



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Instituto de Ciencias Matemáticas

“Desarrollo de una aplicación informática que utilice modelos Probit y Logit para la determinación del efecto de productos sobre *Mycosphaerella fijiensis*”.

TESIS DE GRADO

Previo a la obtención del Título de:

INGENIERO EN ESTADÍSTICA INFORMÁTICA

Presentada por:

Mario Gonzalo Lindao Díaz - Sánchez

GUAYAQUIL – ECUADOR

Año: 2010

AGRADECIMIENTO

A Dios, a la Virgen María y al Espíritu Santo por otorgarme día a día su gran sabiduría, a mi padre que desde el cielo me llena de bendiciones, a mi querida madre por su gran paciencia, perseverancia y oraciones, a toda mi familia y a mis amigos(as) que junto a ellos alcanzamos nuestras valiosas metas.

DEDICATORIA

A Dios, a mi madre María
Díaz-Sánchez, a mi padre
Gonzalo Lindao, a mi familia y
a mis grandes amigos por
siempre.

TRIBUNAL DE GRADUACIÓN

MAT. JOHN RAMÍREZ F.
PRESIDENTE DEL TRIBUNAL

ING. OMAR RUÍZ B.
DIRECTOR DE TESIS

ING. GUILLERMO BAQUERIZO P.
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).

MARIO GONZALO LINDAO DÍAZ – SÁNCHEZ

RESUMEN

La Sigatoka negra es la enfermedad más importante que afecta al banano y que es causada por el hongo *Mycosphaerella fijiensis*. La práctica más utilizada para su control es la aplicación de fungicidas químicos que ayudan a la reducción de los niveles de infección, los que se dividen según su forma de acción en Protectantes, Sistémicos y de Contacto.

Cada vez que se requiere conocer los niveles de tolerancia o resistencia alcanzados por el hongo *Mycosphaerella f.* a determinado fungicida, se realiza el análisis de sensibilidad probando diferentes dosis del fungicida sobre el patógeno, y midiendo o evaluando su crecimiento o desarrollo – en cualquiera de sus fases de vida – bajo un ambiente controlado. Cada una de las unidades de observación son sometidas a algún nivel de concentración del producto y comparadas con el crecimiento o desarrollo del testigo absoluto con nivel 0 de concentración, lo que da como resultado una inhibición promedio por cada dosis. Este resultado lleva a la obtención de la dosis letal media (DL_{50}) o concentración efectiva media (CE_{50}), la cual es una forma de expresar el grado de toxicidad de una sustancia.

El presente trabajo plantea aplicar técnicas Logit y Probit a través de la creación y uso de una aplicación informática para el análisis estadístico en los estudios de sensibilidad desarrollados en el CIBE e incrementar la confiabilidad y reducir el tiempo de los análisis, además de centralizar la

información. Por ello, el objetivo principal de la presente tesis, es desarrollar una aplicación informática que utilizando técnicas estadísticas Logit y Probit ayude en el adecuado análisis de la información relacionada con este tipo de estudios, aplicando técnicas basadas en estadística descriptiva e inferencial y minería de datos; la aplicación de estas herramientas de análisis a través de la informática hará posible estimar las relaciones entre la dosis y la inhibición por ella obtenida, y contrastar la hipótesis para comparar estadísticamente el efecto que tuvieron las dosis entre ellas. Los modelos a usarse para este efecto son los de regresión Logit y Probit, con lo cual se determinó la relación entre la variable independiente “dosis” y su efecto sobre la variable dependiente “inhibición”.

El primer capítulo muestra una explicación detallada sobre el hongo *Mycosphaerella f.* causante de la Sigatoka negra sobre los cultivos de banano. Se hace una revisión de la biología del hongo, los ciclos de vida del hongo, se explica también los síntomas de la enfermedad, además de sus causas y efectos en el banano; adicionalmente se explica brevemente sobre los tipos de fungicidas suministrados y su comportamiento sobre el hongo; formas de evaluar la sensibilidad; el impacto económico que esta enfermedad produce. Finalmente se habla de las herramientas informáticas ó software frecuentemente utilizados para el análisis estadístico.

El segundo capítulo detalla el proceso de sensibilidad, la población objetivo y la descripción de las variables; la metodología estadística, tal como la estadística descriptiva, se complementa con el análisis de varianza y los métodos de regresión. Los métodos de regresión en los que se basa nuestro estudio son: (1) el análisis de regresión lineal, (2) el análisis de regresión Logit y (3) el análisis de regresión Probit; finalmente se define minería de datos como metodología importante para la construcción de ciertos modelos estadísticos.

El tercer capítulo detalla el desarrollo de la herramienta informática, en el cual se muestra la infraestructura y el recurso humano necesario, los componentes y módulos construidos, los paquetes utilitarios utilizados para el desarrollo del sistema y las técnicas estadísticas y de minería de datos utilizadas en el mismo. Posteriormente, se muestra el análisis y resultados obtenidos mediante el uso de la herramienta informática, aplicado a hongos obtenidos de la finca “*El Paraíso*”, y de una serie de ensayos en el área de Fitopatología del CIBE con el fungicida sistémico *Azoxistrobina*, ha diferentes niveles de dosis, que para este estudio serán del 0.1%, 1%, 5% y 10% del fungicida.

Finalmente, el capítulo cuatro establece las conclusiones y recomendaciones.

ÍNDICE GENERAL

	Pág.
RESUMEN.....	II
ÍNDICE GENERAL.....	V
ABREVIATURAS.....	VIII
SIMBOLOGÍA.....	IX
ÍNDICE DE FIGURAS.....	X
ÍNDICE DE GRÁFICOS.....	XI
ÍNDICE DE TABLAS.....	XII
INTRODUCCIÓN.....	1
CAPÍTULO I	
1. SÍNTESIS DEL HONGO <i>MYCOSPHAERELLA FIJENSIS</i>	3
1.1. Biología del hongo <i>Mycosphaerella fijiensis</i>	4
1.2. Ciclo de vida del hongo <i>Mycosphaerella fijiensis</i>	6
1.3. Síntomas de la Sigatoka Negra.....	8
1.4. Causas y efectos en el banano.....	10
1.5. Control Químico.....	11
1.5.1. Fungicidas.....	11
1.5.1.1. Fungicidas Protectantes.....	12
1.5.1.2. Fungicidas Sistémicos.....	12
1.5.1.3. Fungicidas de Contacto.....	13
1.6. Comportamiento de <i>Mycosphaerella fijiensis</i> a los fungicidas.....	13

1.7. Impacto económico de la Sigatoka Negra en la producción bananera del Ecuador.....	14
1.8. Formas de evaluar la sensibilidad en <i>Mycosphaerella fijiensis</i>	15
1.9. Herramientas informáticas para análisis estadísticos.....	17

CAPÍTULO II

2. MARCO TEÓRICO CONCEPTUAL DEL PROCESO DE SENSIBILIDAD Y LOS MÉTODOS ESTADÍSTICOS.....	18
2.1. Descripción del proceso de sensibilidad.....	19
2.2. Población objetivo.....	20
2.3. Descripción de variables.....	20
2.4. Estadística descriptiva.....	21
2.5. Análisis de varianza para un solo factor.....	22
2.6. Regresión Lineal Simple.....	26
2.7. Regresión Logit y Probit.....	27
2.7.1. Tipos y codificación de variables de los modelos.	28
2.7.2. Modelo Logit.....	29
2.7.3. Modelo Probit.....	30
2.7.4. Coeficientes de riesgo.....	30
2.7.5. Estimación de coeficientes.....	32
2.7.6. Pruebas de hipótesis.....	32
2.7.6.1. Criterio de Razón de Verosimilitudes.....	33
2.7.6.2. Criterio de Hosmer-Lemeshow.....	33
2.7.6.3. Criterio de Wald.....	34

2.8. Minería de datos.....	34
2.9. Software utilizado para el desarrollo de la herramienta.....	36

CAPÍTULO III

3 DESCRIPCIÓN, MODELACIÓN Y ANÁLISIS DE LOS RESULTADOS MEDIANTE EL USO DE LA APLICACIÓN INFORMÁTICA.....	37
3.1. Herramienta informática desarrollada para el análisis estadístico del proceso de sensibilidad.....	38
3.1.1. Esquema general.....	38
3.1.2. Infraestructura y recursos humanos.....	39
3.1.3. Técnicas estadísticas desarrolladas en la herramienta.....	39
3.1.4. Componentes y módulos de la aplicación informática.....	40
3.2. Resultados y análisis.....	45
3.2.1. Análisis Descriptivo o Univariado.....	45
3.2.2. Análisis de Regresión Lineal Simple.....	50
3.2.3. Análisis de Varianzas para un solo factor (ANOVA).....	52
3.2.4. Análisis de Regresión Logit y Probit.....	54

CAPÍTULO IV

4. CONCLUSIONES Y RECOMENDACIONES.....	63
4.1. Conclusiones.....	63
4.2. Recomendaciones.....	65

ANEXOS

BIBLIOGRAFÍA

ABREVIATURAS

ANOVA	Análisis of Varianza
CIBE	Centro de Investigaciones Biotecnológicas del Ecuador
CL ₅₀	Concentración letal media
Q1	Cuartil 1
Q2	Cuartil 2
Q3	Cuartil 3
DCA	Diseño completamente aleatorio
DL ₅₀	Dosis letal media
m	Metro
mm	Milímetros
µm	Micras
ppm	Partes por millón
SST	Suma cuadrática total
SStr	Suma cuadrática de los tratamientos
SSe	Suma cuadrática del error

SIMBOLOGÍA

$^{\circ}\text{C}$	Grados centígrados
%	Porcentaje
X	Variable independiente
Y	Variable dependiente
N	Número de observaciones
A	Número de tratamientos
Y_i	Observación dependiente i-ésima
X_i	Observación independiente i-ésima
β_i	Coefficientes de regresión
\hat{b}_i	Estimadores de los coeficientes de regresión
H_0	Hipótesis nula
H_1	Hipótesis Alternativa

ÍNDICE DE FIGURAS

	Pág.
Figura 1. 1 Hoja infectada con Sigatoka Negra en su cuarta fase.	10
Figura 2. 1 Diagrama de caja.	22
Figura 3. 1 Interfaz principal de la aplicación.	38
Figura 3. 2 Modelo lógico de la herramienta.	39
Figura 3. 3 Infraestructura.	39
Figura 3. 4 Módulos de la aplicación informática.	41

ÍNDICE DE GRÁFICOS

		Pág.	
Gráfico	3. 1	Análisis comparativo del porcentaje de inhibición obtenido en las diferentes dosificaciones de Azoxistrobina aplicada a <i>Mycosphaerella f.</i>	49
Gráfico	3. 2	Cálculo de la DL ₅₀ a través del modelo de regresión lineal.	52
Gráfico	3. 3	Diferencias estadísticas significativas entre los promedios de las dosis evaluadas.	54
Gráfico	3. 4	Análisis del comportamiento sobre las probabilidades de cada observación evaluadas en el modelo de regresión LOGIT.	60
Gráfico	3. 5	Análisis del comportamiento sobre las probabilidades de cada observación evaluadas en el modelo de regresión PROBIT.	60

ÍNDICE DE TABLAS

	Pág.
TABLA 2. 1 Observaciones de un análisis de varianza para un solo factor.	23
TABLA 2. 2 Contraste de hipótesis de diferencias de medias.	24
TABLA 2. 3. Tabla de análisis de varianza para un solo tratamiento.	25
TABLA 2. 4. Contraste de hipótesis para el criterio de Razón de Verosimilitudes.	33
TABLA 2. 5. Contraste de hipótesis para el criterio de Hosmer-Lemeshow.	33
TABLA 2. 6. Contraste de hipótesis para el criterio de Wald	34
TABLA 3. 1 Estadística descriptiva de la variable Crecimiento del tubo germinativo.	46
TABLA 3. 2 Estadística descriptiva de la variable Inhibición.	47
TABLA 3. 3 Modelo de regresión lineal simple.	51
TABLA 3. 4 Resumen del modelo LOGIT.	55
TABLA 3. 5 Tabla de variables en la ecuación del modelo LOGIT.	56
TABLA 3. 6 Resumen del modelo PROBIT.	57
TABLA 3. 7 Tabla de variables en la ecuación del modelo PROBIT.	58

INTRODUCCIÓN

A través del tiempo los cultivos de banano han sufrido una serie de enfermedades ó ataques de patógenos que afectan gravemente el desarrollo de la planta si no se toman acciones a tiempo. Las enfermedades más comunes que podemos mencionar son el Mal de Panamá, Moko, Cogollo negro, Sigatoka amarilla y Negra, etc. Para nuestro estudio nos basaremos principalmente en la enfermedad de la Sigatoka Negra.

El hongo *Mycosphaerella f.* que es el causante de la enfermedad de la Sigatoka Negra, ha producido grandes pérdidas económicas y de cultivos en los países de Oceanía, América Central y América del Sur, para contrarrestar dicho mal se han combinado prácticas culturales tales como remover las hojas infectadas, mantener el nivel de humedad y el uso de fungicidas, los productos más utilizados son los fungicidas convencionales o químicos.

El *Centro de Investigaciones Biotecnológicas del Ecuador (CIBE)* es el ente encargado de ejecutar acciones correctivas y preventivas en el sector bananero, ha realizado una serie de ensayos utilizando distintos tipos de fungicidas sobre el hongo *Mycosphaerella f.*, con el fin de controlar la infección y minimizar de esta forma la pérdida de cultivos y de su producción.

Mediante el uso de herramientas informáticas y técnicas estadísticas, el CIBE ha monitoreado el ataque de estas enfermedades, lo cual le permite estar un paso antes que el patógeno. Por este motivo, el CIBE a través de este proyecto de tesis se ha propuesto mejorar sus procesos, desarrollando una herramienta informática estadística que le permita el correcto almacenamiento de la información de los ensayos y de los datos que servirán para el estudio de sensibilidad del hongo *Mycosphaerella f.* a determinados productos y sus análisis estadísticos posteriores, la fusión de las técnicas de la informática y la estadística permitirán constituirse en caminos para obtener de una manera más rápida los resultados para encontrar la dosis adecuada del producto a suministrarse en los cultivos.

CAPÍTULO I

1. SÍNTESIS DEL HONGO *MYCOSPHAERELLA FIJIENSIS*.

Introducción

En el presente capítulo se explica sobre el principal problema fitopatológico que tienen los cultivos de banano, el hongo *Mycosphaerella f.* causante de la enfermedad Sigatoka Negra, enfermedad que ataca el follaje de la hoja durante su ciclo de vida provocando un deterioro sobre ella, produciendo una serie de etapas de destrucción sobre la planta, retrasando de esta forma la producción y exportación del banano. Los síntomas de la enfermedad en la hoja de banano, sus causas y efectos en el banano, además se explica sobre el control químico por medio de fungicidas que se utilizan para la minimización del ataque hasta su eliminación, los tipos y cómo estos actúan, Posteriormente, el comportamiento del hongo a los fungicidas, causado por la creación de resistencias, por condiciones climáticas favorables para la enfermedad o un método no adecuado de aplicación de los productos químicos; también se realiza una breve síntesis del impacto económico que ha tenido el ataque del hongo sobre los cultivos de banano y la producción bananera; las formas de evaluar la sensibilidad en el hongo; finalmente se

describen las herramientas informáticas actualmente utilizadas en el análisis estadístico.

