

ANÁLISIS ESTADÍSTICO DE DATOS DIFUSOS COMPARATIVO DE LA RENTABILIDAD DE LA PRODUCCIÓN DE MAÍZ MORADO EN EL DISTRITO DE HUANTA, PERÚ

Romel Ángel Molina Huaila^{1*}, Ketty Marilú Moscoso Paucarchuco^{*}, Rafael Jesús Fernández Jaime^{**}, Michael Raiser Vázquez Ramírez^{**}

^{*}Universidad Nacional Autónoma de Huanta, Huanta, Perú

^{**}Universidad Peruana los Andes, Huancayo, Perú

RESUMEN

El maíz morado (*Zea mays L.*) es una especie que se consume altamente en el Perú, particularmente en Lima, es por ello que tiene un gran valor comercial y nutritivo. El objetivo de este artículo es estudiar la rentabilidad de la producción del maíz morado (*Zea mays L.*) en las comunidades agrícolas de Maynay y Pampachacra del Distrito de Huanta, Ayacucho, Perú. La rentabilidad es el indicador de que la producción de maíz produce un valor agregado significativo con respecto a las inversiones y los gastos que se necesitan realizar para su producción. Existen indicios de que la rentabilidad de este cultivo no es la adecuada debido a que los cultivadores no aplican las técnicas apropiadas que se necesitan. Para comprobar que exista una mayor rentabilidad significativa en la producción de maíz morado de una comunidad sobre la otra se utiliza la prueba de t-Student, que es una prueba para comparar si dos muestras de datos provienen de la misma población. Las mediciones se hicieron mediante una encuesta que mide el criterio de los productores locales de maíz. En la encuesta se sustituyeron los datos numéricos por números difusos para tener en cuenta la vaguedad de las opiniones de los encuestados. Esto conduce a métodos de números aleatorios difusos para procesar los datos.

PALABRAS CLAVES: Rentabilidad económica, rentabilidad financiera, maíz morado, prueba estadística de t Student para muestras independientes, datos difusos, número aleatorio difuso.

MSC: 60A86, 62A86, 62P05, 62P10, 62P12, 62P20, 62P30.

ABSTRACT

Purple corn (*Zea mays L.*) is a species that is highly consumed in Peru, particularly in Lima, which is why it has great commercial and nutritional value. The purpose of this paper is to study the profitability of the production of purple corn (*Zea mays L.*) in the agricultural communities of Maynay and Pampachacra in the District of Huanta, Ayacucho, Peru. Profitability is the indicator that corn production produces a significant added value with respect to the investments and expenses needed for its production. There are indications that the profitability of this crop is not adequate because the growers do not apply the appropriate techniques that this crop needs. To verify that there is greater enough profitability in the production of purple corn from one community over the other, the t-Student test is used, which is a test to compare whether two samples come from the same population. We carried out the measurement by means of a survey about the criteria of local corn farmers. In the survey, we substituted numerical data by fuzzy numbers for dealing with the vagueness of interviewed's criteria. This leads to the methods of random fuzzy numbers for data processing.

KEYWORDS: Economic profitability, financial profitability, purple corn, t-Student statistical test for independent samples, fuzzy data, random fuzzy number.

1. INTRODUCCIÓN

El maíz morado es ampliamente consumido en Perú, particularmente en la ciudad de Lima es utilizado en la elaboración de bebidas típicas. El color particular de este tipo de maíz es el morado y presenta corontas y mazorcas de este color. La producción anual es de 6000 toneladas, tiene un alto precio en el mercado limeño y también es un producto de exportación ([12]). Hoy día se cultiva aproximadamente 4000 hectáreas, específicamente en los valles occidentales y sierra del Perú ([12]).

En cuanto a la producción mundial de maíz, los investigadores Cruz y otros ([4]) redactaron en Texcoco, México el artículo "Análisis de la producción mundial, nacional y estatal de maíz (*Zea mays L.*)", con el objetivo de obtener la

tasa de crecimiento promedio anual y examinar la producción de maíz a nivel mundial, nacional y local, por ello trata de una investigación que recogió información de los primeros países productores de granos de maíz entre los años 1995-2012 y también sobre los importantes productores de biocombustibles entre los años 2007-2011; de tal manera que en el presente estudio se examinó el progreso de la cantidad de tierra sembrada, utilidades y cantidad de cosecha desde 1995 a 2003 y desde 2004 a 2012.

En el ámbito local se tiene que Álvarez ([2]) redactó en Ayacucho, Perú un artículo con los objetivos de: a) Medir el beneficio de panojas de maíz morado por la utilización de abonos agrícolas y cantidad de plantas, b) estudiar el atributo de panojas afectados por cantidad de plantas y fertilización, y c) Estudiar el índice de tinción de mazorcas. Por otra parte, la rentabilidad está presente en todas las actividades económicas que el ser humano desarrolla, ya que esta mide el beneficio que se puede obtener en una situación determinada como denominador común de las actividades productivas. Quispe y Valdiviezo ([16]) mencionan que la rentabilidad es "el resultado del proceso productivo", es además un indicador relevante para medir el éxito de un negocio.

En el Perú la sequía causa grandes pérdidas en los cultivos y los agricultores de la sierra peruana aún practican la agricultura tradicional en donde los cultivos de maíz son afectados por las plagas, enfermedades y el cambio climático. En la provincia de Huanta, región Ayacucho, Perú, el sector económico tiene mayor arraigo en la agricultura, de tal manera que en el valle de Huanta se realiza la siembra del maíz morado entre los meses de octubre y diciembre, que es un cultivo transcendental para las familias rurales, pero se origina en contextos inconstantes, sin que aparentemente sea rentable para la localidad.

