



**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

**Facultad de Ciencias Naturales y Matemáticas**

**Departamento de Matemáticas**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas en cacao a través de Regresión Avanzada”*

**TESIS DE GRADO**

*Previa la obtención del Título de:*

**INGENIERA EN ESTADÍSTICA INFORMÁTICA**

*Presentado por:*

**Karla Mirella Aguaguña Méndez**

Guayaquil – Ecuador

**AÑO: 2014**

## **AGRADECIMIENTO**

Quiero expresar mi sincero y profundo agradecimiento a mis padres por su apoyo.

A mi amiga Patricia por su colaboración incondicional durante la elaboración de este documento.

Al M.Sc. John Ramírez y al Ph.D. Omar Ruiz por sus recomendaciones que han permitido que culmine con una etapa académica.

A la Ing. Esther L. Peralta, Ph.D. por haberme permitido ser parte del gran talento humano que forma el CIBE.

A CICYT por el financiamiento del proyecto.

# DEDICATORIA

A mis padres;

A mis hermanos;

A mi sobrino David Enrique.

# TRIBUNAL DE SUSTENTACIÓN

---

**Ing. Vanessa Salazar, M.Sc.**

**Presidente del Tribunal**

---

**Mat. John Ramírez Figueroa, M.Sc.**

**Director de Tesis**

---

**Ing. David Matamoros, Ph.D.**

**Vocal**

## DECLARACIÓN EXPRESA

“La responsabilidad del contenido de esta tesis de grado, me corresponden exclusivamente; y el patrimonio intelectual de la misma a la **Escuela Superior Politécnica del Litoral**” (Reglamento de Graduación de la ESPOL).

---

**Karla Mirella Aguaguiña Méndez**

## RESUMEN

La presente tesis es parte de un proyecto semilla CICYT 2011 que se ejecutó en el Centro de Investigaciones Biotecnológicas del Ecuador de la Escuela Superior Politécnica del Litoral (CIBE-ESPOL), el mismo que fue dirigido por el jefe de Bioestadística-Bioinformática del Centro.

La tesis tiene como finalidad determinar mediante un modelo estadístico descriptivo y predictivo de regresión avanzada, el efecto de cada uno de los parámetros de calidad de bioles sobre los parámetros de producción, agronómicos y fitosanitarios de plantas de cacao.

Se han desarrollado seis capítulos en los cuales se presenta el marco teórico relacionado al tema y las diferentes técnicas estadísticas empleadas en el desarrollo de esta investigación.

El Capítulo Uno indica el marco teórico correspondiente a los antecedentes en la producción de cacao en Ecuador y de cómo en la actualidad nuestros productores cacaoteros tienen mayor afinidad al uso de enmiendas orgánicas que sean amigables con el ambiente. El Capítulo Dos explica el marco teórico de las diferentes definiciones básicas en Estadística y las diferentes técnicas estadísticas empleadas para tal finalidad. En el Capítulo Tres se hace una descripción de las variables que se usarán en la investigación y de las localidades de donde provienen los datos que serán analizados.

En el Capítulo Cuatro se presenta la correspondiente estadística descriptiva y diagrama de cajas de las variables que han sido estudiadas. Luego en el Capítulo Cinco se presenta el análisis de Regresión Lineal realizado a las variables dependientes, cuya finalidad es encontrar modelos que expliquen a las variables dependientes en función de los parámetros de los físico – químicos, microbiológicos y de los macro/micro nutrientes, además se presenta un análisis de regresión por Componentes Principales mediante el cual se pretende encontrar modelos con un alto porcentaje de explicación

Finalmente, en el Capítulo Seis se presenta los resultados y conclusiones de la investigación desarrollada.

## ÍNDICE GENERAL

RESUMEN.....	I
INDICE GENERAL.....	III
ABREVIATURAS.....	IX
SIMBOLOGÍA.....	X
ÍNDICE DE CUADROS.....	XI
INTRODUCCIÓN.....	1
CAPÍTULO UNO	
1.1 PRODUCCIÓN DE CACAO EN ECUADOR.....	4
1.2 DISTRIBUCIÓN DE LOCALIDADES PRODUCTORAS DE CACAO.....	7
1.3 VARIEDADES DE CACAO QUE SE PRODUCEN EN ECUADOR.....	8
1.4 CONDICIONES CLIMÁTICAS PARA LA PRODUCCIÓN DE CACAO.....	10
1.5 TIPOS DE PLAGUICIDAS PARA LAS ENFERMEDADES DE CACAO.....	10
1.6 EFECTOS DEL BIOL EN CACAO.....	12
1.7 DEFINICIONES.....	14

1.7.1 Enmienda Orgánica.....	14
1.7.2 Biol.....	14
1.7.3 Parámetros físicos-químicos.....	16
1.7.4 Parámetros microbiológicos.....	17
1.7.5 Macronutrientes / micronutrientes.....	17
1.7.6 Parámetros fitosanitarios.....	18
1.7.7 Parámetros de producción.....	18
1.7.8 Prácticas agronómicas.....	18

## CAPÍTULO DOS

INTRODUCCIÓN.....	20
2.1 MÉTODOS Y PROCEDIMIENTOS PREVIOS AL ANÁLISIS.....	21
2.2 MUESTREO.....	21
2.3 ELEMENTOS DEL MUESTREO.....	22
2.3.1 Población Objetivo.....	22
2.3.2 Marco Muestral.....	22
2.3.3 Muestra.....	22

2.4 MUESTREO DIRIGIDO.....	22
2.5 ANÁLISIS DESCRIPTIVO.....	23
2.6 ANÁLISIS INFERENCIAL.....	23
2.7 ANÁLISIS DE CORRELACIÓN.....	24
2.8 REGRESIÓN LINEAL.....	26
2.9 ANÁLISIS DE VARIANZA DE REGRESIÓN LINEAL.....	28
2.10 REGRESIÓN LINEAL MÚLTIPLE.....	29
2.11 MODELOS NO LINEALES.....	32
2.12 VALIDACIÓN DE MODELOS PREDICTIVOS.....	32
2.13 ANÁLISIS MULTIVARIADO.....	33
2.13.1 Componentes Principales.....	33
CAPÍTULO TRES	
3.1 DESCRIPCIÓN DE VARIABLES.....	35
3.2 LOCALIDADES INVESTIGADAS.....	40
3.2.1 Ubicación de las haciendas investigadas.....	40
3.3 RECOLECCIÓN DE DATOS.....	41

## CAPÍTULO CUATRO

INTRODUCCIÓN.....	42
-------------------	----

### 4.1 ANÁLISIS DESCRIPTIVO DE LOS DATOS

4.1.1 Estadística Descriptiva de los Parámetros Físicos – Químicos.....	44
--	----

4.1.2 Estadística Descriptiva de los Parámetros Microbiológicos.....	62
--	----

4.1.3 Estadística Descriptiva de los Macro/micro nutrientes.....	71
--	----

4.1.4 Estadística Descriptiva de las Variables Dependientes.....	98
--	----

### 4.2 ANÁLISIS MULTIVARIADO DE LOS DATOS

4.2.1 Componentes Principales.....	116
------------------------------------	-----

## CAPÍTULO CINCO

INTRODUCCIÓN.....	120
-------------------	-----

5.1 MODELO DE REGRESIÓN LINEAL SIMPLE PARA LA VARIABLE DEPENDIENTE “ <i>INHIBICIÓN</i> ”.....	123
--	-----

5.2 MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA LA VARIABLE DEPENDIENTE “ <i>PROMEDIO DEL LARGO DE LA HOJA</i> ”.....	127
---	-----

5.3 MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA LA VARIABLE DEPENDIENTE “ <i>PROMEDIO DEL ANCHO DE LA HOJA</i> ”.....	131
---	-----

5.4 MODELO DE REGRESIÓN LINEAL SIMPLE PARA LA VARIABLE DEPENDIENTE “CLOROFILA”.....	135
5.5 MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA LA VARIABLE DEPENDIENTE “NÚMERO DE FRUTOS DE 4 MESES”.....	139
5.6 MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA LA VARIABLE DEPENDIENTE “NÚMERO DE FRUTOS SECOS”.....	143
5.7 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE “INHIBICIÓN”.....	147
5.8 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE “PROMEDIO DEL LARGO DE LA HOJA”.....	149
5.9 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE “PROMEDIO DEL ANCHO DE LA HOJA”.....	151
5.10 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE “CLOROFILA”.....	153
5.11 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE “NÚMERO DE FRUTOS DE 4 MESES”.....	155

5.12 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE “NÚMERO DE FRUTOS SECOS”.....	157
5.13 MODELOS DE REGRESIÓN FINALES.....	158
CAPÍTULO SEIS	
6.1 CONCLUSIONES.....	160
6.2 RECOMENDACIONES.....	161
BIBLIOGRAFÍA.....	162

## ABREVIATURAS

ESPOL	Escuela Superior Politécnica del Litoral
CIBE	Centro de Investigaciones Biotecnológicas del Ecuador
N	Nitrógeno
P <sub>2</sub> O <sub>5</sub>	Fósforo
K <sub>2</sub> O	Potasio
Ca	Calcio
Mg	Magnesio
Mn	Manganeso
Fe	Hierro
Cu	Cobre
Zn	Zinc
ppm	Partes por millón

## SIMBOLOGÍA

%	Porcentaje
ppm	Partes por millón
$\mu$	Media poblacional
$\sigma^2$	Varianza
$R^2$	Coefficiente de determinación

## ÍNDICE DE CUADROS

Cuadro 1.1	Exportaciones ecuatorianas de cacao y sus derivados... 5
Cuadro 1.2	Porcentaje de superficie plantada y producción, según región y provincia..... 7
Cuadro 1.3	Sintomatología encontrada en mazorcas como manchas de chocolate, mazorcas con una capa polvorienta color crema y mazorcas con crecimiento atrofiado..... 12
Cuadro 2.1	Diagramas de dispersión expresando diferentes tipos de relación..... 25
Cuadro 2.2	Tabla de Análisis de Varianza..... 28
Cuadro 3.1	Ubicación de las haciendas investigadas..... 40
Cuadro 4.1	Tabla e histograma de frecuencias del pH..... 44
Cuadro 4.2	Estadística descriptiva y Diagrama de cajas para el pH..... 45
Cuadro 4.3	Tabla e histograma de frecuencias de la Conductividad..... 47

Cuadro 4.4	Estadística descriptiva y diagrama de Cajas para la Conductividad.....	48
Cuadro 4.5	Tabla e histograma de frecuencias de los Solutos Totales.....	50
Cuadro 4.6	Estadística descriptiva y Diagrama de cajas de los Solutos totales.....	51
Cuadro 4.7	Tabla e histograma de frecuencias de los Grados Brix.....	53
Cuadro 4.8	Estadística descriptiva y Diagrama de cajas para los Grados.Brix.....	54
Cuadro 4.9	Tabla e histograma de frecuencias de la Densidad.....	56
Cuadro 4.10	Estadística descriptiva y diagrama de Cajas para la Densidad.....	57
Cuadro 4.11	Tabla e histograma de frecuencias del Total de Materia Orgánica.....	59
Cuadro 4.12	Estadística descriptiva y Diagrama de cajas del Total de Materia Orgánica.....	60

Cuadro 4.13	Tabla e histograma de frecuencias de los aerobios.....	62
Cuadro 4.14	Estadística descriptiva y Diagrama de cajas para los aerobios.....	63
Cuadro 4.15	Tabla e histograma de frecuencias de los anaerobios.....	65
Cuadro 4.16	Estadística descriptiva y diagrama de Cajas para los anaerobios.....	66
Cuadro 4.17	Tabla e histograma de frecuencias de los Hongos y Levaduras.....	68
Cuadro 4.18	Estadística descriptiva y Diagrama de cajas de los Hongos y Levaduras.....	69
Cuadro 4.19	Tabla e histograma de frecuencias del porcentaje de Nitrógeno.....	71
Cuadro 4.20	Estadística descriptiva y Diagrama de cajas para el porcentaje de Nitrógeno.....	72
Cuadro 4.21	Tabla e histograma de frecuencias de las partes por millón de Fósforo.....	74

Cuadro 4.22	Estadística descriptiva y diagrama de Cajas para las partes por millón de Fósforo.....	75
Cuadro 4.23	Tabla e histograma de frecuencias del porcentaje de Potasio.....	77
Cuadro 4.24	Estadística descriptiva y Diagrama de cajas del porcentaje de Potasio.....	78
Cuadro 4.25	Tabla e histograma de frecuencias del porcentaje de Calcio.....	80
Cuadro 4.26	Estadística descriptiva y Diagrama de cajas para el porcentaje de Calcio.....	81
Cuadro 4.27	Tabla e histograma de frecuencias del porcentaje de Magnesio.....	83
Cuadro 4.28	Estadística descriptiva y diagrama de Cajas para el porcentaje de Magnesio.....	84
Cuadro 4.29	Tabla e histograma de frecuencias de las partes por millón de Manganeso.....	86

Cuadro 4.30	Estadística descriptiva y Diagrama de cajas de las partes por millón de Manganeso.....	87
Cuadro 4.31	Tabla e histograma de frecuencias de las partes por millón de Hierro.....	89
Cuadro 4.32	Estadística descriptiva y Diagrama de cajas para las partes por millón de Hierro.....	90
Cuadro 4.33	Tabla e histograma de frecuencias de las partes por millón de Cobre.....	92
Cuadro 4.34	Estadística descriptiva y diagrama de Cajas para las partes por millón de Cobre.....	93
Cuadro 4.35	Tabla e histograma de frecuencias de las partes por millón de Zinc.....	95
Cuadro 4.36	Estadística descriptiva y Diagrama de cajas de las partes por millón de Zinc.....	96
Cuadro 4.37	Tabla e histograma de frecuencias de la Inhibición.....	98
Cuadro 4.38	Estadística descriptiva y Diagrama de cajas para la Inhibición.....	99

Cuadro 4.39	Tabla e histograma de frecuencias del promedio de largo de la hoja.....	101
Cuadro 4.40	Estadística descriptiva y diagrama de Cajas para el promedio del largo de la hoja.....	102
Cuadro 4.41	Tabla e histograma de frecuencias del promedio del ancho de la hoja.....	104
Cuadro 4.42	Estadística descriptiva y Diagrama de cajas del promedio del ancho de la hoja.....	105
Cuadro 4.43	Tabla e histograma de frecuencias de la Clorofila.....	107
Cuadro 4.44	Estadística descriptiva y Diagrama de cajas para la Clorofila.....	108
Cuadro 4.45	Tabla e histograma de frecuencias del número de frutos de 4 meses.....	110
Cuadro 4.46	Estadística descriptiva y diagrama de Cajas para el número de frutos de 4 meses.....	111
Cuadro 4.47	Tabla e histograma de frecuencias del Número de frutos secos.....	113

Cuadro 4.48	Estadística descriptiva y Diagrama de cajas del Número de frutos secos.....	114
Cuadro 4.49	Total de proporción acumulada por el Análisis de Componentes Principales.....	116
Cuadro 4.50	Matriz de Componentes Principales.....	118
Cuadro 5.1	Normalidad: Prueba de Kolmogorov - Smirnov para la variable dependiente "Inhibición".....	123
Cuadro 5.2	Valores residuales versus los valores pronosticados.....	124
Cuadro 5.3	Tabla ANOVA del modelo seleccionado.....	125
Cuadro 5.4	Normalidad: Prueba de Kolmogorov - Smirnov para la variable dependiente "Promedio del largo de la hoja".....	127
Cuadro 5.5	Valores residuales versus los valores pronosticados.....	128
Cuadro 5.6	Tabla ANOVA del modelo seleccionado.....	129
Cuadro 5.7	Normalidad: Prueba de Kolmogorov - Smirnov para la variable dependiente "Promedio del ancho de la hoja".....	131
Cuadro 5.8	Valores residuales versus los valores pronosticados.....	132
Cuadro 5.9	Tabla ANOVA del modelo seleccionado.....	133
Cuadro 5.10	Normalidad: Prueba de Kolmogorov - Smirnov para la variable dependiente "Clorofila".....	135

Cuadro 5.11	Valores residuales versus los valores pronosticados.....	136
Cuadro 5.12	Tabla ANOVA del modelo seleccionado.....	137
Cuadro 5.13	Normalidad: Prueba de Kolmogorov - Smirnov para la variable dependiente "Número de frutos de 4 meses".....	139
Cuadro 5.14	Valores residuales versus los valores pronosticados.....	140
Cuadro 5.15	Tabla ANOVA del modelo seleccionado.....	141
Cuadro 5.16	Normalidad: Prueba de Kolmogorov - Smirnov para la variable dependiente "Número de frutos secos".....	143
Cuadro 5.17	Valores residuales versus los valores pronosticados.....	144
Cuadro 5.18	Tabla ANOVA del modelo seleccionado.....	145
Cuadro 5.19	Tabla ANOVA del modelo seleccionado para la variable dependiente "Inhibición".....	147
Cuadro 5.20	Tabla ANOVA del modelo seleccionado para la variable dependiente "Promedio del largo de la hoja".....	149
Cuadro 5.21	Tabla ANOVA del modelo seleccionado para la variable dependiente "Promedio del ancho de la hoja".....	151

Cuadro 5.22	Tabla ANOVA del modelo seleccionado para la variable dependiente “Clorofila”.....	153
Cuadro 5.23	Tabla ANOVA del modelo seleccionado para la variable dependiente “Número de frutos de 4 meses”.....	155
Cuadro 5.24	Tabla ANOVA del modelo seleccionado para la variable dependiente “Número de frutos secos”.....	157

## **INTRODUCCIÓN**

El uso de productos químicos en la producción agrícola va degradando poco a poco el ambiente dejándolo poco favorable por la capa vegetal de suelo, su estructura, la micro fauna presente, etc. [11, 17]

Estudios realizados revelan el contenido de residuos químicos en frutas, plantas y suelo [14]; por este motivo se han desarrollado en los principales países consumidores [13] estrictas políticas [12] que favorecen el consumo de fruta libre de trazas o residuos químicos, lo que ha llevado a los productores a desarrollar tecnología libre de estos productos favoreciendo a la producción orgánica.

La agricultura orgánica es una visión sistemática de la producción agrícola la cual utiliza como guía los procesos biológicos de los ecosistemas naturales [15]. La agricultura orgánica es un sistema global de gestión de la producción que fomenta y realza la salud de los agros ecosistemas, inclusive la diversidad biológica, los ciclos biológicos y la actividad biológica del suelo [12].

El abono orgánico es un fertilizante que proviene de la degradación y mineralización de materiales orgánicos; existen diferentes tipos de abono o enmiendas orgánicas según la fuente de aporte de nutrientes y el grado de procesamiento. De forma particular, este estudio se enfoca en el “biol” que es un abono líquido a base de materia orgánica procesada de manera

anaeróbica, el cual es utilizado como fertilizante y biofungicida [16], dado que está compuesto básicamente por productos orgánicos, se comporta de manera amigable con el medio ambiente.

El biol tiene como principales componentes la melaza, estiércol, residuos vegetales, roca fosfórica, microorganismos, sulphomag, ceniza de tamo de arroz y agua. Los componentes pueden cambiar dependiendo de la disponibilidad de cada uno en la zona de producción, si alguno o está disponible es reemplazado por algún otro con similares características en su contenido nutricional. Para su preparación, los elementos son depositados en tanques cerrados, los que contarán con una manguera para desalojar los gases desprendidos por el proceso de fermentación. Cada mes se reactiva el proceso agregando microorganismos y melaza, y al cuarto mes el biol está listo para ser aplicado de manera foliar a las plantas [16]. Al final del proceso, se obtiene un biol, al cual se le puede determinar su calidad, en función de sus características generales y determinar su efecto en plantaciones de cacao.

En Ecuador no se tienen referencias sobre estudios realizados sobre el efecto del biol en cacao, salvo el proyecto realizado en el Centro de Investigaciones Biotecnológicas del Ecuador (CIBE) con financiamiento del PL480 en el año 2011, donde se determina que existe un efecto favorable en la producción de cacao, sin llegar a cuantificarse el efecto de cada parámetro de calidad del biol sobre características agronómicas, sanitarias y de

producción de la planta de cacao [18]. Más aún, se sabe que en Ecuador el cultivo de cacao es el producto de exportación más antiguo, la producción de este cultivo representa uno de los principales rubros del Sector Agropecuario del país, al menos hasta cuando se inició el auge bananero (1969) y el petrolero (1972), generando recursos y empleo, contribuyendo al desarrollo y progreso de la economía ecuatoriana, sobre todo en los inicios del siglo XX y hasta el año 1920 cuando las plantaciones de cacao fueron devastadas por la escoba de bruja, que mermó la producción de la denominada “Pepa de oro”. [26]

## **CAPÍTULO UNO**

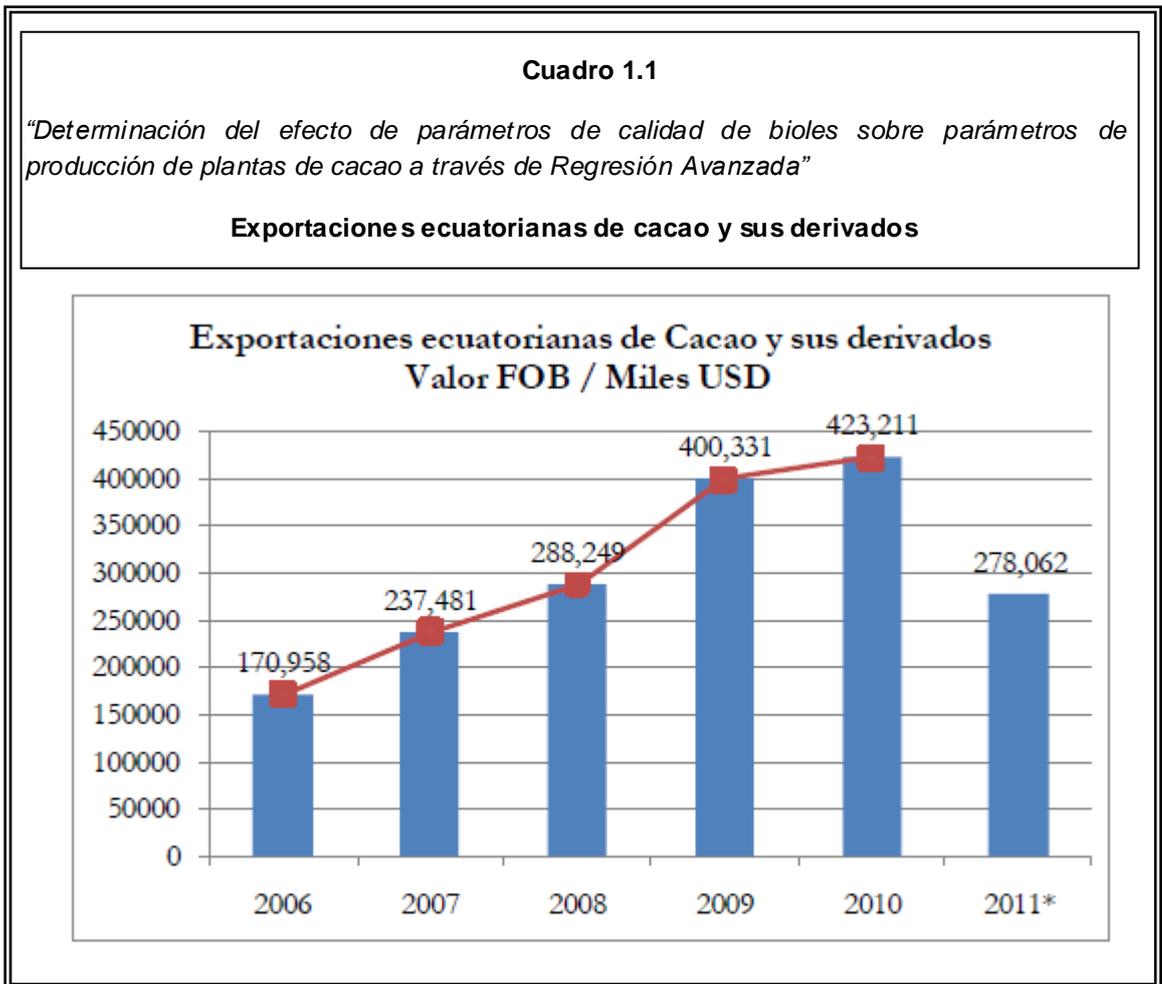
### **1.1 PRODUCCIÓN DE CACAO EN ECUADOR**

Según los resultados obtenidos mediante el Censo Nacional Agropecuario elaborado en el año 2000, se evidencia que en Ecuador existen alrededor de 12'355.831 hectáreas de suelo, los mismos que son destinados a la producción de diferentes cultivos, de este total el 2% representa a las hectáreas que son destinadas exclusivamente al cultivo de cacao; siendo esto 243.146 hectáreas. [1]

Según la distribución a nivel nacional, de acuerdo al tamaño de hectárea que es usado para el cultivo de cacao, se tiene que el 71% representa aquellos productores que cosechan cacao en distancias de hasta 50 hectáreas, seguido del 22% que corresponde a los suelos que tienen entre 50 a 200 hectáreas y finalmente un 7% corresponde a los suelos que tiene 100 o más hectáreas.

En el año 2010, se registró un precio unitario \$3.01 por cada kilo, en los productos primarios de exportación como el cacao; mientras que el valor unitario para los productos industrializados de exportación fue de \$4.34 por kilo, entendiéndose aquellos que son derivados del cacao, como barras de chocolate, cacao en polvo, licor, entre otros.

En este mismo año se registraron a 88 productores clasificados entre empresas, personas naturales, asociaciones, y otros, que producen cacao y/o sus elaborados, siendo para varios la fuente principal de ingresos, y además generan directa e indirectamente empleo para la población ecuatoriana.



**Fuente:** Banco Central del Ecuador

En el **Cuadro 1.1** se observa que para el valor de las exportaciones de cacao y sus derivados existe una tendencia creciente, esto durante el

período 2006-2010, alcanzando una Tasa de Crecimiento Promedio Anual (TCPA) de 25.45%.

En el 2010, las exportaciones ecuatorianas de este producto fueron de \$423.211 y para julio de 2011 ya alcanzan un valor de \$278,062. Con relación al volumen exportado, se tiene 7.70% inferior entre los años 2009 y 2010, registrándose al 2009 un volumen exportado de 143.304 toneladas métricas, mientras que en el 2010 se tiene una cantidad exportada a 133.056 toneladas métricas. [2]

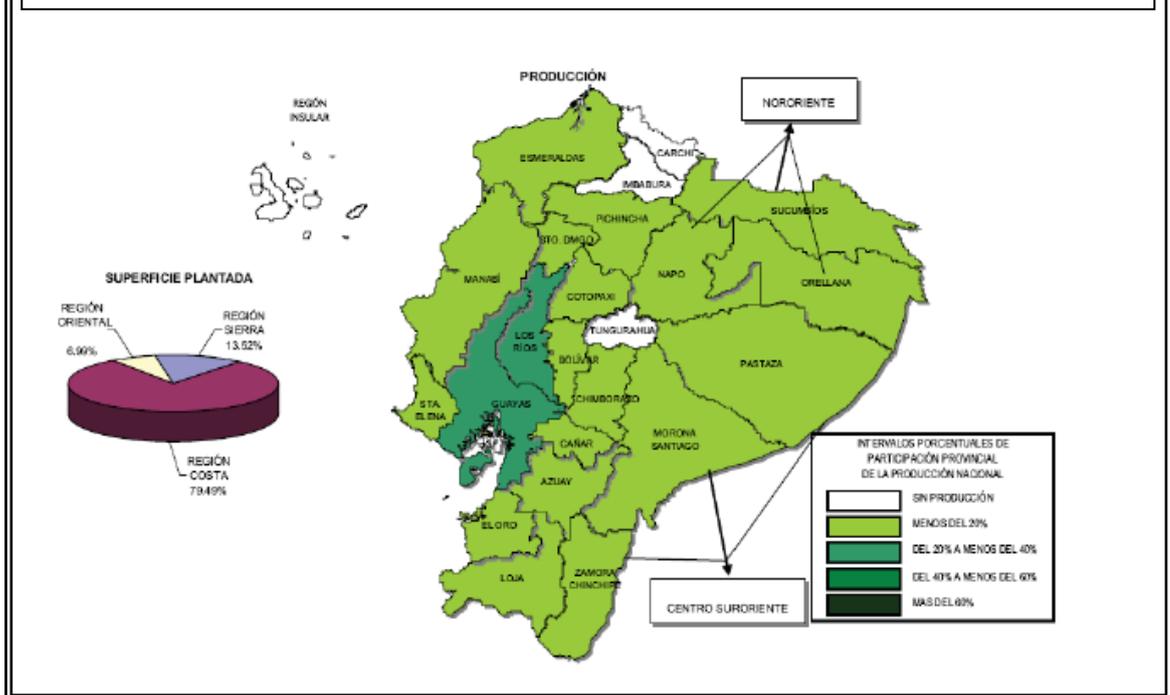
Es importante mencionar, que aunque existe un decremento en las exportaciones de cacao entre el 2009 y 2010, se tiene que el ingreso monetario por exportaciones es superior en un 5,72%.

## 1.2 DISTRIBUCIÓN DE LOCALIDADES PRODUCTORAS DE CACAO

**Cuadro 1.2**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada.”*

**Porcentaje de superficie plantada y producción, según región y provincia**



**Fuente:** Instituto Nacional de Estadísticas y Censos, Encuesta de Superficie y producción agropecuaria Continua ESPAC 2009

En el **Cuadro 1.2** se evidencia que la mayor concentración de la producción de cacao se encuentra en la Región Costa, además se presenta una breve descripción de cómo se encuentran distribuidos las localidades dedicadas a este cultivo, a nivel provincial. Según la ESPAC realizada en 2009 se tiene que un 80% de la superficie total a nivel nacional se encuentra en esta Región.

Las provincias que tienen una mayor superficie cosechada de cacao son Manabí, Los Ríos y Guayas, las mismas que tienen valores de 92.839, 84.222 y 79.768 respectivamente, las mismas que han tenido variaciones crecientes de productividad para el caso de Manabí y Los Ríos y para el Guayas decreciente, respecto al año previo de 3,79%, 1,32% y 1,81% .

En el caso de la región Sierra las principales provincias productoras son Cotopaxi, Bolívar y Cañar. [3]

Para el 2009 las provincias que no tienen producción de cacao son Tungurahua, Carchi, Imbabura y Galápagos, por lo que no generan ningún tipo de ingreso, mediante sus exportaciones.

En este mismo período, la producción de cacao genera ingresos a más de a 55 mil familias, donde el 59% son de pequeños productores que tienen como cultivo menos de 10 hectáreas.

### **1.3 VARIEDADES DE CACAO QUE SE PRODUCEN EN ECUADOR**

En Ecuador se cultivan varios tipos de cacao; pero la variedad conocida comúnmente como “nacional” es la más cotizada entre los fabricantes de chocolate, debido a la calidad y el sabor del producto.

Sin embargo, existen enfermedades severas que deterioran e impiden el normal crecimiento de este fruto, siendo principalmente la moniliasis y la escoba de bruja, estas enfermedades tienen su origen de hace unos 100

años, donde hubo una introducción masiva en el cacao extranjero, previniendo de Venezuela. [4]

Los principales grupos genéticos que han sido cultivados a nivel mundial son **Criollo, Forastero y Trinitario**, siendo la variedad "Forastero" la más cultivada en el mundo; se estima que aproximadamente ocupa alrededor del 80% del área en producción, donde la principal característica se da por la resistencia a ciertas enfermedades y su alta productividad; sin embargo, por la calidad no es considerado como un "cacao fino", por lo que generalmente es usado mezclándolo con otras variedades de mayor calidad.

El cacao "nacional" producido en Ecuador, ha sido clasificado como del tipo "forastero", porque se caracteriza por tener algunas características visiblemente identificadas de la variedad forastero, pero se diferencia en el sabor debido a que posee un sabor y aroma único y característico, debido a esto es preferido por el consumismo, a nivel industrial de todo el mundo. Comúnmente, a nivel internacional, al cacao del país se lo identifica como el "cacao de arriba", debido a que generalmente se lo cultivaba en la zona superior del río Guayas (río arriba), y actualmente se lo denomina como un producto de buen sabor y aroma.

El cacao Trinitario ocupa entre el 10 y 15% de la producción mundial, y está constituido por el cruce del criollo de Trinidad con la variedad

introducida de la Cuenca del Orinoco; se lo considera cacao de calidad. Dentro de esta variedad se ubica el CCN51 que es producto de la investigación realizada en el Ecuador, en la zona de Naranjal, por el Agrónomo Homero Castro. Este clon presenta características de alta producción y tolerancia a las enfermedades pero no tiene el aroma que posee el Nacional. [6]

#### **1.4 CONDICIONES CLIMÁTICAS PARA LA PRODUCCIÓN DE CACAO**

El cacao es un árbol originario de las selvas de América Central y del Sur, su nombre científico es *Theobroma cacao* L, las condiciones climáticas ideales tipo de cultivo es cuando cumple con las siguientes características:

**Altitud:** Entre 0 y 800 msnm

**Lluvia:** 1000 – 2500 mm durante el año

**Luz:** 1000 horas anuales

**Temperatura:** 21 – 28 grados centígrados. Con poca variación entre el día y la noche. [8]

#### **1.5 TIPOS DE PLAGICIDAS PARA LAS ENFERMEDADES DE CACAO**

En Ecuador el cultivo de cacao puede ocasionar serias pérdidas a los productores cuando estos son afectados por plagas como chinches,

áfidos, hongos, etc, es por este motivo que se necesita el uso de un manejo integrado con ayuda de los plaguicidas que ayuden a eliminar este problema que generalmente ocurre en las plantaciones.

En el mercado existen algunas variedades de plaguicidas químicos, la desventaja de estos son las consecuencias que pueden causar al usar de manera frecuente en las plantaciones, puesto que los frutos son consumidos por el ser humano, y puede ocasionar daños a su sistema digestivo.

## 1.6 EFECTOS DEL BIOL EN EL CACAO

En la actualidad los productores de cacao buscan alternativas orgánicas para controlar las diferentes enfermedades que afectan al normal crecimiento del fruto, tales como, la moniliasis (*Moniliophthora roreri*) y la escoba de bruja (*Moniliophthora perniciosa*). Este control se lo ha realizado mediante el uso de bioproductos o bioplaguicidas, los mismos que buscan tener un efecto inhibitorio al avance de estas enfermedades y que a su vez le proporcionen a la planta los variados nutrientes que necesitan para su eficaz desarrollo.

**Cuadro 1.3**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada.”*

**Sintomatología encontrada en mazorcas como manchas de chocolate, mazorcas con una capa polvorienta color crema y mazorcas con crecimiento atrofiado**



**Fuente:** CIBE-ESPOL

Estas dos enfermedades son causadas por dos hongos *Moniliophthora roreri* y *Moniliophthora perniciosa*, los mismos que se caracterizan por dañar al fruto causándole la podredumbre en las semillas del cacao y una necrosis al fruto, estas afecciones se dan especialmente en épocas lluviosas. [27]

Se conoce que durante los meses de Julio a Septiembre se produce la conocida vulgarmente “helada” que son las bajas temperaturas que afectan al crecimiento de los frutos de cacao, esto también ocasiona la pérdida en la producción.

Los bioproductos o bioplaguicidas, en su mayoría, están compuestos por materia orgánica como excremento vacuno, porcino o gallinaza, que son reusados para este tipo de actividad, además estos contienen microorganismos que tienen propiedades que benefician a la planta, y que son propios de la región donde se encuentran los cultivos de cacao, siendo estos capaz de reproducirse de manera anaerobia y aerobia, y también tiene melaza, entre otros activos que ayudan a ser un agente inhibitorio de las enfermedades mencionadas previamente y que afectan a la planta.

El uso de estos bioproductos, al evitar la propagación de ciertas enfermedades, reduce de manera considerable los costos de producción, generando beneficios ya que al disminuir el uso de fungicidas existe una

mejora en la calidad del fruto dado que está libre de químicos, permitiendo de esta manera al consumidor disponer de un producto saludable.

El Centro de Investigaciones Biotecnológicas del Ecuador (CIBE), desde el año 2009 ha trabajado en la estandarización del proceso de producción de estos bioplaguicidas o bioproductos y ha experimentado su efecto inhibitorio *in vitro* obteniendo varios y satisfactorios resultados en la inhibición de las enfermedades fitosanitarias mencionadas anteriormente.

[7, 18]

## **1.7 DEFINICIONES**

### **1.7.1 Enmienda Orgánica**

Las enmiendas orgánicas son materiales derivados de los residuos de naturaleza orgánica (vegetales y animales), que incorporan nitrógeno orgánico y humus al suelo, enriqueciéndolo con carbono orgánico y mejorando sus características físicas, químicas y biológicas. Además promueven el desarrollo de las plantas, las cuales pueden obtener grandes cantidades de nutrimentos. [7]

### **1.7.2 Biol**

Es una fuente de compuestos, que se obtienen como producto del proceso de descomposición anaeróbica de los desechos orgánicos. También se lo

denomina como el residuo líquido sobrenadante que se descarga de un digestor de desechos orgánicos [7]

#### **1.7.2.1 Preparación**

Por lo general, el proceso de elaboración consta de las siguientes partes: recolección del estiércol, adición de agua, melaza y microorganismos eficientes, seguido por la mezcla homogénea de todos los ingredientes. Luego se procede a tapar herméticamente el envase durante el tiempo en que la fermentación va a tener lugar. [7]

#### **1.7.2.2 Aplicación**

El biol puede ser utilizado para realizar aplicaciones foliares y aplicaciones al suelo. Dependiendo del tipo de cultivo, el biol puede ser utilizado puro o en diluciones. [7]

#### **1.7.2.3 Mecanismos de acción**

El biol es capaz de estimular el desarrollo de las plantas. Aumenta y fortalece la base radical, actúa sobre el follaje incrementando su área, mejora la floración y activa el poder germinativo de las semillas.

Su contenido nutricional ayuda al desarrollo de microorganismos los cuales protegen a las plantas y mejoran la estructura de los suelos.

Además los bioles son considerados como potenciales alternativas como fungicidas, aplicados directamente al suelo y follaje de las plantas estos proveen microorganismos y nutrientes para reducir enfermedades. [7]

### 1.7.3 Parámetros físicos químicos

Las enmiendas orgánicas se determinan a través de las propiedades físicas - químicas y biológicas. Según Leblanc *et al.* (2007), la calidad de un abono orgánico se determina a partir de su contenido nutricional y de su capacidad de proveer nutrientes a un cultivo. Este contenido está directamente relacionado con las concentraciones de esos nutrientes en los materiales utilizados para su elaboración. [19]

La calidad en los biofertilizantes líquidos, se refiere a factores tales como madurez, minerales presentes y contenido de microorganismos, los cuales pueden variar de lote en lote debido a la diversidad de materia prima utilizada, como también el tipo de elaboración. La naturaleza de los materiales, la relación C-N y las condiciones físicas- químicas tales como temperatura, humedad, pH y presencia de oxígeno, son factores clave para la obtención del tipo de producto deseado y su calidad.

