

**Omnia** Año 30, No. 2 (julio-diciembre, 2024) pp. 252-266  
Universidad del Zulia. e-ISSN: 2477-9474  
Depósito legal ppi201502ZU4664

## Propuesta de baremo tecno didáctico para el p – valor y el error del tipo I

*Edwin Rafael Carrasquero Cabrera*

### Resumen

Las pruebas estadísticas nacieron para comprobar formalmente las sospechas del razonamiento intuitivo y lógico. Sin embargo, por su afinidad con las matemáticas lucen de gran complejidad para la gran mayoría de la gente, incluyendo al mundo académico. Muchos profesionales, sobre todo de la salud, manifiestan su dominio dogmático de estas técnicas especiales para la investigación científica. Particularmente el p-valor estadístico, como una de las formas de llevar a cabo el ritual de los test de hipótesis, es de gran preocupación cuando asume valores cercanos al nivel de significancia. Ya lo advertía el mismo Fisher, padre del análisis inferencial, a principio del siglo XX, entre 1920 y 1930, sobre las decisiones de falso negativo cuando este p-valor era cercano al valor alfa. En este sentido, se asume la responsabilidad de llevar una propuesta de solución a esta inquietud científica, con el objetivo general de Proponer un baremo tecno didáctico para el p-valor y su aplicación en la investigación científica. En detalles, se busca: i) Categorizar el dominio probabilístico del p-valor; ii) Definir zona de incertidumbre en la zona de rechazo, iii) Precisar intensidad de la fuerza de rechazo de la hipótesis nula en el test, iv) Categorizar la probabilidad de cometer el error Tipo I y v) Presentar virtualmente los elementos del análisis inferencial. Como resultados, se tienen aplicaciones de Excelia que es un recurso tecnológico en ambiente de Microsoft Excel 2010, tomado del Artículo titulado *Excelia: recurso digital tecno-didáctico de valores agregados a Excel*, de autoría propia (Carrasquero, E. 2023).

**Palabras clave:** Baremo, Tecno didáctico, p-valor, error tipo I, Zona de incertidumbre.

\* Profesor de la Facultad de Humanidades y Educación. Universidad del Zulia. Maracaibo-Venezuela. [ercaca093@gmail.com](mailto:ercaca093@gmail.com). <https://orcid.org/000-0001-6418-9365>.

**Recibido:** 18/07/24 • **Aceptado:** 27/11/24

## *Proposal for a Technical didactic scale for the p-value and type I error*

### **Abstract**

Statistical tests were born to formally verify the suspicions of intuitive and logical reasoning. However, due to their affinity with mathematics, they appear highly complex to the vast majority of people, including the academic world. Many professionals, especially health professionals, express their dogmatic mastery of these special techniques for scientific research. Particularly the statistical p-value, as one of the ways to carry out the ritual of hypothesis testing, is of great concern when it assumes values close to the level of significance. Fisher himself, the father of inferential analysis, already warned this at the beginning of the 20th century, between 1920 and 1930, about false negative decisions when this p-value was close to the alpha value. In this sense, the responsibility of bringing a proposal for a solution to this scientific concern is assumed, with the general objective of proposing a techno-didactic scale for the p-value and its application in scientific research. In details, it seeks to: i) Categorize the probabilistic domain of the p-value; ii) Define the area of uncertainty in the area of rejection, iii) Specify the strength of the rejection of the null hypothesis in the test, iv) Categorize the probability of committing the Type I error and v) Virtually present the elements of the inferential analysis. As a result, there are Excelia applications, which are a technological resource in the Microsoft Excel 2010 environment, taken from the Article entitled Excelia: techno-didactic digital resource of added values to Excel, self-authored (Carrasquero, E. 2023).

**Keywords:** Scale, didactic technology, p-value, type I error, uncertainty zone

### **Introducción**

Bajo la presunción de alguna verdad, se mira la realidad y se apoyará o negará tal afirmación hipotética. La *distancia* (espacial, temporal, etc.), juega un papel muy importante en esos momentos de indefinición. Lo que está aquí no puede estar allá y viceversa. Al ver un chico de 12 años se puede pensar que estudia el nivel medio y un joven de 18 años estará en sus estudios universitarios. Nunca se pensaría lo contrario, en que el chico está en la universidad o el joven en la educación media. Lo anterior es solo razonamiento intuitivo y lógico. No es necesario ser matemático o experto estadístico para llegar a esas

conclusiones. Ese es el fundamento esencial de los *test estadísticos* para llegar a inferencias deductivas sobre la veracidad de una proposición.