1.1. Biología del hongo *Mycosphaerella fijiensis*.

Entre los principales problemas fitopatológicos que tienen los cultivos de banano, es el hongo causante de la enfermedad de la Raya Negra de la Hoja, *Mycosphaerella f.* (anamorfología *Paracercospora fijiensis*), el cual pertenece a la clase heterothallic ascomycetes, sub clase Loculoascomycetidae, orden Donthideales, familia Dophideaceae.¹

En la etapa inicial de la enfermedad, se producen los estomas² y células conidiógenas, que brotan a la superficie inferior de la hoja (envés de la hoja), produciendo las primeras rayas o manchas de color café, durante esta etapa aparecen los primeros conidióforos³, los mismos que continúan produciéndose hasta que se vuelve de color negro el centro de la mancha.

Los conidióforos emergen de los estomas⁴ de la hoja, solos o en pequeños grupos de 2 a 6, dentro de los límites de la raya o mancha, estos están conformados por **conidios**⁵, los conidios (células conidiógenas) se forman individualmente en el ápice del conidióforo, más de 4 conidios maduros

¹ Véase en bibliografía [16].

² Estroma.- Capa gelatinoso o coriáceo que cubre la superficie del medio y que lleva las esporas en conidióforos muy cortos o enclavados en los peritecios o picinidios.

³ Conidióforo.- Hongos rectos o geniculados, con cicatrices en la punta, un poco delgados en el ápice, con una base de 8um. de diámetro, 0-5 septos de 16.5-62.5 um x 4-7 um, ligeramente estrechos y extenso en la punta. De color marrón en el centro, tornándose pálidos en la punta.

⁴ Estoma.- Poros de la hoja localizados en la superficie.

⁵ Conidio.- Tipo de esporas de forma obclavada o cilindro-obclavada, rectos o curvados de color verde pálido u oliváceos, con un numero de septos de 1-10 con un hilo basal distintivo (cicatriz), entre 30-132 um x 2.5-5 um.

forman un solo conidióforo, que al momento de desprenderse deja cicatriz visible en la célula conidiógena y otra en la base del conidio.

Durante la etapa sexual del hongo *Mycosphaerella f.*, se producen un gran número de ascosporas⁶ y conidios, durante la etapa de engrosamiento de la raya esta se transforma en una mancha, dando paso al desarrollo de un gran número de espermogonias⁷ en la superficie inferior de la hoja, las que son asociadas con los conidióforos.

El peritecio⁸ es una estructura que se encuentra inmersa en el tejido de la hoja, y se encuentra en gran abundancia en la parte superior de la hoja, las paredes del peritecio contienen capas en forma de celdas poligonales, con una apertura denominada ostíolo⁹, los que son encontrados en ambas superficies de la hoja, dentro del peritecio se encuentran las ascas o sacos llenas de ascósporas y conidios, las que emergen al romperse el peritecio propagándose localmente por medio del aire e infectando otras hojas.

⁶ Ascosporas.- Son hialinas fusiformes obclavadas, con dos células y ligeramente constrictivas en el septo, estas miden de 12.5-16.5um de largo y 2.5-3.8um de ancho.

⁷ Espermogonia.- Son ovalados y en forma de globo, miden 55-88um de largo y 35-50um de ancho.

⁸ Peritecio.- Estructura en forma de globo y de color café oscuro, con un diámetro de 47-58um.

⁹ Ostíolo.- Es una espemata madura, en forma de vara y mide 2.5-5.0um de largo y 1.0-2.5um de ancho.

1.2. Ciclo de vida del hongo *Mycosphaerella fijiensis*.

El ciclo de vida del patógeno se divide en varias etapas, que se muestran a continuación.¹⁰

Ciclo de Infección, en este ciclo las primeras esporas brotan después de 2 a 3 horas de haberse depositado en la superficie húmeda de la hoja, denominada la etapa de la germinación, durante la germinación se desarrollan los primeros cimientos de conidias y ascosporas, emitiendo un tubo germinativo¹¹ que penetran por el estoma, después de 48-72 horas con una temperatura superior a los 20⁰ C.

El desarrollo de la infección por medio de ascosporas se produce con mayor frecuencia cuando la superficie de la hoja contiene capas de agua, mientras que la infección por conidias se produce por niveles elevados de humedad, en donde las hojas más jóvenes son más propensas a infectarse.

Ciclo de Incubación, durante este ciclo emergen del estoma un gran número de células conidiógenas, las que se desarrollan luego en conidióforos, crecen de manera paralela a las venas de la hoja con una distancia mayor a 3mm, luego infectan el estoma adjunto, estas forman pequeñas ramificaciones sobre el estoma. Este comportamiento ocasionará una incubación en cadena, a manera de red formado las

¹⁰ Véase en bibliografía [16].

¹¹ Tubo germinativo.- Son elementos filamentosos cilíndricos característicos de la mayoría de los hongos.

colonias¹², en donde se tendrá como resultado el desarrollo de la raya negra en la totalidad de la hoja, la cual es denominada también como “la invasión inicial sobre el tejido de la hoja”.

El período de incubación aparece entre la etapa de infección y la aparición de la primera mancha en la hoja, entre 10-14 días en condiciones ideales para el desarrollo de la enfermedad, la cual es producida por el tipo de campo de los cultivos de banano o por las condiciones climáticas.

Ciclo de Evolución de los Síntomas, este ciclo se desarrolla entre la aparición del primer síntoma y la aparición de la primera mancha madura con su centro seco y de color gris, entre los 11 y 139 días, causando la dispersión de los conidios por medio del aire o agua, haciendo más severa la invasión sobre el tejido de las hojas, debido a la susceptibilidad del cultivo y las condiciones climáticas, o sea con altos niveles de humedad, temperatura y el desarrollo de resistencias en el hongo. A esta etapa se la denomina también como Ciclo de Transición.

Desarrollo de la Enfermedad, este ciclo se desarrolla entre el período de infección y la formación de manchas maduras, y es la etapa en donde se puede establecer y medir el desarrollo de la enfermedad.

Durante esta etapa se produce la liberación de una gran cantidad de ascosporas desarrolladas en hojas secas o húmedas, las que dependen de los altos niveles de susceptibilidad de los cultivos, severidad de la

¹² Colonias.- Están constituidos por una fila de células alargadas envueltas por la pared celular.

infección y las condiciones climáticas, a esta etapa se la denomina también como la etapa de reproducción sexual de la enfermedad.

1.3. Síntomas de la Sigatoka Negra.

El primer síntoma que presenta la hoja afectada a causa de la enfermedad de la Sigatoka negra¹³, es la aparición de varias rayas en el envés de la hoja, alargadas y ligeramente anchas, con dimensiones menores a 0.25 mm de largo y de un color marrón-pálido, creciendo de manera paralela a las venas de la hoja y no son muy visibles a la luz, no existen síntomas en la parte superior.

En el segundo síntoma, las rayas miden de 2 a 3 mm de longitud, con un color café en la parte inferior y amarillo en la parte superior de la hoja, para luego cambiar a color café y después a color negro, manteniendo el color café en el envés de la hoja durante esta etapa, las que son más visibles a distancias de 1 a 2 m.

En el tercer síntoma, las dimensiones de la raya aumentan a 20-30 mm de longitud y 2 mm de ancho, presentando un cambio de color de marrón a negro, resultando más visible en el haz de la hoja (parte superior de la hoja).

En condiciones favorables para la enfermedad, en donde los niveles de inoculación¹⁴ son elevados y existe un gran número de rayas muy pegadas,

¹³ Véase en bibliografía [16].

¹⁴ Inoculación.- Transmisión de una enfermedad contagiosa.

estas producen una necrosis¹⁵ en la hoja, causando automáticamente la muerte de la hoja en esta etapa.

En el cuarto síntoma, una sola raya no tiende a formar una necrosis, esta aumenta su grosor formando una mancha de color café en el envés y negro en el haz de la hoja y con borde lagrimoso, la misma que será más visible durante una lluvia o rocío.

En el quinto síntoma, la mancha se extiende en forma elíptica o fusiforme, de color negro en ambas superficies de la hoja, el área central de la mancha se vuelve ligeramente aplanada y hundida, por lo que el contorno de la mancha es más pronunciado y angosto de color amarillo.

En el sexto y último síntoma, el centro de la mancha se seca, se marchitan los tejidos, y obtiene un color gris brillante, cada mancha tiene un borde bien definido de color café oscuro o negro y un cerco de color amarillo brillante, como resultado en esta etapa se presenta un deterioro en los tejidos, produciendo la muerte en la hoja de 3 a 4 semanas y persisten los colores brillantes de la mancha.

¹⁵ Necrosis.- Gangrena de los tejidos.

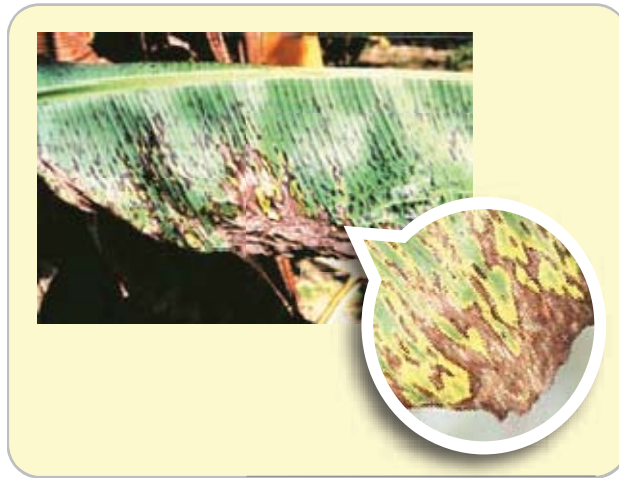


FIGURA 1.1 HOJA INFECTADA CON SIGATOKA NEGRA EN SU CUARTA FASE.

FUENTE: GUILLERMO AROSEMENA – EL COMERCIO (2007)

1.4. Causas y efectos en el banano.

El hongo *Mycosphaerella f.* es el causante principal de la defoliación, muerte de las hojas de banano y gran pérdida de los cultivos, esto se debe a factores climáticos y a la susceptibilidad de los cultivos, que dan como resultado un ataque más severo e incontrolable.¹⁶

Las condiciones climáticas favorables para el hongo serán los niveles de temperatura superiores a 27⁰ C, que contribuyen a la germinación acelerada de conidios y ascosporas, acompañada de altos niveles de humedad, la que es originada durante la época de lluvias en donde se produce la mayor liberación de ascosporas; y de conidios a través del viento, diseminándose hacia las demás hojas, declarándose la etapa óptima para la infección.

La susceptibilidad de los cultivos también beneficia al desarrollo de la enfermedad a una etapa prematura, provocada por el mal mantenimiento que

¹⁶ Véase en bibliografía [16].

se le da a los cultivos, mantenimientos que implican una regulación de la población de plantas, sobre el control de las malezas, limpieza de los canales de riego y drenaje, sobre el control de los desechos y de las inapropiadas prácticas de deshije¹⁷ y deshoje¹⁸.

Estos factores aumentan la severidad del ataque, afectando al crecimiento normal de la planta, esto produce un retraso en la floración del fruto, alterando su proceso normal de maduración, originando una pérdida de peso, vigor y tamaño de los racimos, perdiéndose los cultivos y la producción de banano.

1.5. Control Químico¹⁹.

1.5.1. Fungicidas.

Una de las prácticas más utilizadas para el control del ataque de la Sigatoka negra es la aplicación de fungicidas, que ayudan a la reducción de los niveles de contaminación en las plantaciones de banano. Los fungicidas son aplicados de manera aérea sobre grandes cultivos o plantaciones exportadoras y por medio terrestre con mochilas fumigadoras, comúnmente usados sobre pequeños cultivos para satisfacer los mercados locales.

Los tipos de fungicidas utilizados para combatir el ataque de la enfermedad son los Protectantes, Sistémicos y De Contacto.

¹⁷ Deshije.- Quitar los brotes o retoños de la planta madre a su alrededor.

¹⁸ Deshoje.- Arrancar las hojas de las plantas, quitar las hojas infectadas.

¹⁹ Véase en bibliografía [16].

1.5.1.1. Fungicidas Protectantes.

Los fungicidas protectantes tienen el efecto de permanecer en la superficie de la hoja, sin penetrar en los tejidos, creando una barrera contra las esporas evitando la germinación, tiene un gran alcance de actividad exterminadora y es de mayor duración, el que ejerce un ataque sobre el organismo del patógeno en múltiples partes, por lo general el fungicida es aplicado antes de la etapa de infección. Los de mayor utilidad para el control del hongo son el Ditiocarbamatos y Clorotalonil, que son sustancias solubles en agua o aceite, dependiendo de la sustancia y contra-indicaciones, estos reducirán los niveles de toxicidad del patógeno.

1.5.1.2. Fungicidas Sistémicos.

Los fungicidas sistémicos a diferencia de los protectantes estos penetran el tejido de la hoja, mejorando la protección de los cultivos, debido a que ejerce sobre el patógeno un efecto de inmunidad, a pesar de haber infectado los tejidos, sin embargo este tipo de aplicación no da protección permanente, debido a los cambios estructurales que sufren los cultivos, que afectan la composición del fungicida, inhabilitando su ataque. Existen diferentes tipos de fungicidas sistémicos, el Benzimidazole, Tridemorph, Propiconazole, etc. que son sustancias solubles en aceite y aceite con agua.

1.5.1.3. Fungicidas de Contacto.

Los fungicidas de contacto tienen un efecto parecido a los fungicidas protectantes, este permanece en la superficie de la hoja, sin penetrar en los tejidos, desarrollando una capa o defensa contra las esporas, la diferencia con respecto a los protectantes se encuentra en el modo de ataque del fungicida, el cual no tiene una permanencia continua y gran alcance de actividad exterminadora, ejerciendo un ataque a una sola parte del organismo del patógeno. Los de mayor utilidad para el control del hongo son el mancozeb y propineb, que son sustancias solubles en agua o aceite.

1.6. Comportamiento de *Mycosphaerella fijiensis* a los fungicidas

La eficacia de los químicos utilizados para combatir el hongo *Mycosphaerella f.* se halla en la disminución de los niveles de toxicidad de la enfermedad, hasta su eliminación, pero el uso constante de estos productos químicos, sin la aplicación de otros métodos de defensa sobre los cultivos, resultan innecesarios, debido a ciertas resistencias que el hongo adquiere hacia los químicos, causando un descontrol sobre la producción de conidias y ascosporas, aumentando los niveles de infección y germinación en las plantaciones, además de alargar el tiempo de vida y mejorar su sistema inmunológico.²⁰

²⁰ Véase en bibliografía [16].