El desconocimiento para generar mayor rentabilidad en el cultivo de maíz morado tiene como causa principal la escasa capacitación en temas de comercialización y débil apoyo de los gobiernos locales y nacionales; así como también la poca participación por parte de las instituciones como el Ministerio de Agricultura, generando la incertidumbre de no saber si se está ganando o perdiendo en su producción, es decir, el desconocimiento de si se tiene o no la debida rentabilidad en su comercialización.

El objetivo principal de este artículo es comparar el nivel de rentabilidad de la producción del maíz morado (*Zea mays L.*) en las comunidades de Maynay y Pampachacra del Distrito de Huanta, Ayacucho, Perú. La presente investigación permitirá identificar las causas que limitan la producción de maíz morado y también conocer los factores externos e internos que influyen en la rentabilidad de la producción del maíz morado; así como también dar a conocer la rentabilidad de dicho cultivo en las comunidades de Maynay y Pampachacra en el distrito de Huanta, Provincia de Huanta de la región de Ayacucho.

Para lograr este objetivo se realizó una encuesta a 24 productores de maíz de la Comunidad de Pampachacra y 26 productores de la Comunidad de Maynay con el fin de cuantificar el estado de 5 indicadores de rentabilidad en estas comunidades y el distrito en general, estos son la Rentabilidad Económica, Rentabilidad Financiera, Costos por proceso, Cantidades de Producción y Nivel de rentabilidad de la producción. La escala de evaluación utilizada fue cualitativa con los valores "Bajo", "Medio" y "Alto". Estos datos nominales se convirtieron en Números Difusos Triangulares para no perder información en el procesamiento de estos. Los números difusos son la generalización de los valores numéricos añadiendo una función de pertenencia convexa y normal [7].

Desde el punto de vista estadístico se aplicó un test de t Student que permite determinar si las muestras recogidas en ambas comunidades pertenecen a la misma población ([10]). El test se aplicó a datos difusos, lo cual pertenece a la teoría de Números Difusos Aleatorios, que es la teoría que estudia el procesamiento estadístico de datos difusos ([5]). Los métodos del Análisis Estadístico de Datos Difusos se utilizan usualmente en el procesamiento de encuestas donde los encuestados responden con valores difusos en lugar de con valores numéricos o escalas de Likert. Esta manera tiene la ventaja que los resultados de las encuestas son más exactos en cuanto a lo que quiere expresar el encuestado. Una visión de la teoría de Estadística Difusa se puede encontrar en [3], no obstante en este artículo se usará la teoría desarrollada en [7][8].

El artículo consta de una sección de Materiales y Métodos, donde se explican algunos elementos fundamentales sobre la rentabilidad y las pruebas estadísticas utilizadas. La sección 3 de Resultados resume los resultados obtenidos en la investigación. La sección 4 se dedica a dar las conclusiones. El Apéndice A contiene el código en lenguaje m del algoritmo desarrollado por los autores para realizar los cálculos que aparecen en los resultados.

2. MATERIALES Y MÉTODOS

2.1. Rentabilidad

Se define la rentabilidad como la capacidad que tiene un producto, sistema, servicio, entre otros, para poder crear suficientes ganancias; por ejemplo, una organización es rentable cuando tiene más entradas que salidas. Así mismo

una corporación que genera más ganancias que pérdidas es rentable.

Por ello una de las definiciones más exactas de rentabilidad es que mide la correlación entre la ganancia o beneficio obtenido. Para calcular esa rentabilidad, se debe dividir la ganancia entre la inversión y luego multiplicar el resultado por 100 para que se pueda expresar en porcentaje ([6]).

$$\text{Rentabilidad} = (\text{Utilidad o Ganancia} / \text{Inversión}) \times 100$$

Está compuesto por 2 segmentos: Beneficios obtenidos y Recursos empleados.

Rentabilidad económica:

Estudia la capacidad de los activos de la organización productiva, como por ejemplo las maquinarias e inmuebles, y de esa forma obtener beneficios totales, aquel donde no hay descuentos, intereses e impuestos que se tienen que costear en las diferentes instancias ([15]).

La Rentabilidad económica consta de las dos razones más importantes:

- ROI: Retorno sobre la inversión. Mide la rentabilidad que gana la institución sobre sus activos, así mismo la eficacia del activo en alguna utilización. Se halla con la siguiente fórmula:
Beneficios brutos/Activos totales.
- ROE: Retorno sobre activos propios. Prana ([15]) menciona que es el beneficio neto logrado, en paralelo con la inversión de los asociados, se halla como:
Beneficio neto/ Activos propios o patrimonio propio.

Rentabilidad financiera:

En este tipo de rentabilidad no se toma en consideración la deuda utilizada para generar ganancias, es por ello que se define como el rendimiento que se genera al realizar una inversión ([15]).

Se halla el beneficio final tras quitar los intereses e impuestos. La razón más significativa para su cálculo es:

- ROA: Retorno sobre activos. Deduce la correspondencia entre rendimiento y los activos totales de la compañía, se halla de la siguiente forma:
Beneficios netos/Activos totales.

Algunos conceptos econométricos que se utilizan en este artículo son los siguientes:

Producción: Es el trabajo de producir algo útil, puede ser un bien o un servicio, el fin de la producción es tener un producto que tenga un valor añadido ([19]).