Con esas competencias intrínsecas del razonamiento intuitivo y lógico como inicio, surge la creatividad de la utilización del concepto de cantidad y así precisar las diferencias cuando son pequeñas, p.e. si la persona está por alrededor de los 16 años de edad, cabría la duda del nivel educativo alcanzado: ¿medio, universitario?. Entonces se hace necesario formalizar un procedimiento que conduzca a la validación de verdades supuestas.

Formalmente las *pruebas de hipótesis* surgen en las décadas de 1920 y 1930 como resultado del trabajo de dos grupos o escuelas de pensamiento: por un lado, Ronald Fisher (1890-1962), y por el otro, Jerzy Neyman (1894-1981), y Egon Pearson (1895-1980). (Díaz, L y Ríos, F. 2018). Ya en ese entonces, afirman estos autores, advirtió el mismo Fisher sobre la incertidumbre de rechazar o no rechazar la hipótesis nula cuando la evidencia real es un valor cercano a la frontera (tal como el chico de 16 años de edad) entre las zonas de aceptación y de rechazo.

Así, por ejemplo, para los p-valor:  $P_1 = 0,049$  y  $P_2 = 0,051$ ; uno por encima y uno por debajo del valor crítico,  $p. e. \alpha = 0,05$ , donde con el primero no se rechaza la  $H_0$  y con el segundo sí, surge la gran inquietud de estar cometiendo los errores del tipo I y II, respectivamente. Y la gran pregunta, ¿Qué hacer si  $p = 0,50$ ?

Pareciera exigencia necesaria, declarar zona de incertidumbre las cercanías del valor crítico, ya que luce como un fino vidrio sin grosor, una frontera débil entre los errores Tipo I (“rechazar lo verdadero” o “falso positivo”) y Tipo II (“aceptar lo falso” o “falso negativo”) al rechazar o no la hipótesis nula. Es necesario, entonces, declarar una zona de incertidumbre alrededor de 0,05 para disminuir la probabilidad de evitar decisiones erradas.

Para Molina Arias (2017), una  $p < 0,05$  quiere simplemente decir que es poco probable que la  $H_0$  sea cierta, luego se rechaza, pero siempre con cierta probabilidad de cometer lo que se denomina un error de Tipo I. Por otra parte, el valor de  $p > 0,05$  no afirma que la  $H_0$  sea verdadera, ya que puede ocurrir que la diferencia sea real y el estudio no tenga potencia para detectarla. Se está entonces ante el error de Tipo II.

Otro detalle, es el señalado por Prieto y Herranz (2015), quienes aluden el error de no considerar la diferencia entre dos valores de  $p$  que califican de significativos por ser menores que alfa. Para ello citan el ejemplo de  $p = 0,039$  y  $p = 0,000003$ , donde ambos son menores que 0,05 pero el menor indica mucho mayor confiabilidad en el rechazo que minimiza la probabilidad

de cometer el error de Tipo I, por el contrario, el otro valor presenta una mayor incertidumbre de cometer dicho tipo de error. Para estos autores, (Prieto, y Herranz, 2015), no tiene sentido en la investigación científica frases como “se decide considerar significativos todos los resultados con  $p < 0,05$ ” sin valorar en qué medida los resultados obtenidos constituyen evidencia a favor de cierta hipótesis.

Entonces, además de declarar una zona de incertidumbre alrededor del valor 0,05, se hace necesario tomar en consideración el alejamiento del valor p de esta zona de incertidumbre: mientras más alejado esté hacia el extremo más próximo, mayor confianza en el rechazo y menor probabilidad de cometer el error Tipo I. Esto significa categorizar este alejamiento del valor p, en su rango probabilístico de rechazo: (p. e. para una cola: 0 - 0,05). Así mismo, para una investigación futura, se sugiere baremar la zona de aceptación en su intervalo: (0,05; 0,50), para pruebas de dos colas en este caso, que permita medir la probabilidad del error Tipo II.