Algunos productores o distribuidores de fungicidas químicos indican que limitando el uso de fungicidas de una misma categoría y realizando combinaciones entre los tipos de fungicidas, se evitará que el hongo adquiera suficientes niveles de resistencias, desarrollando un combate más efectivo contra la enfermedad.²¹

1.7. Impacto Económico de la Sigatoka negra en la producción bananera del Ecuador.

Dentro de las enfermedades que más han afectado a la producción bananera ecuatoriana desde los años 50 fue la Sigatoka negra (*Núñez R. 1989*), el cual ha ocasionado pérdidas entre 50% y 100% de las plantaciones y entre 10% y 30% de la cosecha, obteniendo un rendimiento por debajo de los 500 racimos por hectárea al año, afectando más a las pequeñas plantaciones de agricultores. Esto hace más complejo el control de la enfermedad, realizando grandes cambios en los métodos de manejo de los cultivos, junto con los programas de riego y el aumento de los ciclos de fumigación aérea y terrestre, de 17 a 24 aplicaciones por año de los diferentes tipos de fungicidas, control químico que representa el 45% de los costos de producción, además de afectar la salud de los agricultores.

²¹ Véase en bibliografía [19].

1.8. Formas de evaluar la sensibilidad en *Mycosphaerella fijiensis*.

Uno de los principales métodos para combatir el hongo *Mycosphaerella f.* han sido los fungicidas, que por muchos años han protegido a los cultivos de diferentes especies, a través del tiempo el uso repetitivo o el uso de otros fungicidas químicamente relacionados dan la oportunidad al patógeno a crear ciertas resistencias, que es la capacidad de sobrevivir en la naturaleza, estas resistencias de no ser reconocidas a una etapa temprana conllevaría al fracaso del control de la enfermedad y consecuentemente pérdidas de las cosechas.

Desde el año 1969 se detectó el primer caso de resistencias en ciertos cultivos, de 1976 a 1980 en los cultivos de mango, lo cual conlleva a proponer maneras de evaluar los diferentes tipos de resistencias, donde DuPont en 1980 evaluó la resistencia cuando el hongo *Mycosphaerella f.* sobrevive a cierto nivel de dosificación, indicando que a un 5% de germinación de ascosporas en 10 ppm de benomil es suficiente para detener el uso del producto.²²

En 1983 se evaluó la resistencia al medir el tubo germinativo después que el hongo ha infectado las hojas con *Mycosphaerella f.*, a diferentes concentraciones de fungicidas. Otra forma de evaluación es medir el diámetro de la colonia de igual manera a diferentes dosis de fungicidas

²² Véase en bibliografía [14].

determinado por *Romero en 1995*, ambos métodos son fáciles de realizar y se obtienen resultados relativamente rápidos.²³

Con el fin de evitar problemas fitosanitarios y un aporte ecológico en 1994 se planteó utilizar extractos vegetales con efectos fungicidas, por ejemplo ajo, cola de caballo, ortiga, manzanilla, lo cual estas segregan sustancias tóxicas por sus raíces y hojas causando un efecto de inhibición.²⁴

En el 2001 la evaluación de nuevos productos de fungicidas con pocos efectos nocivos al ambiente y salud humana son prioritarios para la búsqueda de nuevas alternativas de manejo de Sigatoka negra, dentro de esta nueva tendencia se encuentra la Azoxistrobina, fungicida utilizado en las muestras para estudio del presente proyecto.

El caso del hongo *Mycosphaerella f.* es el más complejo debido a su tipo de reproducción que puede ser sexual y asexual, lo cual conlleva a cambiar la metodología de evaluación de sus resistencias, el uso de tecnologías a disminuido los factores de riesgo, pero en el campo todavía existe la inadecuada fertilización, las prácticas de sanidad mal realizadas que incrementan el desarrollo de la enfermedad. Para esto se han tomado medidas en la cual se recolectan un número de hojas infectadas, en las que se mide el crecimiento del hongo según el tubo germinativo, para luego procesar dicha información, mediante el uso de técnicas estadísticas e informáticas.

²³ Véase en bibliografía [22].

²⁴ Véase en bibliografía [29].

1.9. Herramientas informáticas para análisis estadísticos.

Las herramientas informáticas ó software utilizados para el análisis estadístico se han aplicado a una serie de situaciones que van desde el almacenamiento de los resultados hasta el análisis y toma de decisiones en las empresas.

Hoy en día existe una gran variedad de herramientas en el mercado, que ayudan al estudiante, profesor y a las empresas a resolver incógnitas o encontrar características en una población. Podemos mencionar algunas de las más utilizadas: SPSS, MATLAB, PROJECT-R entre otras, que ayudan a resolver incógnitas y encontrar características en una población. Sin embargo no existe una herramienta informática que ayude con el almacenamiento del histórico de los ensayos, requerimientos del cliente y análisis estadísticos que ofrezca información relevante o de interés específico para el investigador.

Los programas antes mencionados, tienen la capacidad de procesar una gran cantidad de información en poco tiempo, además cuenta con una diversidad de opciones en el campo de la estadística y cálculos matemáticos. Por esta diversidad de opciones resulta altamente costosa su adquisición por asuntos de licencias y la interpretación de los resultados requieren de un investigador especializado en el área, por eso las empresas se están dedicando a desarrollar sus propias herramientas con la misma capacidad y confiabilidad.

CAPÍTULO II

2. MARCO TEÓRICO CONCEPTUAL DEL PROCESO DE SENSIBILIDAD Y LOS MÉTODOS ESTADÍSTICOS

Introducción

En este capítulo se indica la descripción del proceso de sensibilidad para la determinación del efecto de los productos químicos sobre *Mycosphaerella f.*, en el cual se explica la selección y recolección de los datos, procesamiento en laboratorio y resultados; la población objetivo y la descripción de las variables a utilizar.

En el marco teórico se indica la metodología estadística necesaria para realizar el presente estudio, donde se detallan los objetivos, características e interpretación de cada método utilizado.

Se tratará como punto inicial la estadística descriptiva, continuando con la estadística inferencial para ello se realizan pruebas de hipótesis, uso de modelos lineales como la regresión simple y finalmente el uso de modelos no lineales de clasificación la regresión Logit y la regresión Probit.

Es importante definir el concepto de minería de datos que es parte de la metodología utilizada para el desarrollo de ciertos modelos estadísticos

usados; y finalmente, qué tipos de software fueron utilizados durante la construcción de la aplicación.

2.1. Descripción del proceso de sensibilidad.

El proceso de sensibilidad proporcionado por el CIBE, comienza con las entrevistas sobre los requerimientos de necesidad del cliente, donde se proveen los productos a utilizar y dosis a suministrar; para el estudio se seleccionó el fungicida sistémico Azoxistrobina (Bankit).

La siguiente etapa del proceso es la selección y recolección de manera dirigida de hojas afectadas con Sigatoka negra; luego en laboratorio las muestras pasan al proceso de incubación y colocadas en cajas petri con medios de cultivo por 2 horas; después de 24 horas bajo el lente del microscopio se mide el tamaño de los tubos germinativos para evaluar su crecimiento, es necesario mencionar que en este proceso también se mide el tamaño de la colonia de *Mycosphaerella f.* como parte del método utilizado y que es de igual importancia en los análisis que se realizan en el laboratorio.

Los datos obtenidos son almacenados en hojas electrónicas, previo al análisis estadístico, y se procede a realizar el cálculo de la Dosis Letal Media mediante técnicas estadísticas, que se ampliará la explicación de dichas técnicas mediante el uso de la herramienta informática más adelante.

Posteriormente se prepara el informe, el cual es revisado y entregado al cliente para su satisfacción, y finalmente se respaldan los resultados en formato digital.¹

2.2. Población Objetivo.

Se seleccionó y recopiló un número de hojas infectadas según la escala Stover modificada con grado de infección 4 y 5 (*Gauhl, 1989*), dichas hojas fueron recolectadas de manera dirigida de la finca “*El Paraíso*”, Santa Ana – Sto. Domingo de los Colorados en el 2005. El número de observaciones es de 150 unidades, las cuales fueron obtenidas en los laboratorios del CIBE, esto sucede durante el proceso de sensibilidad descrito en el literal 2.1.

Es necesario mencionar que el análisis de sensibilidad dependerá generalmente de los requerimientos del cliente, en cuanto a producto, dosis y muestras.

2.3. Descripción de las variables.

En los ensayos se utilizan tanto productos orgánicos como productos químicos ó convencionales, cabe recalcar que se utilizan dos tipos de medida para las dosis suministradas, tal como el porcentaje (%) del producto utilizado, por ejemplo: dosis al 0.5%, 1%, 10%; y partes por millón (ppm) del

¹ Véase en bibliografía [24].

producto, por ejemplo 0.001 ppm, 0.00001 ppm; con esto se obtendrán los diferentes tratamientos ó concentraciones para el análisis.

Las variables de estudio serán el crecimiento del tubo germinativo y el diámetro de la colonia de *Mycosphaerella f.*, las que tienen como escala de medición micras y milímetros respectivamente; estas variables de origen o iniciales son necesarias para determinar la variable principal para el análisis y es en la que se fundamenta nuestro estudio, la variable porcentaje de inhibición.

El porcentaje de inhibición es obtenido a través de la siguiente fórmula:

$$\text{Porcentaje de Inhibición} = \left(\frac{T' - Z}{T'} \right) 100\%$$

Donde T' es el promedio de la medida de control ó dosis 0; Z es igual a la medida de la unidad de observaciones a la que se le suministró cierta dosis del producto químico en estudio.²

2.4. Estadística Descriptiva

El análisis descriptivo de los datos, tiene como objetivo obtener una visión general del comportamiento de las variables, para ello se obtienen medidas de tendencia central: media y moda; medidas de dispersión: varianza, desviación estándar, rango y cuartiles; distribución de los datos: sesgos y curtosis, que indica que tan distantes se encuentran los datos respecto al promedio; y percentiles que es la división de los datos ordenados y divididos

² Véase en bibliografía [24].

en cien partes, los que pueden ser representados gráficamente mediante la gráfica denominada Ojiva y su presentación más común es el diagrama de cajas, donde se observan los percentiles (P_{25} , P_{50} , P_{75}).

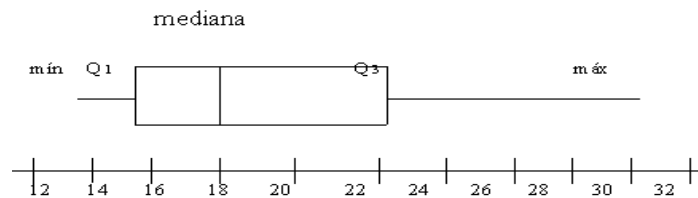


FIGURA 2.1 DIAGRAMA DE CAJA.

FUENTE: N.M. DOWNIE Y R.W. HEATH (1986)

El análisis descriptivo se utilizó para conocer la distribución de los datos, la existencia de valores atípicos, sesgos, que pudieran dificultar el análisis inferencial y si es necesario realizar alguna transformación matemática.³

2.5. Análisis de varianza para un solo factor.

El análisis de varianza es una de las técnicas usada en el diseño experimental, la cual ayuda a probar si existen diferencias significativas de al menos uno de los promedios de la variable de estudio, obtenido de los tratamientos de diferentes poblaciones.

Diseño experimental es una técnica estadística, donde la primera fase es planificar un ensayo de tal manera que este provea de la información pertinente sobre el problema investigado. Existen varios tipos de diseños tales como el completamente al azar, de bloques completos al azar,

³ Véase en bibliografía [8].

experimentos factoriales, entre otros, los cuales son utilizados de acuerdo a los objetivos de cada experimento.

Para el propósito de la investigación, el análisis y resultados estarán fundamentados en un diseño completamente aleatorio DCA, el mismo que se examinará a través del análisis de varianza (ANOVA) de un solo factor, donde sus diferentes niveles son las dosis probadas, y se tratará de explicar la variable cuantitativa a través de una variable cualitativa.

La Tabla 2.1 muestra la matriz de observaciones, donde se tienen a tratamientos diferentes de un misma población que quieren compararse, las n observaciones tomadas del experimento en cada tratamiento, constituyendo una matriz de datos de axn , donde y_{ij} es la observación j -ésima tomada del i -ésimo tratamiento.⁴

Tabla 2.1				
Observaciones de un análisis de varianza de un solo factor				
Tratamiento (nivel)	Observaciones			
1	y_{11}	y_{12}	...	y_{1n}
2	y_{21}	y_{22}	...	y_{2n}
⋮	⋮	⋮	...	⋮
a	y_{a1}	y_{a2}	...	Y_{an}

FUENTE: MONTGOMERY D. (2003)

En este contexto se describirá mejor las observaciones del experimento mediante el modelo siguiente

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases} \quad (2.1)$$

⁴ Véase en bibliografía [18].

Donde μ es la media total de experimento, τ_i es el efecto del i -ésimo tratamiento y ε_{ij} es el error aleatorio por cada observación. Siendo el supuesto para los errores, estos deben seguir una distribución normal con media igual a cero y varianza constante y la suma de los efectos por tratamiento debe ser igual a cero $\sum_{i=0}^a \tau_i = 0$

Se define la hipótesis en el Tabla 2.2, para probar la igualdad de las a medias de los tratamientos, donde el contraste de hipótesis es:

Tabla 2.2 Contraste de Hipótesis de diferencias de medias
$H_0: \mu_1 = \mu_2 = \dots = \mu_a$ Vs. $H_1: \mu_i \neq \mu_j$ para al menos un par (i, j)

FUENTE: MONTGOMERY D. (2003)

Se procede a la partición de la variabilidad total, que ayudará al análisis correcto de los datos en cada tratamiento, la Tabla 2.3 muestra los componentes de la varianza total denotada por SST suma cuadrática total, con $N-1$ grados de libertad, siendo esta la misma medida de variabilidad global de los datos; la suma cuadrática de los tratamientos (SS_{tr}), que mide la variabilidad que existe entre los tratamientos, con $a-1$ grados de libertad; y la suma cuadrática del error (SS_e), que es una medida de variabilidad dentro de los tratamientos, que se atribuye al error aleatorio, con $N-a$ grados de libertad; formando la ecuación descrita a continuación

$$SS_T = SS_{tr} + SS_e \quad (2.2).$$

Tabla 2.3				
Tabla de análisis de varianza para un solo tratamiento				
Fuente de variación	Suma de los Cuadrados	Grados de libertad	Medias de los Cuadrados	F0
Entre los tratamientos	$SS_{tr} = n \sum_{i=1}^a (\bar{y}_i - \bar{y}_{..})^2$	a-1	$MS_{tr} = \frac{SS_{tr}}{a-1}$	$F_0 = \frac{MS_{tr}}{MS_e}$
Dentro de los tratamientos (Error)	$SS_e = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2$	N-a	$MS_e = \frac{SS_e}{N-a}$	
Total	$SS_T = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$	N-1		

$$*y_i = \sum_{j=1}^n y_{ij} \quad ; \bar{y}_i = y_i/n \quad , i=1,2,\dots,a$$

$$y_{..} = \sum_{i=1}^a \sum_{j=1}^n y_{ij} \quad ; \bar{y}_{..} = y_{..}/N$$

FUENTE: MONTGOMERY D. (2003)

Bajo los supuestos de normalización e independencia de los errores ($\varepsilon_{ai} \sim N(0, \sigma^2)$) y las observaciones ($y_{ai} \sim N(\mu + \tau_a, \sigma^2)$), se usa el Teorema de Cochran en el Anexo 3, que establece que la Suma de los cuadrados de los tratamientos y de los errores son variables aleatorias Ji-cuadrada con una distribución independiente, por lo tanto se postula el estadístico de prueba $F_0 = \frac{MS_{tr}}{MS_e}$ sobre la hipótesis propuesta en Cuadro 2.1, que se distribuye con F , con $a-1$ y $N-a$ grados de libertad, donde con $(1-\alpha)\%$ de confianza se rechaza la hipótesis nula a favor de la alternativa, si $F_0 > F_{\alpha, a-1, N-a}$ o tomando el *P-Valor* como criterio de decisión. Al rechazar la hipótesis nula, se asegura que existe evidencia estadística para afirmar que al menos uno de los promedios es diferente y que por lo tanto existe un efecto del factor, diferente de cero, en la variable de estudio.