Costo de producción: Tapia ([20]) señala a los costos de producción u operación como aquellos gastos que se originan para preservar un proyecto o para mantener una línea de equipo en buen funcionamiento. Por ejemplo, en una empresa estándar el beneficio bruto es la resta entre el costo de producción y el ingreso. Es por ello que el ingreso está muy relacionado con el destino económico de una empresa o compañía.

Rentabilidad económica: Sánchez ([17]) lo define como una medida alusiva a una determinada etapa del tiempo de los rendimientos de los activos de una compañía con libertad de financiación de los mismos. Por tanto, se define como una medida de la capacidad de los activos de una compañía para crear valor, es eso lo que facilita la asimilación de la rentabilidad entre empresas.

Rentabilidad financiera: Sánchez ([17]) explica que es una medida de los ingresos obtenidos por el capital propio en un determinado período de tiempo, y generalmente no considera la distribución de los resultados. Por lo tanto, la rentabilidad financiera puede pensarse como una medida más cercana a la rentabilidad de los accionistas o propietarios que a la rentabilidad económica, por lo que, teóricamente, según la visión más común, es lo que buscan los gerentes en beneficio de los propietarios para maximizar las métricas de rentabilidad.

Costos por procesos: Según Sánchez ([17]) los costos por procesos se definen como un sistema que agrupa todos los gastos que se realizan con respecto a un proceso conocido, además se utiliza cuando la producción es en cantidad o de forma continua. Así menciona el costo en la actividad agrícola.

Costo en la actividad agrícola: Estos costos son considerados básicos ya que se gastan en materia prima o elementos básicos como por ejemplo: semillas, insumos, mano de obra. Están los costos indirectos de producción como las maquinarias, riego, depreciación de materiales y maquinarias, entre otros.

Cantidad de producción: Tapia ([20]) menciona que la producción es la suma de un valor ya sea a un producto o un servicio, esto puede ser por un proceso. Es por ello que producir es darle un valor añadido a los bienes

con el fin de convertirlos en consumibles. Además, la producción puede clasificarse en producción de bienes físicos o materiales o servicios.

2.2. Números aleatorios difusos

Definición 1. ([7]) Un *número difuso* es una aplicación $\bar{U}: \mathbb{R} \rightarrow [0, 1]$ que es normal, convexa y tiene niveles compactos. Esto significa que los niveles- α dados por $\bar{U}_\alpha = \{x \in \mathbb{R}: \bar{U}(x) \geq \alpha\}$ si $\alpha \in (0, 1)$, $\bar{U}_0 = cl\{x \in \mathbb{R}: \bar{U}(x) > 0\}$, donde cl denota la clausura del correspondiente conjunto; estos son intervalos acotados y cerrados. $\bar{U}(x)$ se interpreta como el grado de compatibilidad de x con la propiedad que define a \bar{U} .

Un caso particular de número difuso es el *número difuso triangular*, denotado por $A = Tri(a, b, c)$ que tiene la forma que se aprecia en la Figura 1. Donde:

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a}, & \text{si } a \leq x \leq b \\ \frac{c-x}{c-b}, & \text{si } b \leq x \leq c \\ 0, & \text{en otro caso} \end{cases}$$

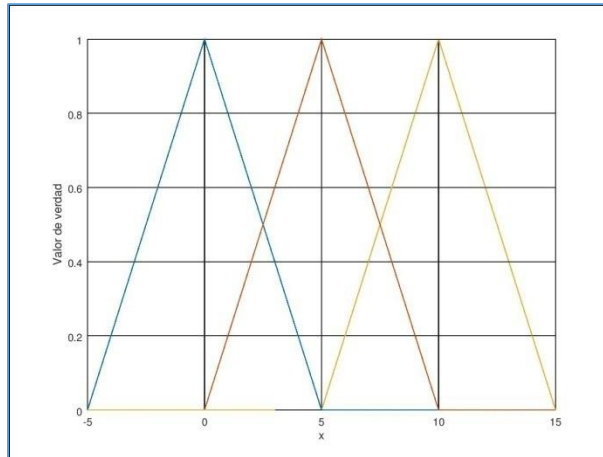


Figura 1. Ejemplo de tres números difusos triangulares. De izquierda a derecha: $Tri(-5,0,5)$, $Tri(0,5,10)$ y $Tri(5,10,15)$.

En el caso de los números difusos triangulares se cumple:

- $\bar{U}_\alpha = [\alpha b + (1 - \alpha)a, \alpha b + (1 - \alpha)c]$, $\alpha \in [0, 1]$,
- $\bar{U}_1 = \{b\}$ que se considera el valor que satisface completamente la propiedad \bar{U} ,
- $\bar{U}_0 = [a, c]$ que son todos los valores compatibles con \bar{U} con algún valor de verdad.

Definición 2. ([7]) Sean \bar{U} y \bar{V} dos números difusos. La *suma* de ambos se denota por el número difuso $\bar{U} + \bar{V}$ y se define a partir de sus niveles- α como: $(\bar{U} + \bar{V})_\alpha = [inf\bar{U}_\alpha + inf\bar{V}_\alpha, sup\bar{U}_\alpha + sup\bar{V}_\alpha] \forall \alpha \in [0, 1]$.

Definición 3. ([7]) Sea \bar{U} un número difuso y γ un número real. El *producto* \bar{U} por γ , que se denota por $\gamma\bar{U}$ es un número difuso que se define a partir de sus niveles- α como: $(\gamma\bar{U})_\alpha = \begin{cases} [\gamma inf\bar{U}_\alpha, \gamma sup\bar{U}_\alpha], & \text{si } \gamma \geq 0 \\ [\gamma sup\bar{U}_\alpha, \gamma inf\bar{U}_\alpha], & \text{si } \gamma < 0 \end{cases} \forall \alpha \in [0, 1]$.