**Materia prima.-** Tipo de estiércol y calidad del compost, este último para el caso de los lixiviados (fertilizante líquido orgánico) y té de compost. La utilización de estiércol fresco o sólido debe ser regulado así como las técnicas de elaboración de compost de que se extraerán té y lixiviados.

**Aditivos.-** Como melaza, microorganismos eficientes, entre otros deben estar presentes en cantidades necesarias y suficientes. Es decir, el éxito en la obtención de un bioproducto consistente, efectivo y de alta calidad

depende del balance de los ingredientes presentes en la elaboración del mismo.

#### **1.7.4 Parámetros microbiológicos**

Los microorganismos eficientes son una combinación de microorganismos benéficos de origen natural. Los principales organismos que forman parte de este complejo son bacterias fototrópicas, levaduras, bacterias productoras de ácido láctico y hongos de fermentación. Estos microorganismos al entrar en contacto con la materia orgánica secretan sustancias benéficas como vitaminas, ácidos orgánicos, minerales quelatados y sustancias antioxidantes. [28]

La efectividad de los biofertilizantes estará, entonces, determinada por la presencia y calidad de microorganismos eficientes en la solución en proceso.

#### **1.7.5 Macronutrientes / Micronutrientes**

Los macronutrientes se necesitan en grandes cantidades, y tienen que ser aplicadas si el suelo es deficiente en uno o más de ellos. Los suelos pueden ser naturalmente pobres en nutrientes, o pueden llegar a ser deficientes debido a la extracción de los nutrientes por los cultivos a lo largo de los años, o cuando se utilizan variedades de rendimientos altos, las cuales son más demandantes en nutrientes que las variedades locales.

Los micronutrientes son requeridos sólo en cantidades ínfimas para el crecimiento correcto de las plantas y tienen que ser agregados en cantidades muy pequeñas cuando no pueden ser provistos por el suelo [20, 21]

#### **1.7.6 Parámetros fitosanitarios**

Como parámetros fitosanitarios se analizará lo relacionado a la inhibición de enfermedades como resultado del uso de la enmienda orgánica (biol) en las plantaciones, esto implica la reducción de enfermedades como monilia, escoba de bruja que son las enfermedades más comunes en las plantaciones de cacao; se hará un promedio del número de frutos enfermos con estas enfermedades respectivamente.

#### **1.7.7 Parámetros de producción**

Se analizarán como parámetros de producción el contenido de clorofila en las hojas, el promedio del largo y ancho de las hojas y el número promedio de mazorcas de 4 meses producidas por hectárea. Estos parámetros fueron determinados en base a los criterios de los investigadores de Técnicas Agrícolas del CIBE.

#### **1.1.8 Prácticas agronómicas**

Se basan en el empleo de técnicas de ordenación que reducen el número de insectos de las plagas, modificando el medio ambiente.

Estas prácticas incluyen la destrucción del rastrojo después de la cosecha, para eliminar el alimento de las plagas y la destrucción de los lugares de producción.

También son parte de este método: la rotación de cultivos; la roturación de la tierra para reducir el número de las malas hierbas; el cultivo de plantas que atraen a los insectos plagas para poder eliminarlos; la diversificación del hábitat para albergar a un mayor número de enemigos naturales, siembras tempranas para evitar el surgimiento de las plagas y el riego controlado para destruir los agentes patógenos del suelo. [15]

## **CAPÍTULO DOS**

### **INTRODUCCIÓN**

En la actualidad es necesario el uso de diferentes técnicas estadísticas conjuntamente con las diversas ciencias existentes, esto se da con el fin de tomar la mejor decisión o tener un criterio que se sustente los diferentes supuestos en las investigaciones.

Desde la perspectiva ambiental se está investigando sobre el uso de nuevas tecnologías que sean amigables con el ambiente, es decir el uso de enmiendas orgánicas que ayuden a reducir las enfermedades en los cultivos para este caso de estudio en el cacao tales como la moniliasis y escoba de bruja que son las principales enfermedades que atacan a estos cultivos, en esta investigación se pretende estudiar modelos de regresión que expliquen el efecto inhibitorio de las enmiendas orgánicas a las enfermedades del cacao a través de variables aleatorias con características fitosanitarias, microbiológicas y de micro-macro nutrientes.

Para esto se debe tener un amplio conocimiento que sustente los análisis que se realicen y bajo qué criterios se los establecen. Es por eso que en este capítulo se definirán las técnicas usadas y los supuestos que se efectúen para los análisis de las variables involucradas.

## **2.1 MÉTODOS Y PROCEDIMIENTOS PREVIOS AL ANÁLISIS**

A continuación se mencionan aquellos métodos que se han usado para la recopilación de datos previos al análisis, además de las técnicas estadísticas que se aplican en esta tesis, los mismos que son necesarios para obtener resultados y generar conclusiones.

Cada metodología tanto descriptiva, univariada, multivariada además de la estadística inferencial, la regresión lineal simple, múltiple, de una y más variables predictoras son detalladas para conocer los supuestos y requerimientos necesarios, con la finalidad de que puedan ser aplicadas a los datos que han sido recopilados para realizar este proyecto.

## **2.2. MUESTREO**

Es una técnica estadística que se realiza con la finalidad de elegir subconjunto de elementos a la cual se la denomina muestra y que se encuentra en la población que se desea investigar. Muchas veces el tamaño de la Población Objetivo (población a la que pertenecen los entes que son investigados) suele ser “grande” o incluso desconocido por ser infinita, lo que imposibilita que todos sus elementos sean investigados, ya sea porque no se dispone con los recursos necesarios, siendo estos, económicos, logísticos o por falta de tiempo.

Esta técnica que forma parte de la Estadística, determina una cantidad de elementos que son representativos para conocer las diferentes características de los elementos del estudio y que son referentes de una determinada *población objetivo*.

Para esta investigación se hará énfasis en el *Muestreo por estratos*, *Muestreo aleatorio simple* y el *Muestreo sistemático o dirigido*.

## **2.3 ELEMENTOS DEL MUESTREO**

### **2.3.1 POBLACIÓN OBJETIVO**

Debido que al tener un problema planteado es necesario conocer cuáles son todas aquellas unidades que serán motivo de la investigación, a todo este conjunto de entes se lo denomina Población Objetivo.

Para la presente investigación son todas plantas marcadas en las haciendas de las provincias de la costa ecuatoriana.

### **2.3.2 MARCO MUESTRAL**

Son todos aquellos registros, mapas, bases de datos en donde se representan las unidades de investigación que son parte de la Población Objetivo.

### **2.3.3 MUESTRA**

Es el subconjunto que ha sido seleccionado de una determinada población (Población Objetivo).

## **2.4 MUESTREO SISTEMÁTICO O DIRIGIDO**

Es un tipo de muestreo no aleatorio, en donde la muestra dirigida está conformada por los elementos de una población objetivo que el investigador considera es representativa del universo.

## 2.5 ANÁLISIS DESCRIPTIVO

Este es un análisis básico para identificar el comportamiento estadístico de cada una de las variables cuantitativas, la distribución de frecuencias, medidas de tendencia central, medidas de dispersión, estadísticos de orden, entre otros.

Las principales medidas se citan a continuación:

***Medidas de Tendencia Central:*** Media aritmética (promedio), Mediana

***Estadísticos de Orden:*** Cuartiles, Deciles, Percentiles

***Otras medidas:*** Sesgo, Curtosis

***Medidas de Dispersión:*** Varianza, Desviación Estándar, Covarianza, Coeficiente de Variación

Además en esta tesis, para cada una de las variables se presenta el Diagrama de Cajas, la tabla de frecuencia y su histograma respectivo, con la finalidad de determinar valores y la distribución de frecuencia de la variable, así como las medidas de tendencia central y la dispersión.

## 2.6 ANÁLISIS INFERENCIAL

Este es un análisis más avanzado, en comparación con el análisis descriptivo, debido a que se empieza a inferir en base a los elementos de la muestra, respecto a la población, aquí se generan las hipótesis y se determina la existencia de correlación entre las variables, así mismo se determina los diferentes supuestos que requieren las técnicas aplicadas, tales como las pruebas de homogeneidad y normalidad de los errores para generar modelos de regresión.

## 2.7 ANÁLISIS DE CORRELACIÓN

El concepto de relación lineal o correlación se refiere al grado de variación lineal conjunta que existe entre dos variables.

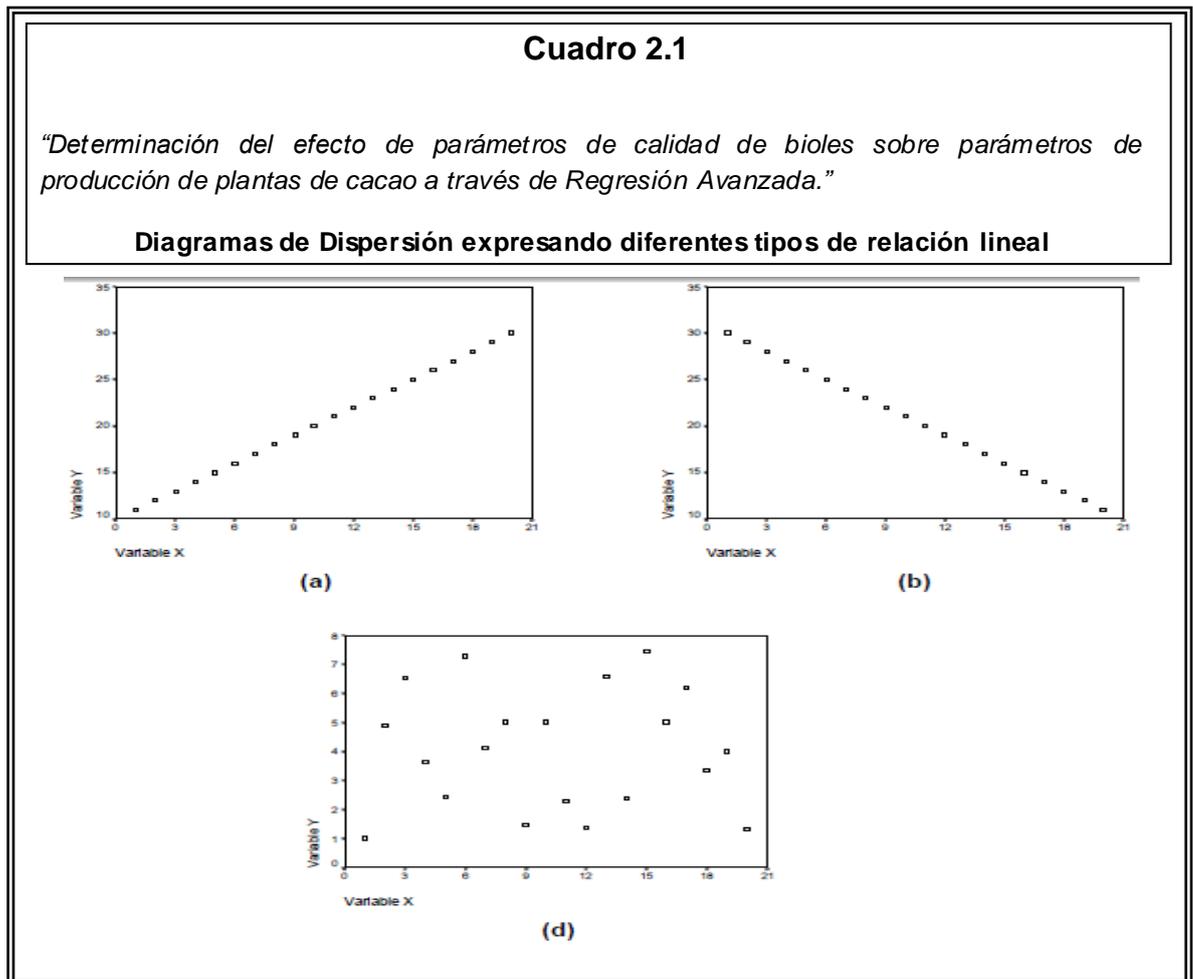
El coeficiente de correlación lineal,  $\rho_{ij}$  entre  $X_i$  y  $X_j$  es un número entre -1 y 1, se lo define como  $E[(X_i - \mu_i)(X_j - \mu_j)] = \rho_{ij}$  y se lo calcula mediante la

siguiente fórmula  $\hat{\rho}_{ij} = r_{ij} = \frac{S_{ij}}{S_i S_j}$ , cuyo resultado mide la fuerza de la

relación lineal entre el par de variables estudiadas.

Cuando  $\hat{\rho} = 0$  se entiende que las dos variables no están relacionadas linealmente, cuando  $\hat{\rho}$  es igual a 1 ó -1 se dice que hay una relación lineal perfecta entre  $X_i$  y  $X_j$ . si el coeficiente de correlación lineal  $\hat{\rho}$  es positivo se dice que estos están directamente relacionados y si es negativo, este par

de variables están inversamente relacionados; es decir, que una variable es creciente y la otra decreciente. [22, 23]



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 2.1** se observa un gráfico con los diagramas de dispersión los tipos de relaciones lineales que puede existir entre dos variables.

Una vez calculado el valor del coeficiente de correlación se puede determinar si existe o no dependencia lineal, y en el caso de existir, que

tipo de relación existe entre las variables X e Y. Al hablar de la significancia de dicho coeficiente de correlación se puede decir, con una cierta probabilidad, que es diferente de cero. Más estrictamente, en términos estadísticos, preguntarse por la significancia de un cierto coeficiente de correlación no es otra cosa que preguntarse por la probabilidad de que tal coeficiente proceda de una población cuyo valor sea de cero, siendo la siguiente hipótesis con la que determina si existe o no dependencia lineal:

$$H_0: r_{xy} = 0;$$

Es decir que el coeficiente de correlación obtenido procede de una población cuya correlación es cero ( $\rho = 0$ ).

Vs.

$$H_1: r_{xy} \neq 0$$

Es decir que el coeficiente de correlación obtenido procede de una población cuyo coeficiente de correlación es distinto de cero ( $\rho \neq 0$ ).

## 2.8 REGRESIÓN LINEAL

La regresión lineal es un método estadístico que modela la relación entre una variable dependiente Y, las variables independientes  $X_j$  y un término aleatorio  $\varepsilon$ . Este modelo puede ser expresado como:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i + \varepsilon_i$$

$Y_i$ : Variable dependiente o explicada.

$X_1, X_2, \dots, X_i$ : Variables explicativas o independientes.

$\beta_0, \beta_1, \beta_2, \dots, \beta_i$ : Parámetros, miden la influencia que las variables explicativas tienen sobre la variable independiente.

Donde  $\beta_0$  es la intersección o término "constante", las  $\beta_i$  ( $i > 0$ ) son los parámetros respectivos a cada variable independiente.

## 2.9 ANÁLISIS DE VARIANZA EN REGRESIÓN LINEAL

La tabla de Análisis de Varianza (ANOVA) en regresión se usa para analizar estadísticamente la validez del modelo  $Y = \beta_0 + \beta_1 X_1 + \dots + \beta_i X_i + \varepsilon_i$  y los supuestos del mismo. [8]

<b>Cuadro 2.2</b>				
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada.”</i>				
<b>Tabla de Análisis de Varianza</b>				
<b>Fuentes de Variación</b>	<b>Grados de Libertad</b>	<b>Sumas Cuadráticas</b>	<b>Medias Cuadráticas</b>	<b>Estadístico de Prueba F</b>
<b>Regresión</b>	$p - 1$	$SCR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SCR}{p - 1}$	$F = \frac{MCR}{MCE}$
<b>Error</b>	$n - p$	$SCE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SCE}{n - p}$	
<b>TOTAL</b>	$n - 1$	$SCT = \sum_{i=1}^n (y_i - \bar{y})^2$		

**Elaborado por:** Karla M. Aguaguiña Méndez

El **Cuadro 2.2** muestra la tabla ANOVA que consiste en un arreglo rectangular cuyas componentes son: las Fuentes de Variación (Regresión, Error, Total); sus grados de libertad; las Sumas y Medias Cuadráticas; y un

valor adicional que es el Estadístico de Prueba F, además del valor de significancia o “valor p” que es con el cual decidimos en si se acepta o no la hipótesis de que todos los parámetros betas son cero.

## 2.10 REGRESIÓN MULTIPLE

Generalizando la notación usada para el modelo de regresión lineal simple, a partir de una muestra de  $n$  observaciones de la cual se dispone los valores de una variable respuesta  $Y$  además de los valores de  $i$  variables explicativas  $X_1, X_2, \dots, X_i$ .

El modelo de regresión lineal múltiple con  $i$  variable intenta aproximar la variable de respuesta mediante una función lineal de las variables explicativas de la siguiente forma:

$$(2.10.1) \quad Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_n X_{in} + \varepsilon_i; \quad i = 1, 2, \dots, n$$

De modo que los coeficientes (parámetros) de la ecuación se estiman siguiendo el criterio de mínimos cuadrados:

$$(2.10.2) \quad \text{Mín} \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i1} - \beta_2 X_{i2} - \dots - \beta_k X_{ik})^2$$

La obtención de las expresiones de los estimadores mínimos cuadráticos de dichos coeficientes exigen reescribir la expresión (2.1) utilizando notación matricial quedaría así:

$$(2.10.3) \quad Y = X\beta + \varepsilon$$

Donde:

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad X = \begin{pmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Resolviendo este problema mediante el cálculo diferencial, se obtienen los estimadores de mínimos cuadrados a través de la siguiente expresión matricial:

$$(2.10.4) \beta = (X'X)^{-1} X'Y$$

Y mantienen una interpretación análoga al caso de la regresión lineal simple (representa el incremento por término medio en la variable respuesta por cada unidad adicional en la variable).

Estimación de sigma:

$$(2.10.5) s^2 = \frac{SCE}{n-p}$$

Donde:

$n$  es el número de observaciones

$p$  el número de parámetros ( $k+1$ )

**Suma Cuadrática del Error**

$$(2.10.6) \text{SCE} = (Y - X\hat{\beta})^T (Y - X\hat{\beta})$$

### Matriz de covarianza

$$(2.10.7) \sum = S^2 (X^T X)^{-1}$$

### Coefficiente de determinación

$$(2.10.8) R^2 = \frac{SCR}{SCT}, \quad 0 \leq R^2 \leq 1$$

Para realizar un análisis de regresión lineal se hacen las siguientes consideraciones sobre los datos:

**a) Linealidad:** los valores de la variable dependiente están generados por el siguiente modelo lineal:

$$Y = X\beta + \varepsilon$$

**b) Homocedasticidad:** todas las perturbaciones tienen la misma varianza:

$$V(\varepsilon_i) = \sigma^2$$

**c) Independencia:** las perturbaciones aleatorias son independientes entre sí:

$$E(\varepsilon_i, \varepsilon_j) = 0; \quad \forall i \neq j$$

**d) Normalidad:** la distribución de la perturbación aleatoria tiene distribución Normal:

$$\varepsilon_i \sim N(0, \sigma^2)$$

## 2.11 MODELOS NO LINEALES

Otros modelos no lineales, al tener  $Y$ , una variable a ser explicada que toma valores dicotómicos 0 ó 1, y al suponer como consecuencia de esto que la Función de Respuesta  $E(Y)$  no es rectilínea sino más bien sigmoideal, esto nos llevó, a adoptar como Función de Respuesta el denominado *Modelo Logit*, esto es

$$(2.11.1) E(Y) = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)}$$

Dado que en esta tesis no hay variables de tipo dicotómicas no se hará uso de este tipo de modelo.

En particular el modelo lineal  $Y = X\beta + \varepsilon$  no debe ser confundido con el **Modelo Lineal Generalizado**, puesto que estos efectúan una amplia combinación de modelos lineales y no lineales. [9]

## 2.12 VALIDACIÓN DE LOS MODELOS PREDICTIVOS

Una vez hallados los modelos de regresión se procede a la validación de estos mediante el uso del Coeficiente de Determinación  $R^2$ , mediante este

indicador podemos determinar el mejor modelo que explique las variables dependientes en función de las variables independientes.

Donde SSR representa la suma de cuadrados debido a la regresión, y SST representa la suma de cuadrados del total.

Zurita (2009) muestra que el coeficiente de determinación es el cuadrado del coeficiente de correlación. El coeficiente de determinación varía entre 0 y 1, aunque es bastante común expresarlo en porcentaje. De forma empírica, se puede afirmar que un  $R^2$  mayor del 70 % indica una buena asociación lineal entre las variables, luego la variable X puede usarse para predecir Y. Hay que tener presente que el  $R^2$  es afectado por la presencia de valores atípicos.

También  $R^2$  indica qué porcentaje de la variabilidad de la variable de respuesta Y es explicada por su relación lineal con X, mientras más alto sea este valor mejor es la predicción de Y usando X. [9]

## **2.13 ANÁLISIS MULTIVARIADO**

En Estadística se analizan diferentes técnicas multivariadas que permitirán la reducción de la dimensión en la que estamos trabajando sin afectar a la información relevante contenida en la muestra investigada.

### **2.13.1 COMPONENTES PRINCIPALES**

Es una técnica estadística que se utiliza para la reducción de la dimensionalidad de los datos, transformando el conjunto de p variables

originales en otro conjunto  $q$  de variables no correlacionadas ( $q \leq p$ ) llamadas componentes principales.

Las nuevas variables son obtenidas como combinaciones lineales de las variables originales, las componentes principales se ordenan en función del porcentaje de varianza explicada. En este sentido, el primer componente será el más importante por ser el que explica mayor porcentaje de la varianza de los datos. Queda a criterio del investigador decidir cuántos componentes se elegirán en el estudio. [25]

Entre los usos más frecuentes del Análisis de Componentes Principales esta:

- Como técnica de análisis exploratorio que permite descubrir interrelaciones entre los datos y de acuerdo con los resultados, proponer los análisis estadísticos más apropiados.
- Reducir la dimensionalidad de la matriz de datos con el fin de evitar redundancias y destacar relaciones. En la mayoría de los casos, tomando sólo los primeros componentes, se puede explicar la mayor parte de la variación total contenida en los datos originales.

## CAPÍTULO TRES

### 3.1 DESCRIPCIÓN DE LAS VARIABLES

A continuación se indican las definiciones de las variables cuantitativas, las mismas que representan características físico-químicas, microbiológicas y de los macro-micro nutrientes de bioles y de las características fitosanitarias y de producción en las plantaciones de cacao.

#### **Parámetros Físicos - Químicos**

*i) Variable 1: pH (PF1)*

Es una medida de acidez o alcalinidad de una solución, en este caso de los bioles que se lo obtiene usando el peachímetro.

*ii) Variable 2: Conductividad eléctrica (PF2)*

Es una medida de la capacidad de un material (biol) de dejar pasar la corriente eléctrica.

*iii) Variable 3: Solutos Totales (PF3)*

Es una medida correspondiente a la concentración de sólidos en los bioles analizados.

*iv) Variable 4: Grados Brix (PF4)*

Es una medida correspondiente a los niveles de azúcares contenida en las muestras de bioles analizadas.

v) *Variable 5: Densidad (PF5)*

Es una medida correspondiente a la cantidad de masa contenida en las muestras de bioles analizadas.

vi) *Variable 6: Materia Orgánica (PF6)*

Es una medida referente al porcentaje de materia orgánica seca contenida en las muestras de bioles analizadas.

**Parámetros microbiológicos**

i) *Variable 7: Aerobios (MB1)*

Corresponde al número total de microorganismos aerobios contenidos en las muestras de bioles analizadas.

ii) *Variable 8: Anaerobios (MB2)*

Corresponde al número total de microorganismos anaerobios contenidos en las muestras de bioles analizadas.

iii) *Variable 9: Hongos y Levaduras (MB3)*

Corresponde al número total de hongos y levaduras contenidos en las muestras de bioles analizadas.

**Macros – micros nutrientes***i) Variable 10: Porcentaje de nitrógeno (N) (MM1)*

Corresponde al porcentaje total de nitrógeno contenido en las muestras de bioles analizadas.

*ii) Variable 11: Ppm de Fósforo (P)(MM2)*

Corresponde a la cantidad de partes por millón de Fósforo contenido en las muestras de bioles analizadas.

*iii) Variable 12: Porcentaje de Potasio (K) (MM3)*

Corresponde al porcentaje de Potasio contenido en las muestras de bioles analizadas.

*iv) Variable 13: Porcentaje de Calcio (Ca) (MM4)*

Corresponde al porcentaje de Calcio contenido en las muestras de bioles analizadas.

*v) Variable 14: Porcentaje de Magnesio (Mg) (MM5)*

Corresponde al porcentaje de Magnesio contenido en las muestras de bioles analizadas.

*vi) Variable 15: Ppm de Manganeso (Mn) (MM6)*

Corresponde a la cantidad de partes por millón de Manganeso contenido en las muestras de bioles analizadas.

*vii) Variable 16: Ppm de Hierro (Fe) (MM7)*

Corresponde a la cantidad de partes por millón de hierro contenida en las muestras de enmienda orgánica (biol) analizadas.

*viii) Variable 17: Parte por millón de Cobre (Cu) (MM8)*

Corresponde a la cantidad de partes por millón de cobre contenida en las muestras de enmienda orgánica (biol) analizadas.

*ix) Variable 18: Ppm de Zinc (Zn) (MM9)*

Corresponde a la cantidad de partes por millón de zinc contenida en las muestras de enmienda orgánica (biol) analizadas.

### **Parámetros de Producción y Fitosanitarios**

*i) Variable 19: Porcentaje de Inhibición (P1)*

Esta variable nos indica el porcentaje de inhibición de la enmienda orgánica (biol) a las enfermedades fitosanitarias.

*ii) Variable 20: Promedio del largo de la hoja (P2)*

Indica el promedio en centímetros del largo de la hoja de la planta de cacao por hacienda.

*iii) Variable 21: Promedio del ancho de la hoja (P3)*

Indica el promedio en centímetros del ancho de la hoja de la planta de cacao por hacienda.

*iv) Variable 22: Clorofila (P4)*

Indica el porcentaje de contenido de clorofila en la hoja de la planta de cacao por hacienda.

*v) Variable 23: Frutos de 4 meses (P5)*

Indica el número promedio de frutos de 4 meses producidos por hacienda durante el período de noviembre 2009 a abril 2010.

*vi) Variable 24: Frutos secos (P6)*

Indica el número promedio de frutos secos producidos por hacienda durante el período de noviembre 2009 a abril 2010.

*vii) Variable 24: Frutos con escoba de bruja (P7)*

Indica el número promedio de frutos enfermos con escoba de bruja producidos por hacienda durante el período de noviembre 2009 a abril 2010.

### 3.2 LOCALIDADES INVESTIGADAS

En el proyecto que realizó el CIBE se consideraron las provincias de la costa ecuatoriana: Esmeraldas, Manabí, Guayas, Los Ríos y El Oro. En cada una de las provincias se escogieron dos fincas pilotos con plantaciones de cacao fino y de aroma, las mismas que cedieron una hectárea para la realización del experimento.

#### 3.2.1 UBICACIÓN DE LAS HACIENDAS INVESTIGADAS



FUENTE: CIBE - ESPOL

En el **Cuadro 3.1** se observa la ubicación de las haciendas que fueron parte del proyecto realizado por CIBE-ESPOL

### **3.3 RECOLECCIÓN DE DATOS**

La recolección de datos se realizó junto a los investigadores de Técnicas Agrícolas del CIBE – ESPOL durante la ejecución del proyecto PL480 desarrollado durante el período de septiembre 2008 – noviembre 2010; luego en enero de 2012 se realizaron dos visitas a dos haciendas que aún mantienen la tecnología implementada y se volvieron a recolectar datos de los parámetros en estudio.

Las haciendas investigadas están ubicadas en las provincias de la costa ecuatoriana, dentro de cada una de ellas se tiene un hectárea de la cual se seleccionaron de manera dirigida 10 plantas, a las que se les midieron los parámetros de interés.

## **CAPÍTULO CUATRO**

### **INTRODUCCIÓN**

En el presente capítulo se realiza el análisis descriptivo de los datos, este procedimiento se debe realizar en todo estudio que implique el uso de la estadística.

Considerando lo descrito anteriormente también se complementa con el uso de gráficas como los histogramas de frecuencias que son barras en las cuales se cuantificación de valores que pertenecen a las variables, también diagramas de cajas que sirve para identificar datos considerados como “aberrantes” y así poder visualizar de cómo están distribuidos los datos.

El análisis descriptivo que se presenta en este capítulo se basa en el cálculo de medidas de tendencia central tales como: media y mediana. Además se calculan medidas de dispersión para conocer que tan distribuidos se encuentran los datos con respecto a la media, también se ha realizado el cálculo de los cuartiles, rango intercuartil, coeficiente de variación, sesgo, varianza y su respectiva desviación estándar.

Se realiza el análisis univariado para cada una de las variables de interés del proyecto en este caso los parámetros físicos-químicos, microbiológicos, de micros-macros nutrientes y los de producción.

#### **4.1 ANÁLISIS DESCRIPTIVO DE LOS DATOS**

Las variables analizadas corresponden a los parámetros físicos-químicos, microbiológicos, y macro-micro nutrientes de las enmiendas orgánicas líquidas (bioles); también se analizan los parámetros fitosanitarios y de producción.

Se realizó un análisis descriptivo univariado a cada una de las variables que forman parte del estudio haciendo uso del software estadístico INFOSTAT y SPSS V.13, además se presentan los respectivos diagramas de cajas que permiten observar “datos aberrantes” e histogramas de frecuencias relativas que permiten observar la distribución de los datos.

También se muestra el cálculo de algunas medidas de tendencia central tales como la media y la mediana; se han calculado las medidas de dispersión tales como la varianza, la desviación estándar, el coeficiente de variación, el sesgo, la curtosis, y otras medidas como los cuartiles y los percentiles 25, 50 y 75.

Luego se realiza el análisis multivariado y el de regresión para obtener un modelo que explique a las variables de interés en función de los parámetros mencionados.

### 4.1.1 Estadística Descriptiva de los Parámetros Físicos-Químicos

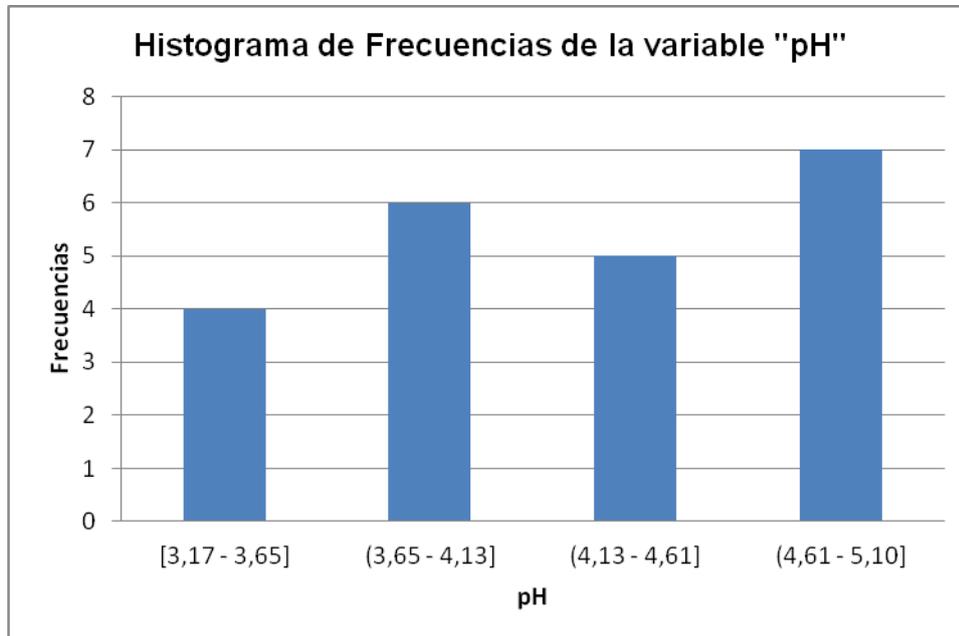
**Variable: pH**

**Cuadro 4.1**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

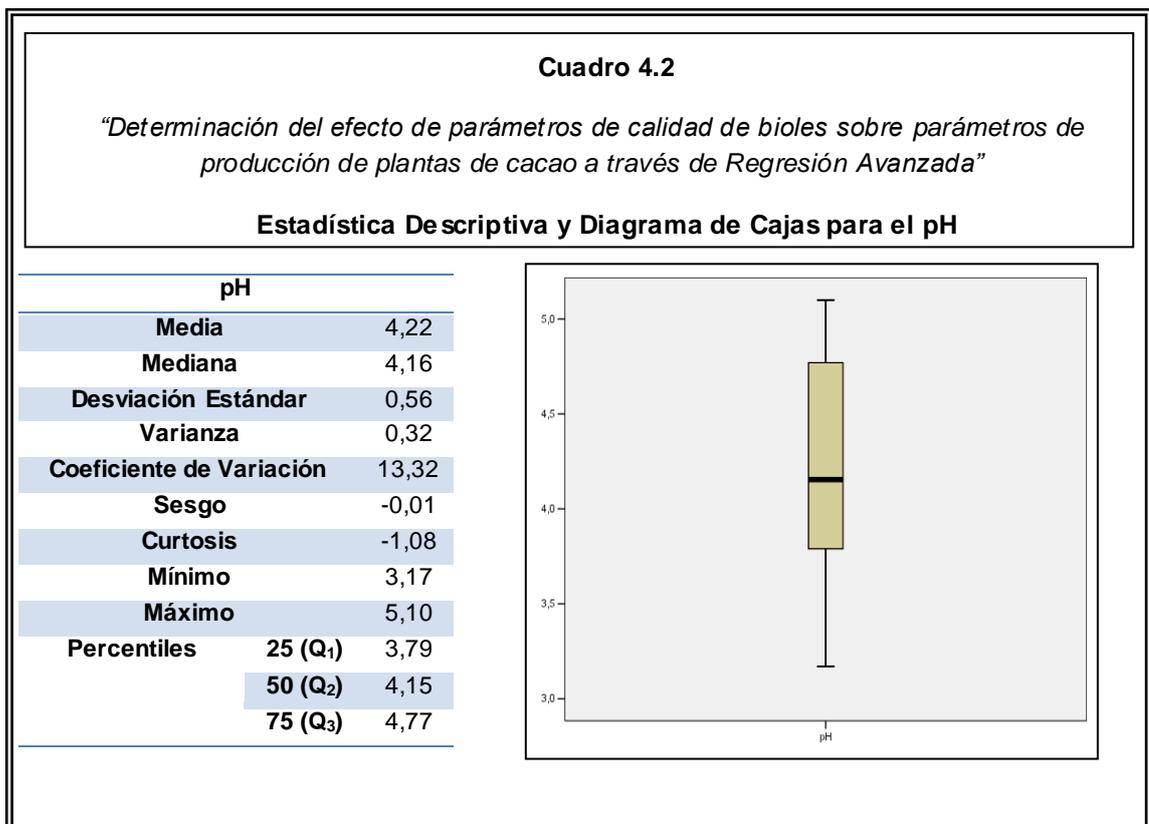
**Tabla e Histograma de frecuencias del pH**

Tabla de Frecuencias de la variable pH			
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[3,17 - 3,65]	4	0,18	0,18
(3,65 - 4,13]	6	0,27	0,45
(4,13 - 4,61]	5	0,23	0,68
(4,61 - 5,10]	7	0,32	1
<b>Total</b>	<b>22</b>	<b>1</b>	



**Elaborado por:** Karla M. Aguaguiña Méndez

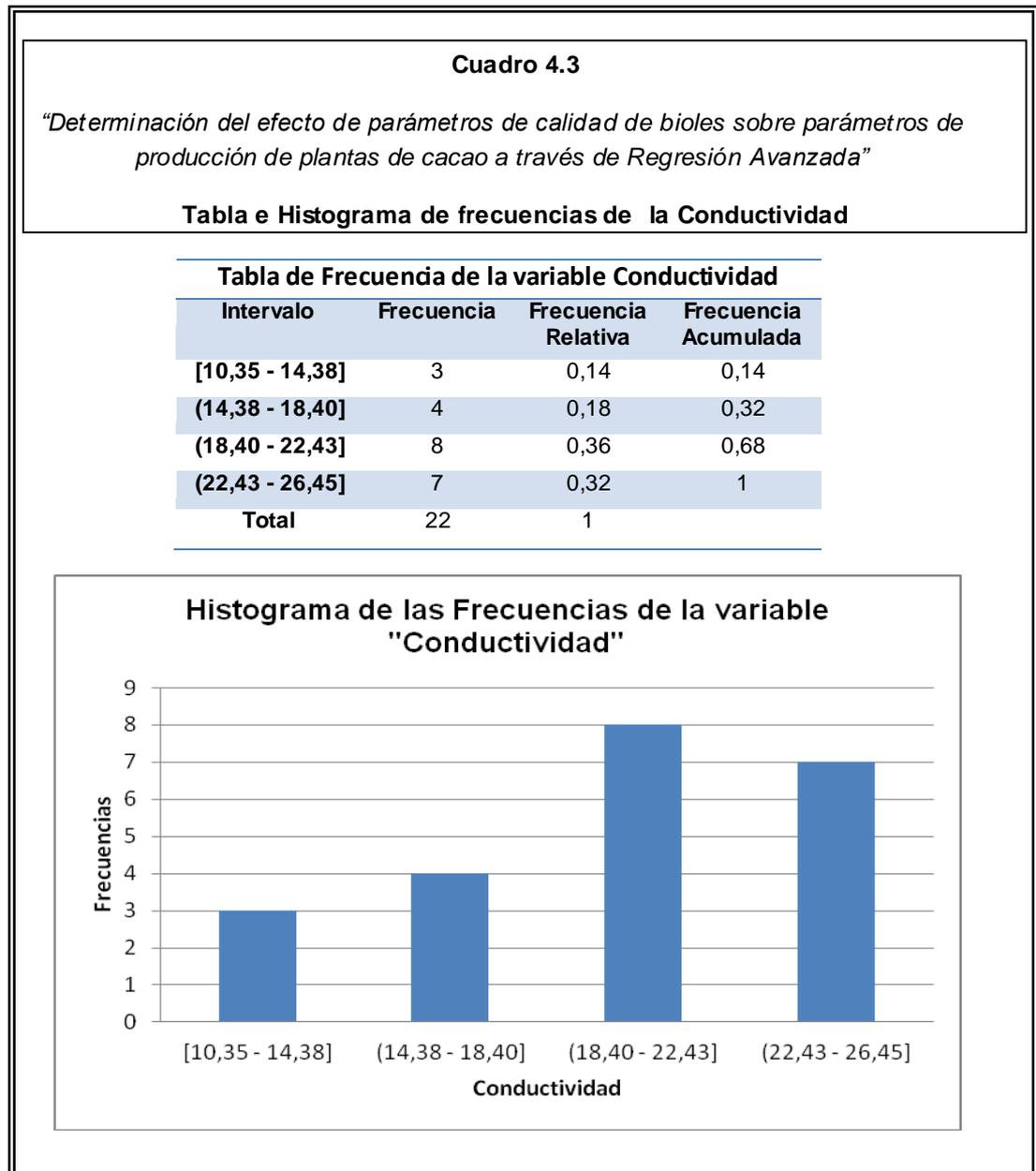
En el **Cuadro 4.1** podemos apreciar la tabla de frecuencias de la variable pH, se analizan cuatro marcas de clases o intervalos siendo el primer grupo los bioles que tienen un grado de acidez entre 3.17 y 3.65 son 4 del total de haciendas muestreadas esto es un 18%; el segundo grupo corresponde a los bioles que tienen un grado de acidez entre 3.65 y 4.13 son 6 del total de haciendas muestreadas esto es un 27%, el tercer grupo corresponde a los bioles que tienen un grado de acidez entre 4.13 y 4.61 son 5 del total del haciendas muestreadas esto es un 23% y finalmente el cuarto grupo corresponde a los bioles que tienen un grado de acidez entre 4.61 y 5.10 son 7 del total de haciendas muestreadas esto es un 32%.