2.6. Regresión Lineal Simple.

Uno de los modelos de regresión a desarrollar es el de regresión lineal simple, el cual tiene como objetivo determinar un modelo lineal que se ajuste a los datos de este estudio⁵, el método busca la relación funcional entre dos variables, una variable X llamada independiente ó explicativa y otra variable Y llamada dependiente o de respuesta, el cual se representa por la siguiente ecuación de predicción:

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (2.3)$$

β_0 es el valor donde la línea de regresión se intercepta con el eje Y , β_1 es el coeficiente de regresión ó la pendiente de la línea recta. Para estimar el modelo y los valores de la variable dependiente, se obtienen los estimadores de los parámetros β_i ; y al lado derecho de la ecuación 2.3 se le debe agregar ϵ que es el error aleatorio asociado a la medición de X .

Se supone que el error aleatorio tiene una distribución normal con media cero y varianza constante y la relación entre las variables es lineal.

Mediante el uso del método de los mínimos cuadrados, se procede a calcular los b_0 y b_1 que son los estimadores de β_0 y β_1 respectivamente, partiendo de una muestra de n pares de observaciones de la variable independiente y dependiente tales como $(x_1; y_1), (x_2; y_2), \dots, (x_n; y_n)$, para minimizar el error se calcula la suma cuadrada de los errores mediante la ecuación siguiente:

$$SCE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n [y_i - (\hat{b}_0 + \hat{b}_1 x_i)]^2 \quad (2.4)$$

⁵ Véase en bibliografía [18].

Luego se deriva con respecto a los coeficientes de regresión e igualándolos a cero se obtienen las ecuaciones finales para encontrar dichos coeficientes.

$$b_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.5)$$

$$b_0 = \bar{y} - b_1 \bar{x} \quad (2.6)$$

Donde $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ y $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ que son los promedios de las variables independientes y dependientes. Para el presente caso de estudio, la variable independiente **X** es el logaritmo natural de la dosis (Ln(X)) y la dependiente **Y** es el porcentaje de inhibición, una vez estimados estos valores se los reemplaza en la ecuación de predicción expuesta en (2.3) la que nos servirá como modelo en búsqueda de la Dosis Letal, para el cálculo de la DL₅₀ se iguala dicha ecuación para **Y** = 50, y despejando la variable **X** se obtendrá el valor de la dosis letal media.

2.7. Regresión LOGIT y PROBIT.

Es un análisis mediante el uso de modelos no lineales en la investigación, que tiene como objetivo estimar las relaciones de una o más variables respuestas en términos de otras a través de una función no lineal que haga posible un mejor ajuste de los datos de estudio y que haga posible predecir o explicar las decisiones e influencias tomadas sobre un conjunto de variables. Los modelos de regresión a usarse para efectos de la investigación son Logit y Probit, con los cuales se determinará si tienen o no las variables

independientes X_1, X_2, \dots, X_n , un efecto sobre la variable dependiente, descrita por una respuesta a dos posibles eventos Y_1 y Y_2 , formando la relación funcional entre Y y X ; tal como $Y = f(X_1, X_2, \dots, X_n)$.⁶

2.7.1. Tipos y codificación de Variables de los modelos.

Se definen los tipos de variables que tomarán los modelos no lineales mencionados:

Variable Dependiente: es aquella variable binaria o dicotómica que solo puede tomar dos posibles valores (Éxito o fracaso, ej: Si-No, 0-1, Verdadero-Falso). Este tipo de variable es un subconjunto de las llamadas variables categóricas o cualitativas.

Variables Independientes: pueden ser de los tipos a continuación:

Binaria es aquella que solo puede tomar dos posibles valores, 1 si ocurre y 0 caso contrario.

Variable Categórica Nominal: aquella variable que puede tomar varios valores que representan las categorías sin tener algún orden específico (1=Hombre, 2=Mujer).

Variable Categórica Ordinal: aquella variable que puede tomar varios valores que representan las categorías con un orden específico (1=Bajo, 2=Medio, 3=Alto).

Dummy: Si una variable categórica nominal u ordinal que puede tomar más de dos valores, y se la puede codificar en varias variables

⁶ Véase en bibliografía [4], [28] y [31].

binarias o dicotómicas, esta puede ser llamada variable indicadora dummy.

Cuantitativa: aquella variable que puede tomar un rango numérico de valores $(-\infty, +\infty)$, los cuales puede ser discreta (numerable) o continua ($x \in \mathbb{R}$).

2.7.2. Modelo Logit

Se procede a construir el modelo de regresión logística donde se tiene una variable dependiente dicotómica o predictora de manera

$$y_i = \begin{cases} 1 & \text{Favorece a la ocurrencia del evento.} \\ 0 & \text{Caso contrario.} \end{cases},$$

y un conjunto de variables independientes X_1, X_2, \dots, X_n . Se plantea el siguiente modelo

$$y = b_0 + b_1X_1 + \dots + b_nX_n \quad (2.7)$$

Se estiman los coeficientes b_i del modelo por el procedimiento de mínimos cuadrados, como un modelo de regresión lineal, sin embargo esto llevaría a resultados ilógicos, debido a que los resultados estarán fuera del rango $[0-1]$, causando una contradicción con la variable dependiente planteada.⁷

Se procede a construir el modelo logístico que establece una relación entre la probabilidad de que ocurra un evento, dado un conjunto de variables explicativas, de la siguiente forma

$$\log \left(\frac{P(Y|X_1, \dots, X_n)}{1-P(Y|X_1, \dots, X_n)} \right) = b_0 + b_1X_1 + \dots + b_nX_n \quad (2.8)$$

⁷ Véase en bibliografía [28] y [31]

Ahora se tiene un modelo que puede tener cualquier valor, que mediante transformaciones logarítmicas se obtiene la ecuación equivalente a

$$P(Y|X_1, \dots, X_n) = \frac{e^{(b_0, b_1 X_1, \dots, b_n X_n)}}{1 + e^{(b_0, b_1 X_1, \dots, b_n X_n)}} \quad (2.9)$$

2.7.3. Modelo Probit

El modelo de regresión Probit es otro método para explicar la variable dependiente dicotómica, a través de las variables independientes.⁸

Tomando de la ecuación 2.8 se explica el modelo usando la distribución normal estándar, que se describe a continuación

$$P(Y|X_1, \dots, X_n) = \int_{-\infty}^{b_0 + b_1 X_1 + \dots + b_n X_n} \frac{1}{\sqrt{2\pi}} e^{(-\frac{t^2}{2})} dt \quad (2.10)$$

Donde los coeficientes b_i reflejan el impacto que tiene la variable independiente sobre la probabilidad, coeficientes que son estimados para ambos modelos mediante el método de Máxima Verosimilitud, que se explica con detalle en el literal 2.7.6.1.

2.7.4. Coeficientes de riesgo

Riesgo (probabilidad): se define como una medida de probabilidad de un evento, mediante el cociente del número de veces (casos) que ocurre un evento frente al número total de casos

⁸ Véase en bibliografía [28] y [31]

Odds (oportunidad): se define como una medida de ocurrencia de un evento, mediante el cociente del número de casos que ocurre un evento frente al número de casos que no ocurre, que es una manera de decir lo mismo que con probabilidades.

$$odds = \frac{p}{1-p} \quad (2.11)$$

Siendo p la probabilidad del suceso, mientras más probable se produzca el evento frente a que no se produzca, dicho cociente fue el principio para construir los modelos planteados.

Odds Ratio (oportunidad relativa): se define como el cociente del Odds correspondiente al evento bajo cierta condición entre el Odds del mismo evento bajo otra condición.

$$odds\ ratio = \frac{odds2}{odds1} \quad (2.12)$$

Por ejemplo, la probabilidad de aparición de hipertensión en una persona cuando se consume mucha sal (odds1) respecto a la probabilidad del mismo evento cuando consume poca sal (odds2).

Los Odds ratio resultan de gran utilidad para evaluar la influencia que cada variable independiente tiene sobre la variable de respuesta, donde un Odds ratio igual a uno indica igual probabilidad bajo un evento con diferentes condiciones, mayor a uno indica un aumento en la probabilidad del evento y menor a uno implica disminución sobre dicha probabilidad.⁹

⁹ Véase en bibliografía [28] y [31]

2.7.5. Estimación de coeficientes

Se procede a estimar los coeficientes del modelo, denotados por el vector $b = (b_0, b_1, \dots, b_n)$, partiendo de una muestra aleatoria de tamaño n dada por $\{(x_i, y_i); i = 1, 2, \dots, n\}$, donde $x_i = (x_{i1}, x_{i2}, \dots, x_{ik})$ es el valor de las variables independientes e $y_i \in \{0, 1\}$ es el valor observado de y en el i -ésimo elemento de la muestra. Utilizando el hecho de que $y|(x_1, x_2, \dots, x_n)$ tiene una distribución binomial $(1, p(y|x_1, x_2, \dots, x_n; \hat{b}))$, ya que la variable dependiente toma dos posibles valores éxito y fracaso, cuando el numero de éxitos en n repeticiones tiene un distribución binomial $B(n, p)$. Entonces la función de verosimilitud vendrá dada por

$$L(\hat{b} | (x_1, y_1), \dots, (x_n, y_n)) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} \quad (2.13)$$

Donde $p_i = p(x_i; \hat{b}) = p(x_1, x_2, \dots, x_n; \hat{b}); i = 1, 2, \dots, n$

2.7.6. Pruebas de Hipótesis.

Las pruebas a continuación implican formular y probar una hipótesis estadística para determinar si las variables independientes del modelo se encuentran relacionadas con la variable dependiente, mediante una comparación de los valores observados respecto a los valores que se han predicho por los modelos utilizados.¹⁰

¹⁰ Véase en bibliografía [28] y [31]

2.7.6.1. Criterio de Razón de Verosimilitudes.

Se plantea el siguiente contraste de hipótesis en la Tabla 2.4 para determinar si las variables independientes influyen o aportan significativamente al modelo encontrado y si explican a la variable dependiente, bajo el criterio de *Razón de Verosimilitudes*

Tabla 2.4 Contraste de Hipótesis para el criterio de Razón de Verosimilitudes
$H_0: \beta_0 = \beta_1 = \dots = \beta_n = 0$
<p style="text-align: center;">Vs.</p> $H_1: \text{para algún } \beta_i \neq 0$
Estadístico de Prueba es : $D = -2\text{Log} \left[\frac{L(\beta_1)}{L(\beta_2)} \right]$

FUENTE: MONTGOMERY D. (2003)

Donde b_i son los estimadores máximo verosímiles de los vectores de parámetros del modelo observado y el predictor.

El criterio de decisión será si $D > X_{\alpha,k}^2$; α, k grados de libertad, se concluye que al menos una variable aporta significativamente al modelo.

2.7.6.2. Criterio de Hosmer-Lemeshow.

Se procede a evaluar el modelo seleccionado, mediante el criterio de *Hosmer-Lemeshow*, que indica si el modelo obtenido se ajusta a los valores observados, se muestra el contraste de hipótesis en la Tabla 2.5.

Tabla 2.5 Contraste de Hipótesis para el criterio de Hosmer-Lemeshow.
$H_0: \beta_0 = \beta_1 = \dots = \beta_n = 0$
<p style="text-align: center;">Vs.</p> $H_1: \text{para algún } \beta_i \neq 0$
Estadístico de Prueba es : $X^2 = \sum_{i=1}^n \frac{(y_i - p_i)^2}{p_i(y_i - p_i)}$

FUENTE: MONTGOMERY D. (2003)

El criterio de decisión será si $X^2 > X_{\alpha,k}^2$, no se rechaza la hipótesis nula y se concluye que el modelo se ajusta a los datos observados.

2.7.6.3. Criterio de Wald.

Una vez detectada la existencia de un conjunto de variables explicativas que influyen en el modelo, se procede a identificar cuáles son estas variables, mediante *La Prueba de Wald*, se muestra el contraste de hipótesis en la Tabla 2.6.

Tabla 2.6 Contraste de Hipótesis para el criterio de Wald
$H_0: \beta_i = 0$ Vs. $H_1: \text{para algún } \beta_i \neq 0$ Estadístico de Prueba es : $W = \frac{(\beta_i)^2}{S_{\beta_i}^2}$

FUENTE: MONTGOMERY D. (2003)

El criterio de decisión será si $W > X_{\alpha,k}^2$ entonces se rechaza la hipótesis nula a favor de la alterna, concluyendo que la variable independiente correspondiente al coeficiente evaluado influye a la probabilidad del evento.

2.8. Minería de datos

Una de las tendencias a nivel de pequeñas, medianas y grandes empresas es diseñar y construir sus propias aplicaciones o sistemas informáticos para realizar actividades específicas, que ayudarán al correcto almacenamiento de la información, mejora de los procesos, optimización de los tiempos y confiabilidad de los resultados.

La minería de datos es un término utilizado para la exploración y análisis de los datos con el fin de encontrar patrones y reglas que expliquen el comportamiento de los datos. La minería de datos se encarga principalmente en la construcción de modelos, estos modelos no son más que algoritmos en los que se especifican entradas, procesos y salidas, que servirán para descubrir el conocimiento de los datos.