De manera más sencilla se tienen las siguientes operaciones sobre números difusos triangulares $A_1 = Tri(a_1, b_1, c_1)$ y $A_2 = Tri(a_2, b_2, c_2)$ ([11]):

- $A_1 + A_2 = Tri(a_1 + a_2, b_1 + b_2, c_1 + c_2)$.
- $A_1 - A_2 = Tri(a_1 - c_2, b_1 - b_2, a_2 - c_1)$
- $\gamma A_1 = Tri(\gamma a_1, \gamma b_1, \gamma c_1)$, si $\gamma \geq 0$; $\gamma A_1 = Tri(\gamma c_1, \gamma b_1, \gamma a_1)$, si $\gamma < 0$.

Definición 4. ([11]) Sean \bar{U} y \bar{V} dos números difusos triangulares $\bar{U} = Tri(a_1, b_1, c_1)$ y $\bar{V} = Tri(a_2, b_2, c_2)$. La *distancia* entre ellos se define por la ecuación siguiente:

$$D(\bar{U}, \bar{V}) = \sqrt{\frac{1}{6}[(a_1 - a_2)^2 + 4(b_1 - b_2)^2 + (c_1 - c_2)^2]} \quad (1)$$

Definición 5. ([7]) Dado un experimento aleatorio modelado en un espacio de probabilidades $(\Omega, \mathcal{A}, \mathcal{P})$. Un número aleatorio difuso (NAD) tiene asociada una aplicación χ definida sobre Ω de tipo número difuso y tal que $\forall \alpha \in [0, 1]$ se cumple que las aplicaciones reales $\inf \chi_\alpha$ y $\sup \chi_\alpha$ son variables aleatorias reales. Donde, $\inf \chi_\alpha(\omega) = \inf(\chi(\omega))_\alpha$ y $\sup \chi_\alpha(\omega) = \sup(\chi(\omega))_\alpha$.

Definición 6. ([7]) Dado un experimento aleatorio y un NAD χ asociado, la *esperanza matemática* de χ (de tipo Aumann) es el número difuso $\bar{E}(\chi)$, si este existe, tal que $\forall \alpha \in [0, 1]$ se tiene:

$$(\bar{E}(\chi))_\alpha = [E(\inf \chi_\alpha), E(\sup \chi_\alpha)] \quad (2)$$

Donde E es la esperanza matemática de la variable aleatoria asociada.

En particular, si $\tilde{x}^{(n)} = (\tilde{x}_1, \dots, \tilde{x}_n)$ es una muestra de observaciones del NAD χ , entonces la *media muestral* es el número difuso $\bar{\tilde{x}}^{(n)}$ tal que $\forall \alpha \in [0, 1]$:

$$\bar{\tilde{x}}^{(n)} = \left[\frac{\inf(\tilde{x}_1)_\alpha + \dots + \inf(\tilde{x}_n)_\alpha}{n}, \frac{\sup(\tilde{x}_1)_\alpha + \dots + \sup(\tilde{x}_n)_\alpha}{n} \right] \quad (3)$$

Definición 7. ([7]) Dado un experimento aleatorio y un NAD χ asociado, la *varianza- ρ_2* de χ es el número real $Var_{\rho_2}(\chi)$, si esta existe, se define por:

$$Var_{\rho_2}(\chi) = E \left([\rho_2(\chi, \bar{E}(\chi))]^2 \right) \quad (4)$$

2.3. Pruebas estadísticas

Una de las pruebas t de Student consiste en determinar si las medias entre dos muestras son significativamente iguales, para probar que ambas pertenecen a la misma población ([10][13][14]). En este artículo se tiene que ambas muestras tienen diferentes tamaños, que es a lo que se llama muestras independientes o desapareadas. El estadígrafo de esta prueba se calcula con ayuda de la Ecuación 5.

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)}} \quad (5)$$

donde: $s_j^2 = \frac{\sum_{i=1}^{n_j} (x_i - \bar{x}_j)^2}{n_j - 1}$ con $j = 1, 2$. En este caso se considera que las desviaciones estándar de ambas muestras son diferentes.

La prueba de hipótesis consiste en probar que las medias de ambas poblaciones son iguales como hipótesis nula, o sea, si μ_1 y μ_2 denotan las medias de las dos poblaciones, entonces:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2.$$

En el caso de que ambas muestras tengan diferentes tamaños y diferentes valores de la desviación estándar el grado de libertad de la distribución t de Student se calcula por la fórmula siguiente:

$$g.l. = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\left(\frac{s_1^2}{n_1} \right)^2 / (n_1 - 1) + \left(\frac{s_2^2}{n_2} \right)^2 / (n_2 - 1)} - 2 \quad (6)$$

El valor p se calcula fijando α en este caso $\alpha = 0,05$, se calcula el valor $t_{1-\alpha, g.l.}$ y se toma la decisión de la siguiente manera:

Si $t_{1-\alpha, g.l.} > T$, no se rechaza la hipótesis nula de medias iguales, en caso contrario se rechaza la hipótesis nula de medias iguales.

La hipótesis inicial para aplicar la prueba t de Student es la normalidad de los datos. Esta es conveniente cuando el tamaño de la muestra es pequeño.

3. RESULTADOS

Se encuestaron a 24 cultivadores de maíz de la comunidad de Pampachacra y 26 de la comunidad de Maynay seleccionados aleatoriamente con muestreo aleatorio simple, sobre la rentabilidad del maíz en sus lugares de cultivo, 50 en total. Estos evaluaron la rentabilidad de la producción, la rentabilidad económica, la rentabilidad financiera, el costo por proceso de la producción y las cantidades de producción. Se les pidió que las evaluaciones las hicieran en una escala cualitativa de “Bajo”, “Medio” y “Alto”.