**Elaborado por:** Karla M. Aguaguiña Méndez

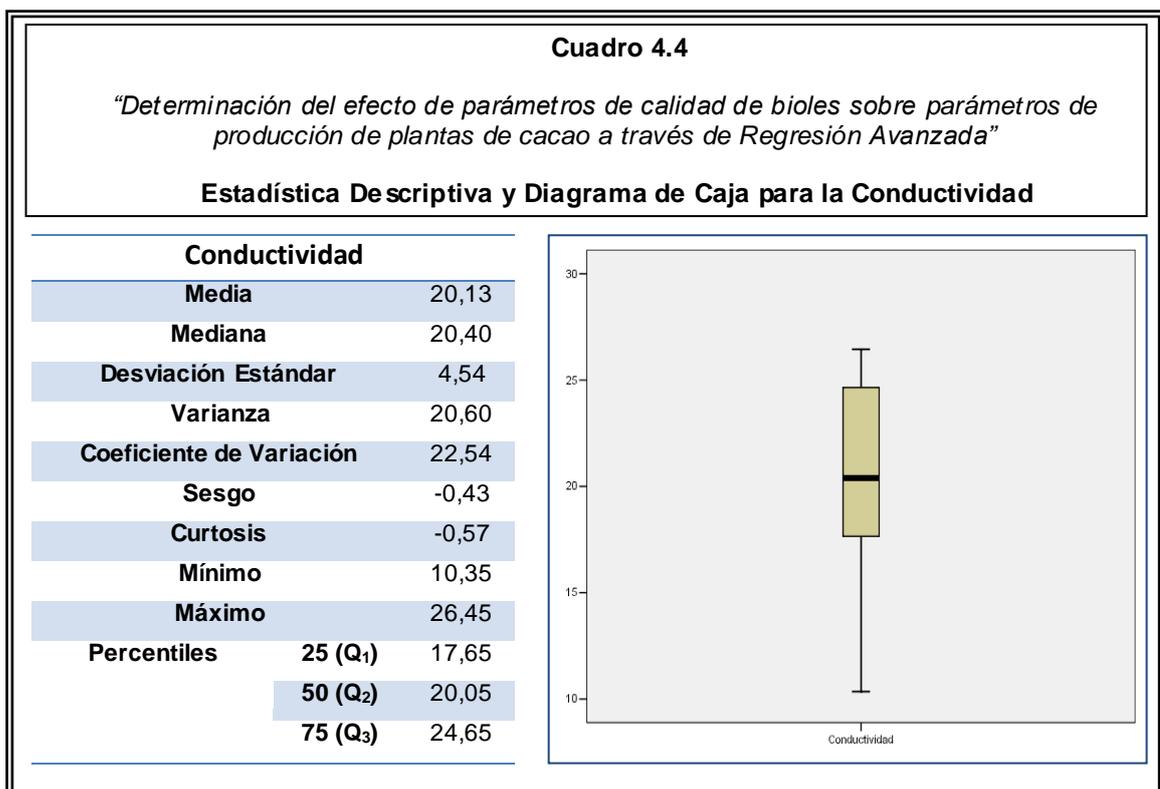
En el **Cuadro 4.2** se observa la estadística descriptiva básica de los datos de la variable pH, se aprecia que en promedio el grado de acidez de los bioles es de 4.22 lo que indica que estas soluciones son ácidas, la variabilidad con respecto a los datos es de 0.32, el valor central (mediana) del grado de acidez del pH de los bioles es de 4.16. El sesgo de la variable es de -0.01 es decir que la distribución de los datos es insesgada, los valores mínimo y máximo son 3.17 y 5.10 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 3.79, 4.15 y 4.77 respectivamente. Podemos indicar que el 25% de los bioles de las haciendas muestreadas tienen un pH menor o igual a 3.79, el 50% del total de bioles tienen un pH menor o igual a 4.15, y finalmente el 75% tiene un pH menor o igual a 4.77.

**Variable: Conductividad**

En el **Cuadro 4.3** tenemos la tabla de frecuencias de la variable Conductividad, se analizan cuatro grupos o intervalos, siendo el primer

grupo aquellos bioles con conductividad entre 10.35 y 14.38 con 3 del total de haciendas muestreadas esto es 14% , el segundo grupo corresponde a los bioles que tienen un rango de conductividad entre 14.38 y 18.40 con 4 del total de haciendas muestreadas esto es 18%, el tercer grupo corresponde a los bioles con conductividad entre 18.40 y 22.43 con 8 del total de haciendas muestreadas esto es un 36% y el cuarto grupo corresponde a los bioles con conductividad entre 22.43 y 26.45 con 7 del total de haciendas muestreadas esto es un 32%.



**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.4** se observa la estadística descriptiva básica de los datos de la variable Conductividad, se aprecia que el promedio de conductividad

de los bioles es de 20.13, la variabilidad con respecto a los datos es de 4.54, el valor central (mediana) de la conductividad de los bioles es de 4.16. El sesgo de la variable es de -0.43 es decir que la concentración de los datos esta a la derecha del promedio (20.13); los valores mínimo y máximo son 10.35 y 26.45 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 17.65, 20.05 y 24.65 respectivamente. Podemos indicar que el 25% de los bioles de las haciendas muestreadas tienen una conductividad menor o igual a 17.65, el 50% del total de bioles tienen una conductividad menor o igual a 20.05, y finalmente el 75% tiene una conductividad menor o igual a 24.65

**Variable: Solutos Totales**

**Cuadro 4.5**

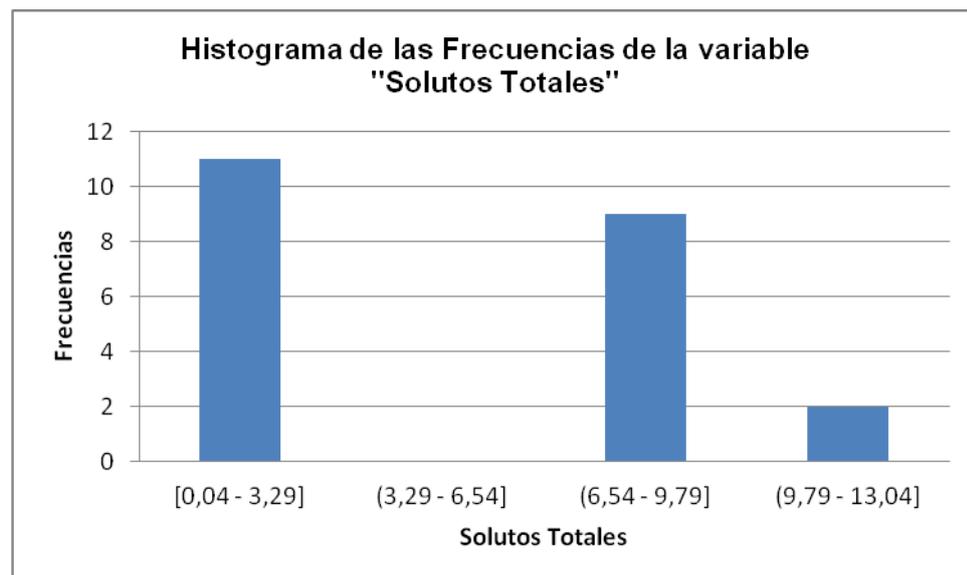
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de frecuencias de los Solutos Totales**

**Tabla de Frecuencia de la variable Solutos Totales**

Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0,04 - 3,29]	11	0,50	0,50
(3,29 - 6,54]	0	0,00	0,50
(6,54 - 9,79]	9	0,41	0,91
(9,79 - 13,04]	2	0,09	1
<b>Total</b>	<b>22</b>	<b>1</b>	

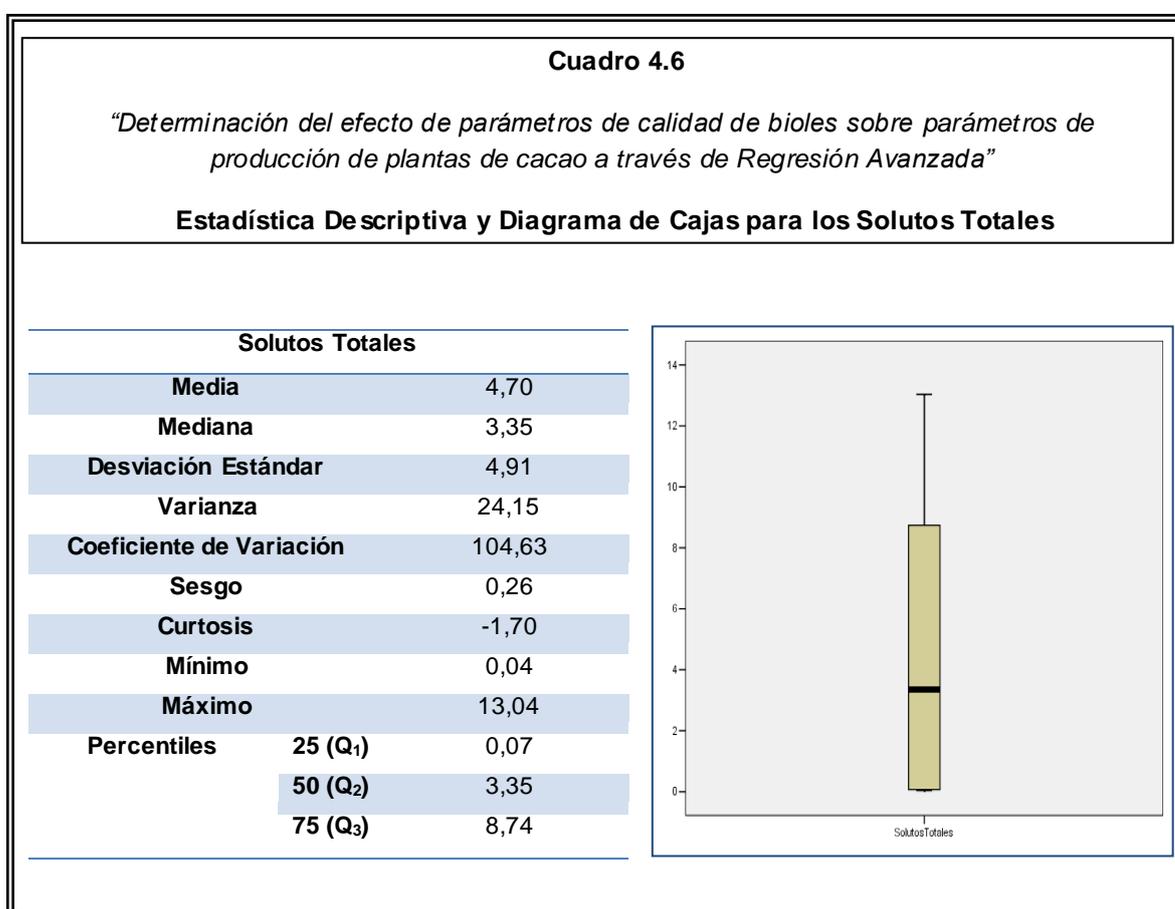
**Histograma de las Frecuencias de la variable "Solutos Totales"**



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.5** podemos apreciar la tabla de frecuencias de la variable Solutos Totales, se analizan cuatro marcas de clases o intervalos, siendo el primer grupo los bioles que tienen solutos totales entre 0.04 y 3.29 son

11 del total de haciendas muestreadas esto es un 50%; el segundo grupo corresponde a los bioles que tienen solutos totales entre 3.29 y 6.54 son 0 del total de haciendas muestreadas esto es un 0%; el tercer grupo corresponde a los bioles que tienen solutos totales entre 6.54 y 9.79 son 9 del total del haciendas muestreadas esto es un 41%; y finalmente el cuarto grupo corresponde a los bioles que tienen solutos totales entre 9.79 y 13.04 son 2 del total de haciendas muestreadas esto es un 9%.



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.6** se observa la estadística descriptiva básica de los datos, se aprecia que el promedio del contenido de solutos totales en los bioles es de 4.70, la variabilidad con respecto a los datos es de 4.71, el valor central (mediana) de los solutos totales en los bioles es de 3.35. El sesgo de la variable es de 0.26 es decir que la concentración de los datos esta a la izquierda del promedio (4.70); los valores mínimo y máximo son 0.04 y 13.04 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 0.07, 3.35 y 8.74 respectivamente. Podemos indicar que el 25% de los bioles de las haciendas muestreadas tienen solutos totales menor o igual a 0.07, el 50% del total de bioles tienen solutos totales menor o igual a 3.35, y finalmente el 75% tienen solutos totales menor o igual a 8.74.

**Variable: Grados Brix**

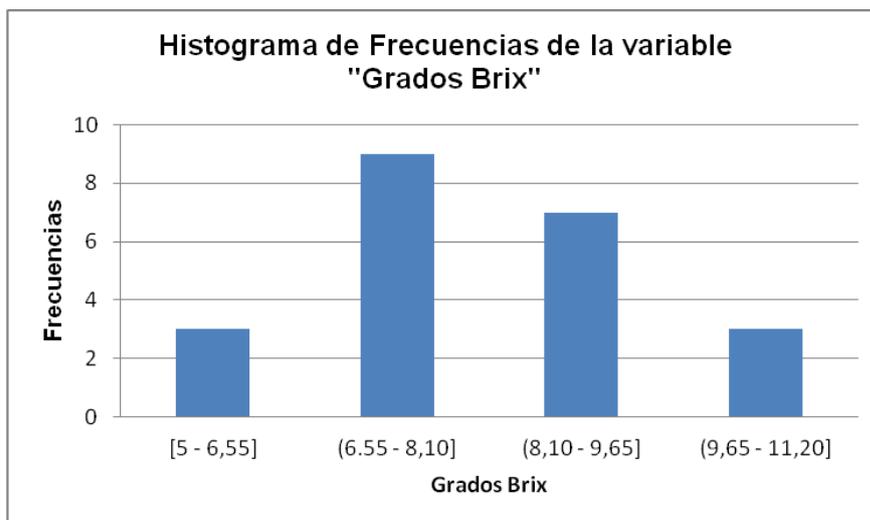
**Cuadro 4.7**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de frecuencias de los Grados Brix**

**Tabla de Frecuencia de la variable Grados Brix**

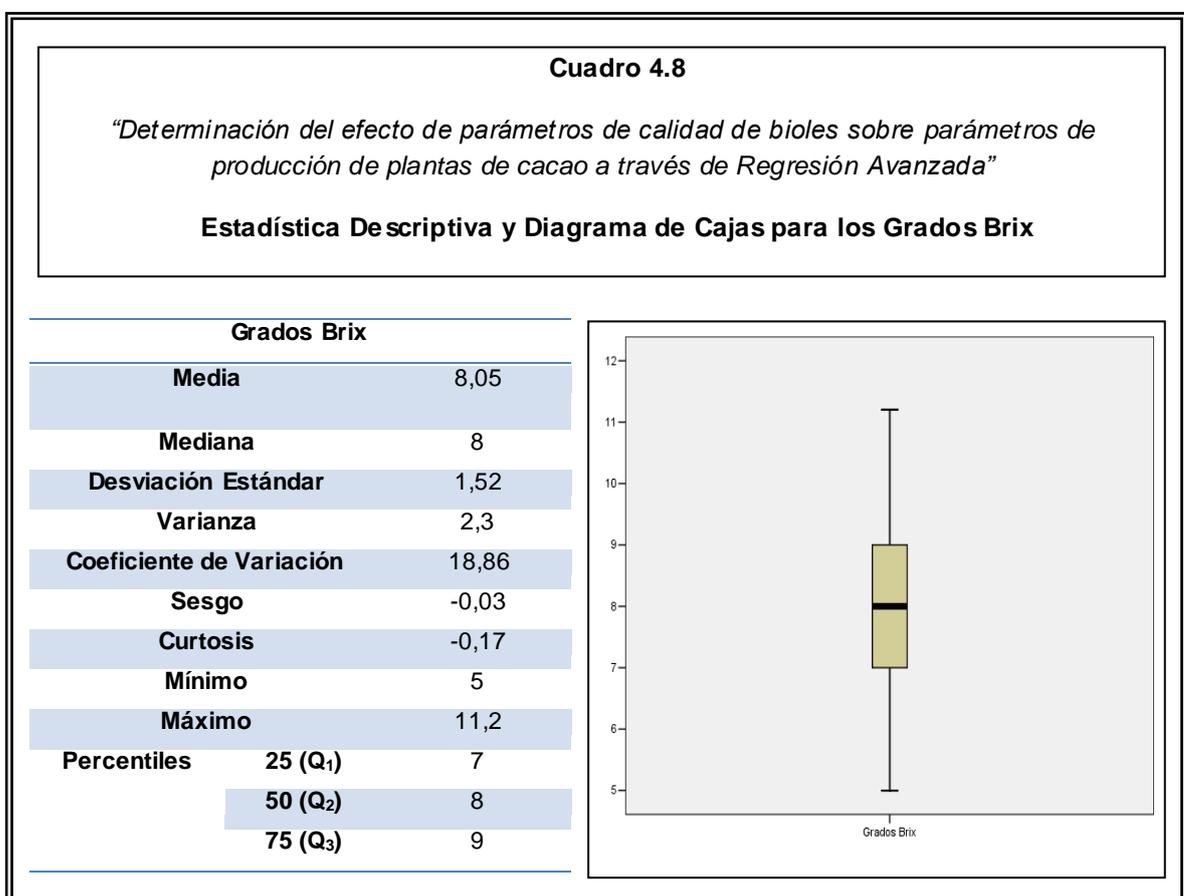
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[5 - 6,55]	3	0,14	0,14
(6,55 - 8,10]	9	0,41	0,55
(8,10 - 9,65]	7	0,32	0,86
(9,65 - 11,20]	3	0,14	1
<b>Total</b>	<b>22</b>	<b>1</b>	



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.7** podemos apreciar la tabla de frecuencias de la variable Grados Brix, se analizan cuatro marcas de clases o intervalos, siendo el primer grupo los bioles que tienen grados brix entre 5 y 6.55 son 3 del

total de haciendas muestreadas esto es un 14%; el segundo grupo corresponde a los bioles que tienen grados brix entre 6.55 y 8,10 son 9 del total de haciendas muestreadas esto es un 41%; el tercer grupo corresponde a los bioles que tienen grados brix entre 8.10 y 9.65 son 7 del total del haciendas muestreadas esto es un 32% y finalmente el cuarto grupo corresponde a los bioles que tienen grados brix entre 9.65 y 11.20 son 3 del total de haciendas muestreadas esto es un 14%.



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.8** se observa la estadística descriptiva básica de los datos de la variable Grados Brix, se aprecia que el promedio el contenido de

grados brix de los bioles es de 8.05, la variabilidad con respecto a los datos es de 2.3, el valor central (mediana) de los grados brix en los bioles es de 8. El sesgo de la variable es de -0.03 es decir que la concentración de los datos esta a la derecha del promedio (8.05); los valores mínimo y máximo son 5 y 11.20 respectivamente.

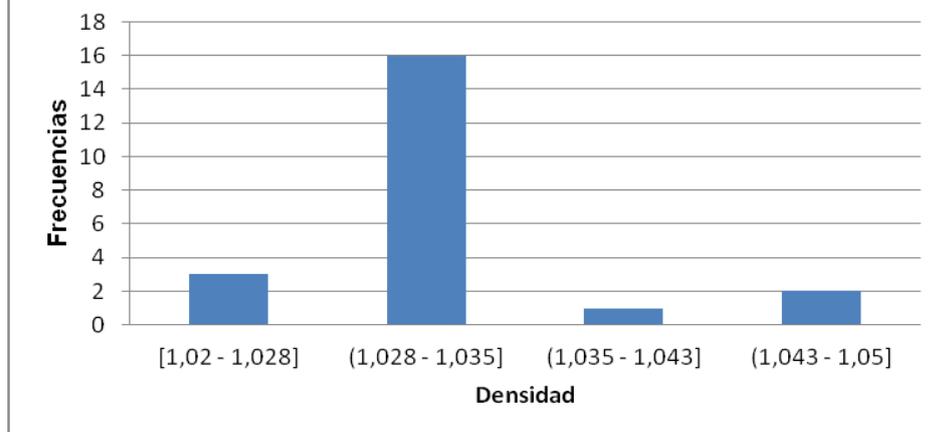
En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 7, 8 y 9 respectivamente. Podemos indicar que el 25% de los bioles de las haciendas muestreadas tienen un contenido de grados brix menor o igual a 7, el 50% del total de bioles tienen un contenido de grados brix menor o igual a 8 y finalmente el 75% tienen un contenido de grados brix menor o igual a 9.

**Variable: Densidad****Cuadro 4.9**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias de la Densidad****Tabla de Frecuencia de la variable Densidad**

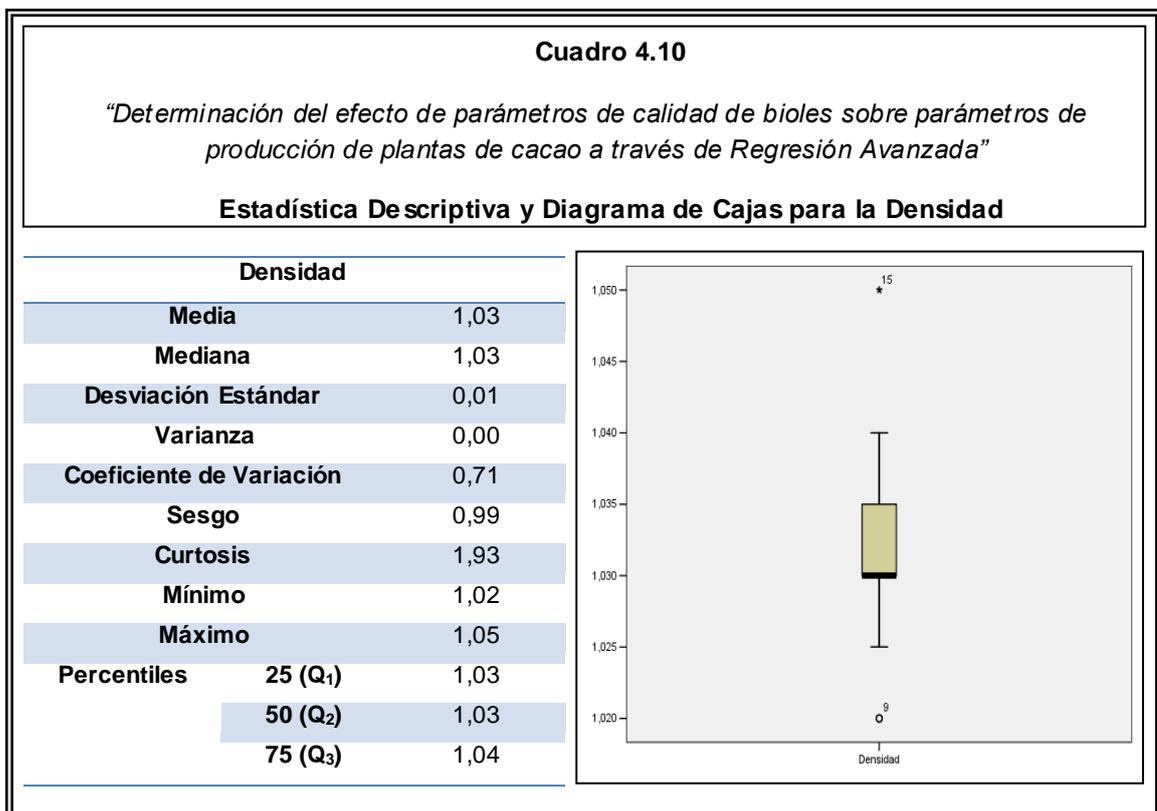
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[1,02 - 1,028]	3	0,14	0,14
(1,028 - 1,035]	16	0,73	0,86
(1,035 - 1,043]	1	0,05	0,91
(1,043 - 1,05]	2	0,09	1
<b>Total</b>	<b>22</b>	<b>1</b>	

**Histograma de Frecuencias de la variable "Densidad"**

**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.9** podemos apreciar la tabla de frecuencias de la variable Densidad se analizan cuatro marcas de clases o intervalos, siendo el

primer grupo los bioles que tienen una densidad entre 1.02 y 1.028 son 3 del total de haciendas muestreadas esto es un 14%; el segundo grupo corresponde a los bioles que tienen una densidad entre 1.028 y 1.035 son 16 del total de haciendas muestreadas esto es un 73%, el tercer grupo corresponde a los bioles que tienen una densidad entre 1.035 y 1.043 son 1 del total del haciendas muestreadas esto es un 5% y finalmente el cuarto grupo corresponde a los bioles que tienen una densidad entre 1.043 y 1.05 son 2 del total de haciendas muestreadas esto es un 9%.



**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.10** se observa la estadística descriptiva básica de los datos de la variable Densidad, se aprecia que el promedio la densidad de

los bioles es de 1.03, la variabilidad con respecto a los datos es de 0; el valor central (mediana) de la densidad en los bioles es de 1.03. El sesgo de la variable es de 0.99 es decir que la concentración de los datos esta a la izquierda del promedio (1.03); los valores mínimo y máximo son 1.02 y 1.05 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 1.03, 1.03 y 1.04 respectivamente. Podemos indicar que el 50% de los bioles de las haciendas muestreadas tienen una densidad menor o igual a 1.03 y finalmente el 75% tienen una densidad menor o igual a 1.04.

**Variable: Total de Materia Orgánica**

**Cuadro 4.11**

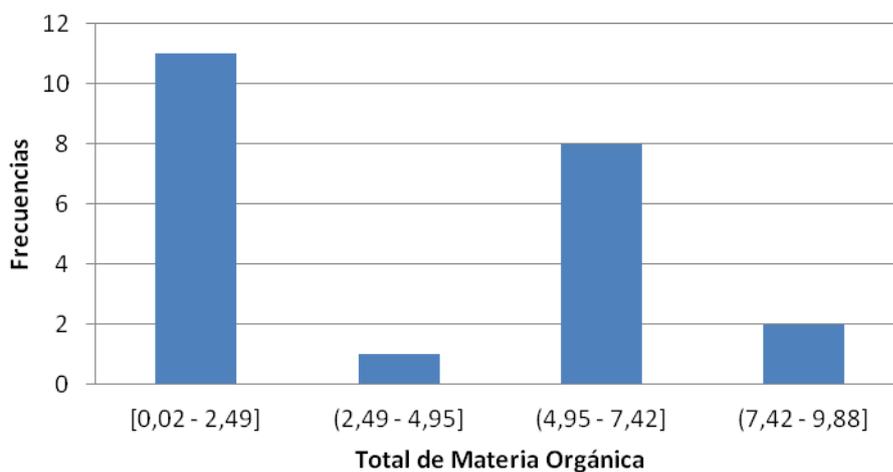
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias del Total de Materia Orgánica**

**Tabla de Frecuencia de la variable Total de Materia Orgánica**

Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0,02 - 2,49]	11	0,50	0,50
(2,49 - 4,95]	1	0,05	0,55
(4,95 - 7,42]	8	0,36	0,91
(7,42 - 9,88]	2	0,09	1
<b>Total</b>	<b>22</b>	<b>1</b>	

**Histograma de las Frecuencias de la variable "Total de materia Orgánica"**



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.11** podemos apreciar la tabla de frecuencias de la variable Total de Materia Orgánica, se analizan cuatro marcas de clases o

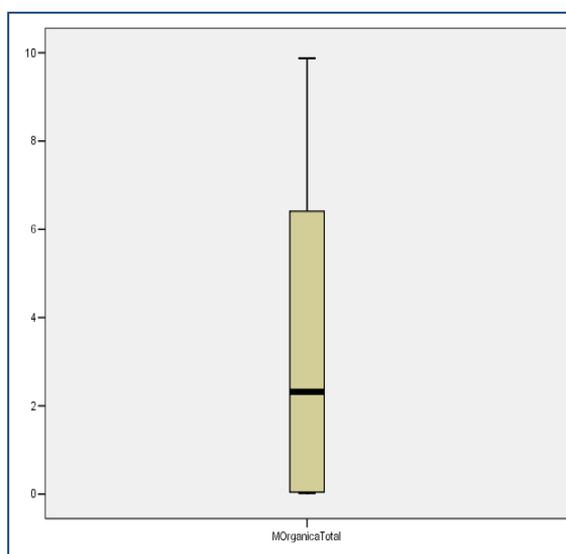
intervalos siendo el primer grupo los bioles que tienen un total de materia orgánica entre 0.02 y 2.49 son 11 del total de haciendas muestreadas esto es un 50%; el segundo grupo corresponde a los bioles que tienen un total de materia orgánica entre 2.49 y 4.95 son 1 del total de haciendas muestreadas esto es un 5%; el tercer grupo corresponde a los bioles que tienen un total de materia orgánica entre 4.95 y 7.42 son 8 del total del haciendas muestreadas esto es un 36% y finalmente el cuarto grupo corresponde a los bioles que tienen un total de materia orgánica entre 7.42 y 9.88 son 2 del total de haciendas muestreadas esto es un 9%.

**Cuadro 4.12**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Estadística Descriptiva y Diagrama de Cajas para el Total de Materia Orgánica**

<b>Total de Materia Orgánica</b>	
<b>Media</b>	3,40
<b>Mediana</b>	2,31
<b>Desviación Estándar</b>	3,60
<b>Varianza</b>	12,96
<b>Coefficiente de Variación</b>	105,96
<b>Sesgo</b>	0,34
<b>Curtosis</b>	-1,51
<b>Mínimo</b>	0,02
<b>Máximo</b>	9,88
<b>Percentiles</b>	
25 (Q <sub>1</sub> )	0,04
50 (Q <sub>2</sub> )	2,31
75 (Q <sub>3</sub> )	6,41



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.12** se observa la estadística descriptiva básica de los datos de la variable Total de Materia Orgánica, se aprecia que el promedio el total de materia orgánica de los bioles es de 3.04, la variabilidad con respecto a los datos es de 12.96, el valor central (mediana) del total de materia orgánica en los bioles es de 2.31. El sesgo de la variable es de 0.31 es decir que la concentración de los datos esta a la izquierda del promedio (3.04); los valores mínimo y máximo son 0.02 y 9.88 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 0.04, 2.31 y 6.41 respectivamente. Podemos indicar que el 25% de los bioles tienen un total de materia orgánica menor o igual a 0.04, el 50% de los bioles de las haciendas muestreadas tienen un total de materia orgánica menor o igual a 2.31 y finalmente el 75% tienen un total de materia orgánica menor o igual a 6.41.

## 4.1.2 Estadística Descriptiva de los Parámetros Microbiológicos

### Variable: Aerobios

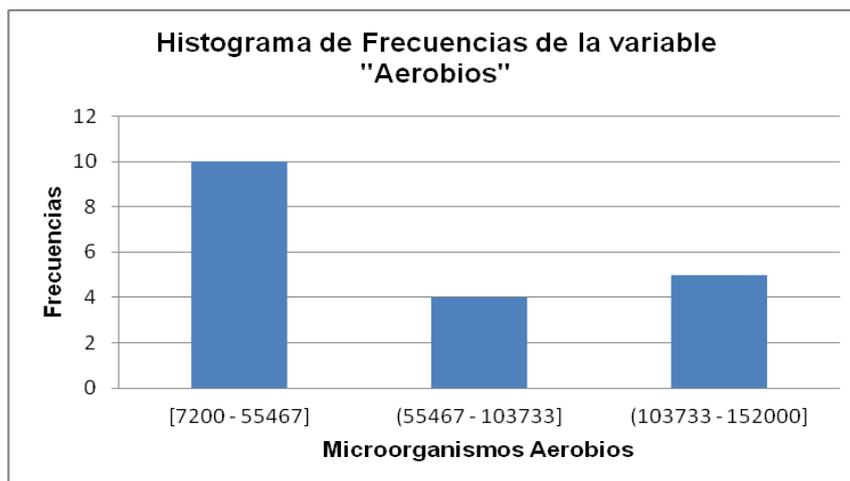
**Cuadro 4.13**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias de los Aerobios**

**Tabla: Frecuencias de la variable Aerobios**

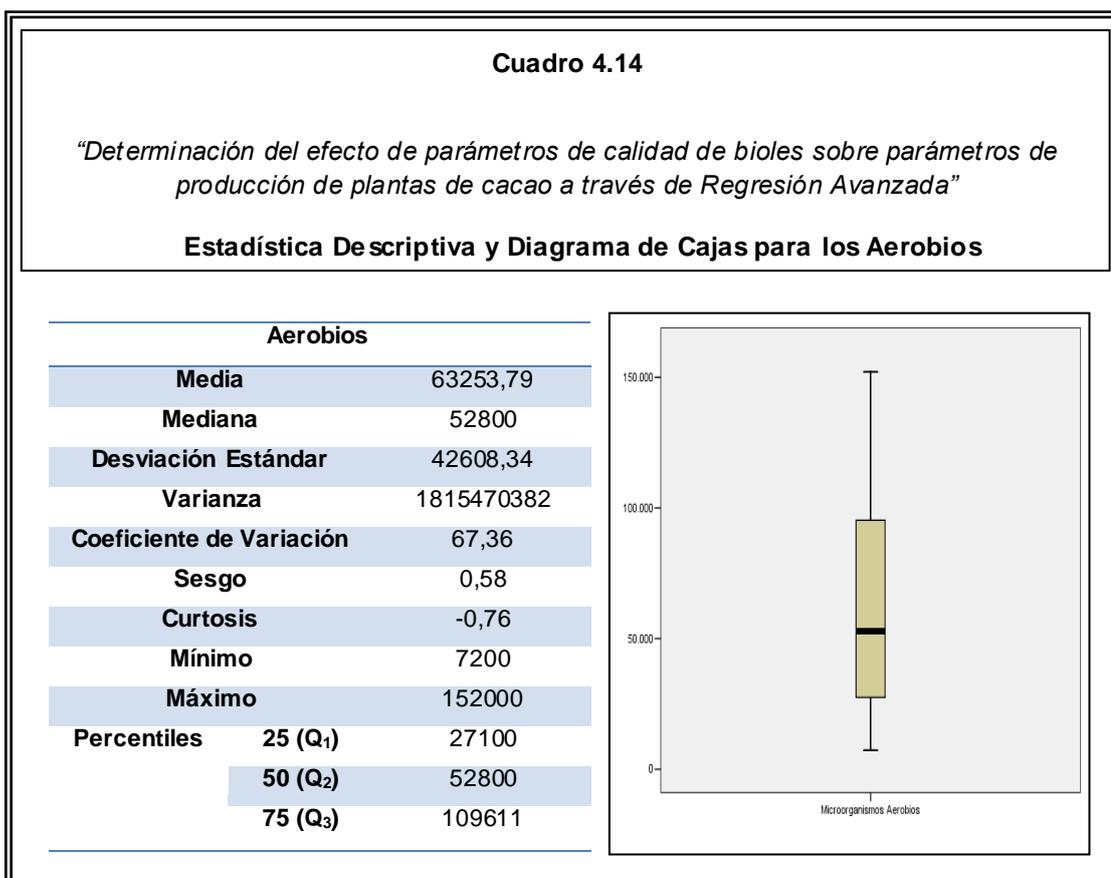
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[7200 - 55467]	10	0,53	0,53
(55467 - 103733]	4	0,21	0,74
(103733 - 152000]	5	0,26	1
<b>Total</b>	<b>19</b>	<b>1</b>	



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.13** podemos apreciar la tabla de frecuencias de la variable Aerobios, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un total de aerobios entre 7200 y 55467 son 10 del total de haciendas muestreadas esto es un 53%; el segundo grupo

corresponde a los bioles que tienen un total de aerobios entre 55467 y 103733 son 4 del total de haciendas muestreadas esto es un 21%, y finalmente el tercer grupo corresponde a los bioles que tienen un total de aerobios entre 103733 y 152000 son 5 del total del haciendas muestreadas esto es un 26%



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.14** se observa la estadística descriptiva básica de los datos de la variable Aerobios, se aprecia que el promedio del total de aerobios contenido en los bioles es de 63253.79, la variabilidad con respecto a los datos es de 1815470382, el valor central (mediana) del total

de aerobios en los bioles es de 52800. El sesgo de la variable es de 0.58 es decir que la concentración de los datos esta a la izquierda del promedio (63253.79); los valores mínimo y máximo son 7200 y 152000 respectivamente.

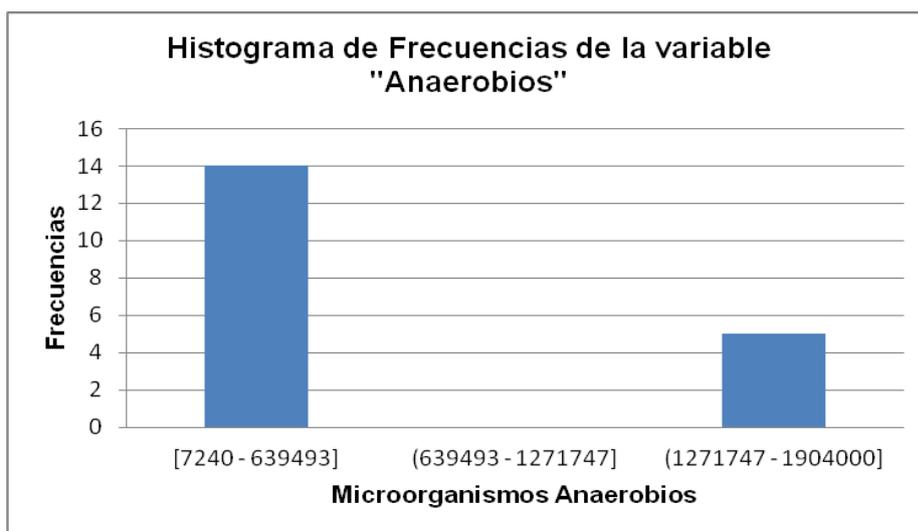
En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 27100, 52800 y 109611 respectivamente. Podemos indicar que el 25% de los bioles tienen un total de aerobios menor o igual a 27100, el 50% de los bioles de las haciendas muestreadas tienen un total de aerobios menor o igual a 52800 y finalmente el 75% tienen un total de aerobios menor o igual a 109611.