El p-valor tiene relación con la fiabilidad del estudio, cuyo resultado será más fiable cuanto menor sea el valor de p: en realidad, muchos autores coinciden en que indicaría la probabilidad de obtener un valor semejante si se realiza el experimento en las mismas condiciones, pero hay muchos factores que pueden intervenir además del hecho de que exista o no diferencia real: el tamaño de la muestra, la varianza de la variable medida, el tamaño del efecto, la distribución de probabilidad empleada, etc. (Molina Arias M. 2017).

## **Enseñanza del p-valor**

La afinidad de la estadística con las matemáticas indudablemente le transmite un blindaje abrumador, al investigador no matemático para su aprendizaje. El razonamiento intuitivo y lógico de su parte, no es suficiente, el profesional no matemático se encuentra necesariamente de nuevo con esta ciencia dura.

Los errores que cometen los investigadores, según Prieto y Herranz (2015), no se debe a una especial incapacidad, sino la deficiente enseñanza que la mayoría de ellos recibió en este campo. Agregan que, por su afinidad matemática, la inferencia estadística es enseñada a profesionales de otras ciencias usando un lenguaje formal que la hace ininteligible para la mayoría, aunque todos tengan sobrada capacidad para entender sin ambigüedad lo que el valor p significa.

Para Prieto y Herranz (2015), aunque los test estadístico son una cuestión matemática y a los estadísticos se debe su desarrollo teórico y su aplicación práctica, estos pueden explicarse y entenderse usándolos correctamente sin recurrir a profundas nociones aritméticas. Realmente, con un poco de intuición y un poco de lógica elemental, pero con predisposición a su captación, estos conceptos son simples, aunque luzcan lo contrario.

Pero, en definitiva, es necesario revisar las administraciones de este tipo de asignatura cuantitativa, en particular las estadísticas que deben ser enseñadas desde la escuela primaria y secundaria, tal como lo reza el Currículo Básico Nacional (2017), pero pareciera que los docentes no están en capacidad o disposición para hacer esta tarea.

No debe quedar por fuera la presencia de la tecnología que, en el caso de la estadística, esta se ha venido beneficiando con la existencia de los software o paquetes estadísticos. Particularmente, con aplicaciones de Excel, es posible diseñar y programar la automatización de la obtención del p-valor, así como su calibración baremal. Indudablemente, esto coadyuvará en la visualización de los conceptos estadísticos inferenciales, como el p-valor, entre otros. El resultado será un recurso tecno didáctico que facilitará, además de las tareas estadística, su enseñanza y aprendizaje.

## **Justificación**

En conclusión, se justifica definir una zona de incertidumbre precisada por un baremo que mida la intensidad del rechazo, así mismo no tomar decisiones cuando la evidencia real se ubique en la mencionada zona de fluctuación, apegada a la débil frontera de cristal. Es menester en esta situación, incrementar el tamaño de la muestra, de manera que, inobjetablemente se acercaría la inferencia hacia el verdadero valor del parámetro poblacional y así fortalecer la intensidad de la decisión del rechazo. Por otro lado, este baremo estaría indicando también la probabilidad de ocurrencia del error Tipo I.

## **Objetivo de la investigación**

Proponer un baremo tecno didáctico para el p-valor y su aplicación en la investigación científica.

## Objetivos específicos

- i) Categorizar el dominio probabilístico del p-valor.
- ii) Definir zona de incertidumbre en la zona de rechazo.
- iii) Precisar intensidad de la fuerza de rechazo de la hipótesis nula en el test estadístico.
- iv) Categorizar la probabilidad de cometer el error Tipo I.
- v) Presentar virtualmente los elementos del análisis inferencial.

## Marco teórico

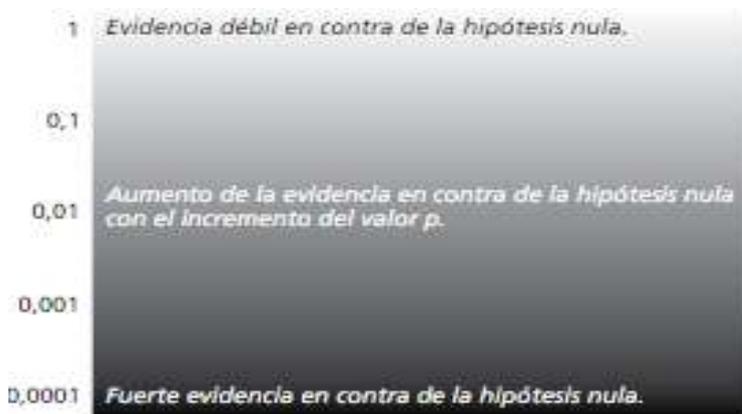
En esta parte se definirán algunos conceptos relacionado con los elementos citados en el objetivo de investigación. Estos son: Baremo: concepto, construcción y aplicación, prueba de hipótesis: su ritual, decisiones y formas, el p-valor: definición, cálculo y categorización, aplicaciones de la informática. Quedan implícitas las utilizaciones de los hipervínculos de Word con Excel en estas páginas.