Los resultados de la encuesta se dan en las tablas 1-5, que en sí mismos son resultados de interés.

		Rentabilidad económica			Total	
		Bajo	Medio	Alto		
Lugar de procedencia	Pampachacra	f 1	19	4	24	
		%	4,2%	79,2%	16,7%	100%
	Maynay	f 1	21	4	26	
		%	3,8%	80,8%	15,4%	100%
Total		f 2	40	8	50	
		%	4,0%	80,0%	16,0%	100,0%

Tabla 1. Nivel de rentabilidad de la producción de las comunidades de Maynay y Pampachacra. Fuente: Elaboración propia.

Se encontró que la mayoría, es decir 40 productores (80%) tienen una rentabilidad media. Como se observa en la Tabla 1, en la comparación, ambas comunidades tienen niveles similares de rentabilidad de producción, predominando el nivel medio con un 79,2% en la comunidad de Pampachacra y 80,8% en la comunidad de Maynay, el nivel bajo con cerca de un 4% en ambas comunidades y un promedio de 16% de nivel de rentabilidad alto en ambas comunidades.

		Rentabilidad económica			Total	
		Bajo	Medio	Alto		
Lugar de procedencia	Pampachacra	f 3	17	4	24	
		%	12,5%	70,8%	16,7%	100%
	Maynay	f 4	20	2	26	
		%	15,4%	76,9%	7,7%	100%
Total		f 7	37	6	50	
		%	14,0%	74,0%	12,0%	100%

Tabla 2. Nivel de Rentabilidad económica de las comunidades de Maynay y Pampachacra. Fuente: Elaboración propia.

En la Tabla 2 se observa que la comunidad Pampachacra tiene más productores con niveles altos de rentabilidad económica (16,7%), mientras que la comunidad Maynay tiene mayor frecuencia de productores con nivel medio (77%) respecto a la comunidad Pampachacra (71%). Es decir, ambas comunidades tienen apreciables niveles de rentabilidad económica, predominando el nivel medio con un 76,9% en Maynay y 70,8% en Pampachacra.

		Rentabilidad financiera			Total
		Bajo	Medio	Alto	
Lugar de procedencia	Pampachacra	f 2	18	4	24
		% 8,3%	75,0%	16,7%	100%
	Maynay	f 1	21	4	26
		% 3,8%	80,8%	15,4%	100%
Total		f 3	39	8	50
		% 6,0%	78,0%	16,0%	100%

Tabla 3. Nivel de Rentabilidad financiera de las comunidades de Maynay y Pampachacra. Fuente. Elaboración propia.

Como se puede apreciar en la Tabla 3, si bien en la comunidad Pampachacra hay menores frecuencias absolutas de niveles medios, no existe una diferencia significativa en los niveles de rentabilidad financiera en la comparación entre los productores de Pampachacra o Maynay. Es decir, en términos porcentuales ambas comunidades tienen apreciables niveles de rentabilidad financiera, predominando el nivel medio con un 80,8% en Maynay y 75,0% en Pampachacra.

		Costos por procesos			Total
		Bajo	Medio	Alto	
Lugar de procedencia	Pampachacra	f 1	19	4	24
		% 4,2%	79,2%	16,7%	100%
	Maynay	f 3	19	4	26
		% 11,5%	73,1%	15,4%	100%
Total		f 4	38	8	50
		% 8,0%	76,0%	16,0%	100%

Tabla 4. Nivel de Costo por proceso de la producción de las comunidades de Maynay y Pampachacra. Fuente. Elaboración propia.

Como se aprecia en la Tabla 4, los niveles medios y altos de costo por proceso son bastante similares en ambas comunidades. La comunidad Maynay presenta una mayor frecuencia de nivel de costo por procesos bajo (11,5%) respecto a la comunidad Pampachacra (4,2%). En términos de frecuencias relativas, el 95,9% del total de productores de Pampachacra cuenta con niveles medios y altos de costos por procesos, mientras que en la comunidad Maynay, estos niveles suponen el 88,5%. Es decir, predomina el nivel medio con un 79,2% para la comunidad de Pampachacra y 73,1% para Maynay.

Como apreciamos en la Tabla 5, en la comunidad de Pampachacra el nivel medio de cantidades de producción supone el 79,2% de los productores, mientras que en Maynay el nivel medio supone solo el 50% del total; por su parte, el nivel bajo en Maynay (7,27%) es mayor que en Pampachacra (2,8%). Esta diferencia sucede también con el nivel alto, representando un 23% en Maynay y un 12,5% en Pampachacra.

El próximo paso es determinar si la diferencia entre los valores de ambas comunidades es significativa. Una de las maneras de ganar en exactitud cuantitativa de los resultados, teniendo en cuenta que la escala de medición es cualitativa y subjetiva, es procesando los datos como si fueran datos difusos en lugar de los métodos tradicionales de la estadística clásica ([21]).

		Cantidades de producción			Total
		Bajo	Medio	Alto	
Lugar de procedencia	Pampachacra	f 2 % 8,3%	19 79,2%	3 12,5%	24 100%
	Maynay	f 7 % 26,9%	13 50,0%	6 23,1%	26 100%
Total		f 9 % 18,0%	32 64,0%	9 18,0%	50 100%

Tabla 5. Nivel de Cantidades de producción de las comunidades de Maynay y Pampachacra. Fuente. Elaboración propia.