**Variable: Anaerobios****Cuadro 4.15**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias de los Anaerobios****Tabla: Frecuencias de la variable Anaerobios**

Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[7240 - 639493]	14	0,74	0,74
(639493 - 1271747]	0	0,00	0,74
(1271747 - 1904000]	5	0,26	1
<b>Total</b>	<b>19</b>	<b>1</b>	

**Histograma de Frecuencias de la variable "Anaerobios"**

**Elaborado por:** karla M. Aguaguiña Méndez

En el **Cuadro 4.15** podemos apreciar la tabla de frecuencias de la variable Anaerobios, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un total de anaerobios entre 7240 y 639493 son

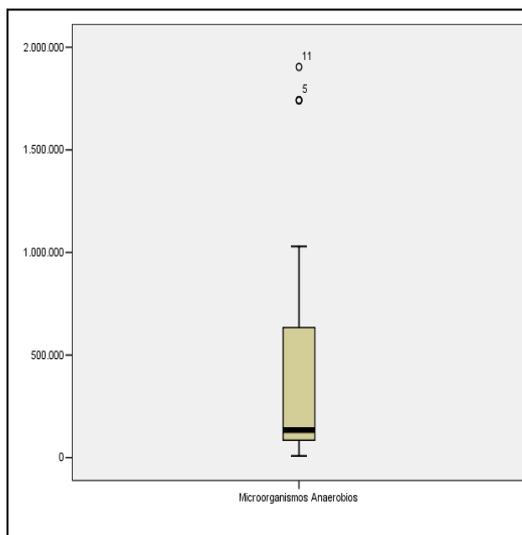
14 del total de haciendas muestreadas esto es un 74%; el segundo grupo corresponde a los bioles que tienen un total de anaerobios entre 639493 y 1271747 son 0 del total de haciendas muestreadas esto es un 0%, y finalmente el tercer grupo corresponde a los bioles que tienen un total de anaerobios entre 1271747 y 1094000 son 5 del total del haciendas muestreadas esto es un 26%.

**Cuadro 4.16**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Estadística Descriptiva y Diagrama de Cajas para los Anaerobios**

Anaerobios	
<b>Media</b>	510568,68
<b>Mediana</b>	135000
<b>Desviación Estándar</b>	709912,77
<b>Varianza</b>	503976138669,01
<b>Coefficiente de Variación</b>	139,04
<b>Sesgo</b>	1,3
<b>Curtosis</b>	-0,16
<b>Mínimo</b>	7240
<b>Máximo</b>	1904000
<b>Percentiles</b>	
<b>25 (Q<sub>1</sub>)</b>	84000
<b>50 (Q<sub>2</sub>)</b>	135000
<b>75 (Q<sub>3</sub>)</b>	1030000



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.16** se observa la estadística descriptiva básica de los datos de la variable Anaerobios, se aprecia que en promedio el total de anaerobios contenido en los bioles es de 510568.68, la variabilidad con

respecto a los datos es de 709912.77, el valor central (mediana) del total de anaerobios en los bioles es de 135000. El sesgo de la variable es de 1.03 es decir que la concentración de los datos esta a la izquierda del promedio (510568.68), los valores mínimo y máximo son 7240 y 1094000 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos de la variable Anaerobios, donde se localizan el primer, segundo y tercer cuartil siendo 84000, 135000 y 1030000 respectivamente. Podemos indicar que el 25% de los bioles tienen un total de anaerobios menor o igual a 84000, el 50% de los bioles de las haciendas muestreadas tienen un total de anaerobios menor o igual a 135000 y finalmente el 75% tienen un total de anaerobios menor o igual a 1030000.

**Variable: Hongos y Levaduras**

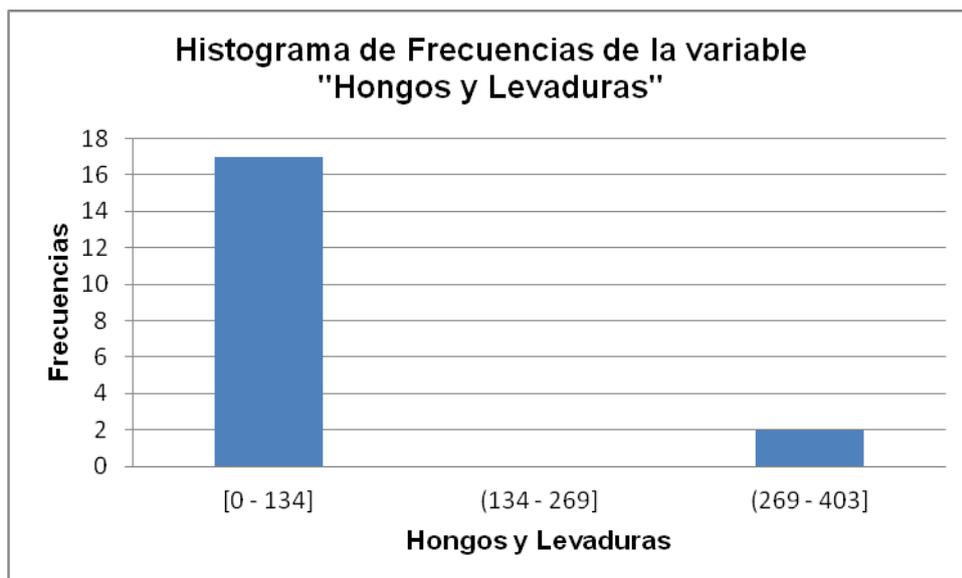
**Cuadro 4.17**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias de los Hongos y Levaduras**

**Tabla: Frecuencias de la variable Hongos y Levaduras**

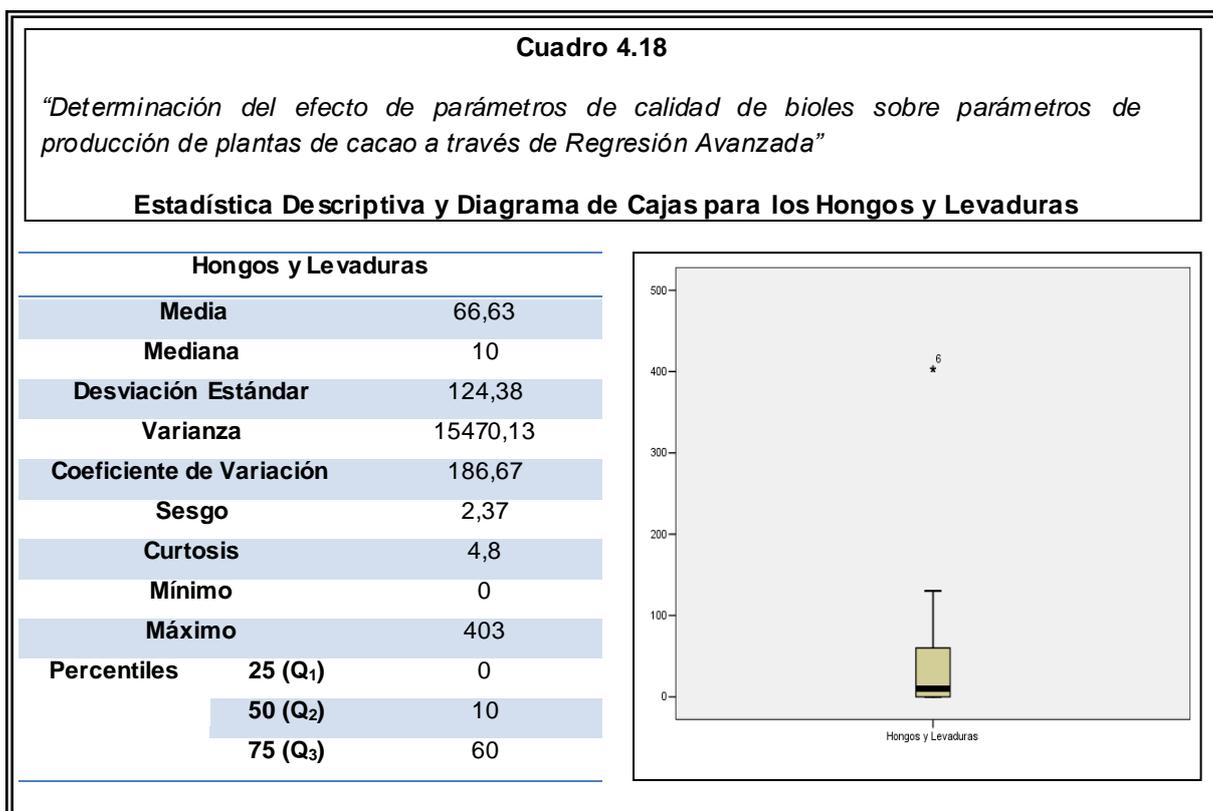
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0 - 134]	17	0,89	0,89
(134 - 269]	0	0,00	0,89
(269 - 403]	2	0,11	1
<b>Total</b>	<b>19</b>	<b>1</b>	



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.17** podemos apreciar la tabla de frecuencias de la variable Hongos y Levaduras, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un total de hongos y levaduras entre 0 y 134 son 17 del total de haciendas muestreadas esto es un 89%; el

segundo grupo corresponde a los bioles que tienen un total de hongos y levaduras entre 134 y 269 son 0 del total de haciendas muestreadas esto es un 0%, y finalmente el tercer grupo corresponde a los bioles que tienen un total de hongos y levaduras entre 269 y 403 son 2 del total del haciendas muestreadas esto es un 11%



**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.18** se observa la estadística descriptiva básica de los datos de la variable Hongos y Levaduras, se aprecia que el promedio del total de hongos y levaduras contenido en los bioles es de 66.63, la variabilidad con respecto a los datos es de 15470.13, el valor central (mediana) del total de hongos y levaduras en los bioles es de 10. El sesgo

de la variable es de 2.37 es decir que la concentración de los datos esta a la izquierda del promedio (66.63), los valores mínimo y máximo son 0 y 403 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 0, 10 y 60 respectivamente. Podemos indicar que el 25% de los bioles tienen un total de hongos y levaduras menor o igual a 0, el 50% de los bioles de las haciendas muestreadas tienen un total de hongos y levaduras menor o igual a 10 y finalmente el 75% tienen un total de hongos y levaduras menor o igual a 60.

### 4.1.3 Estadística Descriptiva de los macros – micros nutrientes

#### Variable: Porcentaje de Nitrógeno (N)

**Cuadro 4.19**

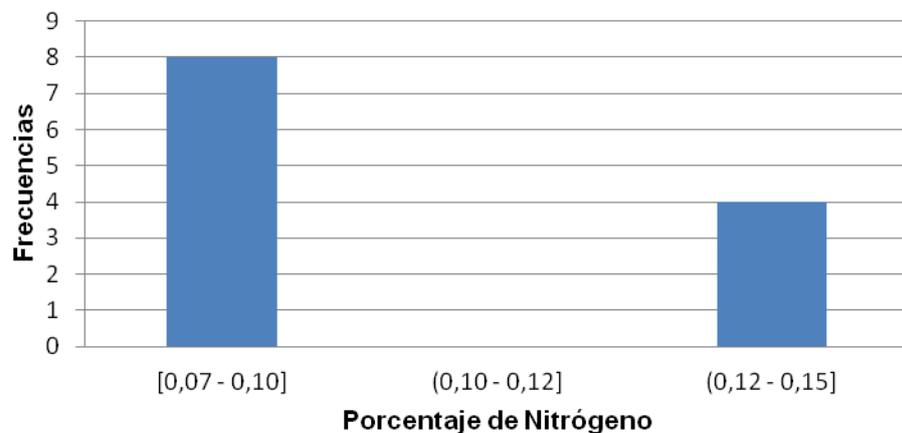
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

#### Tabla e Histograma de Frecuencias del porcentaje de Nitrógeno

**Tabla: Frecuencias de la variable Nitrógeno**

Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0,07 - 0,10]	8	0,67	0,67
(0,10 - 0,12]	0	0,00	0,67
(0,12 - 0,15]	4	0,33	1,00
<b>Total</b>	<b>12</b>	<b>1,00</b>	

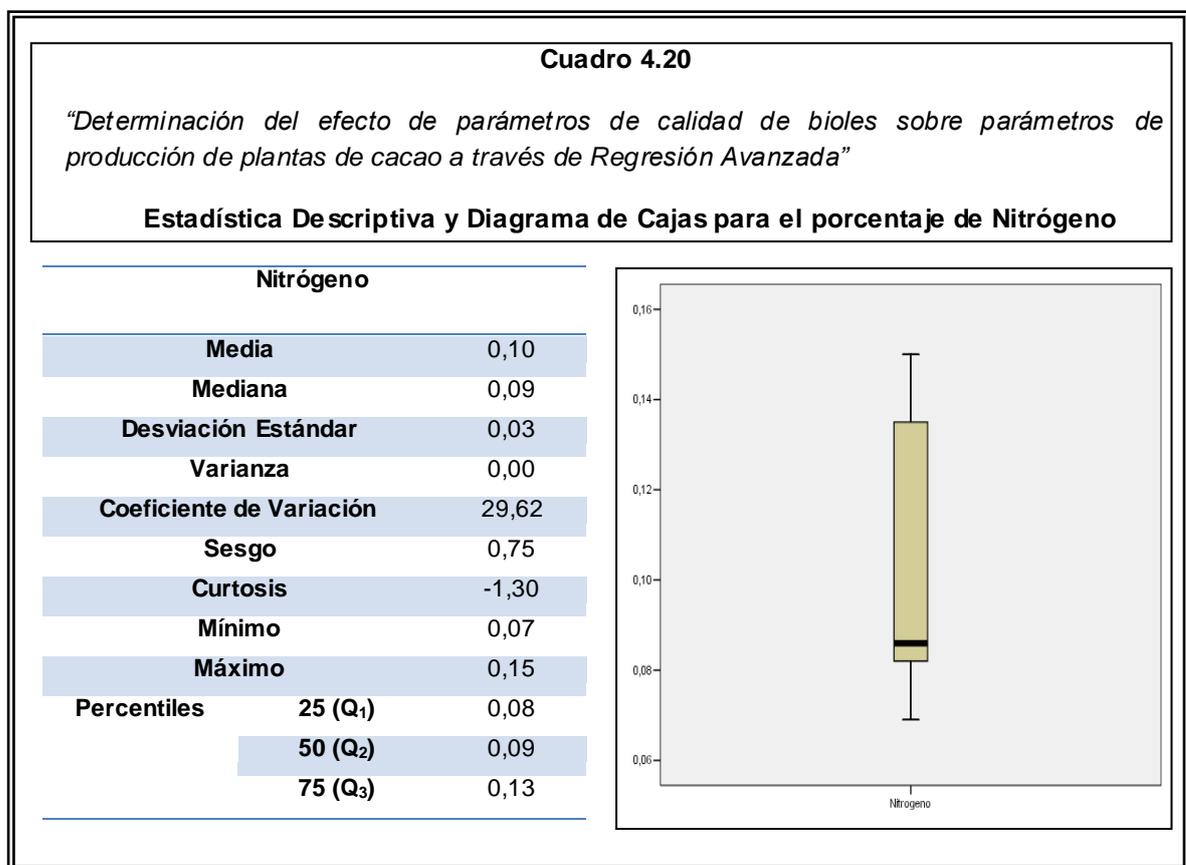
**Histograma de Frecuencias de la variable "Nitrógeno"**



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.19** podemos apreciar la tabla de frecuencias de la variable Porcentaje de Nitrógeno, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un porcentaje de Nitrógeno

entre 0.07 y 0.10 son 8 del total de haciendas muestreadas esto es un 67%; el segundo grupo corresponde a los bioles que tienen un porcentaje de Nitrógeno entre 0.10 y 0.12 son 0 del total de haciendas muestreadas esto es un 0%, y finalmente el tercer grupo corresponde a los bioles que tienen un porcentaje de Nitrógeno entre 0.12 y 0.15 son 4 del total del haciendas muestreadas esto es un 33%.



**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.20** se observa la estadística descriptiva básica de los datos de la variable Porcentaje de Nitrógeno, se aprecia que el promedio del porcentaje de Nitrógeno contenido en los bioles es de 0.10, la

variabilidad con respecto a los datos es de 0.00, el valor central (mediana) del porcentaje de nitrógeno en los bioles es de 0.09. El sesgo de la variable es de 0.75 es decir que la concentración de los datos esta a la izquierda del promedio (0.10); los valores mínimo y máximo son 0.07 y 0.15 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 0.08, 0.09 y 0.13 respectivamente. Podemos indicar que el 25% de los bioles tienen un porcentaje de Nitrógeno menor o igual a 0.08, el 50% de los bioles de las haciendas muestreadas tienen un porcentaje de nitrógeno menor o igual a 0.09 y finalmente el 75% tienen un porcentaje de nitrógeno menor o igual a 0.13.

**Variable: Ppm de Fósforo ( $P_2O_5$ )**

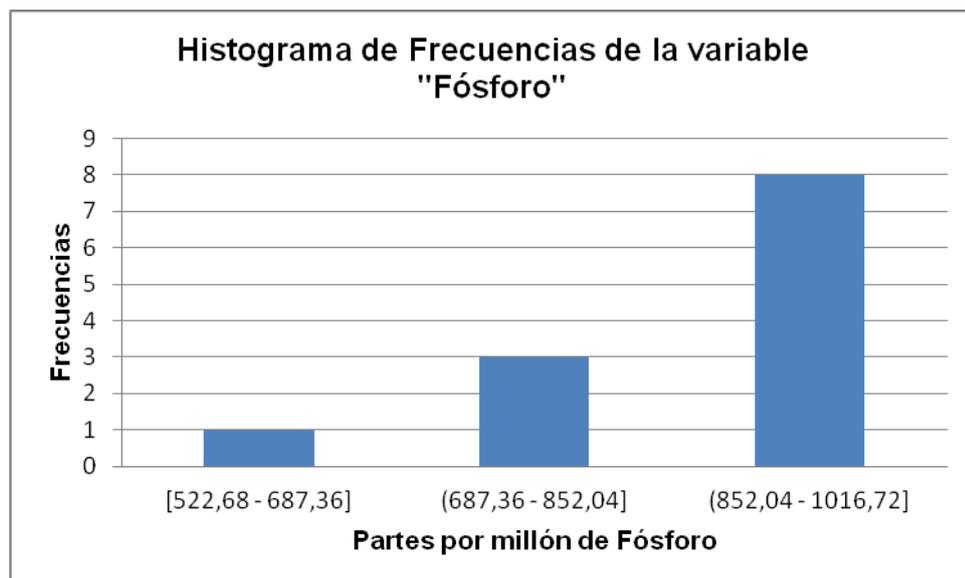
**Cuadro 4.21**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias de las partes por millón de Fósforo**

**Tabla: Frecuencias de la variable Fósforo**

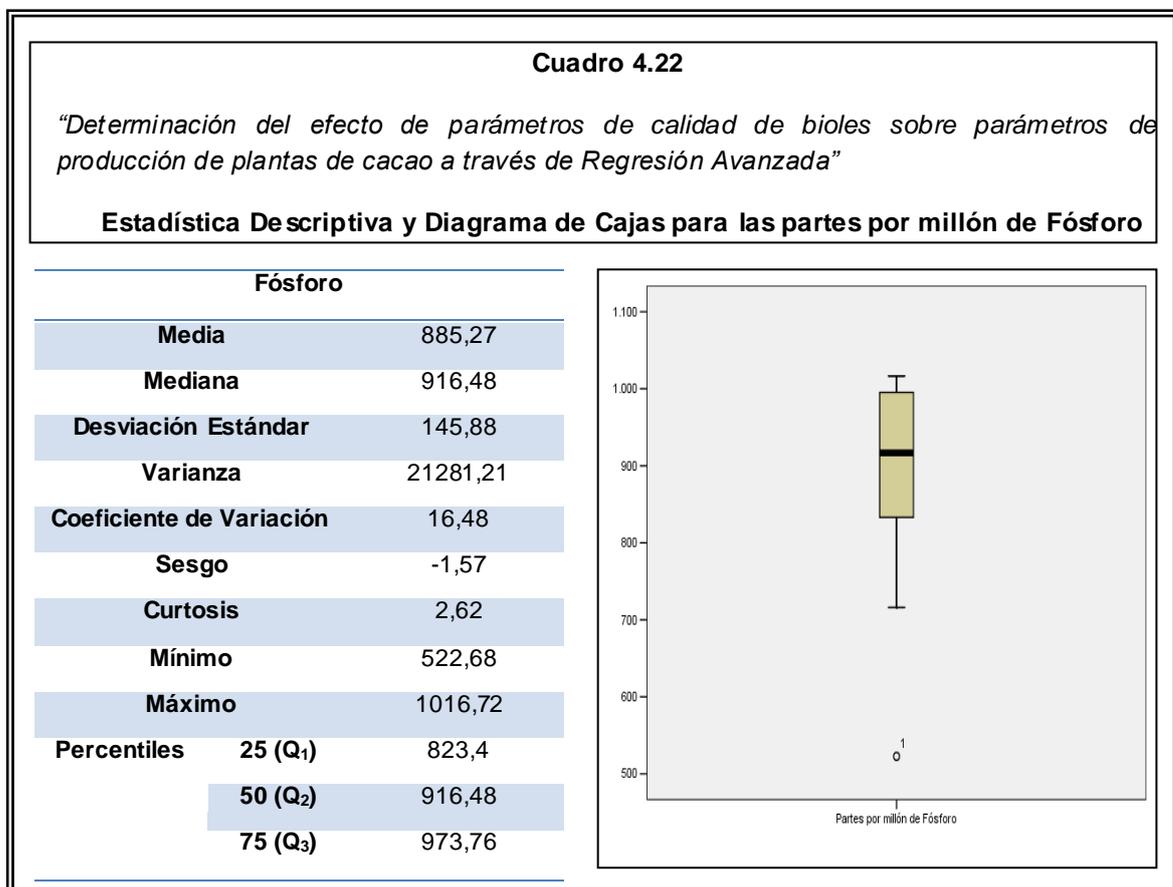
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[522,68 - 687,36]	1	0,08	0,08
(687,36 - 852,04]	3	0,25	0,33
(852,04 - 1016,72]	8	0,67	1
<b>Total</b>	<b>12</b>	<b>1</b>	



**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.21** podemos apreciar la tabla de frecuencias de la variable Partes por millón de Fósforo, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen partes por millón de

Fósforo entre 522.68 y 687.36 son 1 del total de haciendas muestreadas esto es un 8%; el segundo grupo corresponde a los bioles que tienen partes por millón de Fósforo entre 687 y 852.02 son 3 del total de haciendas muestreadas esto es un 25%, y finalmente el tercer grupo corresponde a los bioles que tienen partes por millón de Fósforo entre 852.02 y 1016.72 son 8 del total del haciendas muestreadas esto es un 67%.



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.22** se observa la estadística descriptiva básica de los datos de la variable Ppm de Fósforo, se aprecia que el promedio del

contenido de partes por millón de Fósforo en los bioles es de 885.27, la variabilidad con respecto a los datos es de 21281.21, el valor central (mediana) del contenido de partes por millón de Fósforo en los bioles es de 916.48. El sesgo de la variable es de -1.57 es decir que la concentración de los datos esta a la derecha del promedio (885.27); los valores mínimo y máximo son 522.68 y 1016.72 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 823.4, 916.48 y 973.76 respectivamente. Podemos indicar que el 25% de los bioles tienen partes por millón de Fósforo menor o igual a 823.4; el 50% de los bioles de las haciendas muestreadas tienen partes por millón de Fósforo menor o igual a 916.48 y finalmente el 75% tienen partes por millón de Fósforo menor o igual a 973.76.

**Variable: Porcentaje de Potasio ( $K_2O$ )**

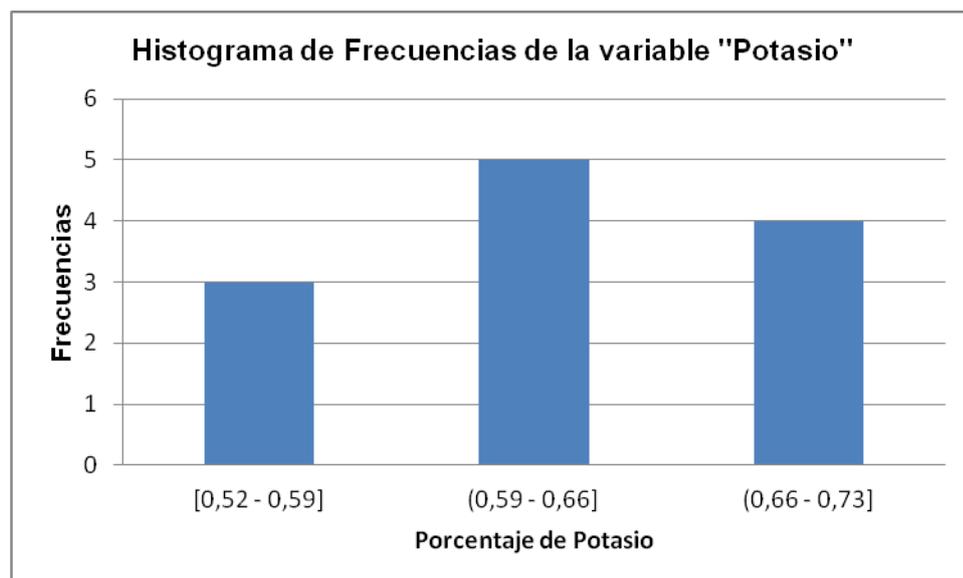
**Cuadro 4.23**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias del porcentaje de Potasio**

**Tabla: Frecuencias de la variable Potasio**

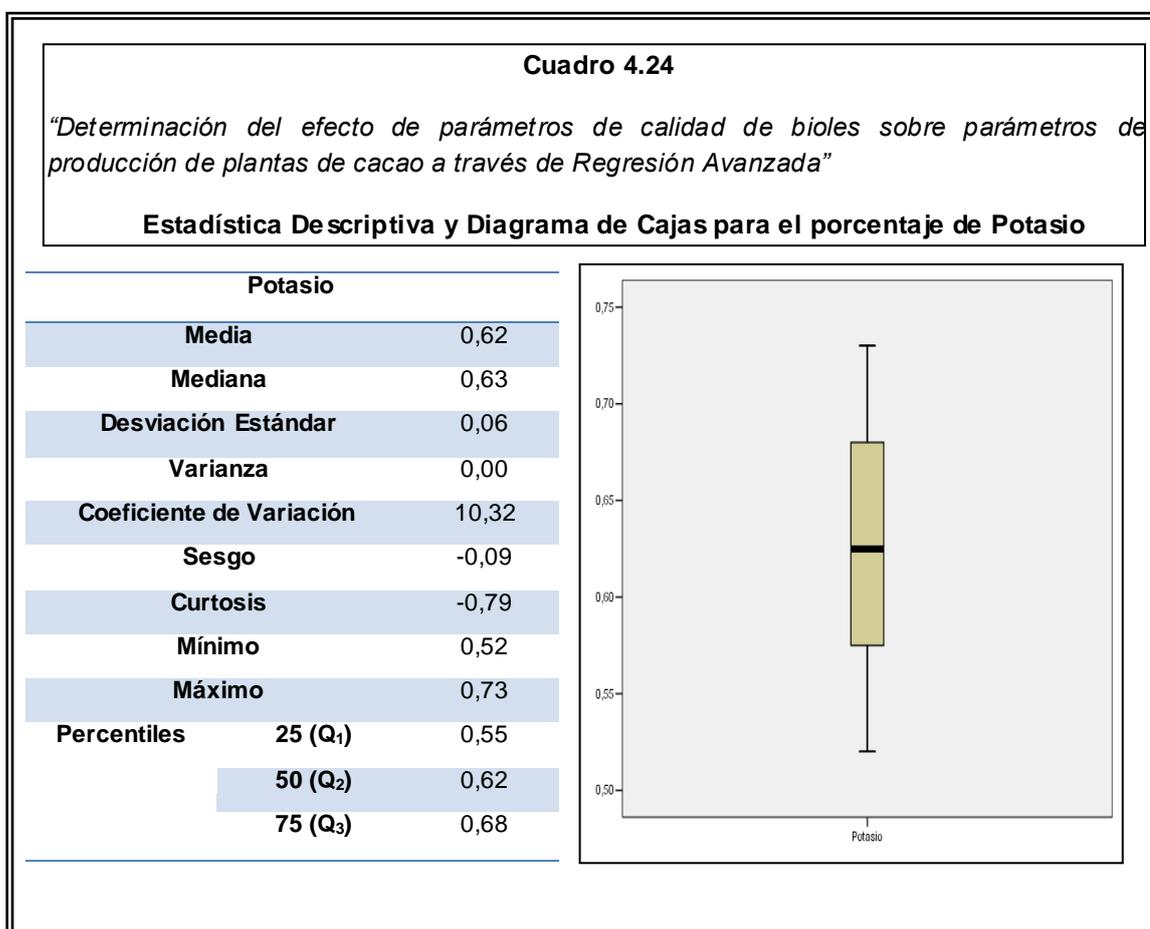
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0,52 - 0,59]	3	0,25	0,25
(0,59 - 0,66]	5	0,42	0,67
(0,66 - 0,73]	4	0,33	1
<b>Total</b>	<b>12</b>	<b>1</b>	



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.23** podemos apreciar la tabla de frecuencias de la variable Porcentaje de Potasio, se analizan tres marcas de clases o intervalos, siendo el primer grupo los bioles que tienen un porcentaje de Potasio

entre 0.52 y 0.59 son 3 del total de haciendas muestreadas esto es un 25%; el segundo grupo corresponde a los bioles que tienen un porcentaje de Potasio entre 0.59 y 0.66 son 5 del total de haciendas muestreadas esto es un 42%, y finalmente el tercer grupo corresponde a los bioles que tienen un porcentaje de Potasio entre 0.66 y 0.73 son 4 del total del haciendas muestreadas esto es un 33%.



**Elaborado por:** Karla M. Aguaguina Méndez

En el **Cuadro 4.24** se tiene se observa la estadística descriptiva básica de los datos de la variable Porcentaje de Potasio, se aprecia que el promedio

del porcentaje de Potasio en los bioles es de 0.62, la variabilidad con respecto a los datos es de 0.00, el valor central (mediana) del porcentaje de Potasio en los bioles es de 0.63. El sesgo de la variable es de -0.09 es decir que la concentración de los datos esta a la derecha del promedio (0.62); los valores mínimo y máximo son 0.52 y 0.73 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 0.55, 0.62 y 0.68 respectivamente. Podemos indicar que el 25% de los bioles tienen un porcentaje de Potasio menor o igual a 0.55, el 50% de los bioles de las haciendas muestreadas tienen un porcentaje de Potasio menor o igual a 0.62 y finalmente el 75% tienen un porcentaje de Potasio menor o igual a 0.68.

**Variable: Porcentaje de Calcio (Ca)**

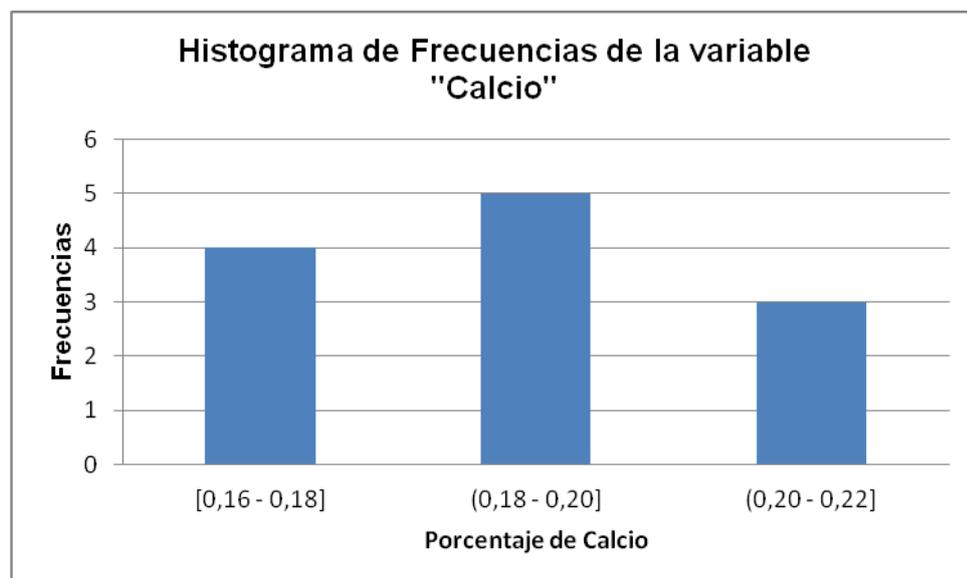
**Cuadro 4.25**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias del porcentaje de Calcio**

**Tabla: Frecuencias de la variable Calcio**

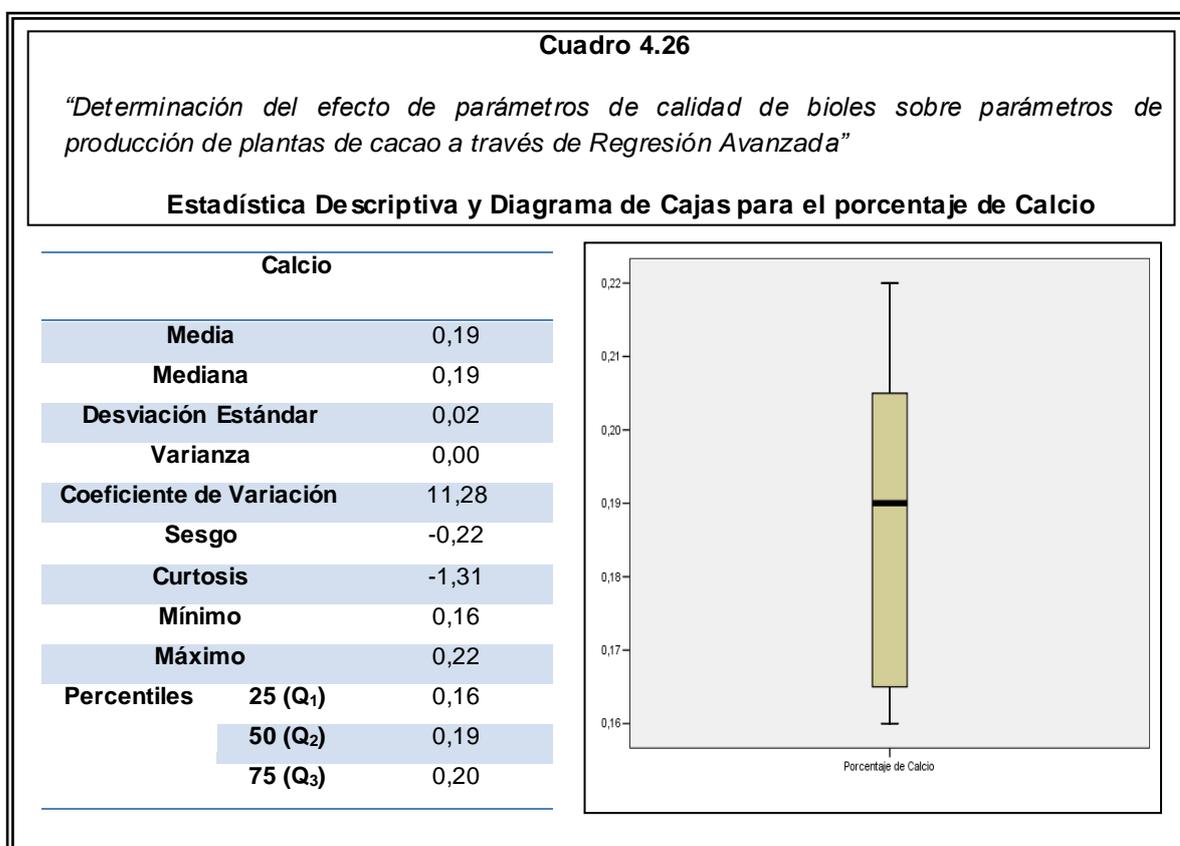
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0,16 - 0,18]	4	0,33	0,33
(0,18 - 0,20]	5	0,42	0,75
(0,20 - 0,22]	3	0,25	1
<b>Total</b>	<b>12</b>	<b>1</b>	



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.25** podemos apreciar la tabla de frecuencias de la variable Porcentaje de Calcio, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un porcentaje de Calcio entre

0.16 y 0.18 son 4 del total de haciendas muestreadas esto es un 33%; el segundo grupo corresponde a los bioles que tienen un porcentaje de Calcio entre 0.18 y 0.20 son 5 del total de haciendas muestreadas esto es un 42%, y finalmente el tercer grupo corresponde a los bioles que tienen un porcentaje de Calcio entre 0.20 y 0.22 son 3 del total del haciendas muestreadas esto es un 25%.