## Baremo

Según la RAE, un *Baremo* es un cuadro gradual para evaluar daños, méritos, solvencia, etc. Esta definición se puede extrapolar al campo de las estadísticas y de la investigación, para ello se agrega que es una categorización que se hace de un rasgo-variable en forma numérica con lectura cualificada. P. e. las calificaciones de los estudiantes, que son numéricas del 1 al 20, se pueden baremar en las siguientes categorías: Excelente (20-19), Distinguido (18-17), Bueno (16-15), Regular (14-10), Deficiente (09-06), Muy deficiente (05-01), o de manera dicotómica: Aprobado (20-10) y Desaprobado (09-01).

Lo citado por Molina (2017), en la introducción, permite desarrollar la idea para las unidades de medición de la fiabilidad medida con el p valor: muy alta fiabilidad, alta, media, baja y muy baja. Dichas categorías quedarían distribuidas de manera exhaustiva y excluyente en todo el rango del p-valor: desde el cero (0), hasta el valor crítico  $\alpha$ : (0;  $\alpha$  ), es decir, dividir la zona de rechazo en categorías de fortaleza para la evidencia en contra de la hipótesis nula. En la misma dirección, Díaz y Ríos (2018), apoyan este criterio mostran-

do la siguiente figura, donde se muestra una fuerte evidencia en contra de la hipótesis nula para valores tendientes a cero, análogamente, una evidencia débil cuando tienden a uno (1).



Fuente: Díaz, L. y Ríos, F. (2018).

Para construir un baremo, intuitiva y lógicamente, se hace necesario conocer el rango de los valores numéricos tomados por la variable en cuestión, p.e. el promedio en una escala tipo Likert, es un número decimal que tiene por rango: (1; 5), de manera que se visualizan 4 espacios (1-2, 2-3, 3-4 y 4-5). Estos 4 espacios se dividen en 5 categorías y se tiene el cociente:  $4 / 5 = 0,8$ . Por lo tanto cada categoría tiene una amplitud de tamaño: 0,8. Finalmente, las categorías serán: 1. (1,00-1,80); 2. (1,81-2,60); 3. (2,61-3,40), 4. (3,41-4,20) y 5. (4,21-5,00). Resta asignarle la calificación a cada categoría. Así, se tiene el baremo final para la escala Likert.

**Tabla 1. Escalas Likert**

Alternativa	Frecuencia	Opinión	Presencia	
1,00-1,80	Nunca	Total desacuerdo	Ausencia total	Ninguno
1,81-2,60	Casi nunca	En desacuerdo	Ausencia parcial	Muy pocos
2,61-3,40	A veces	Ni a favor, ni en contra	Ausencia media	Algunos
3,41-4,20	Casi siempre	De acuerdo	Presencia parcial	Casi todos
4,21-5,00	Siempre	Totalmente de acuerdo	Presencia total	Todos

Fuente: Elaboración Propia (2023).