Para ello se siguieron los pasos siguientes:

1. Las evaluaciones de “Bajo” se identificaron con el número difuso $Tri(-5,0,5)$, $Tri(0,5,10)$ se identificó con “Medio” y $Tri(5,10,15)$ con “Alto”, véase Figura 1.
2. Se calcula el estadígrafo T de la Ecuación 5 de los números difusos de las encuestas. Para ello se aplicaron las fórmulas 3 y 4. Los cálculos se hacen más sencillos porque se trata con números difusos triangulares cuyas operaciones aritméticas se realizan de manera directa.
3. El número difuso obtenido del paso anterior se desfuzzifica con $\hat{x} = \bar{U}_1$. Aquí se utiliza la simplicidad que significa operar con números difusos triangulares. Es un caso particular de los esquemas de desfuzzificación que aparecen en [9]. Este valor coincide con los tres esquemas que son: el mínimo de \bar{U}_1 , el máximo de \bar{U}_1 y el valor medio de \bar{U}_1 .
4. El valor numérico real obtenido se utiliza para realizar las pruebas de hipótesis de la prueba t de Student. El valor del grado de libertad con la Ecuación 6, se realiza con las operaciones algebraicas entre números difusos que indica la fórmula y la manera de desfuzzificación indicada en el paso anterior.

Para los cálculos se utilizó el software Octave 4.2.1. Este es un software libre muy útil para cálculos numéricos y estadísticos, que se puede descargar de manera gratuita y que emula a MATLAB en los cálculos que realiza, puesto que permite programar en lenguaje m ([1]). Para realizar los cálculos usados en este artículo se programó un algoritmo en lenguaje m cuyo código aparece en el Apéndice A.

En este párrafo se trata el tema de la normalidad de cada una de las muestras. Por pruebas realizadas por los autores sobre las muestras, se evidencia que esta hipótesis no se cumple. No obstante, de acuerdo a [18], demostraron que la prueba t de Student es aún robusta cuando no existe normalidad entre los elementos de las dos muestras, pero el tamaño de ambas es lo que denominan “bastante grande”, que son muestras de entre 25 y 30 elementos, lo que se cumple aproximadamente en las muestras utilizadas en los experimentos mostrados en esta investigación. Es por ello que se consideran válidos los resultados siguientes obtenidos de los experimentos.

En las Tablas 6 y 7 se resumen los resultados de la prueba estadística de t Student con respecto a cada uno de los indicadores estudiados.

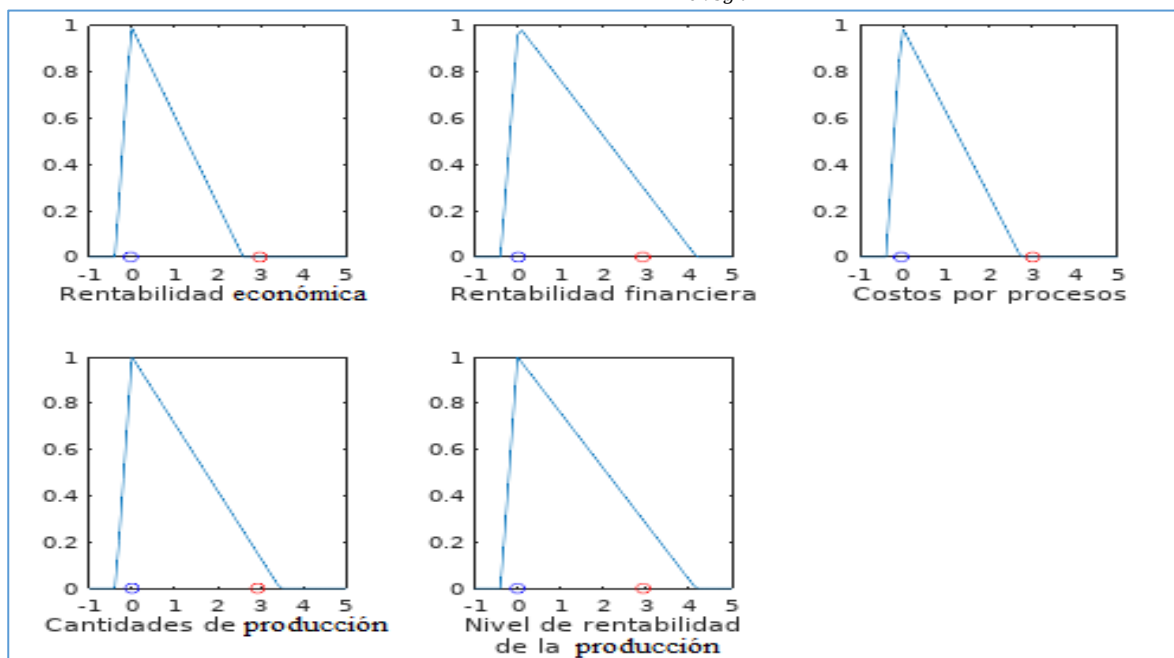
	Rentabilidad económica	Rentabilidad financiera	Costos por procesos
Número Difuso	$[-0.385055,$	$[-0.389210, 0.016041,$	$[-0.395811, -0.042693,$
T	$-0.028010, 2.592725]$	$4.176345]$	$2.745626]$
Grado de libertad desfuzzificado	1.9418	1.9996	1.9027
$t_{0.95g.l.}$	2.980520	2.920356	3.024457

Tabla 6. Pruebas estadísticas con datos difusos para los 3 primeros indicadores. Fuente: Elaboración propia. Como se puede apreciar de las Tablas 6 y 7 se tiene que el valor teórico de la t de Student es mayor que el valor del estadígrafo T desfuzzificado. Es por ello que no se rechaza la hipótesis nula de igualdad de las medias y por tanto se puede decir que no existe diferencia significativa entre los 5 indicadores del maíz, especialmente la rentabilidad, de las comunidades de Pampachacra y Maynay. Esto se puede apreciar en la Figura 2, donde aparecen los números

difusos del estadígrafo T de las 5 pruebas t de Student, las circunferencias azules representan sus valores desfuzzificados y las circunferencias rojas representan los valores desfuzzificados de la distribución teórica $t_{0,95g.l.}$.