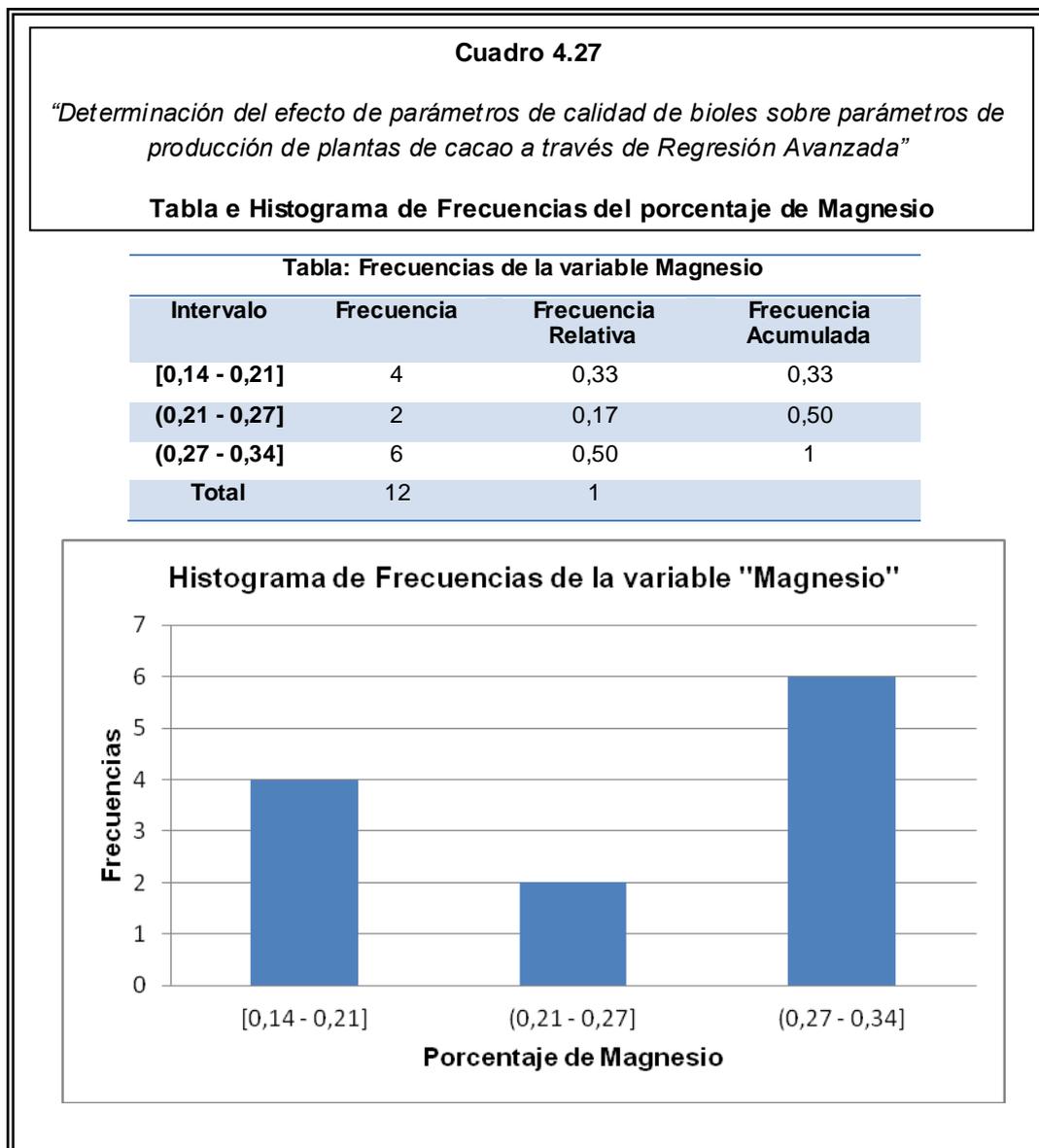
**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.26** se tiene se observa la estadística descriptiva básica de los datos de la variable Porcentaje de Calcio, se aprecia que en promedio el porcentaje de Calcio en los bioles es de 0.19, la variabilidad con

respecto a los datos es de 0.00, el valor central (mediana) del porcentaje de Calcio en los bioles es de 0.19. El sesgo de la variable es de -0.22 es decir que la concentración de los datos esta a la derecha del promedio (0.19); los valores mínimo y máximo son 0.16 y 0.22 respectivamente.

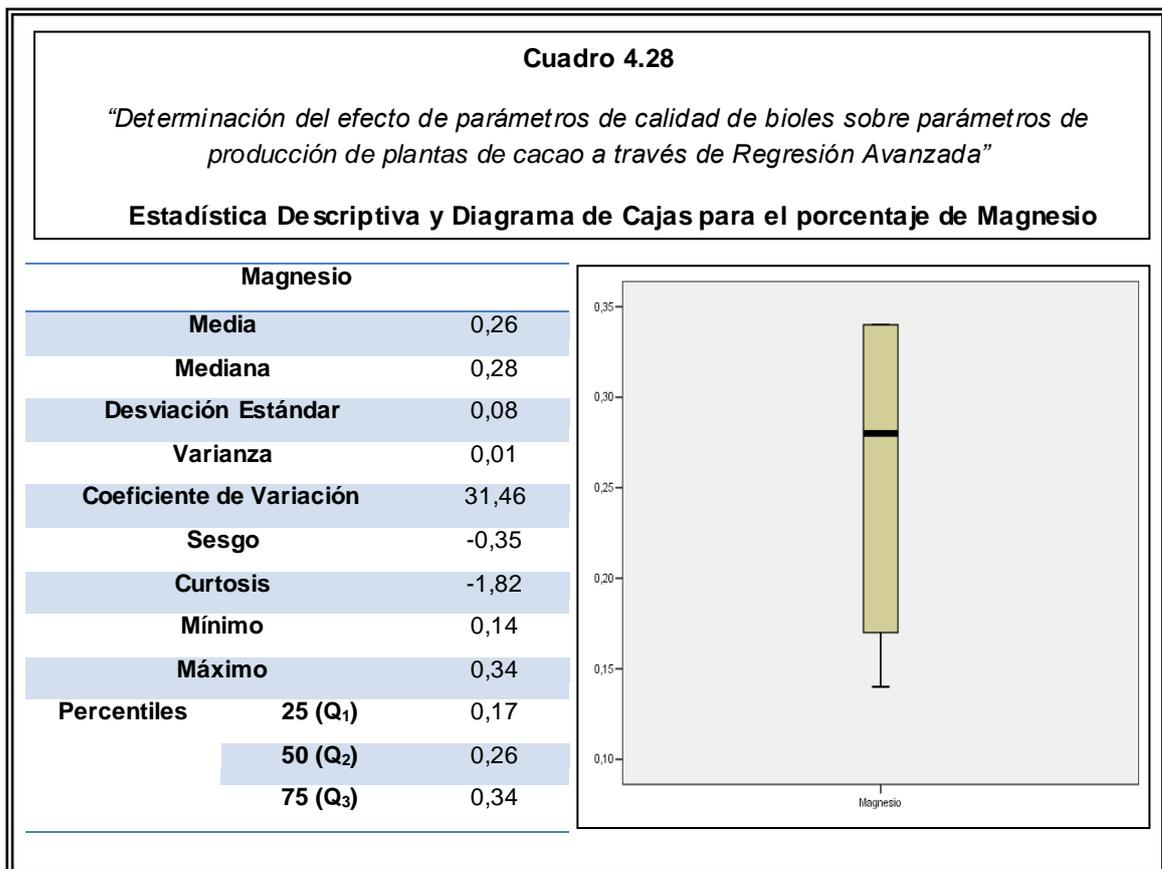
En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 0.16, 0.19 y 0.20 respectivamente. Podemos indicar que el 25% de los bioles tienen un porcentaje de Calcio menor o igual a 0.16, el 50% de los bioles de las haciendas muestreadas tienen un porcentaje de Calcio menor o igual a 0.19 y finalmente el 75% tienen un porcentaje de Calcio menor o igual a 0.20.

**Variable: Porcentaje de Magnesio (Mg)**



En el **Cuadro 4.27** podemos apreciar la tabla de frecuencias de la variable Porcentaje de Magnesio, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un porcentaje de Magnesio entre 0.14 y 0.21 son 4 del total de haciendas muestreadas esto es un

33%; el segundo grupo corresponde a los bioles que tienen un porcentaje de Magnesio entre 0.21 y 0.27 son 2 del total de haciendas muestreadas esto es un 17%, y finalmente el tercer grupo corresponde a los bioles que tienen un porcentaje de Magnesio entre 0.27 y 0.34 son 6 del total del haciendas muestreadas esto es un 50%.



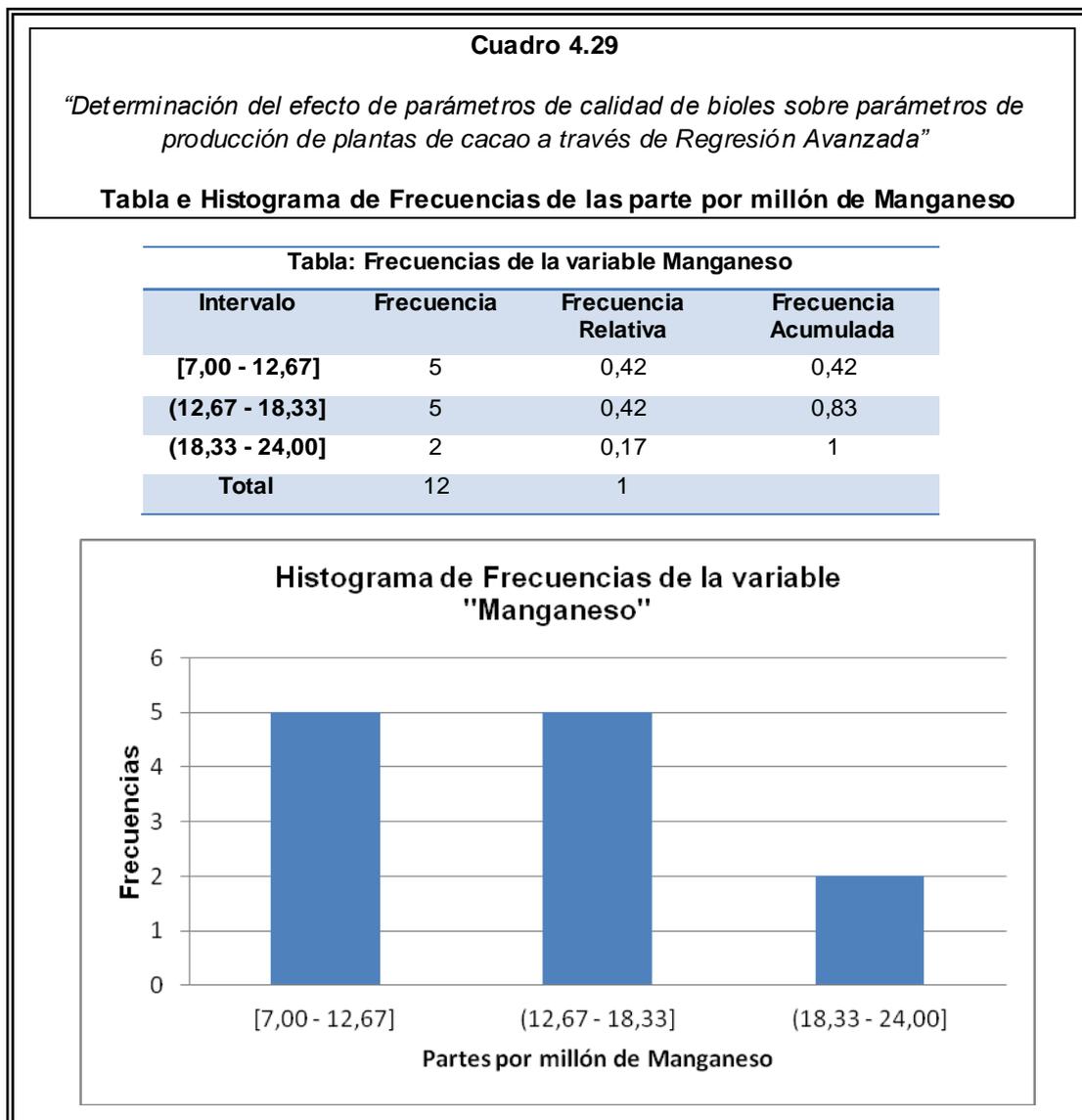
**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.28** se observa la estadística descriptiva básica de los datos Porcentaje de Magnesio, se aprecia que el promedio del porcentaje de Magnesio en los bioles es de 0.26, la variabilidad con respecto a los datos es de 0.01, el valor central (mediana) del porcentaje de Magnesio en

los bioles es de 0.28. El sesgo de la variable es de -0,35 es decir que la concentración de los datos esta a la derecha del promedio (0.26); los valores mínimo y máximo son 0.14 y 0.34 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 0.17, 0.26 y 0.34 respectivamente. Podemos indicar que el 25% de los bioles tienen un porcentaje de Magnesio menor o igual a 0.17, el 50% de los bioles de las haciendas muestreadas tienen un porcentaje de Magnesio menor o igual a 0.26 y finalmente el 75% tienen un porcentaje de Magnesio menor o igual a 0.34.

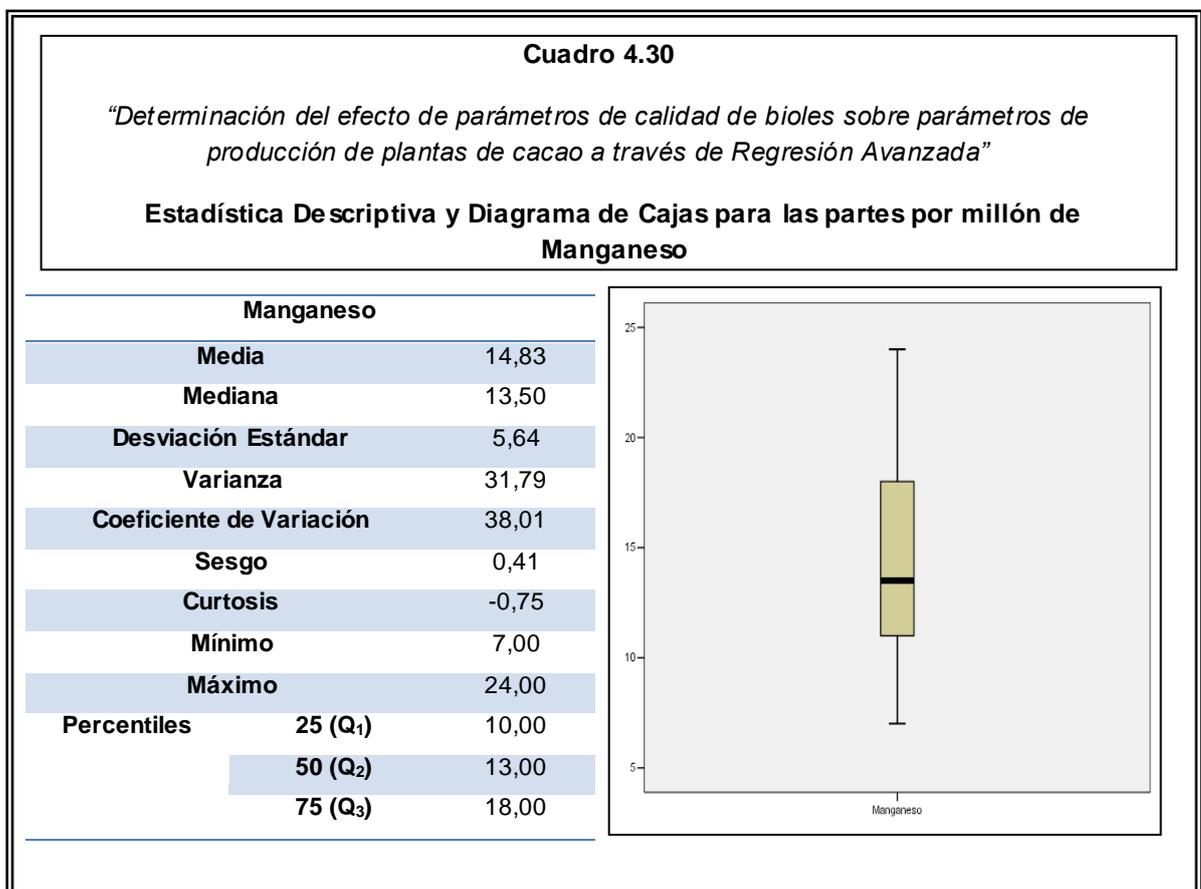
**Variable: Partes por millón de Manganeso (Mn)**



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.29** podemos apreciar la tabla de frecuencias de la variable Partes por millón de Manganeso, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un contenido en partes por millón de Manganeso entre 7 y 12.67 son 5 del total de

haciendas muestreadas esto es un 42%; el segundo grupo corresponde a los bioles que tienen un contenido de partes por millón de Manganeso entre 12.67 y 18.33 son 5 del total de haciendas muestreadas esto es un 42%, y finalmente el tercer grupo corresponde a los bioles que tienen un contenido de partes por millón de Manganeso entre 18.33 y 24 son 2 del total del haciendas muestreadas esto es un 17%.



**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.30** se observa la estadística descriptiva básica de los datos de las partes por millón de Manganeso, se aprecia que el promedio del contenido en partes por millón de Manganeso en los bioles es de

14.83, la variabilidad con respecto a los datos es de 31.79, el valor central (mediana) del contenido en partes por millón de Manganeso en los bioles es de 13.50. El sesgo de la variable es de 0.41 es decir que la concentración de los datos esta a la izquierda del promedio (14.83); los valores mínimo y máximo son 7 y 24 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 10, 13 y 18 respectivamente. Podemos indicar que el 25% de los bioles tienen un contenido en partes por millón de Manganeso menor o igual a 10, el 50% de los bioles de las haciendas muestreadas tienen un contenido en partes por millón de Manganeso menor o igual a 13 y finalmente el 75% tienen un contenido en partes por millón de Manganeso menor o igual a 18.

**Variable: Partes por millón de Hierro (Fe)**

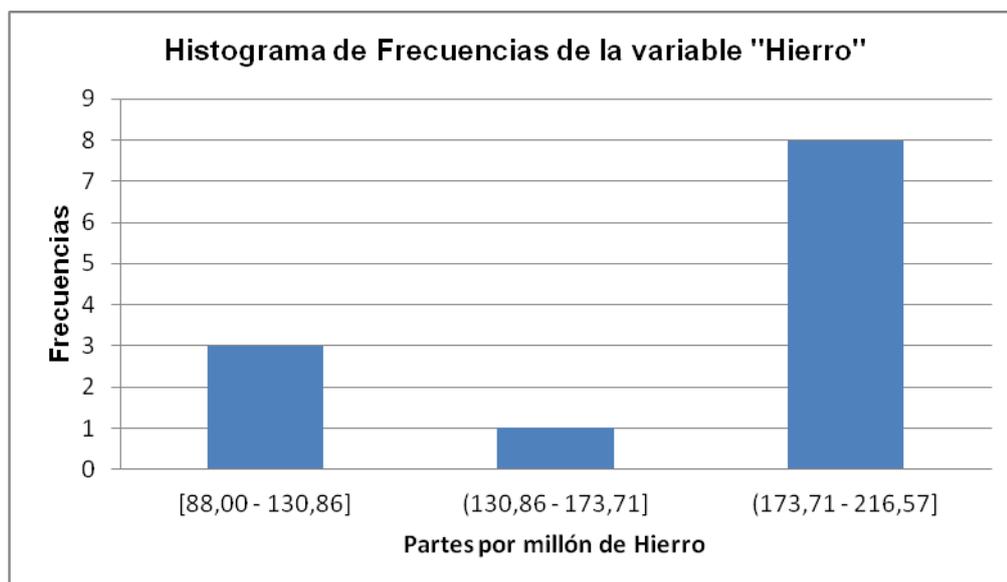
**Cuadro 4.31**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias de las parte por millón de Hierro**

**Tabla: Frecuencias de la variable Hierro**

Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[88,00 - 130,86]	3	0,25	0,25
(130,86 - 173,71]	1	0,08	0,33
(173,71 - 216,57]	8	0,67	1
<b>Total</b>	<b>12</b>	<b>1</b>	



**Elaborado por:** Karla M. Aguaguíña Méndez

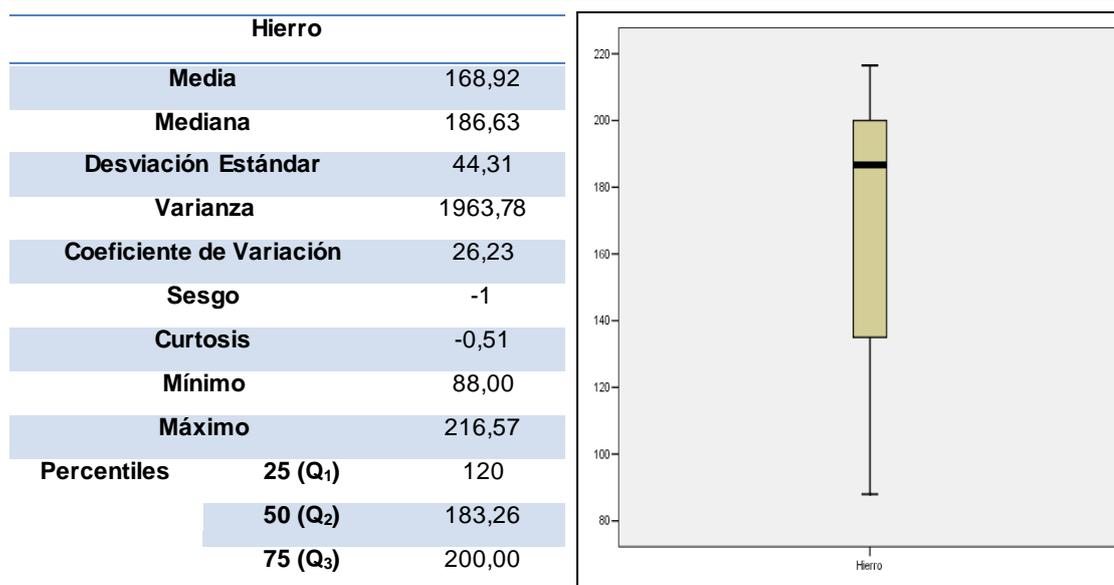
En el **Cuadro 4.31** podemos apreciar la tabla de frecuencias de la variable Partes por millón de Hierro, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un contenido en partes por millón de Hierro entre 88 y 130.86 son 3 del total de haciendas

muestreadas esto es un 25%; el segundo grupo corresponde a los bioles que tienen un contenido de partes por millón de Hierro entre 130.86 y 173.71 son 1 del total de haciendas muestreadas esto es un 8%, y finalmente el tercer grupo corresponde a los bioles que tienen un contenido de partes por millón de Hierro entre 173.71 y 216.57 son 8 del total del haciendas muestreadas esto es un 67%.

**Cuadro 4.32**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Estadística Descriptiva y Diagrama de Cajas para las partes por millón de Hierro**



**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.32** se observa la estadística descriptiva básica de los datos de la variable partes por millón de Hierro, se aprecia que el

promedio del contenido en partes por millón de Hierro en los bioles es de 168.92, la variabilidad con respecto a los datos es de 1963.78, el valor central (mediana) del contenido en partes por millón de Hierro en los bioles es de 186.63. El sesgo de la variable es de -1 es decir que la concentración de los datos esta a la derecha del promedio (168.92); los valores mínimo y máximo son 88 y 216.57 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 120, 183.26 y 200 respectivamente. Podemos indicar que el 25% de los bioles tienen un contenido en partes por millón de Hierro menor o igual a 120, el 50% de los bioles de las haciendas muestreadas tienen un contenido en partes por millón de Hierro menor o igual a 183.26 y finalmente el 75% tienen un contenido en partes por millón de Hierro menor o igual a 200.

**Variable: Partes por millón de Cobre (Cu)**

**Cuadro 4.33**

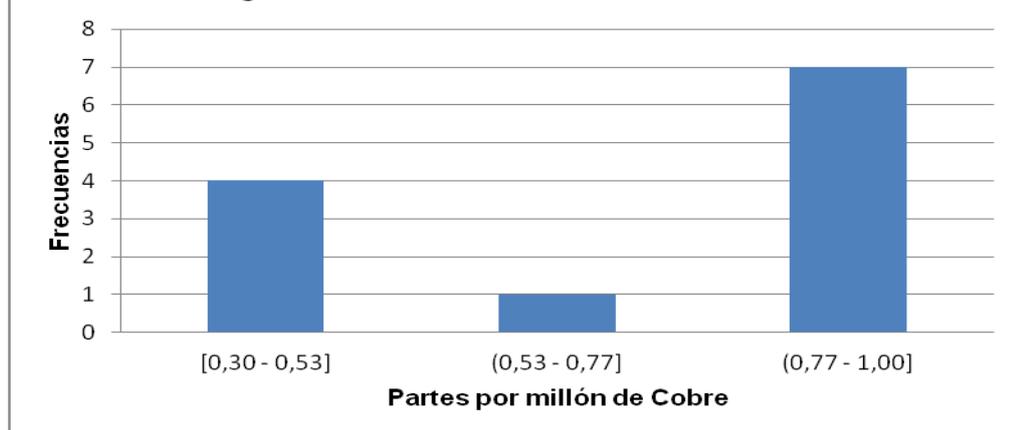
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias de las parte por millón de Cobre**

**Tabla: Frecuencias de la variable Cobre**

Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0,30 - 0,53]	4	0,33	0,33
(0,53 - 0,77]	1	0,08	0,42
(0,77 - 1,00]	7	0,58	1
<b>Total</b>	<b>12</b>	<b>1</b>	

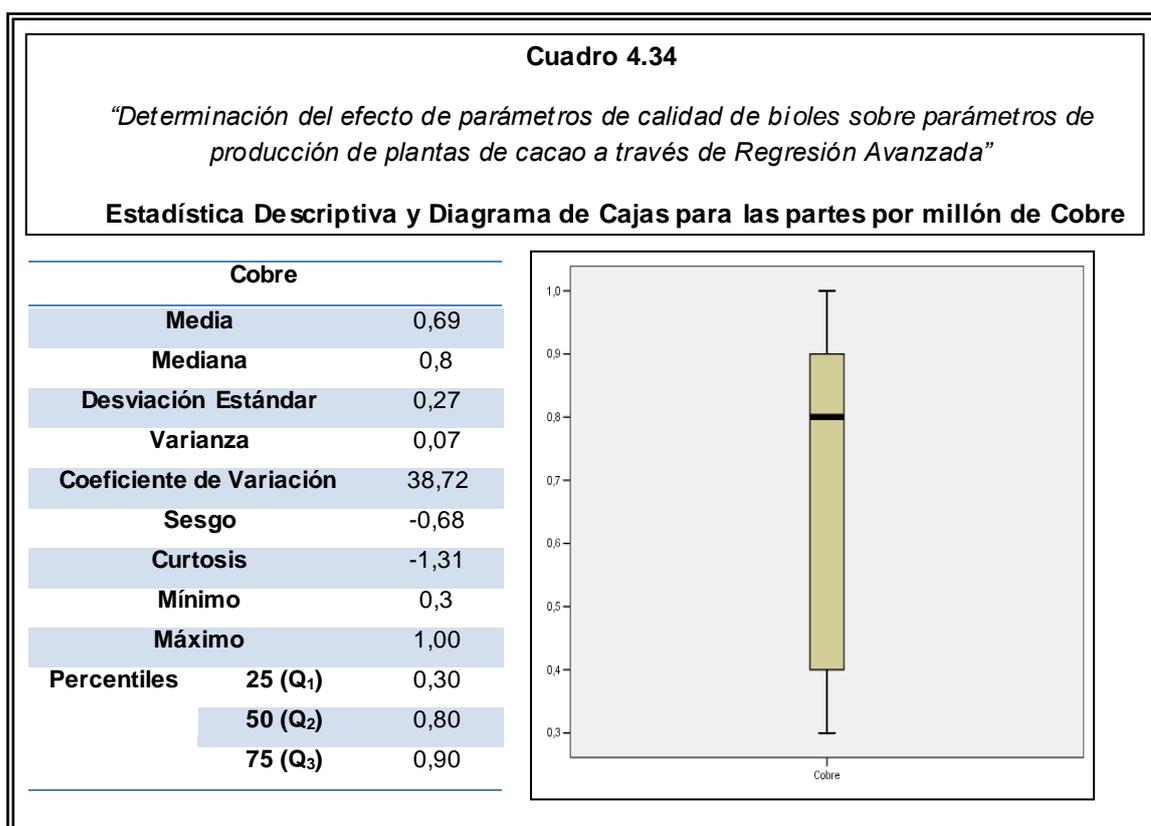
**Histograma de Frecuencias de la variable "Cobre"**



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.33** podemos apreciar la tabla de frecuencias de la variable Partes por millón de Cobre, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un contenido en partes por millón de Cobre entre 0.30 y 0.53 son 4 del total de haciendas muestreadas esto es un 33%; el segundo grupo corresponde a los bioles

que tienen un contenido de partes por millón de Cobre entre 0.53 y 0.77 son 1 del total de haciendas muestreadas esto es un 8%, y finalmente el tercer grupo corresponde a los bioles que tienen un contenido de partes por millón de Cobre entre 0.77 y 1 son 7 del total del haciendas muestreadas esto es un 58%.



**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.34** se observa la estadística descriptiva básica de los datos de la variable partes por millón de Cobre, se aprecia que el promedio del contenido en partes por millón de Cobre en los bioles es de 0.69, la variabilidad con respecto a los datos es de 0.07, el valor central (mediana) del contenido en partes por millón de Cobre en los bioles es de

0.8. El sesgo de la variable es de  $-0.68$  es decir que la concentración de los datos esta a la derecha del promedio ( $168.92$ ), los valores mínimo y máximo son  $0.3$  y  $1$  respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo  $0.30$ ,  $0.80$  y  $0.90$  respectivamente. Podemos indicar que el  $25\%$  de los bioles tienen un contenido en partes por millón de Cobre menor o igual a  $0.30$ , el  $50\%$  de los bioles de las haciendas muestreadas tienen un contenido en partes por millón de Cobre menor o igual a  $0.80$  y finalmente el  $75\%$  tienen un contenido en partes por millón de Cobre menor o igual a  $0.90$ .

**Variable: Partes por millón de Zinc (Zn)**

**Cuadro 4.35**

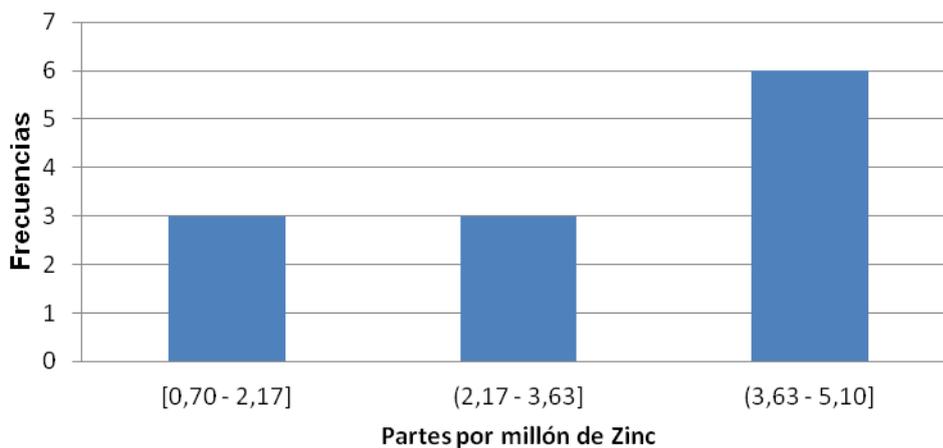
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias de las parte por millón de Zinc**

**Tabla: Frecuencias de la variable Zinc**

Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0,70 - 2,17]	3	0,25	0,25
(2,17 - 3,63]	3	0,25	0,50
(3,63 - 5,10]	6	0,50	1
<b>Total</b>	<b>12</b>	<b>1</b>	

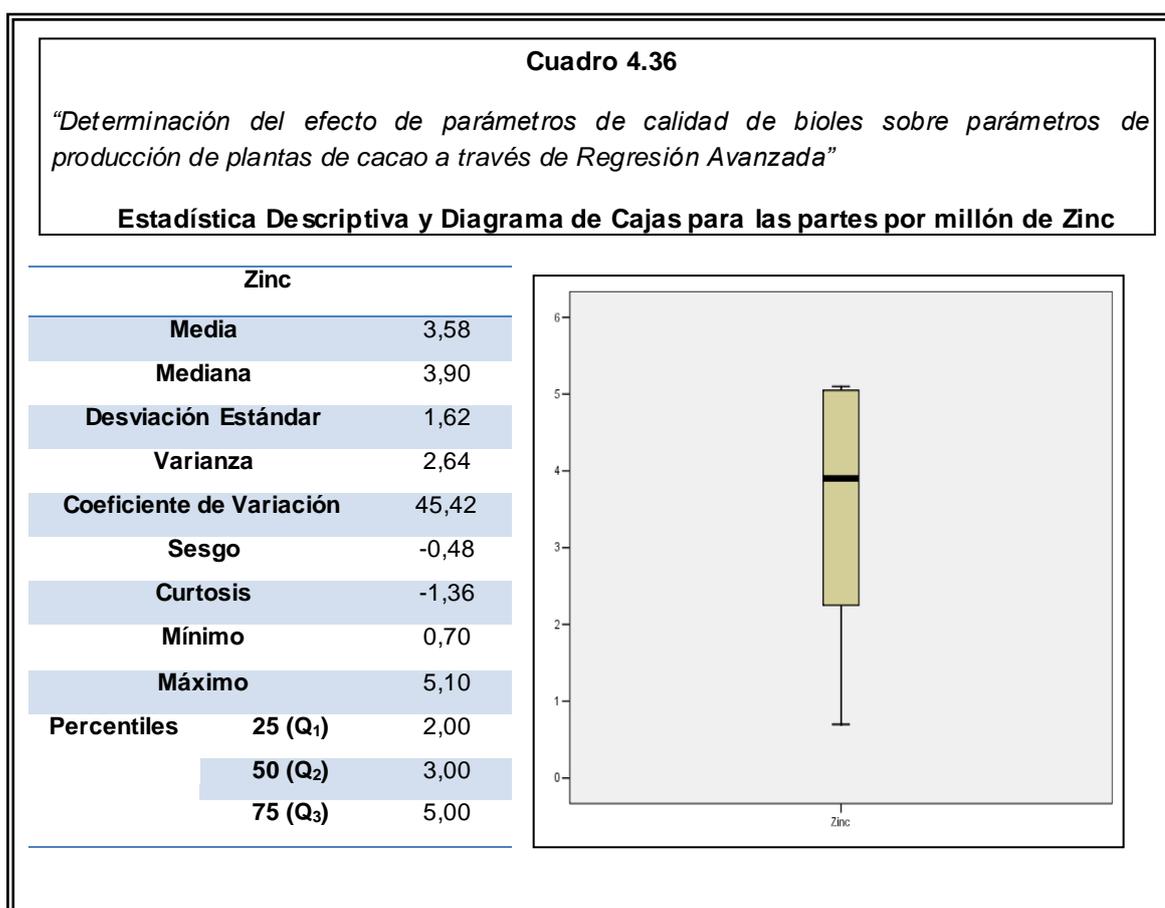
**Histograma de Frecuencias de la variable "Zinc"**



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.35** podemos apreciar la tabla de frecuencias de la variable Partes por millón de Zinc, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que tienen un contenido en partes por millón de Cobre entre 0.70 y 2.17 son 3 del total de haciendas muestreadas esto es un 25%; el segundo grupo corresponde a los bioles

que tienen un contenido de partes por millón de Zinc entre 2.17 y 3.63 son 3 del total de haciendas muestreadas esto es un 25%, y finalmente el tercer grupo corresponde a los bioles que tienen un contenido de partes por millón de Zinc entre 3.63 y 5.10 son 6 del total del haciendas muestreadas esto es un 50%.



**Elaborado por:** Karla M. Aguaguíña Méndez

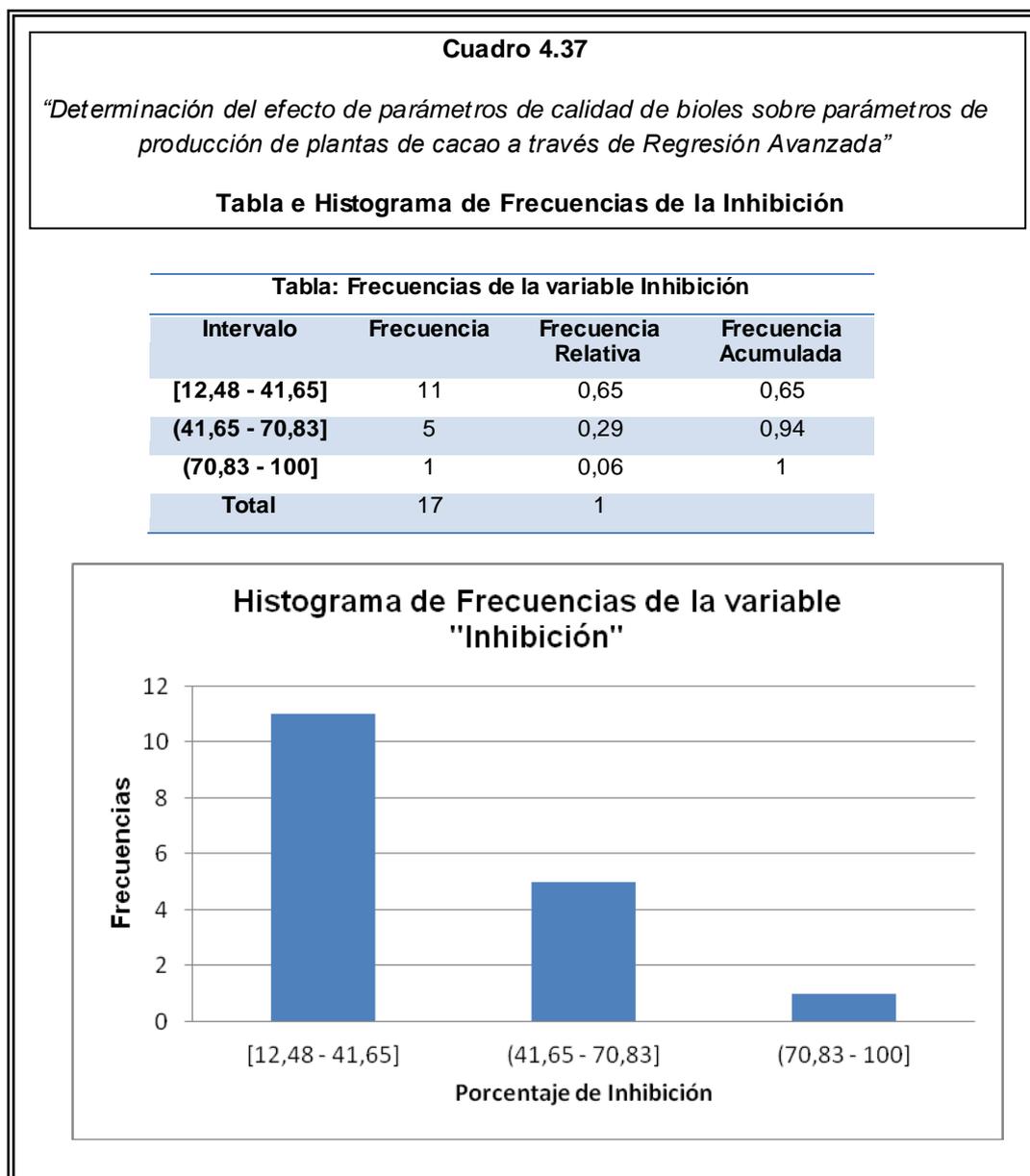
En el **Cuadro 4.36** se observa la estadística descriptiva básica de los datos de la variable partes por millón de Zinc, se aprecia que el promedio del contenido en partes por millón de Zinc en los bioles es de 3.58, la

variabilidad con respecto a los datos es de 2.64, el valor central (mediana) del contenido en partes por millón de Zinc en los bioles es de 3.90. El sesgo de la variable es de -0.48 es decir que la concentración de los datos esta a la derecha del promedio (3.58); los valores mínimo y máximo son 0.70 y 5.10 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 2, 3 y 5 respectivamente. Podemos indicar que el 25% de los bioles tienen un contenido en partes por millón de Zinc menor o igual a 2, el 50% de los bioles de las haciendas muestreadas tienen un contenido en partes por millón de Zinc menor o igual a 3 y finalmente el 75% tienen un contenido en partes por millón de Zinc menor o igual a 5.

