Flores. "Pruebas para comprobar la Normalidad de Datos en procesos productivos: Anderson Darling, Ryan-Joiner, Shapiro-Wilk y Kolmogórov-Smirnov". Societas- Revista de Ciencias Sociales y Humanísticas Universidad de Panamá, Panamá ISSN: 1560-0408 Periodicidad: Semestral vol. 23, núm. 2, 2021.