Muchas de las tareas a cargo de la minería de datos son la clasificación, estimación, predicción, agrupación, reglas de asociación y perfiles, los que son desarrollados en el campo de la ingeniería, economía, biología, modelos de negocios, etc. como por ejemplo el desarrollo de un modelo predictivo del brote de la Sigatoka Negra para las plantaciones de plátano al sur del lago de Maracaibo.¹¹

En fin, se pretende aclarar ciertos conceptos que se encuentran inmersos y hasta de manera oculta pero que son muy importantes para el desarrollo de la herramienta informática, la misma que permitirá acceder a la información que se encuentra almacenada en una base de datos, esta información será procesada por medio de algoritmos que ejecutan técnicas de aprendizaje sobre el 25% de los datos; y luego los algoritmos de regresión Logit y Probit, para la búsqueda de la dosis de fungicida que se aplicará a los cultivos.

¹¹ Véase en bibliografía [7]

2.9. Software utilizado para el desarrollo de la herramienta.

Las herramientas informáticas ó software requeridos para la ejecución de la aplicación fue Excel 2003 como la interfaz principal para el usuario y PostgreSQL 8.2 que es la base de datos para el almacenamiento de la información; cabe recalcar que la base de datos utilizada está en la categoría de software libre, esto significa que no incurre en algún costo por compra o licencias anuales, además es un software completo que cuenta con herramientas y funciones ideales para el desarrollo de este proyecto.

CAPÍTULO III

3. DESCRIPCIÓN, MODELACIÓN Y ANÁLISIS DE LOS RESULTADOS MEDIANTE EL USO DE LA APLICACIÓN INFORMÁTICA.

Introducción

En el presente capítulo se presentan los resultados obtenidos a través de la herramienta creada. El proceso descrito en el capítulo anterior se realiza en dos fases, la manual, que es la recolección, procesamiento y análisis de las muestras en el laboratorio de fitopatología; y la parte automática que es el procesamiento y análisis estadístico de los datos mediante el uso de la aplicación desarrollada, donde se da una explicación acerca de cómo se codifica y almacena la información recolectada, además de la incorporación de las técnicas estadísticas para el procesamiento y análisis de los datos, para esto se utilizarán las técnicas Logit y Probit, por lo que al finalizar el presente proyecto el CIBE contará con esta poderosa herramienta para el cálculo de la dosis ideal.

3.1. Herramienta informática desarrollada para el análisis estadístico del proceso de sensibilidad.

Durante la etapa del análisis estadístico de los datos, se preparan las técnicas utilizadas para el desarrollo de la aplicación informática, en las cuales se trataron técnicas de diferentes áreas del conocimiento en informática, tales como los sistemas de información, minería de datos, bases de datos, programación orientada a objetos y macros.



FIGURA 3.1 INTERFAZ PRINCIPAL DE LA APLICACIÓN
FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

3.1.1. Esquema general

La herramienta desarrollada se mantiene de acuerdo a un esquema basado en sistemas de información, que es un conjunto de pasos que interactúan entre sí, el que realiza las actividades de entradas, almacenamiento, procesamiento y salidas, descritos en el modelo lógico Figura 3.2.



FIGURA 3.2 MODELO LÓGICO DE LA HERRAMIENTA.

FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

3.1.2. Infraestructura y Recursos humanos.

Los requerimientos de la infraestructura para la implementación de la aplicación Figura 3.3 son las de un computador y un servidor, manejado por el personal autorizado tales como un auxiliar y asistente de investigación que son el recurso humano que interactúa con la aplicación.

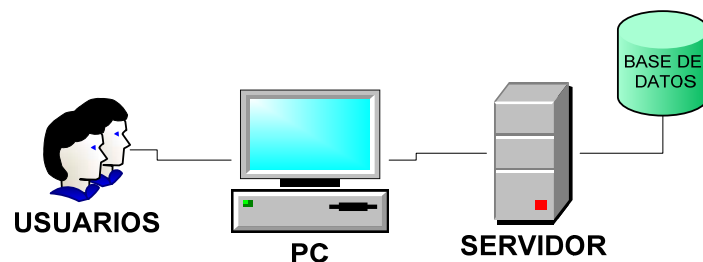


FIGURA 3.3 INFRAESTRUCTURA.

FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

3.1.3. Técnicas estadísticas desarrolladas en la herramienta.

Una de las etapas más importantes de la creación de la herramienta informática es el desarrollo de las técnicas estadísticas para el análisis y selección de la dosis letal media.

Los modelos desarrollados para dichas actividades son:

- ❖ Análisis Descriptivo
- ❖ Regresión Lineal
- ❖ Análisis de Varianza
- ❖ Regresión Logit y Probit

Para el desarrollo de las técnicas estadísticas se tomó como referencia el uso de conceptos estadísticos básicos, uso de algoritmos matemáticos y de la minería de datos, los que ayudaron con el entrenamiento de los datos, en búsqueda del modelo indicado, esto significa que los datos fueron sometidos a evaluación, con la única finalidad de encontrar patrones iniciales, los que sirvieron para describir el modelo final a utilizar; finalmente se muestran los resultados de los análisis.

Se muestra en los Anexos 1 y 2 el diagrama lógico de los algoritmos utilizados para regresión Logit y Probit respectivamente.

3.1.4. Componentes y módulos de la aplicación informática.

Para el mejor desarrollo de la aplicación informática se han dividido en 3 componentes que van de acuerdo a sus funciones, las que son el Componente Transaccional que se encarga de la administración correcta de la aplicación; el Analítico que comprende el análisis de los datos y emisión de reportes de los resultados de la aplicación; y el de Seguridad que comprende el acceso seguro de los usuarios a la aplicación.

Dentro de cada componente existen módulos que son aplicaciones pequeñas que realizan cierto trabajo específico, en la Figura 3.4 se muestra la interacción de los módulos desarrollados en la herramienta, en el manual de usuario Anexo 7 se explicará con detalle cada uno de estos.

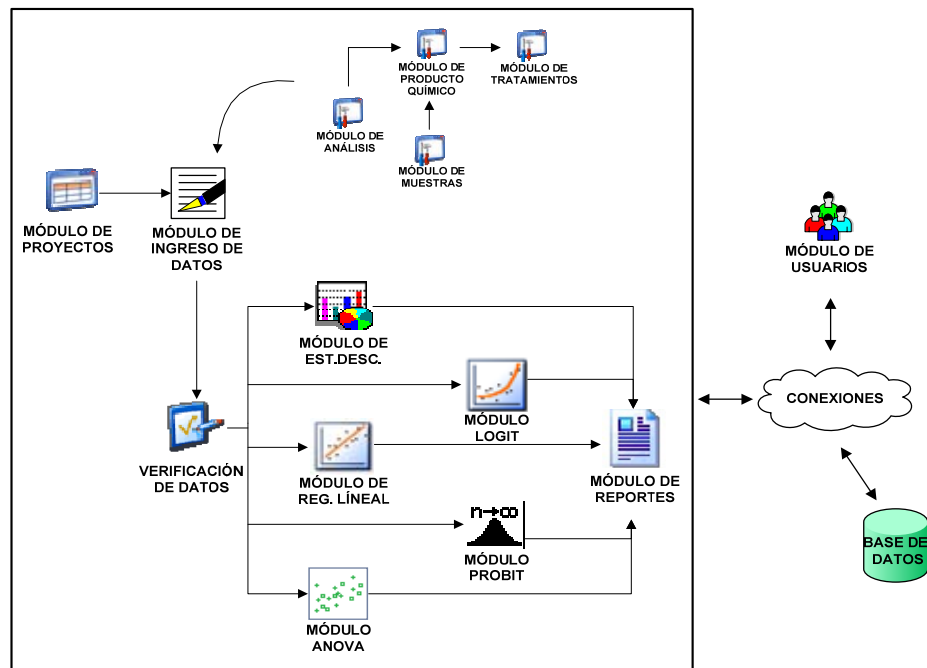


FIGURA 3.4 MÓDULOS DE LA APLICACIÓN INFORMÁTICA.
FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

➤ **Componente Transaccional**

El componente transaccional es el encargado de la operatividad de la aplicación, en donde se administran los proyectos que el centro realiza, además del ingreso correcto de los datos.

Módulo de Administración.

La administración de la aplicación es la encargada de crear, modificar, eliminar y consultar la información registrada en los proyectos.

Módulo de Proyectos.

La creación de un proyecto es el paso inicial para efectuar el registro de un proyecto de investigación que requiera hacer uso de esta herramienta informática; en este módulo se procede a ingresar los detalles correspondientes al proyecto de investigación solicitado por los clientes

Módulo de Ingreso de Datos.

En el siguiente módulo se procede al almacenamiento de la información recolectada del campo o invernadero, en esta sección el encargado del proyecto procede a ingresar los datos numéricos.

➤ **Componente Analítico**

En el siguiente componente se establecen los módulos de informes de los análisis estadísticos realizados para la herramienta.

Módulo de Análisis Descriptivo.

Es una aplicación desarrollada para el análisis exploratorio de la variable de estudio, en el cual se detalla las medidas de tendencia central; las medidas de dispersión y el tipo de distribución de los datos.

Módulo de Regresión Lineal.

El siguiente módulo fue desarrollado utilizando el método de regresión lineal simple, para el cálculo de la dosis letal media (DL_{50}), en el que se determina la ecuación lineal del modelo, que mejor se ajuste a la línea de regresión entre el promedio de inhibición de cada tratamiento

y la transformación logarítmica natural de cada dosis. Cabe recalcar que este módulo es una mejora del modelo que usaba el CIBE en su anterior aplicación.

Módulo de Análisis de Varianza para un solo factor (ANOVA).

El Análisis de Varianza para un solo factor es un método de contraste, el cual compara dos o más tratamientos mostrando si existen diferencias estadísticas significativas entre los tratamientos ó dosis probadas.

Módulo de Regresión Logística (LOGIT).

El módulo de regresión logística es una aplicación construida en base a los algoritmos de la minería de datos, el cual utiliza como medida de predicción o de discriminación el logaritmo natural, obteniendo la función logística, con el cual se clasifica los datos en dos grupos, mediante el uso de la probabilidad al evento 0 = no existe crecimiento del tubo germinativo y 1 = existe crecimiento del tubo germinativo.

El módulo presenta el resumen del modelo encontrado, la tabla de variables de la ecuación Logit, en el cual se calcula los betas de cada dosis en el modelo y también se realizan los cálculos del factor de riesgo (Odds Ratio) para cada dosis, de esta forma determinar cuál será la dosis a ser suministrada.

Módulo de Regresión Probabilística (PROBIT).

El módulo de regresión probabilística PROBIT utiliza como medida de predicción o de discriminación la función de distribución normal estándar, obteniendo de esta forma la función probabilística con el cual clasificaremos los datos en dos grupos, mediante el uso de la probabilidad al evento 0 = no existe crecimiento del tubo germinativo y 1 = existe crecimiento del tubo germinativo.

El módulo presenta un resumen del modelo, la tabla de variables de la ecuación Probit, en el cual se calculan los betas de cada dosis para el modelo y también se realizan los cálculos del factor de riesgo (Odds Ratio) para cada dosis, con este factor se podrá determinar si una dosis tuvo un mayor efecto de inhibición que otra.

➤ **Componente Seguridad**

En el siguiente componente se establecen los módulos de seguridad que permiten que la aplicación sea utilizada de manera correcta y la información sea menos sensible y más confidencial.

Módulo de Usuarios.

Es una aplicación desarrollada para administrar los niveles de acceso a las diferentes actividades que realicen las personas que utilicen el sistema.

3.2. Resultados y Análisis.

Mediante el uso de la aplicación informática construida se realizó el análisis estadístico de la variable Inhibición. Primero se realizó un análisis descriptivo univariado para cada uno de los tratamientos, luego el análisis de regresión lineal determinando la ecuación lineal. Seguidamente el análisis de varianza y finalmente como análisis principal del estudio y de la herramienta informática construida, se realizaron los análisis de regresión Logit y Probit, mediante estos análisis se clasificaron los datos, se comparó que método es más eficaz para la búsqueda de los patrones ó los coeficientes que ayudan a determinar cuál es la dosis más adecuada a suministrarse en los cultivos.

Para los análisis posteriores se utilizará como *P-Valor* de decisión los menores a 0.05 como significativos.

3.2.1. Análisis Descriptivo o Univariado.

En esta etapa se analizó en forma individual la variable denominada de control, en esta variable se mide el crecimiento del tubo germinativo ($\mu\text{m.}$) por cada nivel de dosis suministrada. Es necesario mencionar que los niveles de dosificación están en porcentaje de producto suministrado y los resultados en micras, también cabe recalcar que a partir de la variable de crecimiento se obtiene nuestra variable de objeto de análisis posteriores que será la variable Inhibición.

Haciendo uso de la herramienta informática desarrollada, se obtuvieron los

siguientes resultados descritos en la Tabla 3.1 destacándose lo siguiente: la dosis 0% ó Control es el tratamiento al que no se le suministró el producto químico; tiene un promedio de 80.31 μm , y desviación estándar de 14.42 μm . La dosis 0.1% tiene un promedio de 23.11 μm . Lo que significa que creció 23.11 μm de longitud con respecto a la dosis 0%, siendo la dosis 0.1% con mayor crecimiento a diferencia de las dosis 1% con un promedio de 16.83 μm , dosis 5% con 2.31 μm y dosis 10% con 0.65 μm .

ESTADÍSTICO	DOSIS (%)				
	0	0.1	1	5	10
Cantidad	150	150	150	150	150
Suma	12046.00	3467.00	2525.00	347.00	97.00
Promedio	80.31	23.11	16.83	2.31	0.65
Varianza	207.84	92.02	155.82	12.61	3.89
Desv. Estándar	14.42	9.59	12.48	3.55	1.97
Cuartil 25	70.00	18.00	7.00	0.00	0.00
Cuartil 50	80.00	24.50	16.00	0.00	0.00
Cuartil 75	90.75	28.00	25.00	5.00	0.00

FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

Los cuartiles de la Tabla 3.1 indican en la Dosis 0% que el 25% de los datos son menores a 70 μm , el 50% de los datos están alrededor de 80 μm y el 75% de los datos son mayores a 90.75 μm , y en las demás dosis el porcentaje de los datos en cada Cuartil disminuye debido a que se refleja el crecimiento del tubo germinativo con respecto a la dosis 0%.

Continuando con nuestro análisis descriptivo la Tabla 3.2 indica la variable de inhibición que es la que analizaremos a lo largo de este capítulo, los resultados están medidos en porcentajes.

Tabla 3.2				
Estadística Descriptiva de la Variable "Inhibición".				
ESTADÍSTICO	DOSIS (%)			
	0.1	1	5	10
Cantidad	150	150	150	150
Suma	10682.80	11855.80	14567.91	14879.21
Promedio	71.22	79.04	97.12	99.20
Varianza	142.69	241.61	19.55	6.04
Desv. Estándar	11.95	15.54	4.42	2.46
Error Estándar	0.98	1.27	0.36	0.20
Mínimo	37.74	29.02	81.32	85.06
Máximo	100.00	100.00	100.00	100.00
Curtosis	0.28	-0.20	0.55	13.52
Sesgo	0.11	-0.51	-1.28	-3.54
Cuartil 25	65.13	68.87	93.77	100.00
Cuartil 50	69.49	80.08	100.00	100.00
Cuartil 75	77.59	91.28	100.00	100.00

FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

La dosis 0.1 (0.1% del producto), tiene 71.22% de inhibición, la desviación estándar es 11.95%, el coeficiente de curtosis de 0.28, lo que significa que la distribución es leptocúrtica, además el coeficiente de sesgo de 0.11, lo que indica que la distribución es asimétrica positiva. Los cuartiles permiten determinar que el 25% de los datos tienen porcentajes de inhibición menores ó iguales a 65.13, y el 25% tienen porcentajes de inhibición mayores ó iguales a 77.59.