	Cantidades de producción	Nivel de rentabilidad de la producción
Número Difuso T	$[-0.37787, 0.00000, 3.45214]$	$[-0.39284, 0.00000, 4.15692]$
Grado de libertad desfuzzificado	1.9894	1.9886
$t_{0,95g.l.}$	2.93059	2.93146

Tabla 7. Pruebas estadísticas con datos difusos para los 2 últimos indicadores. Fuente: Elaboración propia. Figura 2. Números difusos triangulares obtenidos de los estadígrafos T de las 5 pruebas. El punto azul es el valor desfuzzificado de T y el punto rojo es el valor desfuzzificado de $t_{0,95g.l.}$.



Fuente: Elaboración propia

4. CONCLUSIONES

Se realizó la medición del nivel de rentabilidad de la producción de maíz morado en dos comunidades de Huanta, involucrando en la ejecución de los instrumentos a 24 productores de Pampachacra y 26 de Maynay. A nivel descriptivo, 8 de cada 10 productores de las comunidades de Pampachacra y Maynay tienen un nivel medio de rentabilidad en la producción de maíz morado (79,2% Pampachacra y 80,8% Maynay). El 4% de productores de ambas comunidades tiene un nivel bajo de rentabilidad y el 16% tiene un nivel alto.

Se llegó a las siguientes conclusiones:

1. El nivel de rentabilidad de la producción del maíz morado en las comunidades de Maynay y Pampachacra es estadísticamente indiferenciable.
2. No hay diferencia, estadísticamente significativa entre el nivel de rentabilidad económica de la producción del maíz morado de las comunidades de Maynay y Pampachacra.
3. El nivel de rentabilidad financiera de la producción del maíz morado es estadísticamente indiferenciable en las comunidades de Maynay y Pampachacra.
4. El nivel de costo por proceso de la producción del maíz morado es estadísticamente indiferenciable en las comunidades de Maynay y Pampachacra del Distrito de Huanta.

5. El nivel de cantidad de producción del maíz morado es estadísticamente indiferenciable en las comunidades de Maynay y Pampachacra.

De acuerdo a las pruebas estadísticas de incidencia del lugar de procedencia sobre los niveles de rentabilidad de las comunidades Pampachacra y Maynay, se concluye que la procedencia de la comunidad de Pampachacra o Maynay no afecta significativamente a la distribución de la variable Nivel de rentabilidad de la producción, ni a las dimensiones Rentabilidad económica o Rentabilidad financiera.

Dado el rechazo de la significancia en la diferencia entre el lugar de procedencia para el nivel de rentabilidad de producción del maíz morado en la población, pensando en posteriores investigaciones se recomienda la intervención de otros factores diferenciadores tales como el nivel de instrucción o la aplicación de implementación tecnológica en los procesos de producción de dicho cultivo.

En el aspecto práctico, se recomienda integrar una comunidad alternativa de participación y asociación para conseguir recursos y asistencia del aparato público, gobiernos locales, regionales y del Ministerio de Agricultura y Riego a través de programas de capacitación técnica e inversión pública en infraestructura productiva para pequeños agricultores y de este modo alcanzar niveles de producción con mayor rentabilidad, que permitan competir en el mercado nacional.

Finalmente, se recomienda regular los precios de manera local, disminuir la venta a intermediarios e integrar economías a escala, pues la venta a intermediarios provoca una reducción en los ingresos de los productores.

RECEIVED: FEBRUARY, 2023.