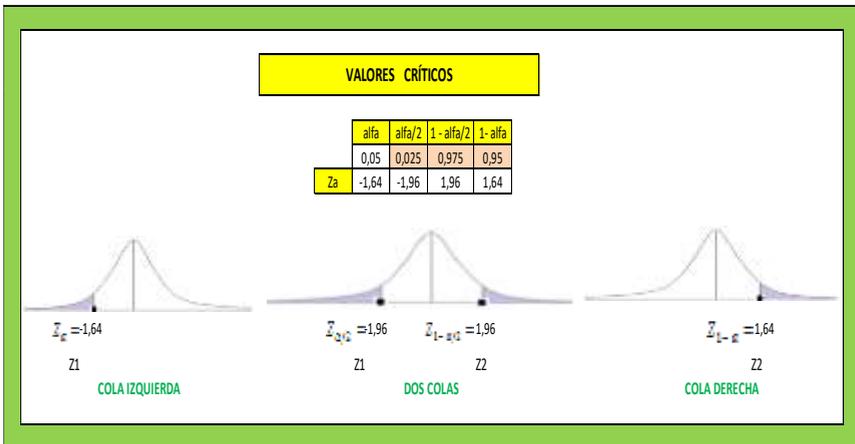
### Prueba de hipótesis

Las pruebas estadísticas se hacen cuando no es posible hacer un censo y se toma una muestra para estimar el *parámetro* de la población, entonces es necesario poder deducir lo observado en la muestra hasta toda la población misma. De aquí que la estadística inferencial reciba el calificativo de estadística deductiva por su capacidad de obtener conclusiones para un conjunto más grande.

Los primeros pasos para llevar a cabo una prueba estadística son: i) definir las hipótesis estadísticas y con ellas el tipo de cola; ii) el nivel de significancia y los valores críticos  $Z$  en función de las colas, iii) el estadístico de prueba o prueba estadística (t de Student, la prueba  $Z$ , chi cuadrado, etc.) y iv) la Regla de Decisión. Seguidamente, con los datos en la mano, se va al paso v) Tomar la decisión.

En detalles, en el ritual de una prueba de hipótesis se debe precisar lo siguiente: i) La *Hipótesis Nula* afirma una *igualdad* que significa “no cambios”, “no diferencias” y se simboliza:  $H_0 : A = B$ . Paralela a ella se define la *Hipótesis Alternativa*, de modo exhaustivo y excluyente, que le lleva la contraria a la anterior, afirmando cambios y diferencias y se simboliza:  $H_1 : A \neq B$ , de manera que cubran de forma definida, la dicotomía de los casos posibles.

**Figura 1: Excelia. Valores críticos. Recurso instruccional**



ii) En el Nivel de error o de significancia, que se simboliza con la letra griega alfa ( $\alpha$ ), se deben determinar los *valores críticos* en cada uno de los tres casos de las colas: a) para las dos colas, los valores críticos son dos:  $Z_{izquierda} = Z_{\alpha/2}$  y  $Z_{derecha} = Z_{1-\alpha/2}$ . b) para una cola derecha, existe un solo valor:  $Z_{derecha} = Z_{1-\alpha}$ , y c) para una cola a la izquierda,  $Z_{izquierda} = Z_{\alpha}$ .

## Formas de rechazo de la hipótesis nula

Hay varias maneras de rechazar la hipótesis nula: a) Cuando el valor muestral cae fuera del *intervalo de confianza* construido alrededor del valor hipotético del parámetro, b) Cuando el estadístico de prueba ( $Z_c$ , por ejemplo) cae fuera del intervalo definido entre los valores críticos: ( $Z_1; Z_2$ ), c) En el Spss, cuando el cero (0) cae fuera del intervalo para la diferencia entre las medias o el Sig es menor que el nivel de significancia, d) Cuando la distancia entre el parámetro hipotético y el estadístico muestral es mayor al radio del intervalo de confianza (p.e.  $R = Z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}}$ ) y e) Cuando la probabilidad que define el p – valor, es menor que el nivel de significancia ( $\alpha$ ). Este último modo es el más usual y más reciente, pero muy dogmatizado para los investigadores. Cuando se lleva a cabo la prueba con el Spss, se debe tomar en cuenta que el p-valor viene identificado con la palabra abreviada Sig.

Debe advertirse el tener cuidado con un valor p muy cercano a 0,05; que implicaría una alta probabilidad de cometer un error del tipo I si se trata de un valor menor y del tipo II si se trata de un valor mayor. Podría no llevar ni al rechazo ni a la aceptación de la hipótesis nula, sino a la decisión de incrementar la muestra para que la decisión salga fortalecida, para que se rechace o no se rechace la hipótesis nula con convicción, que a medida que la zona extrema de rechazo se hace más pequeña (valor p más bajo), la evidencia a favor del rechazo en contra de la hipótesis nula, sea cada vez más contundente.