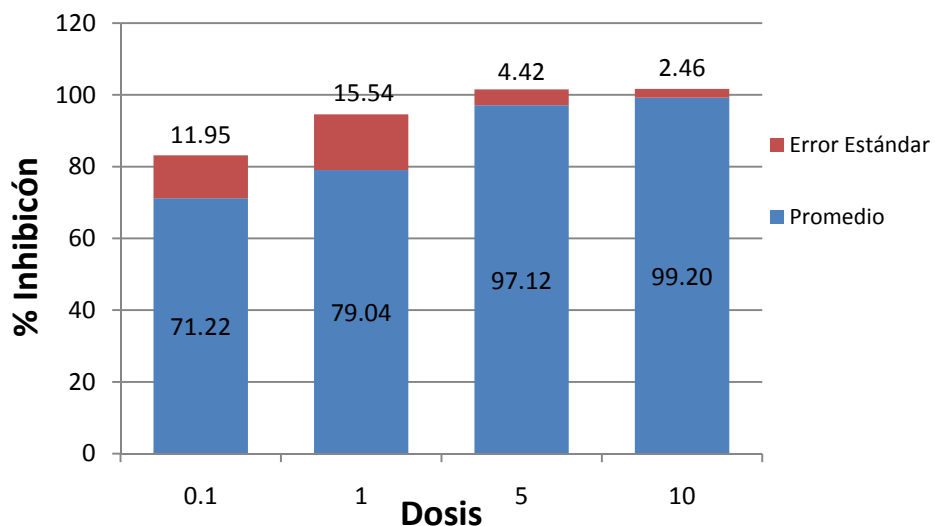
La dosis 1 (1% del producto), tiene 79.04% de inhibición, la desviación estándar es 15.54%, el coeficiente de curtosis de -0.2, lo que significa que la distribución es platicúrtica, además el coeficiente de sesgo de -0.51, lo que indica que la distribución es asimétrica negativa. Los cuartiles permiten determinar que el 25% de los datos tienen porcentajes de inhibición menores ó iguales a 68.87, y el 25% tienen porcentajes de inhibición mayores ó iguales a 91.28.

La dosis 5 o 5% del producto, tiene 97.12% de inhibición, la desviación estándar es 4.42%, el coeficiente de curtosis de 0.55, lo que significa que la distribución es leptocúrtica, además el coeficiente de sesgo de -1.28, lo que indica que la distribución es asimétrica negativa. Los cuartiles permiten determinar que el 25% de los datos tienen porcentajes de inhibición menores ó iguales a 93.77, y el 50% tienen porcentajes de inhibición mayores.

La dosis 10 o 10% del producto, tiene 99.20% de inhibición, la desviación estándar es 2.46%, el coeficiente de curtosis de 13.52, lo que significa que la distribución es leptocúrtica, además el coeficiente de sesgo de -3.54, lo que indica que la distribución es asimétrica negativa. El valor de los cuartiles en este caso es de 100 debido a que el número de observaciones presentaron mayor porcentaje de inhibición.

El gráfico 3.1 muestra el promedio y la desviación estándar de cada dosis del producto suministrado, en el cual podemos observar que el promedio que da un mayor efecto inhibitor es la dosis 10% con 99.20 por ciento, y una variación 2.46 por ciento, seguida de la dosis al 5%; y la dosis que presenta menor efecto inhibitor es la dosis 0.1% y el de mayor variación es la dosis al 1% con 15.54% lo cual indica también que fue el tratamiento más inestable en el ensayo.

GRÁFICO 3.1 ANÁLISIS COMPARATIVO DEL PORCENTAJE DE INHIBICIÓN OBTENIDO EN LAS DIFERENTES DOSIFICACIONES DE AZOXISTROBINA APLICADA A *MYCOSPHAERELLA F.*



FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

Siendo la estadística descriptiva como primer paso de análisis exploratorio de la información, podemos decir que las dosis evaluadas producen un efecto inhibitor sobre el hongo, podemos apreciar que a mayor dosificación resulta una mayor reducción de la infección y al aplicar una menor dosificación resulta una menor reducción de la infección, pero podríamos también atacar

al cultivo y de esta manera matar al mismo, por lo cual es necesario determinar cuál es la dosis letal media.

3.2.2. Análisis de Regresión Lineal Simple.

Se realizó el análisis de regresión lineal simple sobre la variable dependiente Inhibición y la variable independiente Dosis, el cual tuvo como objeto encontrar un modelo lineal, el cual se ajuste a los promedios de inhibición alcanzado en cada dosis probada, y de esta forma calcular la DL_{50} a ser suministrada. La prueba F para el análisis de varianza de la regresión lineal encontrada plantea la siguiente hipótesis:

H₀: No existe relación entre la variable
Inhibición y la Dosis.

Vs.

H₁: No es verdad H₀

En la Tabla 3.3 el análisis de varianzas indican que con un *P-Valor* de 0.033, se rechaza la hipótesis nula a favor de la alternativa que dice que si existe relación lineal entre las variables, por la tanto el modelo encontrado se ajusta a los datos, ver tabla completa en el Anexo 4.

La Prueba t para la significancia de los coeficientes de regresión, plantea la siguiente hipótesis:

H₀: $\beta_i = 0$

Vs.

H₁: No es verdad H₀

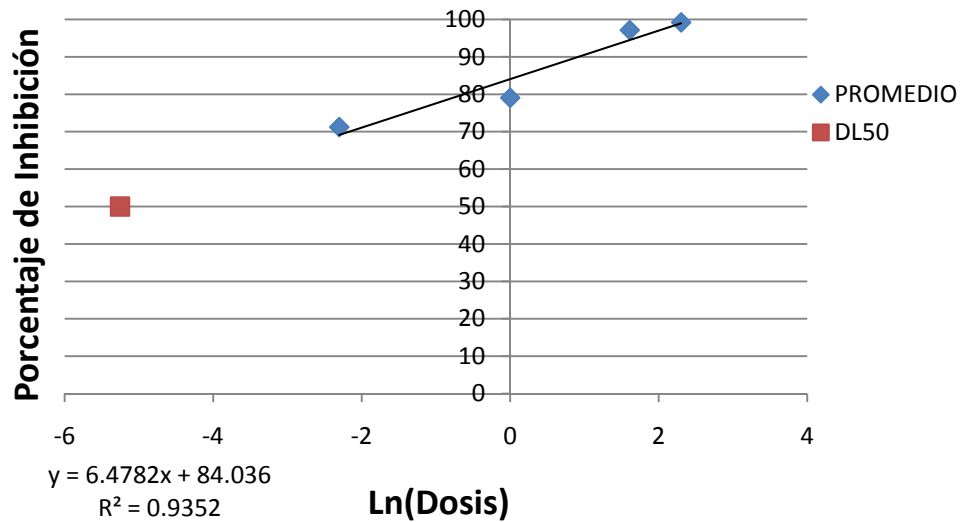
Por lo tanto podemos decir que el coeficiente de regresión b_0 es significativo con *P-Valor* de 0.001 con 1 grado de libertad, y b_1 con *P-Valor* de 0.033 con 2 grados de libertad, esto significa que se rechaza la hipótesis nula de que los coeficientes son iguales a cero.

Tabla 3.3			
Modelo de regresión lineal Simple			
		Prueba F	P-Valor
ANOVA		28.8270	0.033
Coefficientes:		Prueba t	P-Valor
Intersección (b_0)	84.0360	38.3656	0.017
Pendiente (b_1)	6.4780	5.3713	0.033
Dosis	-5.2540		
DL₅₀ estimada	0.0052		
R²	0.935		
Cantidad	150		

FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

El modelo lineal se obtuvo a través del método de mínimos cuadrados el que se muestra en la Tabla 3.3, la ecuación lineal es $y = 84.0360 + 6.4780x$, el objetivo principal del estudio es encontrar la DL₅₀, entonces reemplazando con $y = 50$ procedemos a encontrar la incógnita x , el cual da como resultado -5.2540, mediante el uso de transformaciones logarítmicas, calculando el antilogaritmo de $x = -5.2540$; ($e^{-5.2540}$), se obtuvo que la dosis buscada es de 0.0052, esto quiere decir que la dosis 0.0052% es la DL₅₀ solicitada, El gráfico 3.2 muestra la línea de regresión estimada y la DL₅₀.

GRÁFICO 3.2. CÁLCULO DE LA DL₅₀ A TRAVÉS DEL MODELO DE REGRESIÓN LINEAL.



FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

El R^2 indica que el 93.5% de la variabilidad del porcentaje de inhibición es explicada por el modelo encontrado, esto quiere decir que la ecuación de regresión se ajusta adecuadamente a la variable Inhibición.

3.2.3. Análisis de Varianzas para un solo factor (ANOVA).

Se emplea el análisis de varianza sobre la variable Inhibición, la que tiene como objeto encontrar las diferencias estadísticas significativas entre las dosis evaluadas, el cual ayuda a ampliar el análisis estadístico, en el Anexo 5 se muestra el test de homogeneidad de varianzas que plantea la hipótesis siguiente.

$$H_0: \sigma_1 = \sigma_2$$

Vs.

$$H_1: \text{No es verdad } H_0$$

Entonces con un *P-Valor* de 0.000, significa que se rechaza la hipótesis de igualdad de varianzas a favor de la alternativa, basándonos en estos resultados procedemos a realizar el análisis de las dosis entre sí.

En el Anexo 4 se muestra la suma de los cuadrados entre las dosis es de 84455.52 y dentro de las dosis (error) es de 61072.22, esto indica que la mayor parte de variabilidad fue atribuida a los tratamientos, el estadístico de Prueba *F* de 274.73, da un *P-Valor* de 0.000, por este motivo se puede afirmar que existe al menos una dosis de fungicida que causa un efecto inhibitor sobre las muestras.

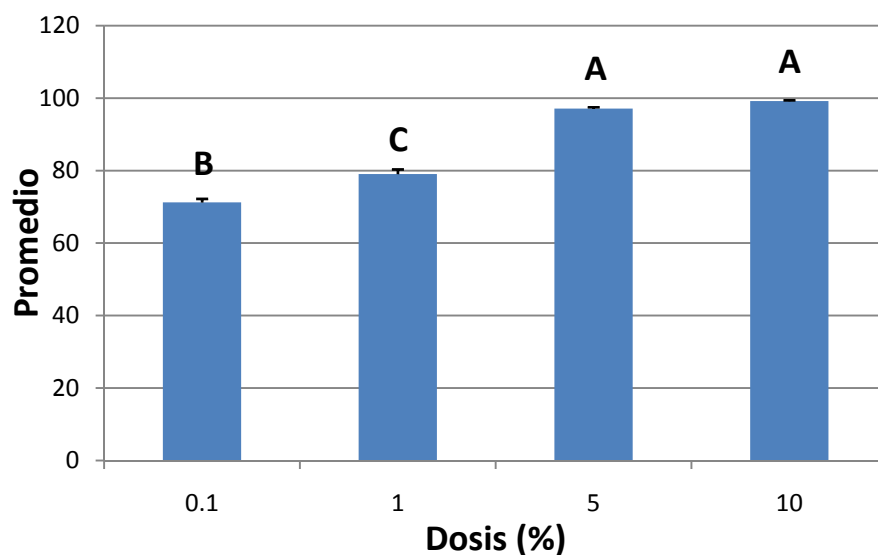
Se aplica el método de LSD de Fisher para el cálculo de las diferencias entre tratamientos, donde encontramos que entre la dosis 5% y 10% con un *P-Valor* de 0.038 no existen diferencias significativas, para los demás casos con un *P-Valor* de 0.000 indica que si existen diferencias significativas, esto significa que la dosis de fungicidas utilizados causan un efecto de inhibición, mostrados en el Anexo 4.

El Gráfico 3.3 indica los promedios de cada dosis y su error estándar en la parte superior, obtenidos de la Tabla 3.2, por la Prueba de *Fisher* se pudo determinar que no existen diferencias significativas entre las dosis 5% y 10%, por lo tanto serán codificadas con un mismo valor, igual a la letra *A*, para las dosis 0.1% y 1% en el que si existe diferencia significativa, se les colocará

las letras *B* y *C* respectivamente, para una mejor interpretación de los resultados se procedió a realizar dicha codificación entre las dosis.

Las dosis 5% y 10% tienen promedios semejantes y errores estándares pequeños, mientras que la barra de las dosis 0.1% y 1% tiene promedios diferentes y sus errores estándares son mayores a las dosis anteriores.

GRÁFICO 3.3. DIFERENCIAS ESTADÍSTICAS SIGNIFICATIVAS ENTRE LOS PROMEDIOS DE LAS DOSIS EVALUADAS.



FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

3.2.4. Análisis de Regresión Logit y Probit.

Al utilizar este tipo de análisis de regresión sobre la variable Inhibición, el interés principal fue calcular las probabilidades de cada observación que esté asociada al grupo con un alto efecto de inhibición ó al grupo que no, mediante la determinación de las dosis más adecuadas que expliquen las

diferencias entre los grupos mencionados; finalmente se determinó mediante los cuantificadores de riesgo que dosis causan en menor o mayor grado un efecto inhibitor.

A continuación se muestra en la Tabla 3.4 los resultados del resumen del modelo de regresión Logit, donde el número de iteraciones antes de encontrar el modelo es de 58 corridas, la variabilidad del modelo indica un error igual a 9.87E-18 ceros de precisión, finalmente se calcula el R^2 del modelo lo que indica que el 58% de la variabilidad es explicada por el modelo encontrado, resultando un buen ajuste al modelo.

Tabla 3.4	
Resumen del Modelo Logit	
Cantidad	150
No. Iteraciones	58
Variabilidad	9.87E-18
R^2	0.580

FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

A continuación procedemos a interpretar los resultados de las dosis evaluadas de la variable Inhibición según el modelo Logit, la Prueba de Wald plantea la siguiente hipótesis:

H₀: La dosis i-ésima no se introduce
al modelo. $\beta_i = 0$

Vs.

H₁: No es verdad H₀

En la Tabla 3.5 podemos observar que la Prueba de Wald resultó con un *P-Valor* de 0.000, esto indica que con tres ceros de precisión las dosis 0.1%, 1%, 5% y 10% influyen significativamente en el modelo y serán introducidas al modelo de regresión.