REVISED: MAY, 2023

REFERENCIAS

- [1]. ALBERTS, C., DOROFEE, A., STEVENS, J. y WOODY, C. (2003): **Introduction to the OCTAVE Approach**. Carnegie.
- [2]. ÁLVAREZ, F. (2016): Fertilización y densidad de plantas en el rendimiento de maíz morado (*Zea mays* L.), Pampa del Arco – Ayacucho, Perú, **Revista Investigación UNSCH**, 24, 35-41.
- [3]. BUCKLEY, J.J. (2004) **Fuzzy Statistics**. Berlin: Springer-Verlag.
- [4]. CRUZ, N., PORTILLO, M., PÉREZ, F., CAAMAL, I. y MARTÍNEZ, M. (2017): Análisis de la producción mundial, nacional y estatal de maíz (*Zea mays* L.) en Texcoco, México. **Agro productividad**, 10, 95-100.
- [5]. DENOEU, T. (2023) Reasoning with fuzzy and uncertain evidence using epistemic random fuzzy sets: general framework and practical models. **Fuzzy Sets and Systems**, 453, 1-36.
- [6]. DF CONSULTORES. Estudio contable notarial. **¿Qué es la rentabilidad y utilidad de un negocio?** Disponible en <https://www.dfconsultores.com.uy/novedades/22-que-es-la-rentabilidad-y-utilidad-de-un-negocio.html>. Consultado: 23 de enero de 2023.
- [7]. GIL, M. A., LUBIANO, M. A., DE LA ROSA DE SÁA, S. y SINOVA, B. (2015) Analyzing data from a fuzzy rating scale-based questionnaire. A case study. **Psicothema**. 27, 182-191.
- [8]. GONZÁLEZ-RODRÍGUEZ, G., COLUBI, A., GIL, M.A. (2012) Fuzzy data treated as functional data: A one-way ANOVA test approach. **Computational Statistics and Data Analysis**, 56, 943–955.
- [9]. JYH-SHING, R.J., SUN, C.-T. and MIZUTANI, E. (1997) **Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence**. Upper Saddle River: Prentice Hall.
- [10]. LUGO ARMENTA, J.G. y PINO FAN, L.R. (2022) Niveles de razonamiento inferencial para el estadístico T-Student. **Bolema: boletim de Educação Matemática**, 35, 1776-1802.
- [11]. MAESE NÚÑEZ, J.D.D., ALVARADO INIESTA, A., VALLES ROSALES, D.J., BÁEZ LÓPEZ, Y.A. (2016) Coeficiente alfa de Cronbach para medir la fiabilidad de un cuestionario difuso. **Culcyt/Instrumentos de medición**, 13, 146-156.
- [12]. MANRIQUE, A. (2000): Maíz morado peruano (*Zea mays* L.) **Instituto nacional de investigación agraria INIA**, 4, 5-18.
- [13]. MOLINA, M. (2022) Paso a paso. Prueba de la t de Student para muestras independientes. **Revista electrónica AnestesiaR**, 14, 4-4.
- [14]. ORTIZ PINILLA, J.E., MORENO, E.C. (2011) ¿Se necesita la prueba t de Student para dos muestras independientes asumiendo varianzas iguales? **Comunicaciones en estadística**, 4, 139-157.
- [15]. PRANA MARKETING GROUP. (27 de octubre de 2020): **¿Qué es la rentabilidad de una empresa y cómo se calcula?** Disponible en: <https://pranagroup.mx/blog/retribucion-para-tu-negocio/que-es-la-rentabilidad-de-una-empresa-y-como-se-calcula-628/>, consultado el 20 de enero de 2022.

- [16]. QUISPE, J. y VALDIVIEZO, W. (2013): Tesis de Grado: **Análisis de un proceso financiero en la determinación de la rentabilidad en la empresa PaperComput**, Universidad Estatal de Milagro, Guayaquil, Ecuador.
- [17]. SÁNCHEZ, J., (2002): **Análisis de Rentabilidad de la empresa**. Universidad de Murcia, España., URL: <http://ciberconta.unizar.es/leccion/anarenta/analisisR.pdf>.
- [18]. SAWIŁOWSKY, S.S. and BLAIR, R.C. (1992) A More Realistic Look at the Robustness and Type II Error Properties of the t Test to Departures From Population Normality. **Quantitative Methods in Psychology**, 3, 352-360.
- [19]. SEVILLA, R y VALDEZ, A (1985): Estudio de la factibilidad del cultivo de maíz morado. **Fondo de promoción y exportación (FOPEX)**: Lima, Perú.
- [20]. TAPIA, T. (2019): Tesis de Grado: **Rentabilidad y costo de producción en el cultivo de zanahoria de los agricultores del distrito de Marca tuna – 2019**, Universidad continental, Huancayo, Perú. URL: https://repositorio.continental.edu.pe/handle/20.500.12394/9129_17
- [21]. VIERTL, R. (2011) **Statistical Methods for Fuzzy Data**. West Sussex: John Wiley & Sons Ltd.

APÉNDICE A

Este apéndice contiene el programa en lenguaje m programado por los autores para realizar los cálculos con ayuda del software Octave 4.2.1.

```

function Prueba_t_fuzzy = Prueba_t_fuzzy(x,y)
%x es el vector con el número de elementos evaluados como “Bajo”, “Medio” y “Alto” en la muestra 1
%y es el vector con el número de elementos evaluados como “Bajo”, “Medio” y “Alto” en la muestra 2

B = [-5,0,5];%Número difuso triangular para “Bajo”
M = [0,5,10]; %Número difuso triangular para “Medio”
A = [5,10,15]; %Número difuso triangular para “Alto”

n1 = length(x);%Tamaño de la muestra 1
n2 = length(y);%Tamaño de la muestra 2

meanX = (x(1)*B+x(2)*M+x(3)*A)/n1;
meanY = (y(1)*B+y(2)*M+y(3)*A)/n2;

mean_num = [meanX(1)-meanY(3), meanX(2)-meanY(2), meanX(3)-meanY(2)];%Diferencia de medias difusas

DE1 = (x(1)*((B-meanX).^2)+x(2)*((M-meanX).^2)+ x(3)*((A-meanX).^2))*(1/(n1-1));%Cuadrado de desviación estándar de la muestra 1
DE2 = (y(1)*((B-meanY).^2)+y(2)*((M-meanY).^2)+ y(3)*((A-meanY).^2))*(1/(n2-1));)%Cuadrado de desviación estándar de la muestra 1

mean_denom = ((DE1/n1)+(DE2/n2)).^(0.5);

T = [mean_num(1)/mean_denom(3),
mean_num(2)/mean_denom(2),mean_num(3)/mean_denom(1)];%Estadígrafo T difuso

gl_num = ((DE1/n1)+(DE2/n2)).^2;

gl_denom = ((DE1/n1).^2)/(n1-1)+((DE2/n2).^2)/(n2-1);

gl = gl_num(2)/gl_denom(2)-2;%Grado de libertad

T

Prueba_t_fuzzy = [tinv(0.95,gl),T(2)];

```