## P – Valor

El *p-value*, nombre original, es la probabilidad, condicionada por la aceptación de la hipótesis nula, de que ocurra nuevamente el valor muestral o uno más extremo que él. El valor p debe entenderse, (Díaz, L y Ríos, F. 2018), como la proporción de veces que el estadístico de contraste (media, desviación estándar, varianza, proporción, etc.) toma un valor más extremo que el resultado del experimento realizado, puede entenderse como la probabilidad de encontrar un valor del estadístico de contraste más alejado o más extremo que lo observado

en la muestra actual, si repitiéramos el experimento en iguales condiciones de forma infinita (Díaz y Ríos, 2018).

Agregan (Díaz y Ríos, 2018), que Fisher fue el primer matemático que propició la utilización de la prueba de significación de la hipótesis nula para la toma de decisiones. Entonces, la hipótesis nula puede o no ser rechazada (es decir, hay diferencias entre los estadísticos de contraste o no hay diferencias, respectivamente). Señalan estos autores que, el valor  $p$ , según Fisher, es la probabilidad de encontrar un resultado igual o más alejado que el hallado en el estudio actual.

## Cálculo del P – Valor

Operacionalmente, el  $p$ -valor es el área bajo la curva de distribución, entre el estadístico de prueba arrojado como evidencia y el extremo más cercano de la curva, valor equivalente a la probabilidad de ocurrencia de los elementos ubicados en dicha área o sector. En este sentido, se tienen tres casos según las colas de los test estadísticos ya definidos.

En la página titulada: *Calcular un valor  $p$  manualmente*, <https://support.minitab.com/es-mx/minitab/20/help-and-how-to/statistics/basic-statistics/supporting-topics/basics/manually-calculate-a-p-value/#:~:text=Para%20una%20prueba%20de%20cola%20superior%2C%20el%20valor%20p%20es,de%20la%20muestra%20es%20negativo>. Se establecen las fórmulas matemáticas para cada uno de los tres casos según las colas. Considerando la distribución normal ( $Z_C$ ) como estadístico de prueba, se tiene que el valor  $p$  para:

- Una prueba de **cola inferior** se especifica mediante:

$$P = p(Z \leq Z_C / H_0), \text{ donde } Z_C = Z_\alpha$$

- Una prueba de **cola superior** se especifica mediante:

$$P = p(Z \geq Z_C / H_0), \text{ donde } Z_C = Z_{1-\alpha}$$

- Una **prueba bilateral** se especifica mediante: valor

$$P = p(Z \geq |Z_C| / H_0), \text{ donde } Z_C = Z_{1-\alpha/2}$$

## Metodología

La ruta metodológica fue aplicada-descriptiva, donde primordialmente

el fin es de aplicación para generar productos de innovación con la tecnología de la inteligencia artificial, incubada en el lenguaje lógico del software Microsoft Excel 2010, coadyuvando en la investigación científica y en la enseñanza aprendizaje de conceptos estadísticos de poca exposición y de poco dominio en nuestros medios académicos. Dícese descriptiva puesto que se hace necesario detallar cada elemento en estudio como lo es el p-valor y el naciente *Baremo para el P-valor*, incluyendo sus extensiones hacia el Error Tipo I y la nueva Zona de Incertidumbre.

El baremo para el p-valor se construyó partiendo del rango probabilístico de valores que éste puede tomar, ordenando cada categoría en niveles, desde un polo de fortalezas, valores cercanos al extremo de la cola, hasta el otro extremo de debilidades, cercanos al punto crítico, definido éste el marco teórico. Se tomó en consideración que el NS toma dos valores, según las colas, a partir del valor prefijado por el investigador. Así se tiene: i) Para una cola:  $cola = \alpha$  ; ii) Para dos colas:  $cada\ cola = \alpha/2$ .

La definición de la zona de incertidumbre dentro de la región de rechazo, fue considerando el criterio de cercanía, declarando las dos categorías cercanas a la frontera crítica como espacio de indecisión y dónde se sugiere ampliar el tamaño de la muestra. Para precisar la intensidad de la fuerza de rechazo de la hipótesis nula, se simuló el criterio de los dos polos de la escala tipo Likert: fuerte cerca del cero y débil cerca del otro extremo.

De igual manera, para categorizar los niveles de probabilidad de cometer el error Tipo I, esta se considera inversamente proporcional a la intensidad de la fuerza de rechazo: Para mayor intensidad de la fuerza de rechazo, menor probabilidad de cometer el error Tipo I; y a mayor debilidad mayor probabilidad para cometer este error.