Tabla 3.5					
Tabla de Variables en la Ecuación del Modelo Logit.					
Dosis	Betas	Prueba de Wald	GL	P-Valor	Odds Ratio
Dosis a 0.1	0.001	4.42E+10	3	0.000	1.001
Dosis a 1	-0.001	3.82E+10	3	0.000	0.999
Dosis a 5	0.011	1.15E+13	3	0.000	1.011
Dosis a 10	0.019	3.49E+13	3	0.000	1.019

FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

Para las Dosis 0.1%, 5% y 10% del producto químico suministrado, se obtuvieron coeficientes positivos $\beta = 0.001$; $\beta = 0.011$ y $\beta = 0.019$; y Odds-Ratios iguales a 1.001, 1.011 y 1.019 respectivamente, es decir que por cada muestra infectada se tiene 1.001 veces más probabilidad de disminuir su infección con la dosis 0.1; 1.011 veces con la dosis 5 y 1.019 veces con la dosis 10.

En la Dosis 1% se obtuvo un coeficiente $\beta = -0.001$ y un Odds-Ratio igual a 0.999, es decir que por cada muestra infectada se tiene 0.999 veces más probabilidad de aumentar su infección con la dosis aplicada.

Podemos decir que el modelo Logit se ajusta de mejor forma a los datos debido a su baja variabilidad, lo cual le da mayor precisión a los resultados encontrados, donde el modelo Logit reafirma al análisis descriptivo e inferencial, en el que se observa que la dosis 10%, es la de mayor efecto inhibitor.

Continuando con el análisis de los modelos desarrollados, procedemos con el de Regresión Probit, la Tabla 3.6 muestra los resultados del resumen del modelo, donde el número de iteraciones antes de encontrar el modelo es de 1001 corridas, la variabilidad del modelo da como resultado un error igual a 2.24E-03 ceros de precisión, esto indica que existe en mayor grado error en los resultados y exige mayor tiempo en búsqueda del modelo solicitado, a diferencia del modelo logístico.

Tabla 3.6	
Resumen del Modelo Probit	
Cantidad	150
No. Iteraciones	1001
Variabilidad	2.24E-03
R²	0.580

FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

Se calcula también el R² del modelo que indica, que el 58% de la variabilidad es explicada por el modelo encontrado, al igual que en el modelo Logit.

De igual manera se interpretan los resultados de las dosis evaluadas según el Modelo Probit sobre la variable Inhibición, la Prueba de Wald plantea la siguiente hipótesis:

H₀: La dosis *i*-ésima no se introduce
al modelo $\beta_i = 0$

Vs.

H₁: No es verdad H₀

En la Tabla 3.7 podemos observar que la Prueba de Wald resultó con un *P-Valor* de 1, esto indica que las Dosis 0.1%, 1%, 5% y 10% no aportan significativamente en el modelo; de igual manera se interpretarán los resultados debido a que su obtención está basada en algoritmos de aprendizaje de la minería de datos, por ende el modelo encontrado a continuación es el que más se ajusta al fenómeno estudiado.

Tabla 3.7					
Tabla de Variables en la Ecuación del Modelo Probit.					
Dosis	Betas	Prueba de Wald	GL	P-Valor	Odds Ratio
Dosis a 0.1	0.004	6.57E-03	3	1.000	1.004
Dosis a 1	0.001	7.01E-04	3	1.000	1.001
Dosis a 5	0.001	7.89E-04	3	1.000	1.001
Dosis a 10	0.010	4.39E-02	3	0.998	1.010

FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

Con el modelo Probit, la Dosis 0.1%, 1%, 5% y 10% de producto químico suministrado, se obtuvieron coeficientes positivos $\beta = 0.004$; $\beta = 0.001$; $\beta = 0.001$ y $\beta = 0.010$ y Odds-Ratios iguales a 1.004; 1.001; 1.001 y 1.010 respectivamente, es decir que por cada muestra infectada se tiene 1.004

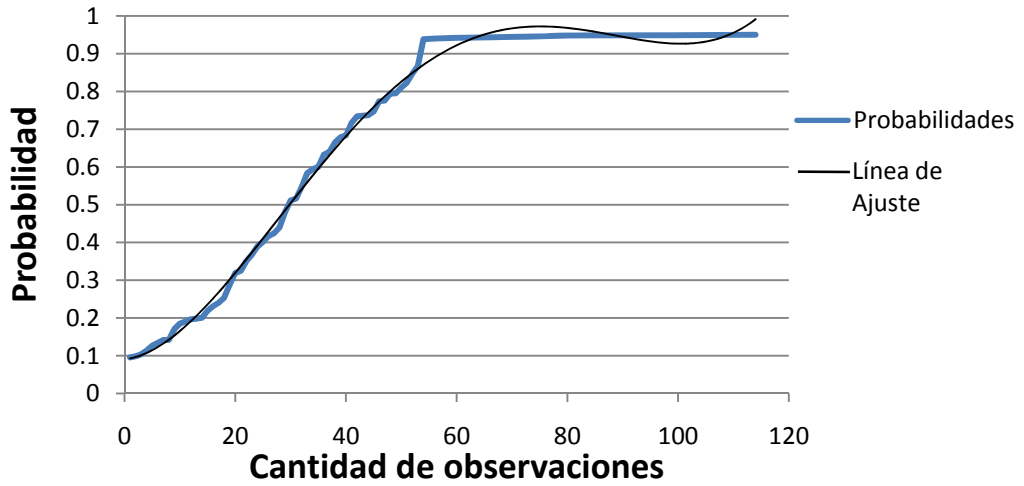
veces más probabilidad de disminuir su infección con la dosis 0.1; 1.001 veces con la dosis 1 y 5; y 1.010 veces con la dosis 10.

Procedemos a realizar la comparación entre las dosis basándonos en los resultados de los Odds Ratios, en donde la Dosis 10% con un Odds Ratio igual a 1.010 significa que causa un mayor efecto inhibitor al igual que en el modelo logístico, seguida de la Dosis 0.1% con 1.004, el resto causa en menor parte un efecto inhibitor e igual sobre los cultivos, que son la Dosis 1% y 5% con 1.001 y 1.001 respectivamente.

El modelo Probit se ajusta de manera satisfactoria a los datos, pero con una mayor variabilidad con respecto al modelo Logit, lo cual le quita precisión a los resultados.

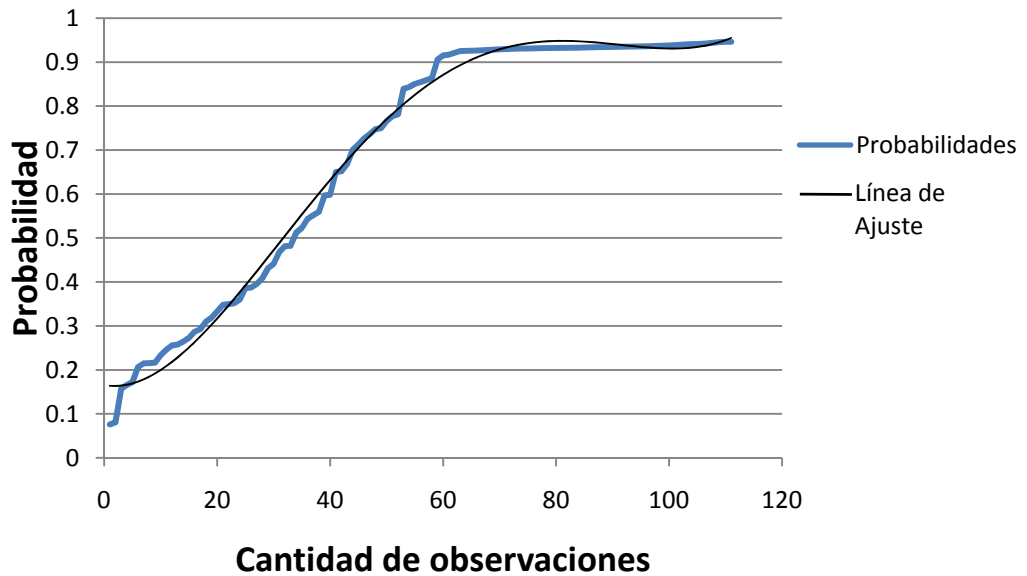
En los gráficos 3.5 y 3.6 se muestran las diferencias entre los modelos desarrollados, para ambos modelos estos indican un efecto inhibitor, pero el modelo Logit presenta mayores probabilidades de inhibición sobre el modelo Probit.

GRÁFICO 3.5. ANÁLISIS DEL COMPORTAMIENTO SOBRE LAS PROBABILIDADES DE CADA OBSERVACIÓN EVALUADAS EN EL MODELO DE REGRESIÓN LOGIT.



FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

GRÁFICO 3.6. ANÁLISIS DEL COMPORTAMIENTO SOBRE LAS PROBABILIDADES DE CADA OBSERVACIÓN EVALUADAS EN EL MODELO DE REGRESIÓN PROBIT.



FUENTE: CIBE – ESPOL AUTOR: MARIO LINDAO

Podemos decir que ambas técnicas de regresión no lineal aportan en la búsqueda de las dosis que contienen mayor efecto inhibitor, donde la regresión Logit muestra mayor precisión de resultados a diferencia del modelo Probit, esto se demuestra con el error generado en el modelo Logit igual a $9.87E-18$ y el modelo Probit con $2.24E-03$.

La variable de crecimiento del tubo germinativo obtuvo su mayor crecimiento al suministrársele 0.1% de fungicida con un promedio de $23.11 \mu\text{m}$ y el de menor crecimiento con 10% de fungicida suministrado con un promedio de $0.65 \mu\text{m}$.

La variable Inhibición ó de interés al suministrársele con 10% de fungicida obtuvo un promedio de 99.20% de inhibición a las ascosporas, y la dosis que obtuvo un menor efecto fue al suministrársele 0.1% de fungicida con un promedio de 71.22% de inhibición.

La dosis 10% de fungicidas suministrado es la de menor desviación estándar con 2.46% de inhibición, y la dosis 1% es la de mayor desviación estándar con 15.54% de inhibición.

El aporte global de las técnicas utilizadas conllevan a encontrar comportamientos en cada una de ellas, lo cual indican que el análisis descriptivo sumado a los modelos no lineales, llevan la misma línea de

decisión en sus resultados, a diferencia del modelo de regresión lineal que expresa decisiones contrarias en cuanto a cuál debería ser la dosis a ser suministrada

En base a estos resultados decimos que los modelos no lineales generan mejores resultados, esto se debe a que estos analizan el comportamiento interno y externo de los datos en búsqueda del mejor ajuste a ambos modelos.

Sobre la herramienta informática podemos decir que tiene una base de datos diseñada para el almacenamiento adecuado y seguridad de los proyectos y de los datos, lo cual lleva a tener una bitácora más amplia y de manera digital. Los resultados que esta herramienta emita son de total confiabilidad, debido a que ha sido probado sus resultados de manera manual y también a través del uso de SPSS software estadístico.

La facilidad de manejo del sistema para el usuario que conoce de herramientas básicas como lo es Excel, permitirá una buena interacción usuario – máquina.

CAPÍTULO IV

4. CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

1. La integración de la informática en el campo estadístico, permitió tener una herramienta especializada en el área de bioestadística, optimizando de esta manera el proceso de análisis de sensibilidad.
2. El módulo de verificación de datos redujo el error de escritura al momento de ingresar los datos, mediante el uso de intervalos de control del promedio \pm una desviación estándar.
3. Mediante el uso de la herramienta se realizó el análisis de varianza de la regresión, el que nos muestra la prueba F con un P -Valor de 0.033 lo que indica que existe relación entre el porcentaje de inhibición y la dosis.
4. El $R^2 = 93.5\%$ del modelo de regresión lineal muestra un buen ajuste del modelo a los datos experimentales, lo que ratifica el uso del modelo.
5. Sobre la técnica de regresión simple sumada a la técnica de ANOVA podemos decir que la dosis 0.1% de fungicida se acerca a la dosis letal media estimada $DL_{50} = 0.005\%$ y es la que debe ser suministrada en los cultivos.

6. De acuerdo al análisis de varianza para las dosis realizado a través de la herramienta nos indica que la prueba F con un P -Valor de 0.000 si existe al menos una dosis que causa un efecto inhibitor.
7. El análisis de las diferencias entre los promedios determinó que no existen diferencias estadísticas significativas entre los promedios de las dosis 5% y 10%; encontrando diferencias significativas solo en las dosis 0.1% y 1% del producto suministrado.
8. Los modelos no lineales utilizaron algoritmos basados en minería de datos lo que conlleva a un análisis más profundo, mejor tratamiento de la información y de sus resultados.
9. Los modelos no lineales expuestos como lo es el de regresión Logit y Probit muestran una precisión de $9.87E-18$ y $2.24E-03$ en los resultados respectivamente, esto indica la confiabilidad y robustez.
10. Para ambos modelos se evaluaron las dosis suministradas en donde se logró determinar las diferencias entre las dosis de acuerdo al valor de su Odds-Ratio en donde la dosis 10% con 1.019 y 1.010 para el modelo Logit y Probit respectivamente, esto significa que por cada muestra infectada se tiene 1.019 y 1.010 veces de probabilidad de disminuir su infección, por lo tanto esta dosis sería considerada como la que causa un mayor efecto inhibitor respecto a las demás.
11. La dosis 1% obtuvo un coeficiente negativo igual a -0.001 , por ende un Odds Ratio menor a 1, lo que significa un aumento en la probabilidad de que se produzca la infección.