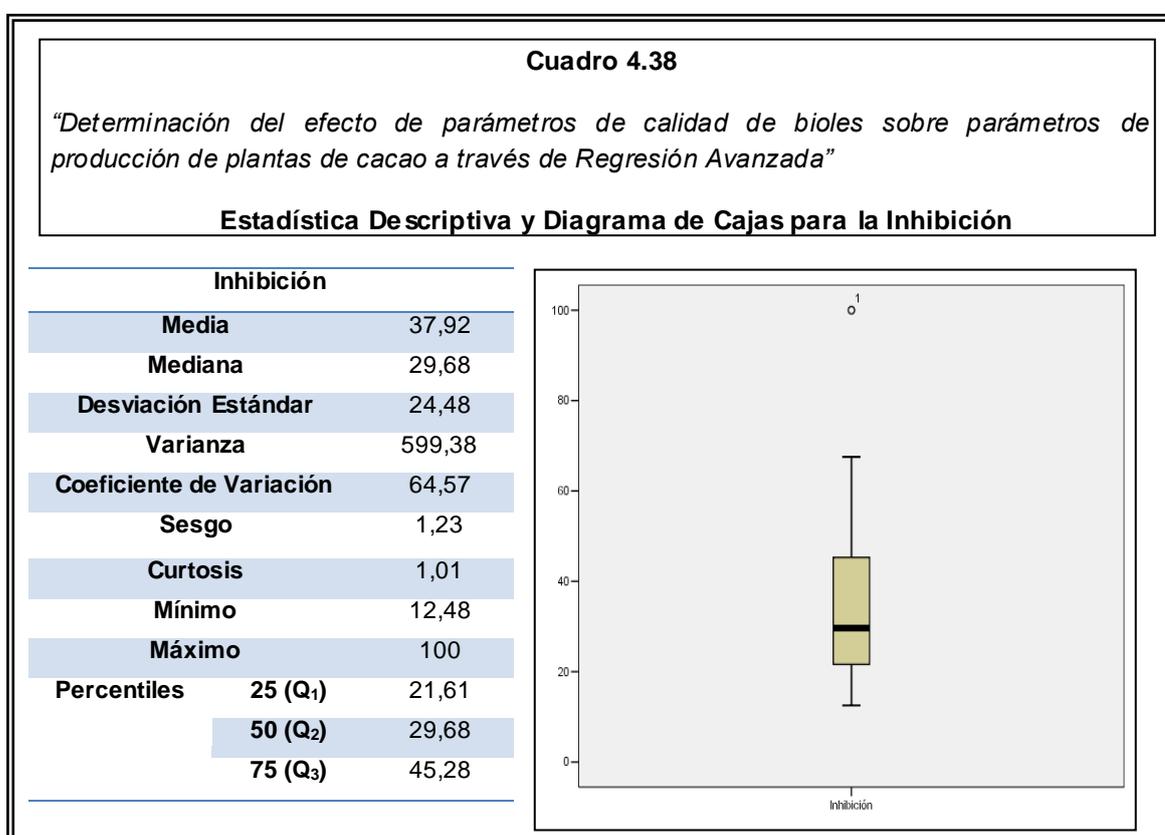
#### 4.1.4 Estadística Descriptiva para las Variables Dependientes

*Variable dependiente: Inhibición*



En el **Cuadro 4.37** podemos apreciar la tabla de frecuencias de la variable Inhibición, se analizan tres marcas de clases o intervalos siendo el primer

grupo los bioles los que tiene un porcentaje de Inhibición entre 12.48 y 41.65 son 11 del total de haciendas muestreadas esto es un 65%; el segundo grupo corresponde a los bioles que tienen un porcentaje de Inhibición entre 41.65 y 70.83 son 5 del total de haciendas muestreadas esto es un 29%, y finalmente el tercer grupo corresponde a los bioles que tienen un porcentaje de Inhibición entre 70.83 y 100 son 1 del total del haciendas muestreadas esto es un 5%.



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.38** se observa la estadística descriptiva básica de los datos de la variable Inhibición, se aprecia que el promedio del porcentaje de Inhibición de los bioles es de 37.92, la variabilidad con respecto a los

datos es de 599.38, el valor central (mediana) del porcentaje de inhibición en los bioles es de 29.68. El sesgo de la variable es de 1.23 es decir que la concentración de los datos esta a la izquierda del promedio (37.92); los valores mínimo y máximo son 12.48 y 100 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 21.61, 29.68 y 45.28 respectivamente. Podemos indicar que el 25% de los bioles tienen un porcentaje de inhibición menor o igual a 21.61, el 50% de los bioles de las haciendas muestreadas tienen un porcentaje de inhibición menor o igual a 29.68 y finalmente el 75% tienen un porcentaje de inhibición menor o igual a 45.28.

Se puede observar la presencia de un “dato aberrante”, este pertenece a una de las haciendas de la provincia de Esmeraldas podemos indicar que aquí el efecto del biol ha sido bueno puesto que alcanza el 100% de inhibición a enfermedades.

**Variable dependiente: Promedio del largo de la hoja**

**Cuadro 4.39**

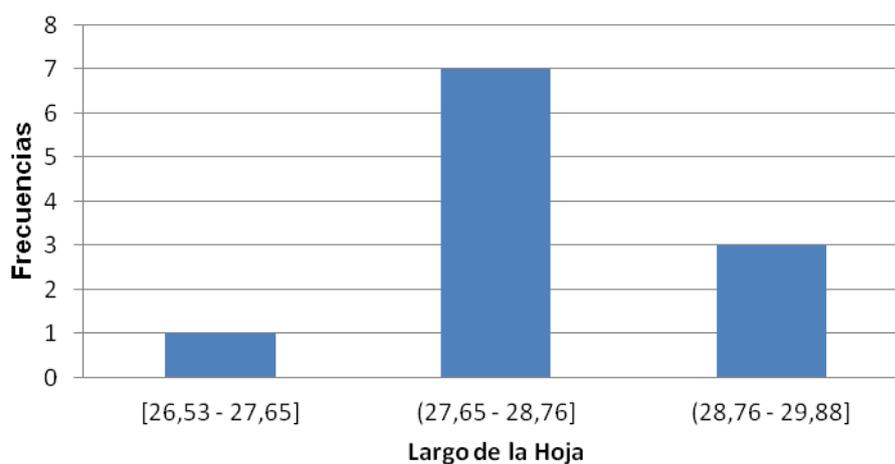
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias del Promedio del largo de la hoja**

**Tabla: Frecuencias de la variable Promedio del largo de la hoja**

Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[26,53 - 27,65]	1	0,09	0,09
(27,65 - 28,76]	7	0,64	0,73
(28,76 - 29,88]	3	0,27	1
<b>Total</b>	<b>11</b>	<b>1</b>	

**Histograma de Frecuencias de la variable "Promedio del largo de la hoja"**



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.39** podemos apreciar la tabla de frecuencias de la variable Promedio del largo de la hoja, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que produjeron que las plantas

de cacao alcancen un promedio de largo de hoja entre 26.53 y 27.65 son 1 del total de haciendas muestreadas esto es un 9%; el segundo grupo corresponde a los bioles que produjeron que las plantas de cacao alcancen un promedio de largo de la hoja entre 27.65 y 28.76 son 7 del total de haciendas muestreadas esto es un 64%, y finalmente el tercer grupo corresponde a los bioles que produjeron que las plantas de cacao alcancen un promedio de largo de la hoja entre 28.76 y 29.88 son 3 del total del haciendas muestreadas esto es un 27%.

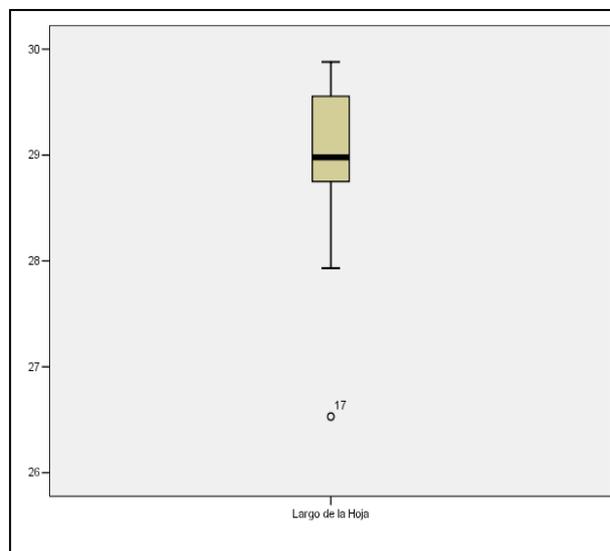
**Cuadro 4.40**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Estadística Descriptiva y Diagrama de Cajas para el Promedio del largo de la hoja**

**Promedio del largo de la hoja**

<b>Media</b>	28,89
<b>Mediana</b>	28,98
<b>Desviación Estándar</b>	0,97
<b>Varianza</b>	0,95
<b>Coefficiente de Variación</b>	3,37
<b>Sesgo</b>	-1,52
<b>Curtosis</b>	2,79
<b>Mínimo</b>	26,53
<b>Máximo</b>	29,88
<b>Percentiles</b>	
<b>25 (Q<sub>1</sub>)</b>	28,56
<b>50 (Q<sub>2</sub>)</b>	28,98
<b>75 (Q<sub>3</sub>)</b>	29,63



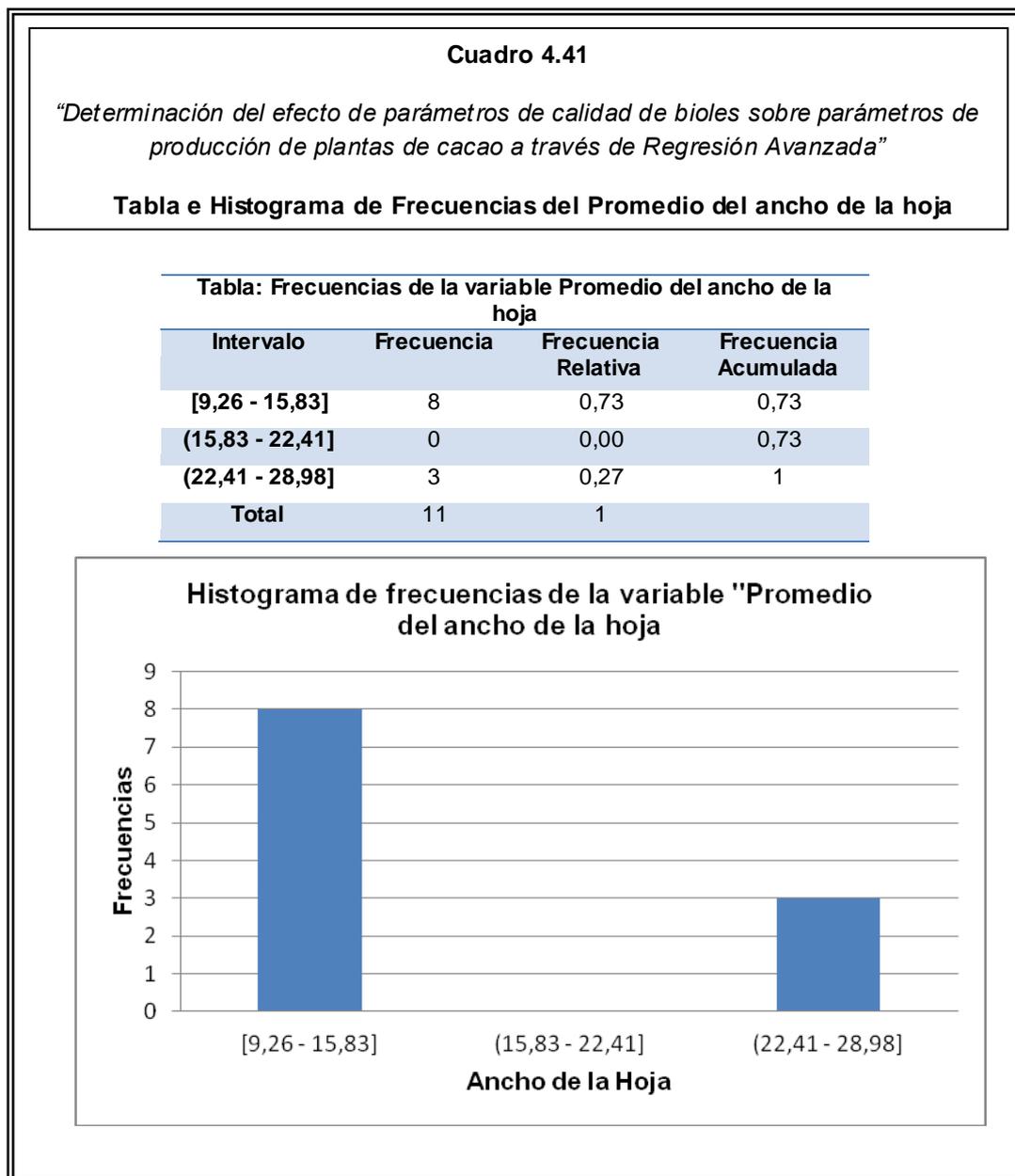
**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.40** se observa la estadística descriptiva básica de los datos de la variable Promedio del largo de la hoja, se aprecia que el promedio del largo de las hojas que alcanza la planta de cacao es de 28.89, la variabilidad con respecto a los datos es de 0.95, el valor central (mediana) del promedio del largo de las hojas es de 28.98. El sesgo de la variable es de -1.52 es decir que la concentración de los datos esta a la derecha del promedio (28.89); los valores mínimo y máximo son 26.53 y 29.88 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 28.56, 28.68 y 29.63 respectivamente. Podemos indicar que el 25% de los bioles hacen que el promedio del largo de la hoja en la planta sea menor o igual a 28.56, el 50% de los bioles hacen que el promedio del largo de la hoja en la planta sea menor o igual a 29.68 y finalmente el 75% tienen un promedio del largo de la hoja en la planta menor o igual a 29.63.

Se puede observar un “dato aberrante”, este valor corresponde a la hacienda ubicada en la provincia de Los Ríos que alcanzó un promedio menor a 28.89 cm en cuanto al largo de la hoja.

**Variable dependiente: Promedio del ancho de la hoja**



En el **Cuadro 4.41** se podemos apreciar la tabla de frecuencias de la variable Promedio del ancho de la hoja, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que produjeron que las

plantas de cacao alcancen un promedio de ancho de hoja entre 9.26 y 15.83 son 8 del total de haciendas muestreadas esto es un 73%; el segundo grupo corresponde a los bioles que produjeron que las plantas de cacao alcancen un promedio de ancho de la hoja entre 15.83 y 22.41 son 0 del total de haciendas muestreadas esto es un 0%, y finalmente el tercer grupo corresponde a los bioles que produjeron que las plantas de cacao alcancen un promedio de ancho de la hoja entre 22.41 y 28.98 son 3 del total del haciendas muestreadas esto es un 27%.

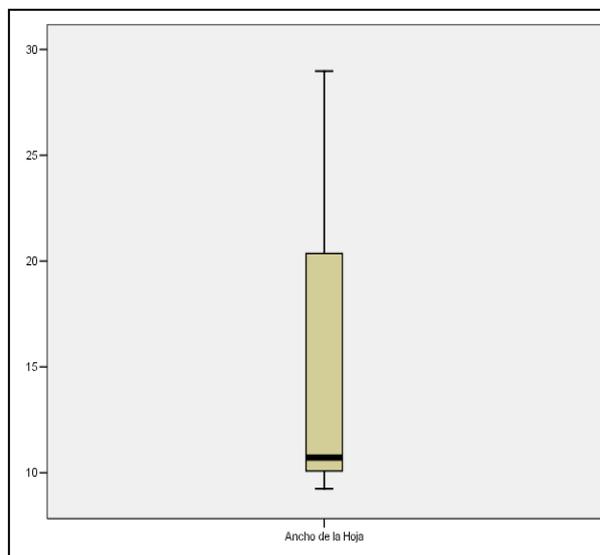
**Cuadro 4.42**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Estadística Descriptiva y Diagrama de Cajas para el Promedio del ancho de la hoja**

**Promedio del ancho de la hoja**

<b>Media</b>	15,45
<b>Mediana</b>	10,72
<b>Desviación Estándar</b>	75,98
<b>Varianza</b>	2,63
<b>Coefficiente de Variación</b>	56,42
<b>Sesgo</b>	1,17
<b>Curtosis</b>	-0,78
<b>Mínimo</b>	9,26
<b>Máximo</b>	28,98
<b>Percentiles</b>	
<b>25 (Q<sub>1</sub>)</b>	9,74
<b>50 (Q<sub>2</sub>)</b>	10,72
<b>75 (Q<sub>3</sub>)</b>	28,98

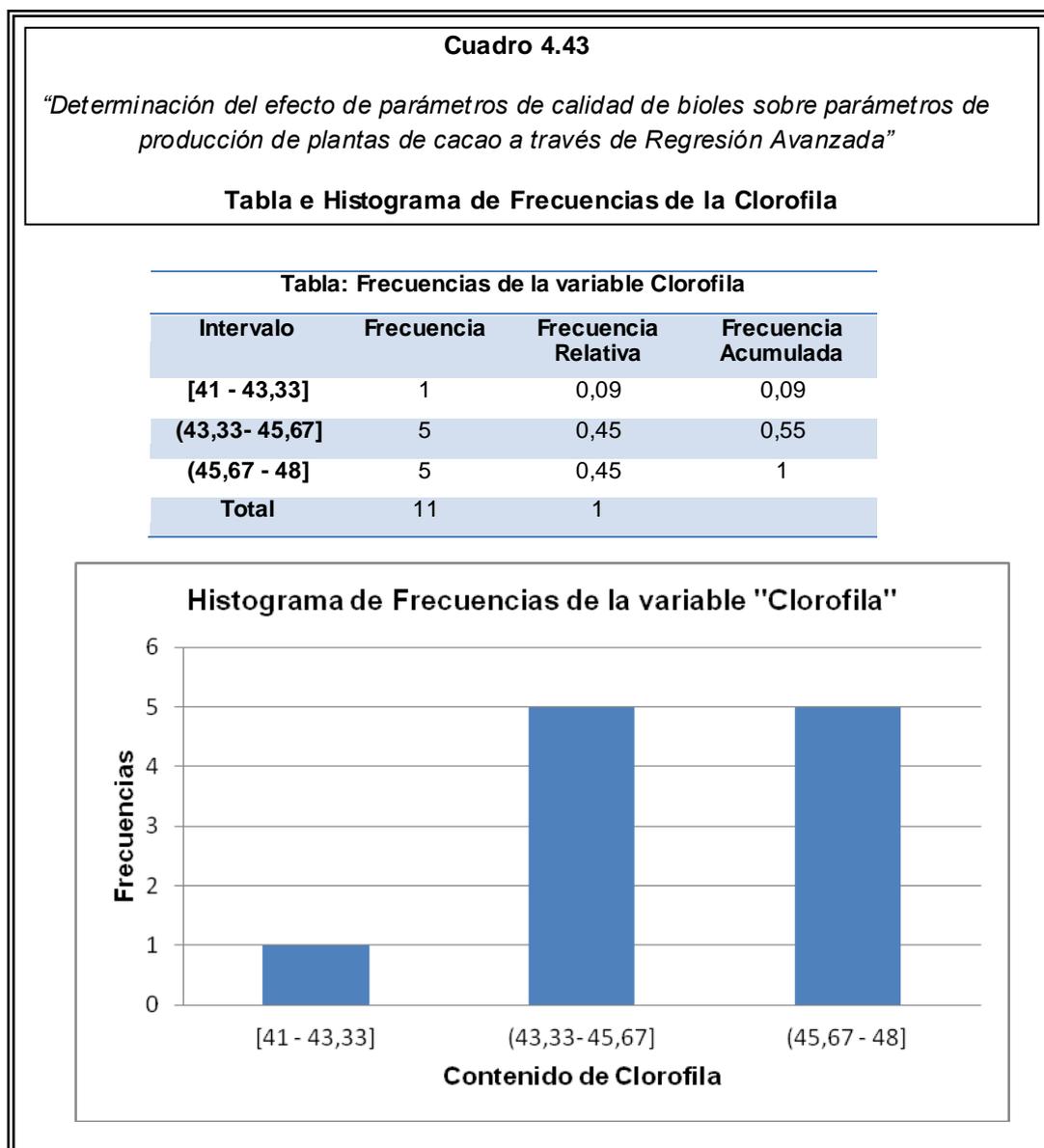


**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.42** se observa la estadística descriptiva básica de los datos de la variable promedio del ancho de hoja, se aprecia que el promedio del ancho de las hojas en las plantas de cacao es de 15.45, la variabilidad con respecto a los datos es de 2.63, el valor central (mediana) del promedio del largo de las hojas es de 10.72. El sesgo de la variable es de 1.17 es decir que la concentración de los datos esta a la izquierda del promedio (15.45); los valores mínimo y máximo son 9.26 y 28.98 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 9.74, 10.72 y 28.98 respectivamente. Podemos indicar que el 25% de los bioles hacen que el promedio del ancho de la hoja en las plantas de cacao sea menor o igual a 9.74, el 50% de los bioles hacen que el promedio del ancho de la hoja en la planta de cacao sea menor o igual a 10.72 y finalmente el 75% tienen un promedio del ancho de la hoja en la planta de cacao sea menor o igual a 28.98.

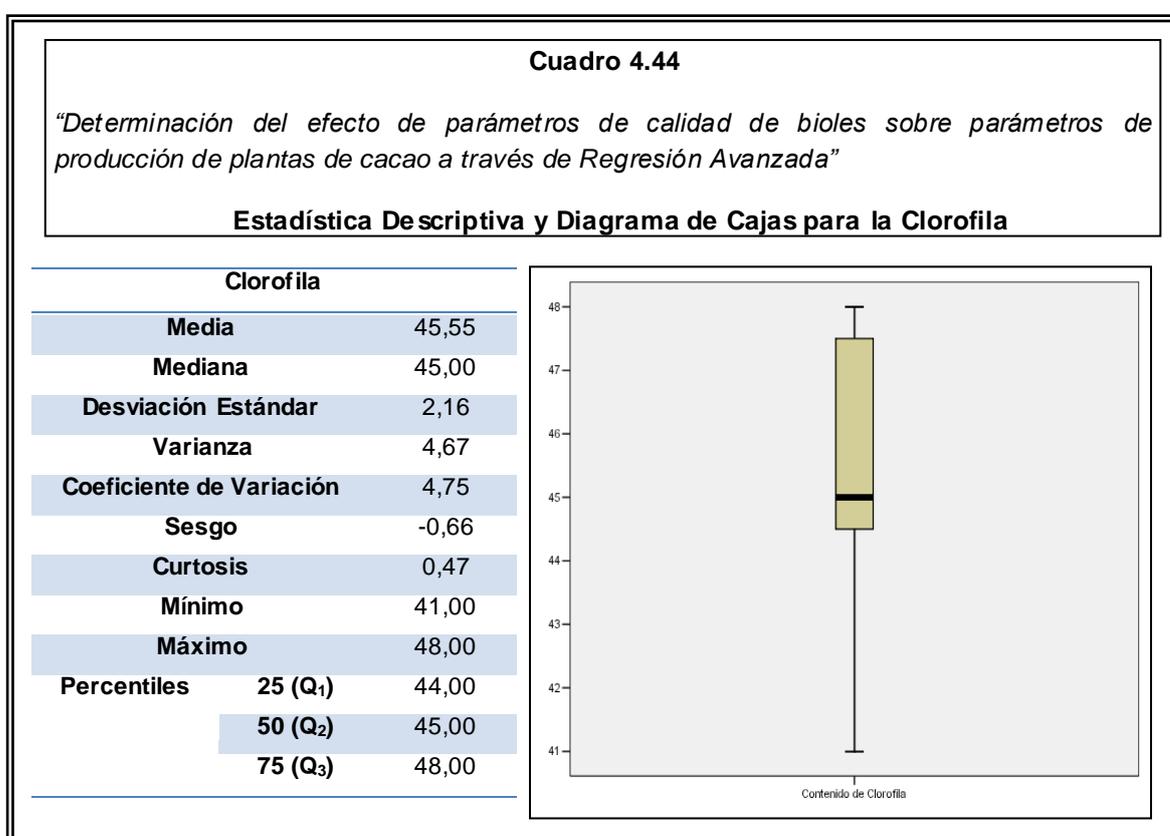
**Variable dependiente: Clorofila**



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.43** podemos apreciar la tabla de frecuencias de la variable Clorofila, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles los que producen un contenido de clorofila entre 41 y

43.33 son 1 del total de haciendas muestreadas esto es un 9%; el segundo grupo corresponde a los bioles que producen un contenido de clorofila entre 43.33 y 45.67 son 5 del total de haciendas muestreadas esto es un 45%, y finalmente el tercer grupo corresponde a los bioles que producen un contenido de clorofila entre 45.67 y 48 son 5 del total del haciendas muestreadas esto es un 45%.



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.44** se observa la estadística descriptiva básica de los datos de la variable Clorofila, se aprecia que el promedio del contenido de clorofila es de 45.55, la variabilidad con respecto a los datos es de 4.67, el valor central (mediana) del contenido de clorofila que se produce en las

hojas de las plantas de cacao es de 45. El sesgo de la variable es de -0,66 es decir que la concentración de los datos esta a la derecha del promedio (45.55); los valores mínimo y máximo son 41 y 48 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 44, 45 y 48 respectivamente. Podemos indicar que el 25% de los bioles producen en las plantas de cacao un contenido de clorofila menor o igual a 44, el 50% de los bioles de las haciendas muestreadas producen en las plantas de cacao un contenido de clorofila menor o igual a 45 y finalmente el 75% de las plantas de cacao tienen un contenido de clorofila menor o igual a 48.

**Variable dependiente: Frutos de 4 meses**

**Cuadro 4.45**

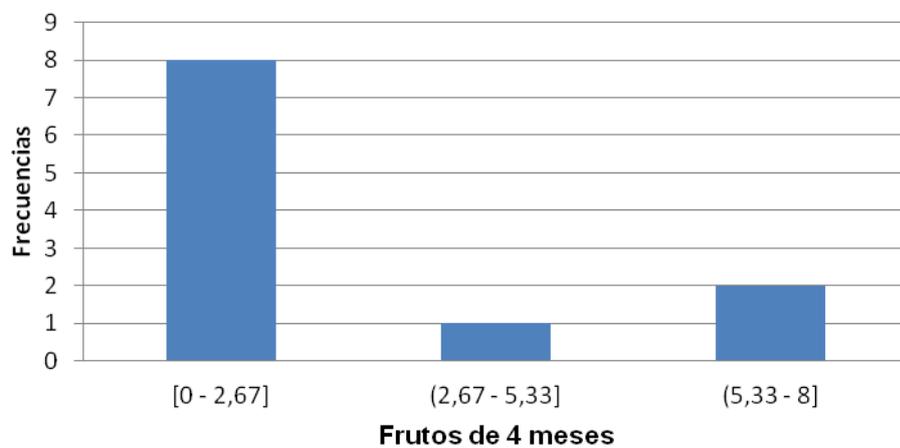
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias del Número de Frutos de 4 meses**

**Tabla: Frecuencias de la variable Frutos de 4 meses**

Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0 - 2,67]	8	0,73	0,73
(2,67 - 5,33]	1	0,09	0,82
(5,33 - 8]	2	0,18	1
<b>Total</b>	<b>11</b>	<b>1</b>	

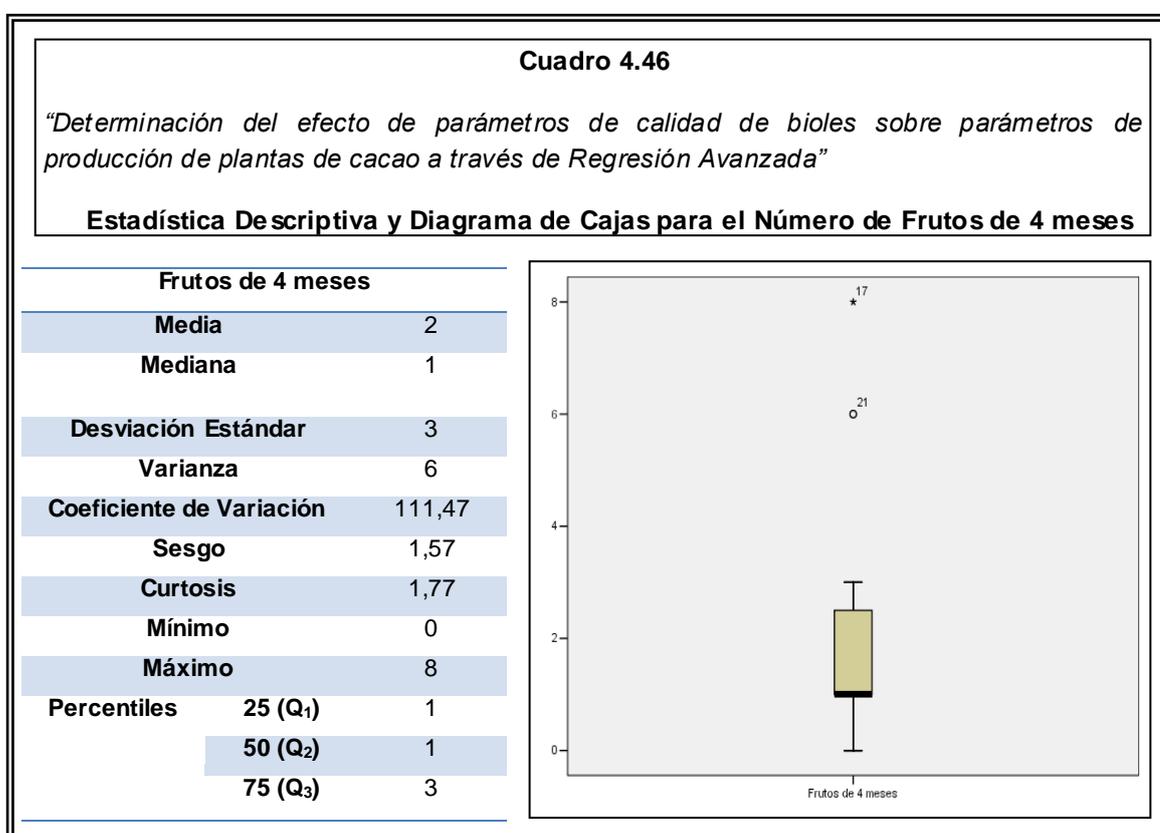
**Histograma de frecuencias de la variable "Frutos de 4 meses"**



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.35** podemos apreciar la tabla de frecuencias de la variable Número de frutos de 4 meses, se analizan tres marcas de clases o

intervalos siendo el primer grupo los bioles que producen un número de frutos de 4 meses entre 0 y 2.67 son 8 del total de haciendas muestreadas esto es un 73%; el segundo grupo corresponde a los bioles que producen un número de frutos de 4 meses entre 2.67 y 5.33 son 1 del total de haciendas muestreadas esto es un 9%, y finalmente el tercer grupo corresponde a los bioles que producen un número de frutos de 4 meses entre 5.33 y 8 son 2 del total del haciendas muestreadas esto es un 18%.



**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.46** se observa la estadística descriptiva básica de los datos, se aprecia que en promedio los bioles producen un número de

frutos de 4 meses es de 3, la variabilidad con respecto a los datos es de 6, el valor central (mediana) del número de frutos de 4 meses que producen los bioles es de 1. El sesgo de la variable es 1.57 es decir que la concentración de los datos está a la izquierda del promedio (3); los valores mínimo y máximo son 0 y 8 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 1, 1 y 8 respectivamente. Podemos indicar que el 50% de los bioles producen un número de frutos de 4 meses menor o igual a 1 y finalmente el 75% producen un número de frutos de 4 meses menor o igual a 8.

En el diagrama de cajas se observan la presencia de "datos aberrantes", estos corresponden a las haciendas asentadas en las provincias de los Ríos y Manabí que tuvieron una producción de frutos de 4 meses superior al promedio estimado.

**Variable dependiente: Frutos secos**

**Cuadro 4.47**

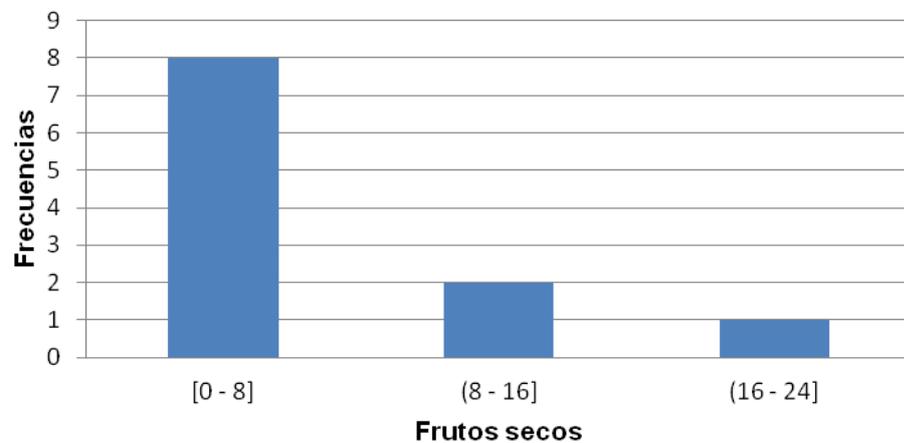
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla e Histograma de Frecuencias del Número de Frutos secos**

**Tabla: Frecuencias de la variable Frutos secos**

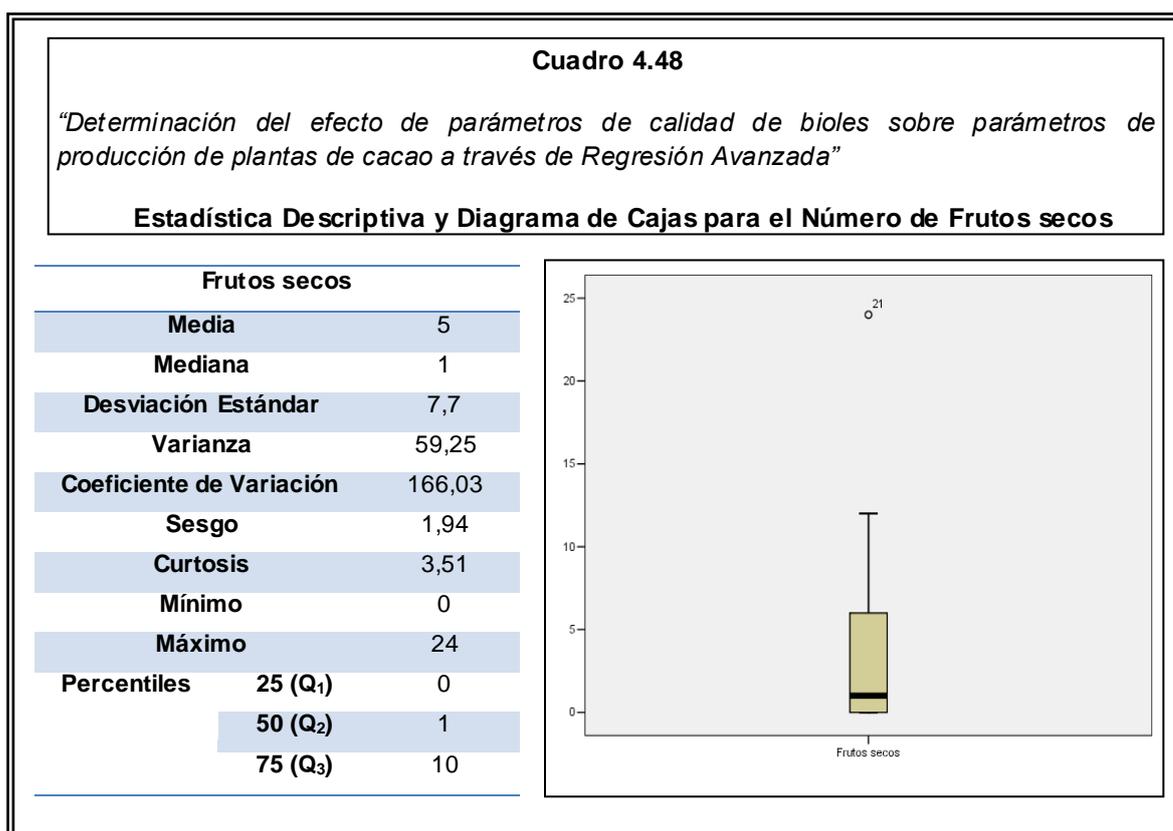
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada
[0 - 8]	8	0,73	0,73
(8 - 16]	2	0,18	0,91
(16 - 24]	1	0,09	1
<b>Total</b>	<b>11</b>	<b>1</b>	

**Histograma de Frecuencias de la variable "Frutos secos"**



**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 4.47** podemos apreciar la tabla de frecuencias de la variable Número de frutos de 4 meses, se analizan tres marcas de clases o intervalos siendo el primer grupo los bioles que producen un número de frutos secos entre 0 y 8 son 8 del total de haciendas muestreadas esto es un 73%; el segundo grupo corresponde a los bioles que producen un número de frutos secos entre 8 y 16 son 2 del total de haciendas muestreadas esto es un 18%, y finalmente el tercer grupo corresponde a los bioles que producen un número de frutos secos entre 16 y 24 son 1 del total del haciendas muestreadas esto es un 9%.



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4.48** se observa la estadística descriptiva básica de los datos de la variable Frutos secos, se aprecia que el promedio del número de frutos secos es de 5, la variabilidad con respecto a los datos es de 59.25, el valor central (mediana) del número de frutos secos que producen los bioles es de 1. El sesgo de la variable es 1.94 es decir que la concentración de los datos esta a la izquierda del promedio (5); los valores mínimo y máximo son 0 y 24 respectivamente.

En el Diagrama de Cajas podemos apreciar la distribución de los datos, donde se localizan el primer, segundo y tercer cuartil siendo 0, 1 y 10 respectivamente. Podemos indicar que el 25% de los bioles producen en las plantas de cacao un número de frutos secos menor o igual a 0, el 50% de los bioles de las haciendas muestreadas producen en las plantas de cacao un número de frutos secos menor o igual a 10 y finalmente el 75% de los bioles producen en las plantas de cacao un número de frutos secos menor o igual a 10.

## 4.2 ANÁLISIS MULTIVARIADO

El uso de la estadística en la actualidad no sólo se queda en el análisis descriptivo que se realiza a los datos para conocer como es la dispersión de estos o conocer su media sino que se puede hacer uso de las diferentes técnicas multivariadas como son las componentes Principales entre otras.

Para esta investigación se desarrolla el análisis de componentes principales que es una técnica estadística multivariada que es usada para reducir la dimensionalidad de los datos a través de nuevas variables que contenga al menos el 75% de la variabilidad de los datos.

Se realizó el análisis multivariado usando el software estadístico INFOSTAT y SPSS V.13.