Se consideró la visualización de todos estos elementos mencionados, de manera que ello permita una mayor captación de los conceptos involucrados en la enseñanza-aprendizaje de cada elemento, estos son: Rango probabilístico, Zona de incertidumbre, intensidad de la fuerza de rechazo y la probabilidad de cometer el error Tipo I.

## Resultados

A continuación, se enlistan la obtención de los resultados propuestos sobre i) un baremo para el p-valor, que ubique su valor obtenido en una escala de intensidad de rechazo de la hipótesis nula, ii) definiendo en la misma una zona de incertidumbre, donde se dude rechazar o no la hipótesis nula dada la

cercanía entre el valor crítico y el estadístico calculado, medida esta corta distancia por el p-valor y el nivel de significancia, que indica además iii) la intensidad del rechazo. También se propuso la creación de una categorización para la iv) probabilidad de cometer el error del Tipo I, la cual se muestra en estos resultados. Finalmente, se muestra el potencial del v) recurso tecno didáctico con valores agregados a Excel

**i) Categorizar el Rango probabilístico del p-valor;**

El Rango probabilístico del p-valor, se refiere al intervalo real de valores decimales de probabilidad, comprendidos entre el cero (0) en el extremo de la cola y el otro extremo definido por el nivel de significancia alfa ( $\alpha$ ). Este último se define según las colas

- Para una cola a la derecha o a la izquierda, el intervalo es:  $(0; \alpha)$ . Por ejemplo, si el nivel de significancia alfa es  $\alpha = 0,05$  entonces el rango es:  $(0 ; 0,05)$ . El baremo resultante se obtiene dividiendo entre 5 este intervalo real.
- Para dos colas, el intervalo, es:  $(0; \alpha/2)$ . Por ejemplo, si el nivel de significancia alfa es  $\alpha = 0,05$  entonces el rango es:  $( 0 ; 0,025)$ .

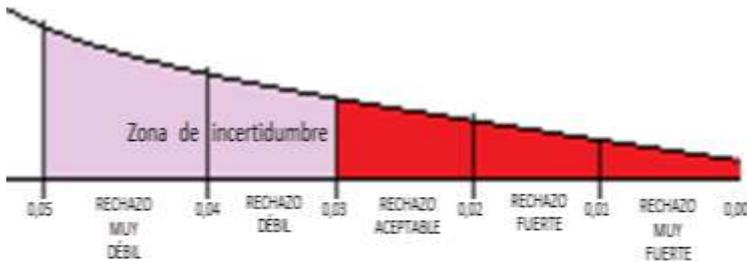
A continuación un *recurso tecno digital* mostrado como un hipervínculo con Excel, que genera dos baremos para cada caso, en función del valor  $\alpha$ , introducido en la celda gris de la esquina superior derecha (Hacer doble clic sobre la figura para activar el hipervínculo).

BAREMOS PARA EL P-VALOR (1 y 2 colas)				
Nivel de Significancia		$\alpha =$		0,05
Una cola		Dos colas		Categoría
0,000	0,010	0,000	0,005	Rechazo Muy fuerte
0,011	0,020	0,006	0,01	Rechazo Fuerte
0,021	0,030	0,011	0,015	Rechazo aceptable
0,031	0,04	0,016	0,02	Rechazo debil
0,041	0,05	0,021	0,025	Rechazo muy debil

Fuente: Diseño Propio

## ii) Definir zona de incertidumbre en la zona de rechazo,

Esta región se refiere a las dos primeras de las cinco categorías, referente a la fortaleza-debilidad del rechazo, en que quedó dividida la region de la cola.



Como puede verse, para el caso de una cola derecha, de área 0,05; en los dos primeros quintiles (0,05 – 0,03), se muestran las decisiones débiles para rechazar la hipótesis nula. En el otro extremo se muestran las fortalezas, quedando en el centro una zona de rechazo aceptable.

## iii) Intensidad de la fuerza de rechazo

Para precisar la intensidad de la fuerza de rechazo de la hipótesis nula, basta con ubicar el valor obtenido del p-valor, en el intervalo ubicado en la zona de rechazo. Por ejemplo, si  $p\text{-valor} = 0,034$ , el rechazo es débil entonces y se está en la zona de incertidumbre por lo cual se recomienda incrementar la muestra. Para un  $p\text{-valor} < 0,03$  se rechaza la hipótesis nula, sin objeción con un rechazo aceptable, fuerte o muy fuerte.