### 4.2.1 ANÁLISIS DE COMPONENTES PRINCIPALES

<b>Cuadro 4.49</b>			
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>			
<b>Total de proporción acumulada por el Análisis de Componentes Principales</b>			
<b>Component</b>	<b>Initial Eigenvalues</b>		
	Total	% of Variance	Cumulative %
<b>1</b>	8,2	34,18	34,18
<b>2</b>	5,64	23,5	57,68
<b>3</b>	4,13	17,23	74,91

**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 4. 49** se muestra la tabla de los valores propios y la proporción acumulada de la variación total explicada por cada componente.

Se observa que el primer componente resume el 34.18% de la variabilidad total, el segundo resume el 23.50%, el tercer resume el 17.23% es decir que la combinación lineal de las variables originales con los tres primeros componentes sintetizan el 75% de la variación total del conjunto de datos.

**Cuadro 4.50**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Matriz de Componentes Principales**

	Component		
	1	2	3
PF1	-0,66	-0,06	0,56
PF2	-0,06	-0,02	0,66
PF3	0,87	0,3	0,2
PF4	0,80	0,29	0,08
PF5	0,50	-0,37	0,55
PF6	0,89	0,25	0,19
MB1	-0,66	-0,35	-0,34
MB2	-0,41	0,50	-0,58
MB3	0,01	0,51	0,15
MM1	-0,19	0,90	0,35
MM2	0,92	-0,16	-0,14
MM3	0,2	-0,66	0,64
MM4	-0,39	0,70	-0,02
MM5	0,71	-0,37	-0,14
MM6	0,57	-0,25	0,64
MM7	0,74	0,39	-0,24
MM8	0,11	-0,98	-0,03
MM9	0,61	-0,15	-0,75
P1	-0,13	0,27	0,37
P2	0,06	0,79	0,4
P3	0,84	-0,46	-0,03
P4	0,27	0,36	0,67
P5	-0,69	-0,55	0,29
P6	-0,79	-0,45	0,3

**Elaborado por:** Karla M. Aguaguiña Méndez

En el **Cuadro 4.50** se observa la matriz de Componentes Principales, en la cual hemos obtenido tres nuevas variables que se detallan a continuación:

- $C_1 = 0.87 PF3 + 0.80 PF4 + 0.89 PF6 + 0.92 MM2 + 0.84 P3$

- $C_2 = 0.90 \text{ MM1} - 0.98 \text{ MM8}$
- $C_3 = - 0.75 \text{ MM9} + 0.67 \text{ P4}$

Podemos notar que las variables que aportan de forma significativa en el Componente Uno ( $C_1$ ) son los solutos totales (PF3) con un 87%, el total de materia orgánica (PF6) con un 89%; además que las partes por millón de Fósforo (MM2) tiene un aporte de 92%, se aprecia que el promedio del ancho de la hoja (P3) esta explicado en un 84%.

Mientras que en el Componente Dos ( $C_2$ ) las variables que tienen un aporte significativo son el porcentaje de Nitrógeno (MM1) tiene un aporte de 90% y las partes por millón de Cobre (MM8) tiene un aporte de 98%.

El Componente Tres ( $C_3$ ) nos muestra que las variables que aportan significativamente son las partes por millón de Zinc (MM9) tiene un aporte de 75% y se aprecia que el contenido de Clorofila (P4) esta explicado en un 67%.

Debido a que las dos primeras componentes principales no acumulan el 60% de información la interpretación de los componentes principales no se ajustan a la realidad.

## CAPÍTULO CINCO

### INTRODUCCIÓN

En este capítulo se aplican modelos de regresión lineal, donde existe una variable a la cual llamaremos dependiente (Y) la misma que será explicada de forma lineal por una o más variables independientes o explicativas ( $x_1, x_2, x_3, \dots, x_n$ ), el número de variables explicativas dependerá del peso que tenga cada una dentro del modelo que se esté generando, y para determinar aquello se usarán varias pruebas estadísticas, previo al cumplimiento de supuestos necesarios para el uso de los modelos de regresión, teniendo que cumplir los principios teóricos de linealidad en los coeficientes, normalidad y homocedasticidad en los residuos, definiéndose como  $\varepsilon \sim N(0, \sigma^2)$ , esto es que los errores se aproximen a una distribución normal con media cero y varianza constante, además de la independencia de los mismos.

Para determinar en los modelos las variables que serán tomados en cuenta primero se define aquellas variables con las que (Y) mantiene dependencia y esto se lo realiza mediante la prueba de Pearson, cuya finalidad es determinar la relación lineal que existe entre dos variables, teniendo seis variables dependientes *Inhibición (P1), largo de la hoja (P2), ancho de la hoja (P3), clorofila (P4), número de frutos de 4 meses (P5), número de frutos*

secos (*P6*) y estas con sus respectivas variables de explicación *pH* (*PF1*), *conductividad eléctrica* (*PF2*), *solutos totales* (*PF3*), *grados brix* (*PF4*), *densidad* (*PF5*), *materia orgánica* (*PF6*), *aerobios* (*MB1*), *anaerobios* (*MB2*), *hongos y levaduras* (*MB3*), *porcentaje de nitrógeno* (*MM1*), *ppm de fósforo* (*MM2*), *porcentaje de potasio* (*MM3*), *porcentaje de calcio* (*MM4*), *porcentaje de magnesio* (*MM5*), *ppm de manganeso* (*MM6*), *ppm de hierro* (*MM7*), *ppm de cobre* (*MM8*), *ppm de zinc* (*MM9*), descartan aquellas variables que entre sí están de alguna manera correlacionadas, esto con la finalidad de evitar la multicolinealidad en el modelo.

Para determinar si la distribución que siguen los residuos es normal, se usa la prueba de Kolmogorov – Smirnov (que indica a qué distribución se ajustan los datos). Además, para comprobar otro supuesto del modelo de regresión que es el de la homocedasticidad<sup>1</sup> se grafican los valores residuales versus los valores predichos para observar si distribución de los residuos son de varianza homogénea.

También usamos la prueba de Durbin-Watson para determinar si existe la correlación entre variables, si el estadístico D está entre 1.5 y 2.5 podemos corroborar que existe independencia entre las variables independientes.

---

<sup>1</sup> **Homocedasticidad:** Es cuando la varianza de los errores es constante.

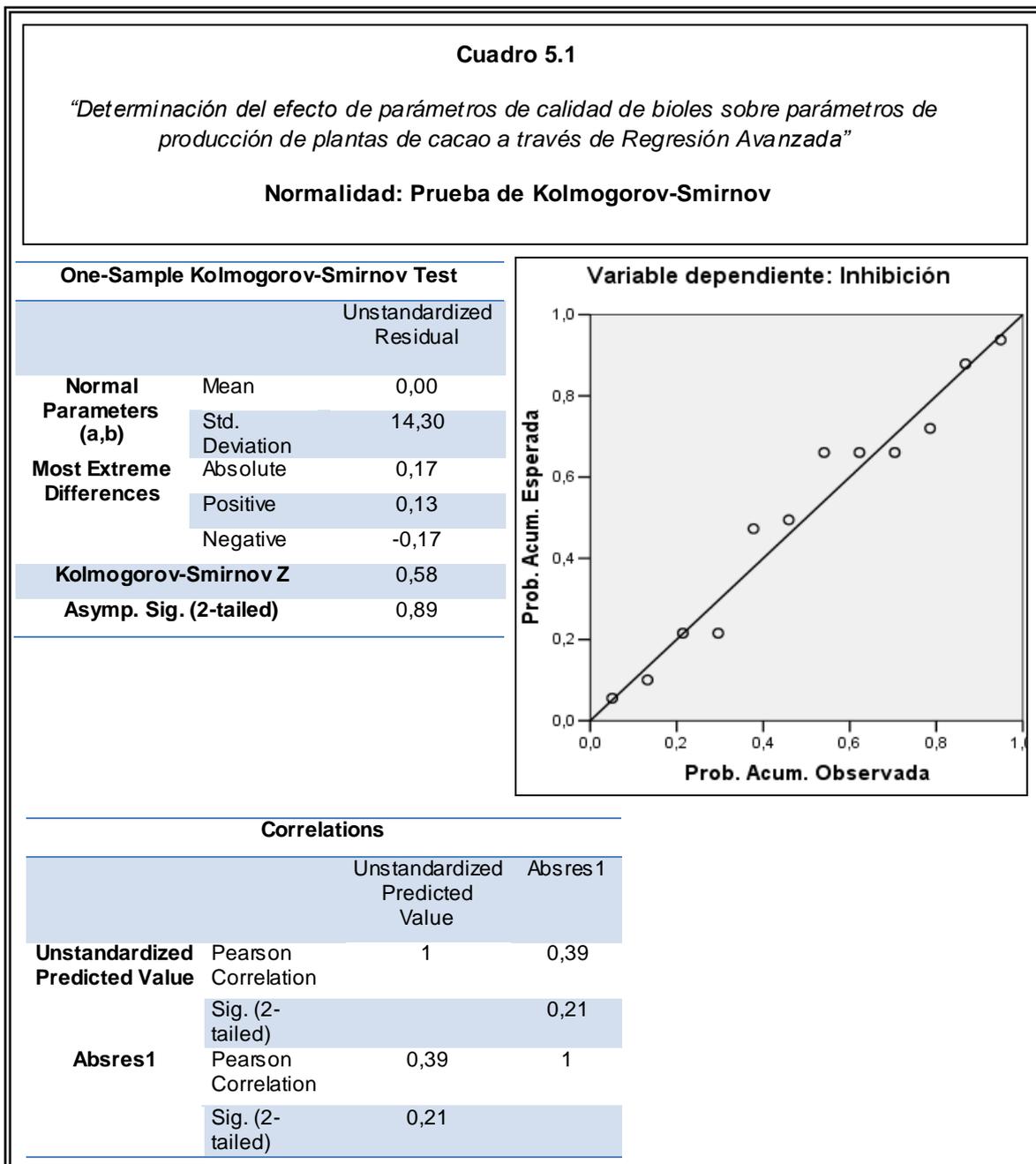
**Heterocedasticidad:** Es cuando la varianza de los errores varía.

En el **Anexo 1** se muestran los resultados de la prueba de Pearson, la misma que permite determinar el grado de dependencia lineal y su respectivo coeficiente de correlación, en este caso entre cada una de las variables dependientes con las variables independientes consideradas en este estudio. Se resaltan aquellas correlaciones cuyo “valor p” es inferior a 0.10 siendo rechazada la hipótesis nula que indica que el coeficiente de correlación es igual a cero (rechazo  $H_0: \rho=0$ ), por lo que se tiene una dependencia entre variables, siendo esta la base para empezar a construir el modelo de regresión para cada una de las variables dependientes.

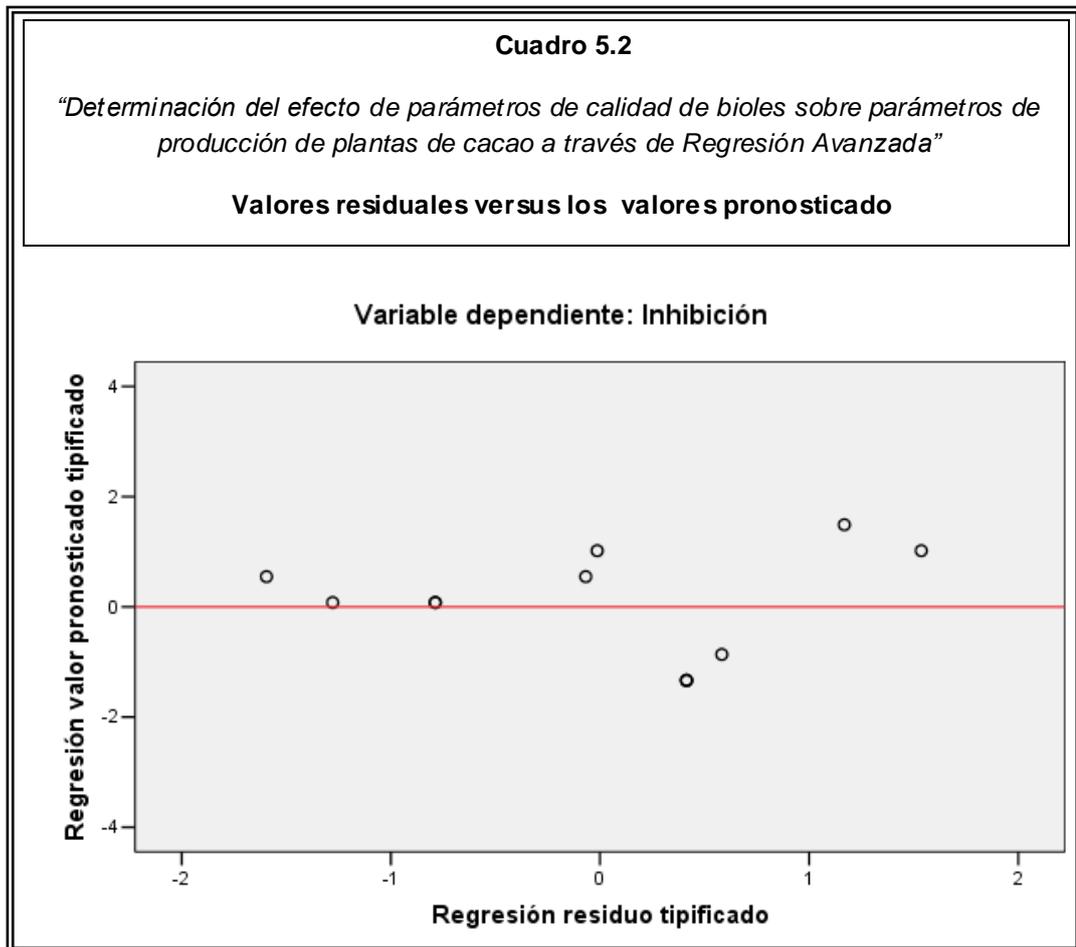
Con la finalidad de evitar la multicolinealidad se realiza la misma prueba pero únicamente entre variables independientes, como se muestra en el **Anexo 2**, incluyendo así aquellas variables que no tengan algún tipo de correlación, lo que implica que no se rechace la hipótesis nula, siendo esto que el “valor p” sea mayor o igual a 0.10, lo cual indica que el coeficiente de correlación es igual a cero, por lo que resultan ser variables independientes entre sí; si se da el caso contrario en que la hipótesis nula sea rechazada se excluye a dicha variable del modelo, debido a que no tendrá aporte alguno; sin embargo, podría ser considerada en otro modelo.

## 5.1 MODELO DE REGRESIÓN LINEAL SIMPLE PARA LA VARIABLE DEPENDIENTE P1: INHIBICIÓN

$$P1 = - 71.41 + 551.76 \text{ MM4}$$



Elaborado por: Karla M. Aguaguña Méndez



**Elaborado por:** Karla M. Aguaguíña Méndez

En este caso se tiene un modelo de regresión lineal simple puesto que existe una sola variable de explicación Porcentaje de calcio (MM4) y otra variable dependiente Inhibición (P1). En el **Cuadro 5.1** se verifican los resultados de las hipótesis de las pruebas de Kolmogorov y Smirnov con un valor p igual a 0.89 podemos decir que los errores siguen una distribución normal con media igual a cero y varianza constante; también se verifica la hipótesis de independencia de errores usando la correlación de Pearson que con un valor p de 0.21 por lo cual no se rechaza esta

hipótesis. En el **Cuadro 5.2** se observa el gráfico de los valores residuales versus los valores pronosticado, se puede observar que la nube de puntos sigue una forma lineal lo cual es un indicio de la homocedasticidad. El estadístico Durbin – Watson (DW=2.21) indica que no existe correlación, que es lo deseable en regresión

<b>Cuadro 5.3</b>					
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>					
<b>Tabla ANOVA del modelo seleccionado</b>					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0,63	0,40	0,34	15,00	2,21
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1512,04	1	1512,04	6,72	0,03
Residual	2250,92	10	225,09		
Total	3762,96	11			
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-71,41	40,33		-1,77	0,11
MM4	551,76	212,89	0,63	2,59	0,03

Elaborador por: Karla M. Aguaguiña Méndez

En el **Cuadro 5.3** se muestra el análisis de la varianza, donde se rechaza la hipótesis nula ( $H_0 = \beta_0 = \beta_1 = \dots = \beta_p = 0$ ), lo que significa que existe al menos un beta que es diferente a cero, siendo rechazada  $H_0$  con un “valor p” igual a 0.03; además se tiene un valor constante de -71.41 y un coeficiente que acompaña a la variable independiente de 551.76 respectivamente, en ambos casos, se rechaza la hipótesis nula de que estos valores son iguales a cero teniendo un valor p muy inferior a 0.10. La potencia de explicación del modelo es de 40%, por lo que la variable involucrada no tiene suficiente aporte.

## 5.2 MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA LA VARIABLE DEPENDIENTE P2: PROMEDIO DEL LARGO DE LA HOJA

$$P2 = 27.73 - 0.000007 MB1 + 15.49 MM1 + 0.08 MM8$$

**Cuadro 5.4**

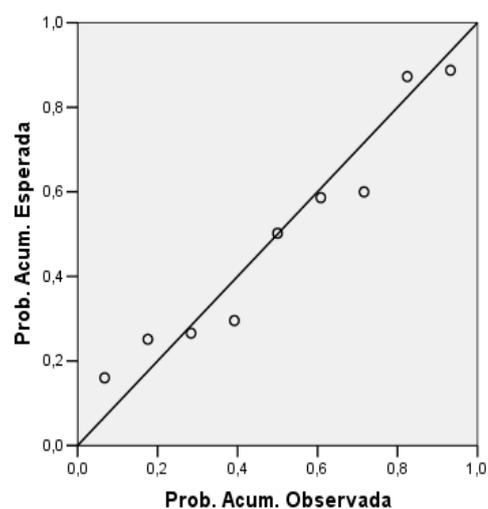
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

### Normalidad: Prueba de Kolmogorov-Smirnov

#### One-Sample Kolmogorov-Smirnov Test

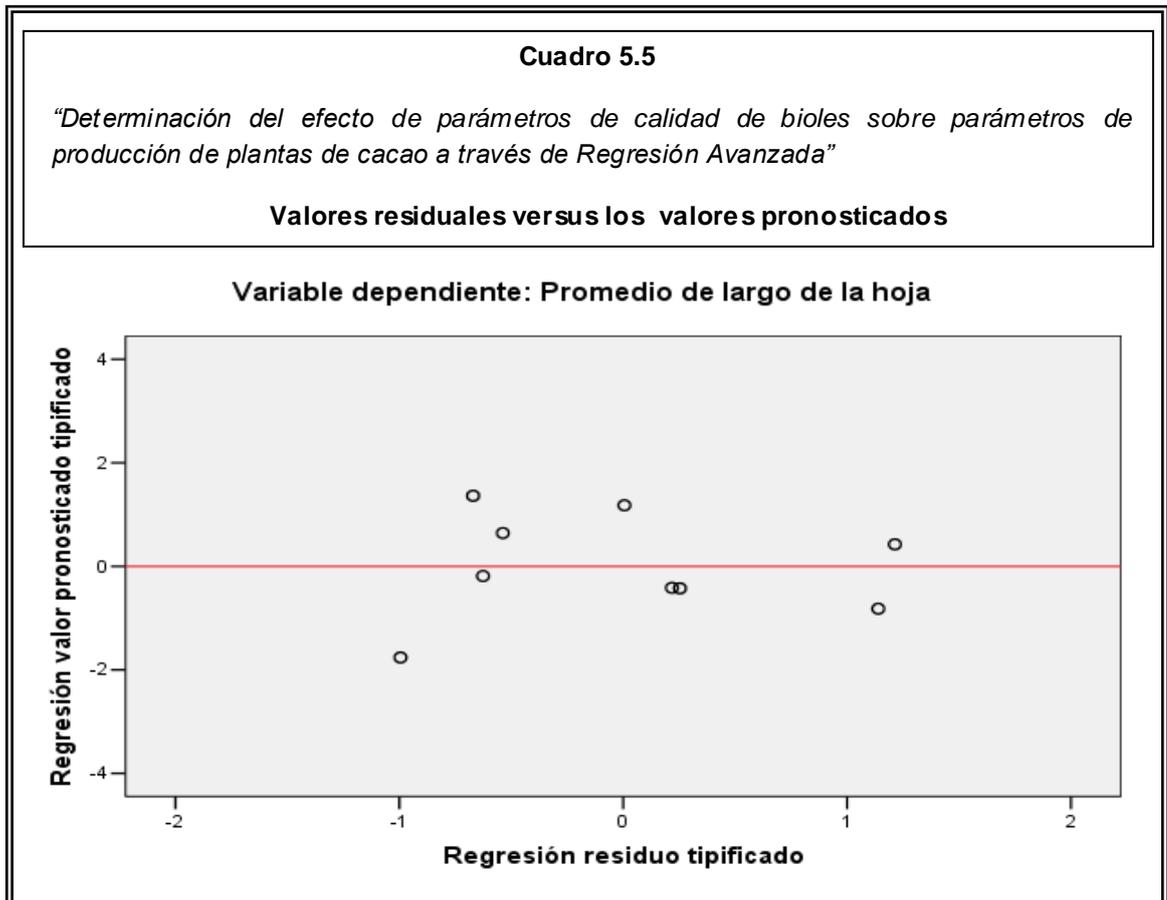
		Unstandardized Residual
<b>Normal Parameters(a,b)</b>	Mean	0,00
	Std. Deviation	0,13
	<b>Most Extreme Differences</b>	Absolute
Positive		0,20
Negative		-0,15
<b>Kolmogorov-Smirnov Z</b>		0,59
<b>Asymp. Sig. (2-tailed)</b>		0,88

#### Variable dependiente: Promedio del largo de la hoja



#### Correlations

		Unstandardized Predicted Value	Absres2
<b>Unstandardized Predicted Value</b>	Pearson Correlation	1	-0,36
	Sig. (2-tailed)		0,34
<b>Absres2</b>	Pearson Correlation	-0,36	1
	Sig. (2-tailed)	0,34	



**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 5.4** se puede verificar la comprobación de la hipótesis del test de Kolmogorov – Smirnov que con un valor p igual a 0.88 nos indica que los errores siguen una distribución normal con media igual a cero y varianza constante, también mediante el test de Pearson podemos corroborar que con un valor p igual a 0.34 se comprueba la independencia de los residuos o errores. En el **Cuadro 5.5** tenemos el gráfico de los valores residuales versus los pronosticados que de forma visual podemos ver que siguen una forma lineal, lo cual corrobora el indicio de

homocedasticidad. Mediante el uso del estadístico de Durbin – Watson (DW=0.601) notamos la presencia de una autocorrelación positiva.

<b>Cuadro 5.6</b>					
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>					
<b>Tabla ANOVA del modelo seleccionado</b>					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0,98	0,96	0,93	0,16	0,60
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	2,742	3	0,91	36,32	0,00
Residual	0,126	5	0,03		
Total	2,868	8			
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	27,73	0,96		28,79	0,00
MB1	0,00	0,00	-0,43	-4,29	0,01
MM1	15,49	5,48	0,79	2,83	0,04
MM8	0,08	0,54	0,04	0,14	0,89

**Elaborado por:** Karla M. Aguaguíña Méndez

En el modelo de regresión lineal múltiple para la variable dependiente Promedio del largo de la hoja (P2) existe más de una variable de explicación que son anaerobios (MB1), porcentaje de Nitrógeno (MM1) y ppm de Cobre (MM8) por lo que en el **Cuadro 5.6** se muestra el análisis

de la varianza podemos ver que el valor p es igual a 0.00 lo cual rechazamos la hipótesis nula ya que existe al menos un valor de beta que es diferente de cero. Tenemos los coeficientes que acompañan a las variables independientes cuyos valores son diferentes a cero, siendo estos 27.73, - 0.000007, 15.49 y 0.08 respectivamente, en estos casos.

La potencia de explicación del modelo es de 96%, por lo que las variables involucradas tienen suficiente aporte en el mismo. Cabe indicar que para este modelo se incluyeron tres variables de explicación donde únicamente dos aportan al modelo, la variable ppm de Cobre (MM8) tiene un valor p igual a 0.89 que es superior a 0.10 por lo que no se rechaza la hipótesis nula de que los betas son iguales a cero, resultando coeficientes que no son estadísticamente significativos.

### 5.3 MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA LA VARIABLE DEPENDIENTE P3: PROMEDIO DEL ANCHO DE LA HOJA

$$P3 = -35.41 - 1.01 PF1 + 7.30 PF3 + 73.12 PF5 - 9.09 PF6 - 0.0001 MB1 - 276.85 MM1 + 0.01 MM2$$

Cuadro 5.7

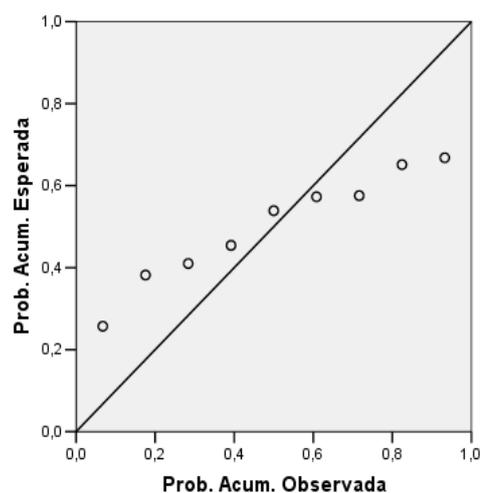
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Normalidad: Prueba de Kolmogorov-Smirnov**

#### One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
<b>Normal Parameters(a,b)</b>	Mean	0,00
	Std. Deviation	1,13
	<b>Most Extreme Differences</b>	Absolute
	Positive	0,11
	Negative	-0,16
<b>Kolmogorov-Smirnov Z</b>		0,49
<b>Asymp. Sig. (2-tailed)</b>		0,97

Variable dependiente: Promedio del ancho de la hoja



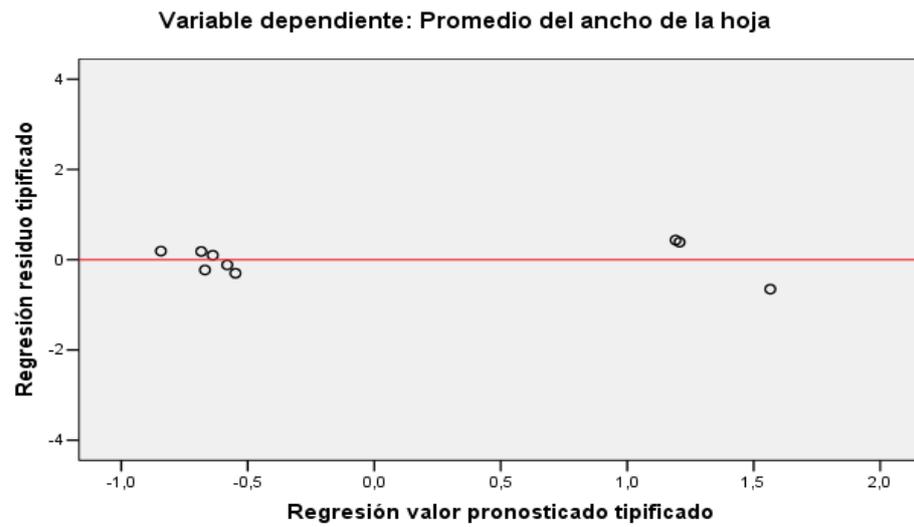
#### Correlations

		Unstandardized Predicted Value	Absres3
<b>Unstandardized Predicted Value</b>	Pearson Correlation	1	0,89
	Sig. (2-tailed)		0,00
<b>Absres3</b>	Pearson Correlation	0,89	1
	Sig. (2-tailed)	0,00	

**Cuadro 5.8**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Valores residuales versus los valores pronosticados**



**Elaborado por:** Karla M. Aguaguña Méndez

Este modelo es de regresión lineal múltiple, debido a que para la variable dependiente ( $P3$ ) existe más de una variable de explicación ( $PF1$ ,  $PF3$ ,  $PF5$ ,  $PF6$ ,  $MB1$ ,  $MM1$ ,  $MM2$ ), aquí se cumple con lo supuestos exigidos esto se observa en el **Cuadro 5.7**, donde se describe el gráfico y se verifica mediante los resultados del test de Kolmogorov-Smirnov en el cual los datos de los errores siguen una distribución normal con media cero (Valor  $p = 0.97$ ). En el **Cuadro 5.8** está el gráfico de los valores residuales versus los pronosticados que de forma visual nos da un indicio que sigue una forma lineal, lo que

corroborar el indicio de homocedasticidad. Usando el estadístico de Durbin – Watson (DW=2.85) se indica la presencia de una autocorrelación negativa.

<b>Cuadro 5.9</b>					
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>					
<b>Tabla ANOVA del modelo seleccionado</b>					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0,99	0,99	0,88	3,21	2,85
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	693,19	7	99,03	9,64	0,24
Residual	10,28	1	10,28		
Total	703,47	8			
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-35,41	204,97		-0,17	0,89
PF1	-1,01	3,40	-0,07	-0,30	0,82
PF3	7,30	11,88	3,68	0,61	0,65
PF5	73,12	224,88	0,07	0,33	0,80
PF6	-9,09	15,16	-3,44	-0,60	0,66
MB1	0,00	0,00	-0,51	-1,05	0,48
MM1	-276,85	127,24	-0,90	-2,18	0,27
MM2	0,01	0,05	0,09	0,23	0,85

Elaborado por: Karla M. Aguaguiña Méndez

En el **Cuadro 5.9** se muestra el análisis de la varianza, donde con un “valor p” igual a 0.24 no se rechaza la hipótesis nula, lo que significa que

todos los valores de beta son cero, por lo que ninguna variable genera algún aporte en el modelo seleccionado, siendo esto que las variables elegidas no explican el comportamiento del promedio ancho de la hoja.

## 5.4 MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA LA VARIABLE DEPENDIENTE P4: CLOROFILA

$$P4 = 36.09 + 0.25 PF2 + 0.53 PF4$$

**Cuadro 5.10**

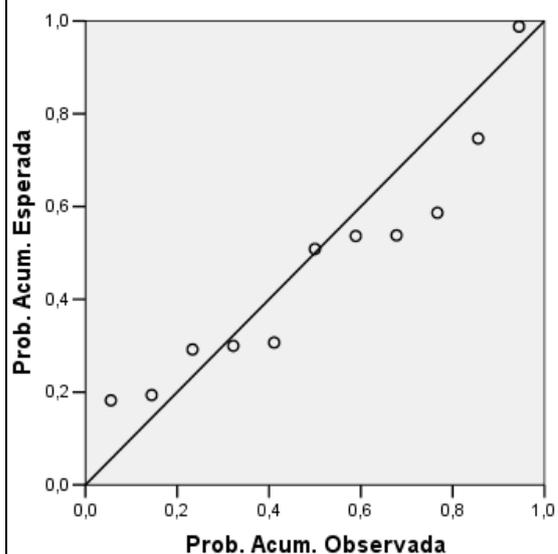
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Normalidad: Prueba de Kolmogorov-Smirnov**

### One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
<b>Normal Parameters(a,b)</b>	Mean	0,00
	Std. Deviation	1,36
	<b>Most Extreme Differences</b>	
	Absolute	0,22
	Positive	0,22
	Negative	-0,16
<b>Kolmogorov-Smirnov Z</b>		0,74
<b>Asymp. Sig. (2-tailed)</b>		0,65

### Variable dependiente: Clorofila



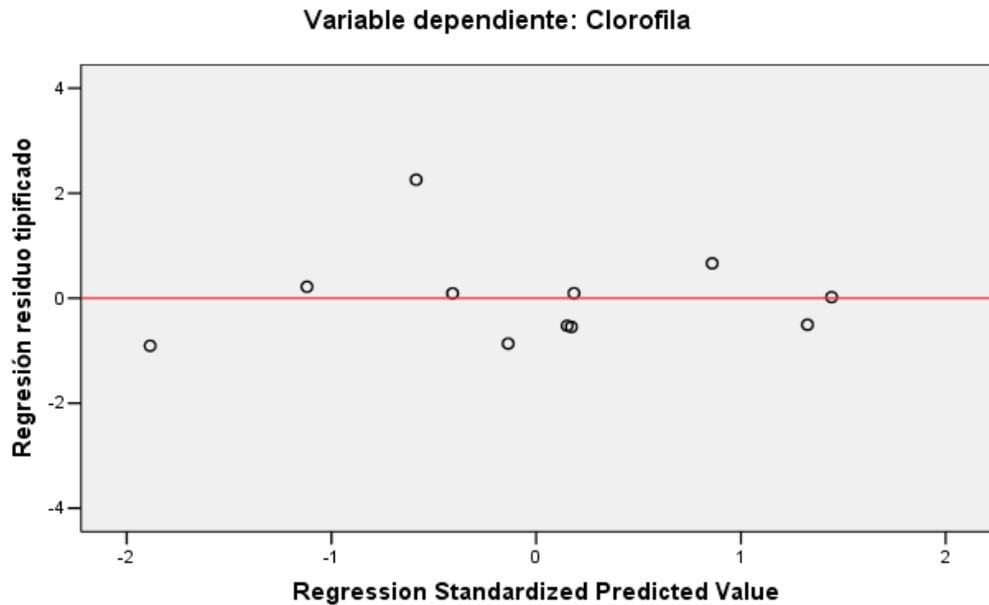
### Correlations

		Unstandardized Predicted Value	Absres4
<b>Unstandardized Predicted Value</b>	Pearson Correlation	1	-0,31
	Sig. (2-tailed)		0,35
<b>Absres4</b>	Pearson Correlation	-0,31	1
	Sig. (2-tailed)	0,35	

Cuadro 5.11

“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”

Valores residuales versus los valores pronosticados



Elaborado por: Karla M. Aguaguña Méndez

Este modelo es de regresión lineal múltiple, debido a que para la variable dependiente ( $P4$ ) es explicada por las variables ( $PF2$  y  $PF4$ ), aquí se cumple con lo supuestos exigidos esto se observa en el **Cuadro 5.10**, donde se describe el gráfico y se verifica mediante los resultados del test de Kolmogorov-Smirnov en el cual los datos de los errores siguen una distribución normal con media cero (Valor  $p = 0.65$ ), y la varianza constante se la determina mediante el comportamiento de los residuos, esto puede ser visualizado en el **Cuadro 5.11** que muestra los residuos frente a los valores pronosticados, en el cual se puede observar de forma visual que la nube de puntos sigue una forma lineal, lo cual nos da indicio de homocedasticidad. Mediante el estadístico de Durbin-Watson ( $D=2$ )

nos indica ausencia de autocorrelación; que es lo esperado en un modelo de regresión. Para verificar la independencia de residuos se elabora la prueba de Pearson, entre el valor residual y el pronosticado donde se verifica la independencia de datos (valor  $p = 0.35$ ).

**Cuadro 5.12**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Tabla ANOVA del modelo seleccionado**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0,78	0,60	0,50	1,52	2

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	28,16	2	14,08	6,07	0,02
Residual	18,56	8	2,32		
Total	46,73	10			

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	36,09	2,88		12,54	0,00
PF2	0,25	0,10	0,56	2,44	0,04
PF4	0,53	0,30	0,41	1,76	0,12

**Elaborado por:** Karla M. Aguaguíña Méndez

En el **Cuadro 5.12** se muestra el análisis de la varianza, donde se rechaza la hipótesis nula, lo que significa que existe al menos un valor de beta que es diferente a cero, siendo rechazada con un valor p igual a 0.02; además se tiene un valor constante y unos coeficientes que acompañan a las variables independientes, siendo estos 36.09, 0.25 y 0.53 respectivamente, en ambos casos, se rechaza la hipótesis de que estos valores son iguales a cero teniendo un valor p muy inferior a 0,10. La potencia de explicación del modelo es de 60%, aunque se lo puede considerar como un aporte “bajo” en algo se explica el comportamiento de la clorofila. Esto implica que a medida que se incremente una unidad en los valores de la conductividad, el valor de la clorofila incrementará en 0.25 décimas, de acuerdo al modelo que se obtiene a partir del cálculo de regresión.

## 5.5 MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA LA VARIABLE DEPENDIENTE P5: NÚMERO DE FRUTOS DE 4 MESES

$$P5 = 5.65 + 0.95 PF1 - 0.47 PF4 + 0.01 MM2 - 0.05 MM7$$

**Cuadro 5.13**

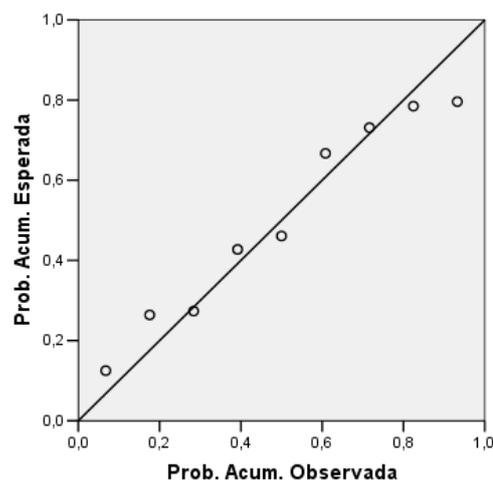
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Normalidad: Prueba de Kolmogorov-Smirnov**

### One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
<b>Normal Parameters(a,b)</b>	Mean	0,00
	Std. Deviation	1,30
	<b>Most Extreme Differences</b>	Absolute
	Positive	0,14
	Negative	-0,17
<b>Kolmogorov-Smirnov Z</b>		0,52
<b>Asymp. Sig. (2-tailed)</b>		0,95

Variable dependiente: Número de frutos de 4 meses



### Correlations

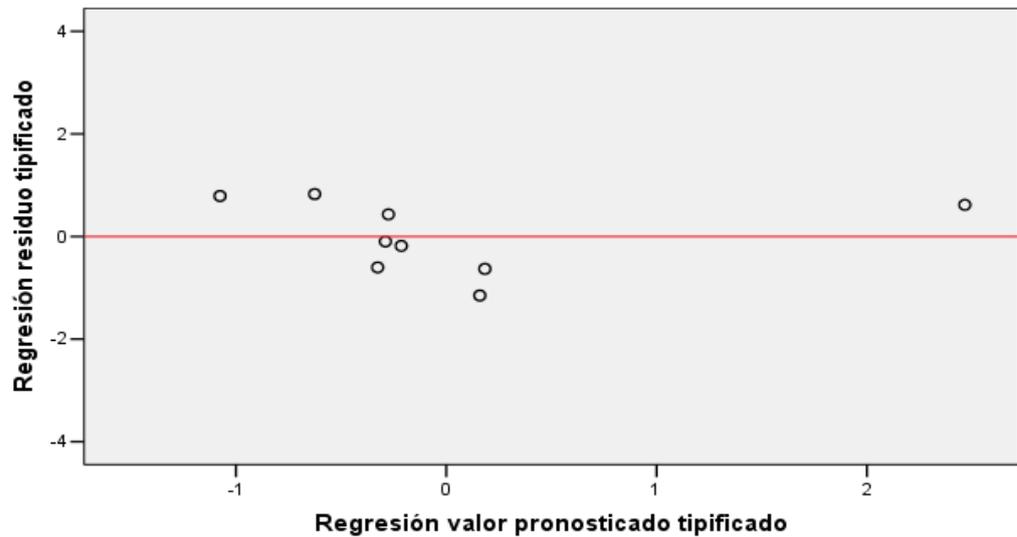
		Unstandardized Predicted Value	Absres5
<b>Unstandardized Predicted Value</b>	Pearson Correlation	1	0,03
	Sig. (2-tailed)		0,95
<b>Absres5</b>	Pearson Correlation	0,03	1
	Sig. (2-tailed)	0,95	

**Cuadro 5.14**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Valores residuales versus los valores pronosticados**

**Variable dependiente: Número de frutos de 4 meses**



**Elaborado por:** Karla M. Aguaguña Méndez

Este modelo es de regresión lineal múltiple, debido a que para la variable dependiente ( $P5$ ) existe más de una variable de explicación ( $PF1$ ,  $PF4$ ,  $MM2$ ,  $MM7$ ), aquí se cumple con lo supuestos exigidos esto se observa en el **Cuadro 5.13**, donde se describe el gráfico y se verifica mediante los resultados del test de Kolmogorov-Smirnov en el cual los datos de los errores siguen una distribución normal con media cero (Valor  $p = 0.95$ ). En el **Cuadro 5.14** se visualiza el gráfico de los residuos frente a los valores pronosticados que de forma visual se nota que la nube de puntos sigue una forma lineal, lo cual es un indicio de homocedasticidad; además esto se confirma con el estadístico Durbin-Watson ( $DW = 2.44$ ), cuyo valor nos

indica la ausencia de autocorrelación que es lo que se desea en regresión. Para verificar la independencia de residuos se elabora la prueba de Pearson, entre el valor residual y el pronosticado donde se verifica la independencia de las observaciones (valor  $p = 0.95$ ).

<b>Cuadro 5.15</b>					
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>					
<b>Tabla ANOVA del modelo seleccionado</b>					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0,85	0,72	0,44	1,83	2,44
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	34,13	4	8,53	2,54	0,19
Residual	13,42	4	3,36		
Total	47,56	8			
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	5,65	15,08		0,37	0,73
PF1	0,95	1,28	0,23	0,74	0,50
PF4	-0,47	0,61	-0,34	-0,76	0,49
MM2	0,01	0,02	0,16	0,33	0,76
MM7	-0,05	0,04	-0,55	-1,19	0,30

Elaborado por: Karla M. Aguaguña Méndez

En el **Cuadro 5.15** se muestra el análisis de la varianza, donde con un “valor p” igual a 0.19 no se rechaza la hipótesis nula, lo que significa que todos los valores de beta son cero, por lo que ninguna variable genera algún aporte en el modelo seleccionado, siendo esto que las variables elegidas no explican el comportamiento del número de frutos de 4 meses.

## 5.6 MODELO DE REGRESIÓN LINEAL MÚLTIPLE PARA LA VARIABLE DEPENDIENTE P6: NÚMERO DE FRUTOS SECOS

$$P6 = 15.54 + 1.02 PF1 + 0.00001 MB1 - 0.02 MM1 - 0.07 MM7 - 0.11 MM9$$

**Cuadro 5.16**

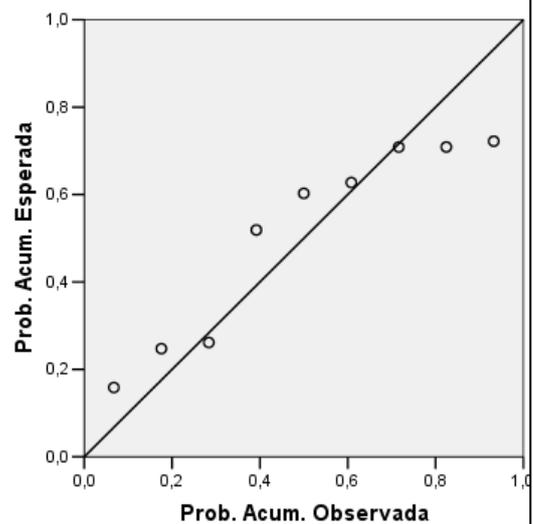
*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Normalidad: Prueba de Kolmogorov-Smirnov**

### One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
<b>Normal Parameters(a,b)</b>	Mean	0,00
	Std. Deviation	1,45
<b>Most Extreme Differences</b>	Absolute	0,22
	Positive	0,18
	Negative	-0,22
<b>Kolmogorov-Smirnov Z</b>		0,66
<b>Asymp. Sig. (2-tailed)</b>		0,78

**Variable dependiente: Número de frutos secos**



### Correlations

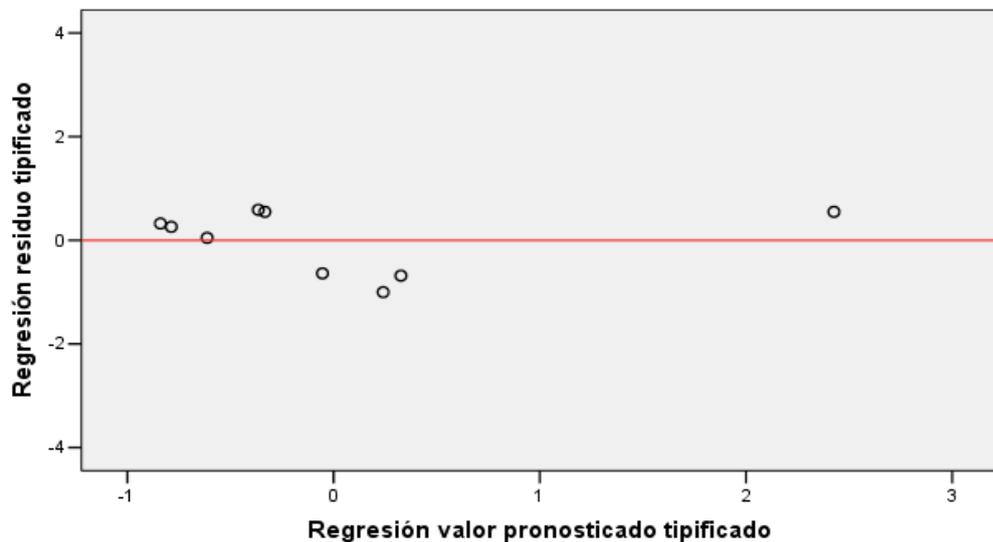
		Unstandardized Predicted Value	Absres6
<b>Unstandardized Predicted Value</b>	Pearson Correlation	1	0,39
	Sig. (2-tailed)		0,30
<b>Absres6</b>	Pearson Correlation	0,39	1
	Sig. (2-tailed)	0,30	

**Cuadro 5.17**

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Valores residuales versus los valores pronosticados**

**Variable dependiente: Número de frutos secos**



**Elaborado por:** Karla M. Aguaguña Méndez

Este modelo es de regresión lineal múltiple, debido a que para la variable dependiente (*P6*) existe más de una variable de explicación (*PF1*, *MB1*, *MM1*, *MM7*, *MM9*), aquí se cumple con los supuestos exigidos esto se observa en el **Cuadro 5.16**, donde se describe el gráfico y se verifica mediante los resultados del test de Kolmogorov-Smirnov en el cual los datos de los errores siguen una distribución normal con media cero (Valor  $p = 0.78$ ). En el **Cuadro 5.17** se observa el gráfico de los residuos frente a los valores pronosticados que de forma visual se observa que la nube de puntos sigue una forma lineal, lo cual es un indicio de homocedasticidad y se corrobora mediante el estadístico de Durbin-Watson ( $DW = 2$ ) que

indica la ausencia de autocorrelación lo cual es necesario en regresión. Para verificar la independencia de residuos se elabora la prueba de Pearson, entre el valor residual y el pronosticado donde se verifica la independencia de datos (0.30).

Cuadro 5.18

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

Tabla ANOVA del modelo seleccionado

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0,89	0,80	0,47	2,36	2

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	67,24	5	13,45	2,41	0,25
Residual	16,76	3	5,59		
Total	84	8			

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	15,54	18,68		0,83	0,47
PF1	1,02	2,44	0,19	0,42	0,70
MB1	0,00	0,00	0,17	0,52	0,64
MM2	-0,01	0,02	-0,12	-0,26	0,81
MM7	-0,07	0,05	-0,57	-1,35	0,27
MM9	-0,11	1,21	-0,05	-0,09	0,93

Elaborado por: Karla M. Aguaguña Méndez

En el **Cuadro 5.18** se muestra el análisis de la varianza, donde con un “valor p” igual a 0.25, no se rechaza la hipótesis nula, lo que significa que todos los valores de beta son cero, por lo que ninguna variable genera algún aporte en el modelo seleccionado, siendo esto que las variables elegidas no explican el comportamiento del número de frutos secos.

## 5.7 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE P1: *INHIBICIÓN*

$$P1 = 27.72 + 18.82 PF1 + 64.27 PF3 - 81.77 PF6 - 0.07 MB3 - 349.48$$

$$MM1 - 0.07 MM2$$

Cuadro 5.19

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

Tabla ANOVA del modelo seleccionado

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
4	0,96	0,93	0,84	7,39	2

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	3489,61	6	581,60	10,64	0,01
Residual	273,35	5	54,67		
Total	3762,96	11			

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	27,72	24,38		1,14	0,31
PF1	18,82	6,15	0,53	3,06	0,03
PF3	64,27	12,08	17,57	5,32	0,00
PF6	-81,77	16,35	-16,56	-5,00	0,00
MB3	-0,07	0,02	-0,58	-3,42	0,02
MM1	-349,48	102,82	-0,58	-3,40	0,02
MM2	-0,07	0,02	-0,53	-2,93	0,03

Elaborado por: Karla M. Aguaguña Méndez

En el **Cuadro 5.19** se muestra el análisis de la varianza, donde se rechaza la hipótesis nula, lo que significa que existe al menos un valor de beta que es diferente a cero, siendo rechazada con un valor p igual a 0.01; además se tiene un valor constante y unos coeficientes que acompañan a las variables independientes, siendo estos 27.72, 18.82, 64.27, - 81.77, - 0.07, - 349.48 y - 0.07 respectivamente, en ambos casos, se rechaza la hipótesis de que estos valores son iguales a cero teniendo un valor p muy inferior a 0.10. La potencia de explicación del modelo es de 93%, se lo puede considerar como un aporte “alto” puesto que en algo se explica el comportamiento de la inhibición.

**5.8 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD  
PARA LA VARIABLE DEPENDIENTE P2: PROMEDIO DEL LARGO DE LA  
HOJA**

$$P2 = 27.79 + 16.14 MM1 + 0.07 PF3 - 0.11 PF4$$

Cuadro 5.20					
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>					
Tabla ANOVA del modelo seleccionado					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
4	0,97	0,95	0,92	0,17	2
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	2,72	3	0,91	30,89	0,001
Residual	0,15	5	0,03		
Total	2,87	8			
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	27,79	0,40		69,79	0,00
MM1	16,14	2,02	0,82	7,98	0,00
PF3	0,07	0,02	0,56	3,82	0,01
PF4	-0,11	0,05	-0,33	-2,26	0,07

Elaborado por: Karla M. Aguaguña Méndez

En el **Cuadro 5.20** se muestra el análisis de la varianza, donde se rechaza la hipótesis nula, lo que significa que existe al menos un valor de beta que

es diferente a cero, siendo rechazada con un valor p igual a 0.01; además se tiene un valor constante y unos coeficientes que acompañan a las variables independientes, siendo estos 27.72, 18.82, 16.14, 0.07 y - 0.11 respectivamente, en ambos casos, se rechaza la hipótesis de que estos valores son iguales a cero teniendo un valor p muy inferior a 0.10. La potencia de explicación del modelo es de 95%, se lo puede considerar como un aporte “alto” puesto que explica el comportamiento del promedio del largo de las hojas en plantas de cacao.

**5.9 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD  
PARA LA VARIABLE DEPENDIENTE P3: *PROMEDIO DEL ANCHO DE LA  
HOJA***

$$P3 = - 116.11 + 0.11 MM2 + 72.37 MM3 - 0.71 PF2$$

<b>Cuadro 5.21</b>					
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>					
<b>Tabla ANOVA del modelo seleccionado</b>					
<b>Model</b>	<b>R</b>	<b>R Square</b>	<b>Adjusted R Square</b>	<b>Std. Error of the Estimate</b>	<b>Durbin-Watson</b>
<b>4</b>	0,96	0,93	0,89	3,11	2,67
<b>Model</b>	<b>Sum of Squares</b>	<b>df</b>	<b>Mean Square</b>	<b>F</b>	<b>Sig.</b>
<b>Regression</b>	655,08	3	218,36	22,56	0,002
<b>Residual</b>	48,39	5	9,68		
<b>Total</b>	703,47	8			
<b>Model</b>	<b>Unstandardized Coefficients</b>		<b>Standardized Coefficients</b>	<b>t</b>	<b>Sig.</b>
	<b>B</b>	<b>Std. Error</b>	<b>Beta</b>		
<b>(Constant)</b>	-116,11	17,77		-6,53	0,00
<b>MM2</b>	0,11	0,02	0,75	6,35	0,00
<b>MM3</b>	72,37	16,24	0,54	4,46	0,01
<b>PF2</b>	-0,71	0,24	-0,36	-2,99	0,03

**Elaborado por:** Karla M. Aguaguña Méndez

En el **Cuadro 5.21** se muestra el análisis de la varianza, donde se rechaza la hipótesis nula, lo que significa que existe al menos un valor de beta que es diferente a cero, siendo rechazada con un valor p igual a 0.002; además se tiene un valor constante y unos coeficientes que acompañan a las variables independientes, siendo estos  $-116.11$ ,  $0.11$ ,  $72.37$  y  $-0.71$  respectivamente, en ambos casos, se rechaza la hipótesis de que estos valores son iguales a cero teniendo un valor p muy inferior a 0.10. La potencia de explicación del modelo es de 93%, se lo puede considerar como un aporte “alto” puesto que explica el comportamiento del promedio del ancho de las hojas en plantas de cacao.

## 5.10 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE P4: CLOROFILA

$$P4 = - 160.05 - 29.56 MM5 + 1.43 PF4 + 192.93 PF5 + 0.00002 MB1$$

Cuadro 5.22					
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>					
Tabla ANOVA del modelo seleccionado					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
4	0,97	0,94	0,89	0,75	2,53
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	37,32	4	9,33	16,73	0,009
Residual	2,23	4	0,56		
Total	39,56	8			
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-160,05	36,74		-4,36	0,01
MM5	-29,56	4,68	-1,09	-6,32	0,00
PF4	1,43	0,22	1,14	6,53	0,00
PF5	192,93	35,21	0,83	5,48	0,01
MB1	0,00	0,00	0,38	2,40	0,07

Elaborado por: Karla M. Aguaguña Méndez

En el **Cuadro 5.22** se muestra el análisis de la varianza, donde se rechaza la hipótesis nula, lo que significa que existe al menos un valor de beta que es diferente a cero, siendo rechazada con un valor p igual a 0.009; además se tiene un valor constante y unos coeficientes que acompañan a las variables independientes, siendo estos - 160.05, - 29.56, 1.43, 192.93 y 0.00 respectivamente, en este caso, se rechaza la hipótesis de que estos valores son iguales a cero teniendo un valor p muy inferior a 0.10. La potencia de explicación del modelo es de 94%, se lo puede considerar como un aporte “alto” puesto que explica el comportamiento del contenido de Clorofila en las plantas de cacao.

### 5.11 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE P5: NÚMERO DE FRUTOS DE 4 MESES

$$P5 = - 16.13 + 7.23 PF1 - 0.59 PF2 - 4.34 PF3 + 5.98 PF6$$

Cuadro 5.23					
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>					
Tabla ANOVA del modelo seleccionado					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
3	0,95	0,90	0,84	1,02	2,93
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	57,97	4	14,49	13,99	0,003
Residual	6,21	6	1,04		
Total	64,18	10			
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-16,13	3,51		-4,59	0,00
PF1	7,23	1,12	1,82	6,44	0,00
PF2	-0,59	0,12	-1,15	-5,15	0,00
PF3	-4,34	1,39	-8,25	-3,13	0,02
PF6	5,98	1,90	8,47	3,15	0,02

Elaborado por: Karla M. Aguaguña Méndez

En el **Cuadro 5.23** se muestra el análisis de la varianza, donde se rechaza la hipótesis nula, lo que significa que existe al menos un valor de beta que

es diferente a cero, siendo rechazada con un valor p igual a 0.009; además se tiene un valor constante y unos coeficientes que acompañan a las variables independientes, siendo estos - 16.13, 7.23, - 0.59, - 4.34 y 5.98 respectivamente, en este caso, se rechaza la hipótesis de que estos valores son iguales a cero teniendo un valor p muy inferior a 0.10. La potencia de explicación del modelo es de 90%, se lo considera como un aporte “alto” puesto que explica el comportamiento del promedio de número de frutos de 4 meses producidos.

## 5.12 MODELO DE REGRESIÓN LINEAL MÚLTIPLE USANDO BACKWARD PARA LA VARIABLE DEPENDIENTE P6: NÚMERO DE FRUTOS SECOS

$$P6 = 203.88 - 1.38 PF4 - 180.12 PF5 + 0.52 MM6 - 0.06 MM7$$

Cuadro 5.24					
<i>“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”</i>					
Tabla ANOVA del modelo seleccionado					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
3	1,00	0,99	0,99	0,37	2,75
Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	83,45	4	20,86	152,35	0,000
Residual	0,55	4	0,14		
Total	84	8			
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	203,88	19,90		10,25	0,00
PF4	-1,38	0,13	-0,75	-10,58	0,00
PF5	-180,12	19,39	-0,53	-9,29	0,00
MM6	0,52	0,04	0,83	12,18	0,00
MM7	-0,06	0,01	-0,52	-8,41	0,00

Elaborado por: Karla M. Aguaguiña Méndez

En el **Cuadro 5.24** se muestra el análisis de la varianza, donde se rechaza la hipótesis nula, lo que significa que existe al menos un valor de beta que es diferente a cero, siendo rechazada con un valor p igual a 0.000; además se tiene un valor constante y unos coeficientes que acompañan a las variables independientes, siendo estos 203.88, - 1.38, - 180.12, 0.52 y - 0.06 respectivamente, en este caso, se rechaza la hipótesis de que estos valores son iguales a cero teniendo un valor p muy inferior a 0.10. La potencia de explicación del modelo es de 99%, el cual tiene un aporte “alto” puesto que explica el comportamiento del promedio de número de frutos secos.

### 5.13 MODELOS DE REGRESIÓN FINALES

Los modelos finales son:

- $P1 = 27.72 + 18.82 PF1 + 64.27 PF3 - 81.77 PF6 - 0.07 MB3 - 349.48 MM1 - 0.07 MM2$
- $P2 = 27.79 + 16.14 MM1 + 0.07 PF3 - 0.11 PF4$
- $P3 = - 116.11 + 0.11 MM2 + 72.37 MM3 - 0.71 PF2$
- $P4 = - 160.05 - 29.56 MM5 + 1.43 PF4 + 192.93 PF5 + 0.00002 MB1$

- $P5 = -16.13 + 7.23 PF1 - 0.59 PF2 - 4.34 PF3 + 5.98 PF6$
- $P6 = 203.88 - 1.38 PF4 - 180.12 PF5 + 0.52 MM6 - 0.06 MM7$

## CAPÍTULO SEIS

### 6.1 CONCLUSIONES

- Los modelos seleccionados bajo el método Backward fueron los que presentaron mejores resultados, el criterio de selección fue el de coeficiente de determinación ( $R^2$ ), el mismo que indica que tan bueno son los modelos para explicar a las variables en estudio. En el presente estudio todos los coeficientes de determinación fueron mayores a 90%.
- Se observó en los modelos finales la presencia significativa estadística de las variables referentes a los parámetros físico-químicos. En los modelos elegidos hay la presencia de al menos uno de estos parámetros, que están relacionadas con la producción a través de los parámetros fitosanitarios y de producción.
- Se observó en los modelos que hay presencia de las variables que representan los micro/macro nutrientes, en especial el Nitrógeno, Calcio, Magnesio, Zinc y más están influenciando en la producción que tiene relación con los parámetros fitosanitarios en las plantas de cacao.

## 6.2 RECOMENDACIONES

- Se recomienda para futuros trabajos de investigación seleccionar un diseño experimental o muestreo previo al ensayo en campo con el propósito de tener una estimación cercana a la realidad.
- Se recomienda realizar un estudio químico meticuloso del suelo y a las haciendas donde se realizarán futuras investigaciones como esta para de esta forma conocer el efecto que ocasionan las enmiendas orgánicas líquidas en las plantas y de alguna forma ver si el suelo también logra absorber estos nutrientes y así enriquecerlo.
- Es recomendable mantener una periodicidad en la recolección de datos, de esta forma se podrá garantizar un mejor análisis de la información y así el uso de más técnicas estadísticas.

## **BIBLIOGRAFÍA**

- [1] Censo Nacional Agropecuario 2000, INEC. Disponible en [http://www.inec.gob.ec/estadisticas/?option=com\\_content&view=article&id=11&Itemid=126&TB\\_iframe=true&height=400&width=931](http://www.inec.gob.ec/estadisticas/?option=com_content&view=article&id=11&Itemid=126&TB_iframe=true&height=400&width=931)
- [2] Banco Central del Ecuador. Exportaciones de cacao y sus derivados. Disponible en [www.bce.fin.ec](http://www.bce.fin.ec) visitado el 25 de febrero de 2013
- [3] Análisis Sectorial del cacao y sus elaborados- PROECUADOR. Disponible en [www.proecuador.gob.ec](http://www.proecuador.gob.ec) visitado el 18 de junio de 2013
- [4] Asociación Nacional de Exportadores de Cacao. Disponible en [www.anecacao.com](http://www.anecacao.com) visitado el 25 de octubre de 2012
- [5] Identificación Varietal de 41 Plantas Seleccionadas de Cacao (*Theobroma cacao* L.) Provenientes de Cuatro Cultivares Distintos de la Región Amazónica Ecuatoriana, Mediante el Uso de Marcadores Microsatélites, Revista Tecnológica ESPOL RTE, Vol. 23, Nº 1, 121–128, (Diciembre 2010)
- [6] Agricultura Tropical de Ecuador. Disponible en <http://agricultura-tropical-ecuador.blogspot.com/2010/11/variedades-de-cacao.html>, visitado el 18 de agosto de 2013

- [7] Magdama, F. Estudio del efecto de bioles y cepas de *Trichoderma* sp. aisladas de zonas cacaoteras, como alternativas de control de *Moniliophthora roreri*, en condiciones *in vitro*. Tesis de grado – ESPOL. 2010
- [8] Cacao: Suelo, variedades, preparación del terreno, siembra, labores culturales, control de malezas, labores fitosanitarias, riegos, cosecha, fermentación, secado y almacenamiento. Disponible en <http://www.crystal-chemical.com/cacao.htm> consultado el 28 de noviembre de 2013.
- [9] Zurita, G. (2009), Probabilidad y Estadística - fundamentos y aplicaciones, Instituto de Ciencias Matemáticas, ESPOL
- [10] Rojo, J.M. Regresión Lineal Múltiple, Instituto de Economía y geografía, Madrid 2007. Disponible en [http://humanidades.cchs.csic.es/cchs/web\\_UAE/tutoriales/PDF/Regresion\\_lineal\\_multiple\\_3.pdf](http://humanidades.cchs.csic.es/cchs/web_UAE/tutoriales/PDF/Regresion_lineal_multiple_3.pdf) visitado en 28 de noviembre de 2013
- [11] Ministerio de Agricultura y Ganadería (2000). Protocolo de requisitos fitosanitarios para la exportación de bananos de Costa Rica a China. Costa Rica
- [12] FAO - Comité de problemas de productos básicos. Grupo intergubernamental sobre el banano y las frutas tropicales; España, 2003
- [13] Huamán Martha. Diagnóstico de la cadena de valor del banano en el valle del Piura. Catholic Relief Services. Perú, 2005
- [14] Jaén Gilberto; Rodríguez Luis; Lic. Brunilda Ulloa. Comisión de Libre Competencia y Asuntos del Consumidor. Diagnóstico de la problemática relacionada con el contenido residual de plaguicidas en los alimentos. 2000.

- [15] Suquilanda, M. Agricultura orgánica: alternativa tecnológica del futuro. Ecuador 1996.
- [16] Chávez, Eduardo. Determinación de la calidad de Biofertilizantes líquidos y estudio del potencial para la inhibición, como alternativa en el manejo de Sigatoka negra en sistema de producción orgánica. Tesis de grado – ESPOL. 2009.
- [17] Programa Nacional de Agricultura Orgánica, Costa Rica. Introducción a la agricultura orgánica. 2000. Disponible en [http://eva.utpl.edu.ec/openutpl/ocw/file.php/469/Agricultura\\_Organica.pdf](http://eva.utpl.edu.ec/openutpl/ocw/file.php/469/Agricultura_Organica.pdf)
- [18] Centro de Investigaciones Biotecnológicas del Ecuador. Orientaciones técnicas para el establecimiento y utilización de enmiendas orgánicas líquidas (bioles). 2009
- [19] Caracterización Física-Química y Biológica de enmiendas orgánicas aplicadas en la producción de cultivos en República Dominicana. Revista de la Ciencia del Suelo y Nutrición Vegetal, Vol. 8, N° 3, Temuco 2008. Disponible en [http://www.scielo.cl/scielo.php?pid=S0718-27912008000300002&script=sci\\_arttext](http://www.scielo.cl/scielo.php?pid=S0718-27912008000300002&script=sci_arttext)
- [20] Cerveñansky A. Micronutrientes – Fertilidad. 2011
- [21] FAO, 2002. Los fertilizantes y su uso. Una guía de bolsillo para los oficiales de extensión. Cuarta Edición.
- [22] Chóez, Geovanny. Determinación y Uso de Métodos Estadísticos en la obtención de un modelo bioclimático para el manejo de la Sigatoka Negra

presente en haciendas bananeras de la Provincia del Guayas. Tesis de Grado – ESPOL. 2006

[23] Crow, Pamela. Aplicación de Métodos Estadísticos Multivariados en el estudio de calidad de enmiendas orgánicas sólidas y líquidas preparadas en las provincias de Guayas, Los Ríos y El Oro. Tesis de Grado – ESPOL. 2006

[24] Vega, José; Guzmán, Josué. Regresión PLS y PCA como solución al problema de multicolinealidad en regresión múltiple. Revista de Matemáticas: Teoría y Aplicaciones 2011 18(1): 9-20.

[25] Rancher, Alvin (2002). Methods of Multivariate Analysis. Second Edition. Wiley Series in Probability and Statistic.

[26] La ventaja comparativa del cacao ecuatoriano. Apuntes de Economía N° 20. Junio 2002. Disponible en <http://www.bce.fin.ec/documentos/PublicacionesNotas/Competitividad/Estudios/ae20.pdf>

[27] Actualización de la Micobiota Patogénica del Cacao “arriba” (*Theobroma cacao*) presente en la Costa Ecuatoriana. Revista Tecnológica ESPOL RTE, Vol. 23, N° 1, 21-26 (Diciembre 2010). Disponible en <http://www.rte.espol.edu.ec/index.php/tecnologica/article/view/8/3>

[28] Ingham, E. 2005. Citado por Quito D., 2007. Estudio comparativo de dos biofertilizantes líquidos en condiciones in vitro e invernadero en plantas de banano y su efecto en el desarrollo de la Sigatoka Negra (*Mycosphaerella fijiensis* Morelet). Tesis de grado – Espol. Ecuador.p113.

# ANEXOS

## MATRIZ DE CORRELACIÓN

### ANEXO 1

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

**Matriz de correlación de las variables dependientes e independientes**

		PF1	PF2	PF3	PF4	PF5	PF6	MB1	MB2	MB3	MM1	MM2	MM3	MM4	MM5	MM6	MM7	MM8	MM9
<b>P1</b>	Pearson Correlation	0,25	0,10	-0,11	0,22	-0,09	-0,14	0,09	-0,14	0,13	0,08	-0,33	-0,03	0,63	0,24	0,06	0,08	-0,20	-0,22
	Sig. (2-tailed)	0,33	0,72	0,67	0,41	0,74	0,60	0,74	0,58	0,61	0,80	0,30	0,91	0,03	0,45	0,86	0,81	0,53	0,49
<b>P2</b>	Pearson Correlation	-0,20	-0,15	0,43	0,13	0,00	0,40	-0,66	0,09	0,43	0,89	-0,27	-0,13	0,40	-0,46	-0,02	0,13	-0,84	-0,38
	Sig. (2-tailed)	0,55	0,67	0,19	0,70	1,00	0,22	0,03	0,80	0,18	0,00	0,48	0,74	0,29	0,21	0,97	0,73	0,00	0,31
<b>P3</b>	Pearson Correlation	-0,62	-0,24	0,61	0,38	0,60	0,65	-0,57	-0,44	-0,05	-0,58	0,78	0,53	-0,76	0,68	0,50	0,36	0,52	0,63
	Sig. (2-tailed)	0,04	0,48	0,05	0,24	0,05	0,03	0,07	0,18	0,88	0,10	0,01	0,14	0,02	0,04	0,17	0,34	0,15	0,07
<b>P4</b>	Pearson Correlation	0,31	0,67	0,13	0,55	0,28	0,12	-0,09	-0,33	0,23	0,49	0,15	0,20	0,03	-0,14	0,43	0,10	-0,30	-0,46
	Sig. (2-tailed)	0,35	0,02	0,71	0,08	0,40	0,72	0,79	0,32	0,49	0,18	0,70	0,60	0,93	0,71	0,25	0,81	0,44	0,21
<b>P5</b>	Pearson Correlation	0,66	0,16	-0,48	-0,58	-0,20	-0,48	0,50	-0,31	-0,36	-0,24	-0,65	0,48	-0,10	-0,32	-0,05	-0,79	0,42	-0,54
	Sig. (2-tailed)	0,03	0,65	0,13	0,06	0,56	0,13	0,12	0,35	0,27	0,53	0,06	0,19	0,80	0,41	0,91	0,01	0,26	0,13
<b>P6</b>	Pearson Correlation	0,67	0,30	-0,33	-0,19	-0,30	-0,34	0,55	-0,27	-0,29	-0,13	-0,75	0,39	0,00	-0,44	-0,13	-0,84	0,31	-0,64
	Sig. (2-tailed)	0,02	0,36	0,32	0,57	0,37	0,31	0,08	0,42	0,39	0,75	0,02	0,30	1,00	0,23	0,73	0,00	0,41	0,07

## MATRIZ DE CORRELACIÓN DE LAS VARIABLES INDEPENDIENTES

### ANEXO 2

*“Determinación del efecto de parámetros de calidad de bioles sobre parámetros de producción de plantas de cacao a través de Regresión Avanzada”*

#### Matriz de correlación de las variables independientes

		PF1	PF2	PF3	PF4	PF5	PF6	MB1	MB2	MB3	MM1	MM2	MM3	MM4	MM5	MM6	MM7	MM8	MM9
<b>PF1</b>	Pearson Correlation	1,00	0,70	-0,51	-0,42	-0,25	-0,52	0,36	0,01	0,11	0,15	-0,14	0,20	0,31	-0,36	0,17	-0,24	0,05	-0,65
	Sig. (2-tailed)		0,00	0,02	0,05	0,25	0,01	0,13	0,98	0,66	0,64	0,67	0,54	0,33	0,25	0,59	0,45	0,88	0,02
<b>PF2</b>	Pearson Correlation	0,70	1,00	-0,47	-0,07	0,04	-0,47	0,18	-0,19	-0,05	0,18	0,16	0,20	0,04	-0,06	0,53	-0,07	0,07	-0,37
	Sig. (2-tailed)	0,00		0,03	0,77	0,88	0,03	0,45	0,44	0,84	0,58	0,62	0,54	0,90	0,86	0,08	0,83	0,83	0,24
<b>PF3</b>	Pearson Correlation	-0,51	-0,47	1,00	0,56	0,48	1,00	-0,72	-0,05	0,28	0,28	0,53	0,31	-0,27	0,35	0,24	0,72	-0,22	0,30
	Sig. (2-tailed)	0,02	0,03		0,01	0,03	0,00	0,00	0,85	0,24	0,37	0,07	0,33	0,39	0,27	0,45	0,01	0,49	0,34
<b>PF4</b>	Pearson Correlation	-0,42	-0,07	0,56	1,00	0,39	0,56	-0,25	-0,17	0,14	0,14	0,43	0,09	-0,04	0,48	0,35	0,62	-0,15	0,24
	Sig. (2-tailed)	0,05	0,77	0,01		0,07	0,01	0,30	0,48	0,56	0,66	0,17	0,77	0,91	0,12	0,26	0,03	0,65	0,45
<b>PF5</b>	Pearson Correlation	-0,25	0,04	0,48	0,39	1,00	0,49	-0,34	-0,41	0,12	-0,11	0,38	0,63	-0,50	0,38	0,48	0,29	0,34	0,08
	Sig. (2-tailed)	0,25	0,88	0,03	0,07		0,02	0,15	0,08	0,63	0,74	0,22	0,03	0,10	0,23	0,11	0,36	0,28	0,80

<b>PF6</b>	Pearson Correlation	-0,52	-0,47	1,00	0,56	0,49	1,00	-0,72	-0,07	0,27	0,26	0,55	0,33	-0,32	0,36	0,25	0,71	-0,18	0,32
	Sig. (2-tailed)	0,01	0,03	0,00	0,01	0,02		0,00	0,78	0,26	0,42	0,07	0,29	0,31	0,25	0,43	0,01	0,57	0,31
<b>MB1</b>	Pearson Correlation	0,36	0,18	-0,72	-0,25	-0,34	-0,72	1,00	-0,02	-0,19	-0,46	-0,72	-0,42	0,36	0,11	-0,40	-0,60	0,26	0,00
	Sig. (2-tailed)	0,13	0,45	0,00	0,30	0,15	0,00		0,95	0,44	0,13	0,01	0,17	0,24	0,73	0,19	0,04	0,41	0,99
<b>MB2</b>	Pearson Correlation	0,01	-0,19	-0,05	-0,17	-0,41	-0,07	-0,02	1,00	0,23	0,35	0,14	-0,67	0,28	-0,24	-0,57	0,28	-0,47	0,25
	Sig. (2-tailed)	0,98	0,44	0,85	0,48	0,08	0,78	0,95		0,33	0,26	0,67	0,02	0,38	0,46	0,05	0,37	0,13	0,44
<b>MB3</b>	Pearson Correlation	0,11	-0,05	0,28	0,14	0,12	0,27	-0,19	0,23	1,00	0,07	-0,17	-0,03	0,22	-0,24	-0,13	0,00	-0,23	-0,25
	Sig. (2-tailed)	0,66	0,84	0,24	0,56	0,63	0,26	0,44	0,33		0,84	0,59	0,93	0,49	0,46	0,68	1,00	0,48	0,43
<b>MM1</b>	Pearson Correlation	0,15	0,18	0,28	0,14	-0,11	0,26	-0,46	0,35	0,07	1,00	0,29	-0,23	0,31	-0,38	-0,06	0,42	-0,88	-0,15
	Sig. (2-tailed)	0,64	0,58	0,37	0,66	0,74	0,42	0,13	0,26	0,84		0,36	0,47	0,32	0,22	0,85	0,18	0,00	0,64
<b>MM2</b>	Pearson Correlation	-0,14	0,16	0,53	0,43	0,38	0,55	-0,72	0,14	-0,17	0,29	1,00	0,30	-0,45	0,17	0,42	0,82	0,02	0,26
	Sig. (2-tailed)	0,67	0,62	0,07	0,17	0,22	0,07	0,01	0,67	0,59	0,36		0,34	0,15	0,60	0,18	0,00	0,95	0,42
<b>MM3</b>	Pearson Correlation	0,20	0,20	0,31	0,09	0,63	0,33	-0,42	-0,67	-0,03	-0,23	0,30	1,00	-0,64	0,05	0,55	0,02	0,58	-0,27
	Sig. (2-tailed)	0,54	0,54	0,33	0,77	0,03	0,29	0,17	0,02	0,93	0,47	0,34		0,03	0,89	0,06	0,95	0,05	0,41
<b>MM4</b>	Pearson Correlation	0,31	0,04	-0,27	-0,04	-0,50	-0,32	0,36	0,28	0,22	0,31	-0,45	-0,64	1,00	-0,20	-0,09	-0,16	-0,56	-0,39
	Sig. (2-tailed)	0,33	0,90	0,39	0,91	0,10	0,31	0,24	0,38	0,49	0,32	0,15	0,03		0,52	0,77	0,61	0,06	0,21
<b>MM5</b>	Pearson Correlation	-0,36	-0,06	0,35	0,48	0,38	0,36	0,11	-0,24	-0,24	-0,38	0,17	0,05	-0,20	1,00	0,13	0,42	0,34	0,70
	Sig. (2-tailed)	0,25	0,86	0,27	0,12	0,23	0,25	0,73	0,46	0,46	0,22	0,60	0,89	0,52		0,69	0,18	0,28	0,01

<b>MM6</b>	Pearson Correlation	0,17	0,53	0,24	0,35	0,48	0,25	-0,40	-0,57	-0,13	-0,06	0,42	0,55	-0,09	0,13	1,00	0,11	0,29	-0,40
	Sig. (2-tailed)	0,59	0,08	0,45	0,26	0,11	0,43	0,19	0,05	0,68	0,85	0,18	0,06	0,77	0,69		0,74	0,36	0,20
<b>MM7</b>	Pearson Correlation	-0,24	-0,07	0,72	0,62	0,29	0,71	-0,60	0,28	0,00	0,42	0,82	0,02	-0,16	0,42	0,11	1,00	-0,26	0,49
	Sig. (2-tailed)	0,45	0,83	0,01	0,03	0,36	0,01	0,04	0,37	1,00	0,18	0,00	0,95	0,61	0,18	0,74		0,41	0,11
<b>MM8</b>	Pearson Correlation	0,05	0,07	-0,22	-0,15	0,34	-0,18	0,26	-0,47	-0,23	-0,88	0,02	0,58	-0,56	0,34	0,29	-0,26	1,00	0,06
	Sig. (2-tailed)	0,88	0,83	0,49	0,65	0,28	0,57	0,41	0,13	0,48	0,00	0,95	0,05	0,06	0,28	0,36	0,41		0,86
<b>MM9</b>	Pearson Correlation	-0,65	-0,37	0,30	0,24	0,08	0,32	0,00	0,25	-0,25	-0,15	0,26	-0,27	-0,39	0,70	-0,40	0,49	0,06	1,00
	Sig. (2-tailed)	0,02	0,24	0,34	0,45	0,80	0,31	0,99	0,44	0,43	0,64	0,42	0,41	0,21	0,01	0,20	0,11	0,86	