## iv) Categorizar la probabilidad de cometer el error Tipo I

Esta categorización de la probabilidad de cometer el error Tipo I, se construyó en cinco (5) niveles de probabilidad, de manera inversamente proporcional a la intensidad de la fuerza de rechazo, esto es, a mayor fuerza de rechazo, menor es la probabilidad de cometer este tipo de error. Para  $p\text{-valor} > 0,03$  se tiene alta o muy alta probabilidad de cometer este error puesto que se está en la zona de incertidumbre y se recomienda no decidir sobre la hipótesis nula y decidir ampliar el tamaño de la muestra.

BAREMO PARA EL ERROR TIPO I	
Rechazo Muy fuerte	Muy poco probable
Rechazo Fuerte	Poco Probable
Rechazo aceptable	Medianamente Probable
Rechazo débil	Alta probabilidad
Rechazo muy débil	Muy alta probabilidad

**v) La enseñanza virtual de los elementos del análisis inferencial.**

Para la enseñanza virtual del  $p$ -valor y los detalles que le acompañan, se diseñó el siguiente complejo, especie de calculadora-procesadora de datos y a la vez un recurso tecno didáctico, que permita visualizar estos elementos y la relación notable en su comportamiento. Se diseñaron tres unidades en concordancia con los tres tipos de colas en las pruebas estadísticas, acompañada cada una de su respectivo gráfico.

- En cada una de las tres unidades hay 5 módulos: en el lado izquierdo se dispone de un espacio, de celdas grises, para la data.
- En la parte central y arriba, está el módulo de la prueba de hipótesis. En él están todos los elementos del test, incluyendo el  $p$ -valor y la decisión automatizada sobre el rechazo no, de la hipótesis nula.
- Seguidamente y debajo, está un cuadro teórico con el baremo para la intensidad del  $p$ -valor y probabilidad del error Tipo I. En el mismo se identifica la zona de incertidumbre, donde se sugiere ampliar la muestra.
- Finalmente, se muestra en un pequeño recuadro el sistema de hipótesis estadísticas.

PRUEBAS DE HIPÓTESIS PARA LA MEDIA											
DATA			1. PRUEBA DE COLA DERECHA								
1	167		$\mu_0$	Nivel de Significancia	$Z_{\alpha}$	$\bar{x}$	desvst	n	Zc	P	
2	169		165,58	0,01	2,33	167,47	4,42	Muestral	15	1,65	0,049145
3	176		Decisión: Aceptar								
4	170										
5	169										
6	168										
7	177		BAREMO PARA EL P-VALOR								
8	165		pt	Rango del P-valor	Categoría del rechazo	Evidencia	P (Error Tipo I)				
9	164		1	0,000 - 0,002	Rechazo Muy fuerte	Zona de Certidumbre	Muy poco probable				
10	167		2	0,003 - 0,004	Rechazo Fuerte		Poco Probable				
11	167		3	0,005 - 0,006	Rechazo aceptable	Mediamente Probable					
12	164		4	0,007 - 0,008	Rechazo debil	Zona de Incertidumbre	Alta probabilidad				
13	165		5	0,009 - 0,010	Rechazo muy debil		Muy alta probabilidad				
14	163		RESULTADOS								
15	161		$H_0$	$H_1$	$P$	Categoría del rechazo	Zona de Evidencia	P (Error Tipo I)			
						0	0	0			



## Referencias bibliográficas

Campos Souto, Mar; Pérez Pascual, José Ignacio (2020). **El Diccionario de la Real Academia Española: ayer y hoy**. Universidad da Coruña, Servicio de Publicacións.

Carrasquero, Edwin (2023). Excelia: recurso para digital tecno-didáctico de valores. Agregados a Excel. **Revista encuentro Educativo**. FHE de LUZ. Mbo. Venezuela

Díaz, Ladislao y Ríos, Fernando (2018). El valor p. Interpretación, orígenes y su utilización. **Revista Argentina de Terapia Intensiva**. Vol. 35 Núm. 3.

Molina Arias, Manuel (2017). ¿Qué significa realmente el valor de p? Rev. **Pediatría Atención Primaria** [online]. vol.19, n.76, pp.377-381. ISSN 1139-7632.