4.2. Recomendaciones

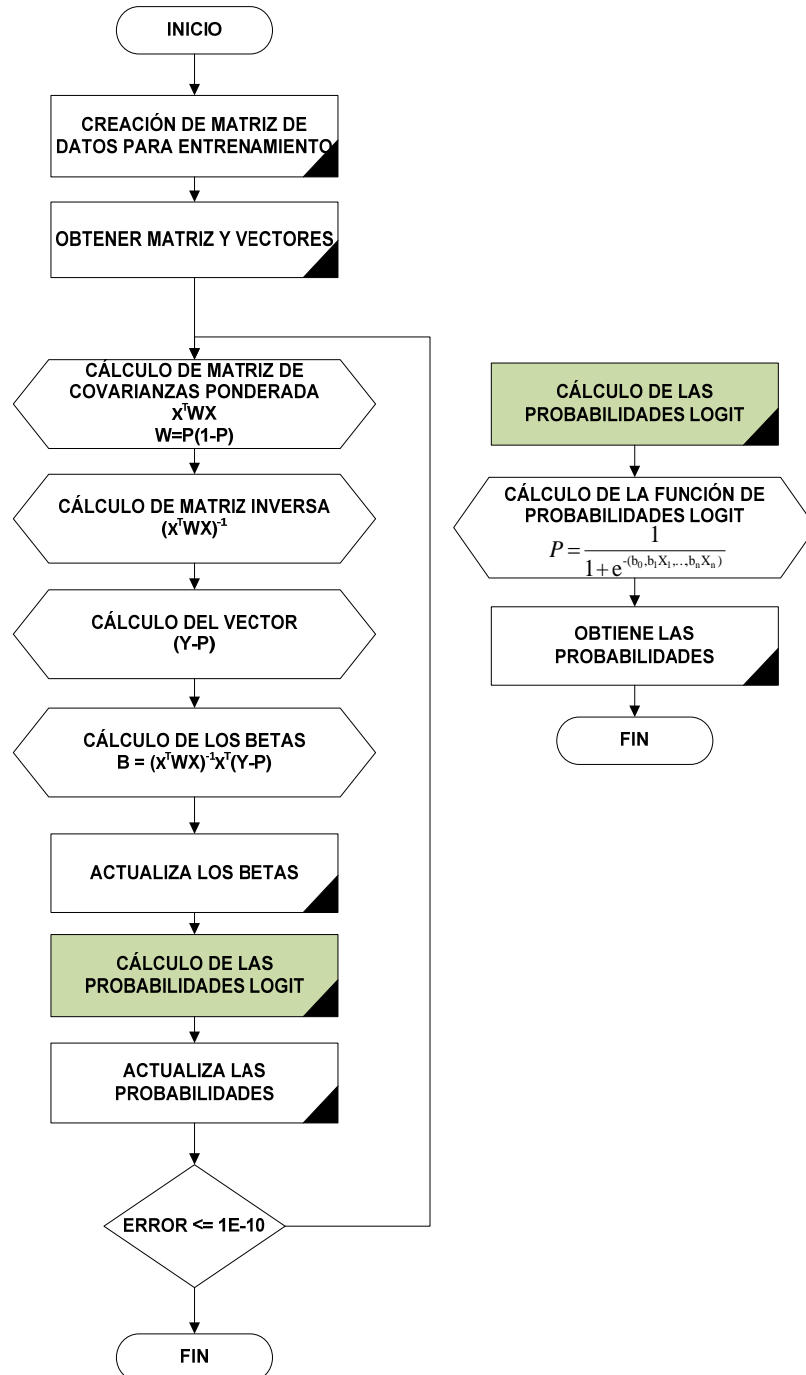
1. Se recomienda desarrollar más herramientas en el área informática, una forma es creando módulos pequeños que permitan realizar tareas específicas en un solo campo, y de esta forma disminuir el costo de licencias por software comerciales.
2. El desarrollo de nuevas herramientas utilizando interfaz Web, lo que permitirá ampliar el uso de la misma a través de cualquier computador conectado a Internet y a su vez optimizar la entrega de resultados.
3. Para el uso de la herramienta informática es necesario revisar el manual de usuario, el cual nos guiará con el uso fácil del mismo, determinando siempre los usuarios que van a utilizar dicha herramienta, de esta forma se evitarán pérdida de archivos y manipulación de los datos del proyecto.
4. La herramienta construida puede ser utilizada en posteriores estudios para creación de sistemas en el área de la biotecnología, debido a la evolución diaria de la tecnología alrededor del mundo.
5. Se recomienda realizar análisis sobre los datos en búsqueda de cuál sería la cantidad de muestras necesarias a ser recolectadas del campo, lo que ayudaría en cierto modo aumentar la precisión de los resultados.

6. Para el caso donde la dosis encontrada se encuentra fuera del rango de las dosis evaluadas, se debería realizar otros ensayos en donde se consideren dosis inferiores a las utilizadas, y de esta manera determinar con mayor precisión cual dosis deberá ser suministrada.

Anexos

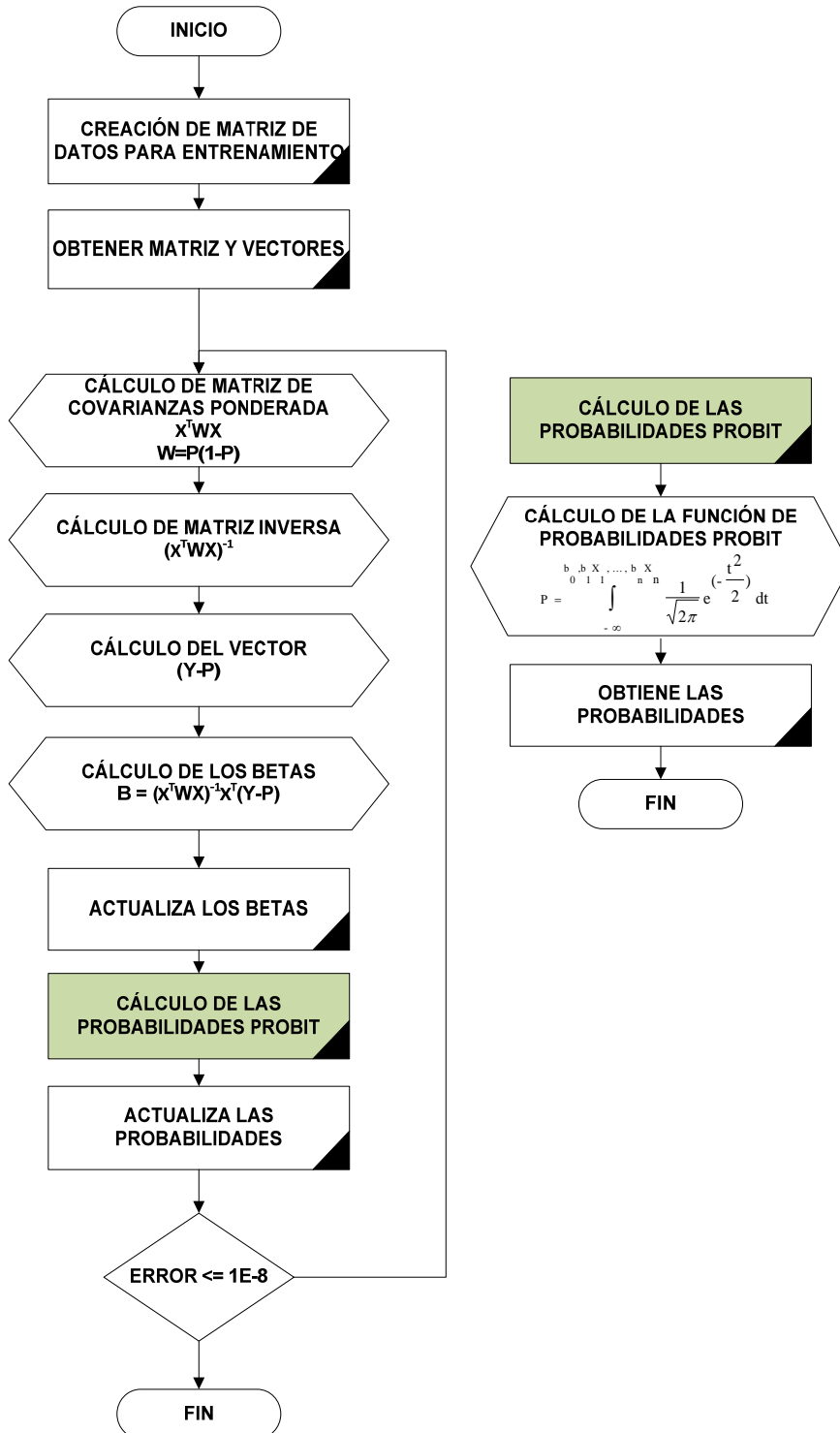
ANEXO 1

Modelo Lógico de la Regresión Logit.



ANEXO 2

Modelo Lógico de la Regresión Probit.



ANEXO 3

TEOREMA DE COCHRAN.

Sea Z_i igual a $NID(0,1)$ para $i=1,2,\dots,v$. y

$$\sum_{i=1}^v Z_i^2 = Q_1 + Q_2 + \dots + Q_s$$

Donde $s \leq v$ y Q_i tiene v_i grados de libertad ($i=1,2,\dots,s$). Entonces Q_1, Q_2, \dots, Q_s son variables aleatorias ji-cuadrado independientes con v_1, v_2, \dots, v_s grados de libertad, respectivamente, si y sólo si

$$v = v_1 + v_2 + \dots + v_s$$

ANEXO 4

Resultados del análisis de varianzas para un solo factor.

Tabla 1.			
Prueba de Homogeneidad de Varianzas			
Estadístico Levene	GL1	GL2	P-Valor
95.6770	3	596	0.000

Tabla 2.					
Tabla de Análisis de Varianza para un solo factor					
Fuente de Variación	SC	GL	SCM	F	P-Valor
Entre Dosis	84455.52	3	28151.84	274.73	0.000
Dentro Dosis (Error)	61072.22	596	102.47		
Total	145527.73	599			

Tabla 3. Comparación de medias - LSD de Fisher		
Dosis (%)	Diferencias	P-Valor
5 Vs. 1	18.08	0.000
5 Vs. 10	-2.08	0.038
5 Vs. 0.1	25.90	0.000
1 Vs. 10	-20.17	0.000
1 Vs. 0.1	7.82	0.000
10 Vs. 0.1	27.98	0.000

ANEXO 5

Como parte de nuestro análisis encontramos la matriz de correlaciones que se muestra a continuación en la Tabla 4 la que nos indica que existe una relación lineal fuerte positiva entre las Dosis 5% y 10% con 0.824, seguida de las Dosis 5% y 1% con 0.778, el resto de igual manera se encuentran fuertemente relacionadas pero en menor valor.

Tabla 4. Matriz de Correlación				
Dosis	Dosis a 0.1	Dosis a 1	Dosis a 5	Dosis a 10
Dosis a 0.1	1	0.623	0.717	0.707
Dosis a 1	0.623	1	0.778	0.729
Dosis a 5	0.717	0.778	1	0.824
Dosis a 10	0.707	0.729	0.824	1

ANEXO 6

MANUAL DE USUARIO DE LA HERRAMIENTA INFORMÁTICA

El presente manual para el usuario proporciona la documentación sobre la manera de interactuar con la aplicación y de la interfaz gráfica bajo la plataforma de Microsoft Excel. La herramienta descrita ayuda a optimizar el proceso de almacenamiento de información y de los análisis que se efectúan en el proceso de sensibilidad para el CIBE. Los usuarios autorizados para el manejo de la herramienta son un auxiliar y el asistente de investigación.

Requerimientos técnicos de hardware y software.

Los requerimientos mínimos de hardware son:

- ❖ Procesador Pentium IV (Compatible) o superior.
- ❖ Mínimo 128 MB de memoria RAM.
- ❖ Mínimo 200 MB libres en disco duro.

La herramienta es compatible con los siguientes sistemas:

- ❖ Windows Xp ó superiores.
- ❖ Microsoft Excel 2003 ó superiores.
- ❖ Base de datos PostgreSQL 2.8.

Diagrama general de la herramienta.

El Figura 1 muestra el diagrama de interacción de la herramienta construida, el cual detalla todas las actividades a realizarse durante el desarrollo de algún proyecto.

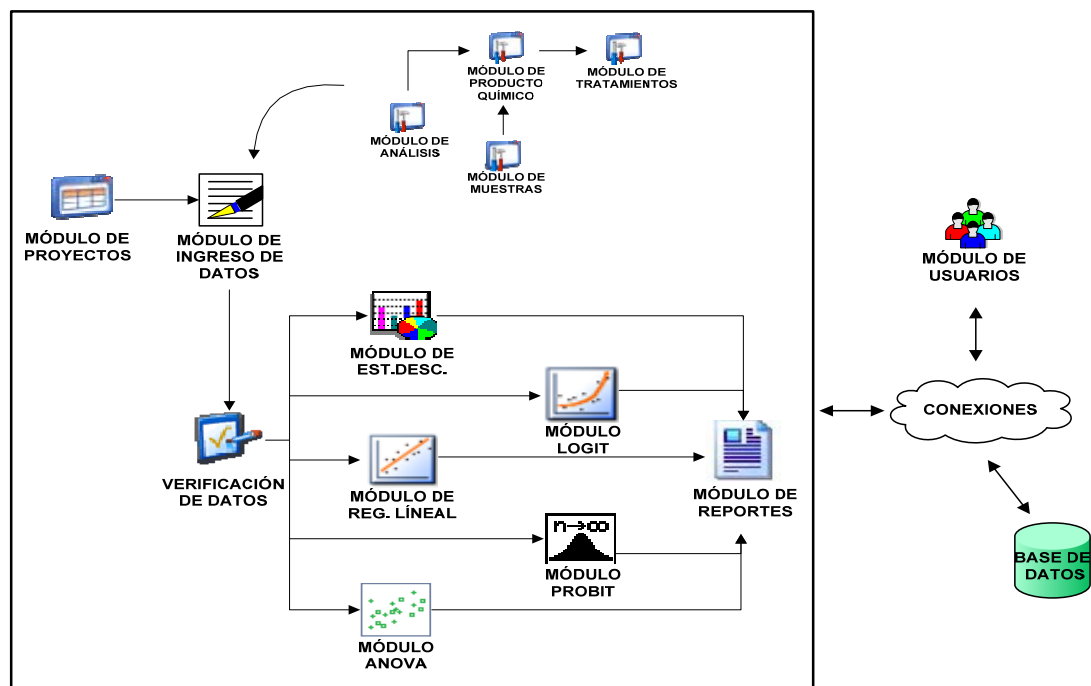


FIGURA 1. APLICACIÓN INFORMÁTICA.

Componentes y módulos de la herramienta informática.

Para el mejor desarrollo de la aplicación informática se han dividido en 3 componentes que van de acuerdo a sus funciones, las que son el Componente Transaccional, Analítico y Seguridad. En cada componente existen módulos que son aplicaciones pequeñas que realizan cierto trabajo específico.

➤ **Componente Transaccional**

El componente transaccional es el encargado de la operatividad de la aplicación, en donde se administrarán los proyectos que el centro realice, además del ingreso correcto de los datos. El detalle de los módulos se muestra a continuación:

Módulo de Administración.

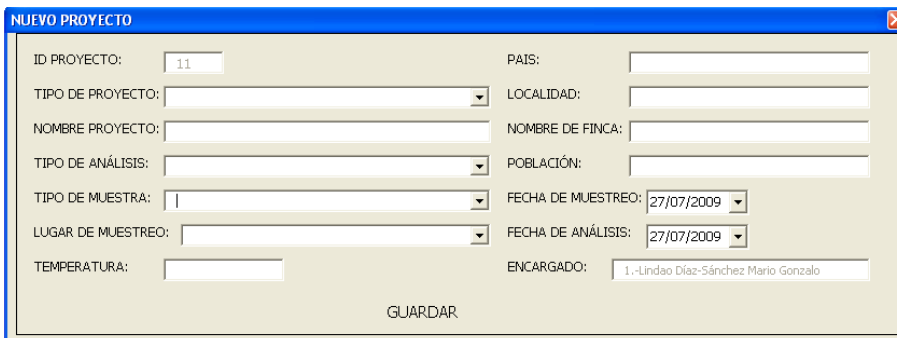
La administración de la aplicación es la encargada de crear, modificar, eliminar y consultar la información registrada en los proyectos del centro, en donde se ingresan las características de un proyecto a ejecutar, los que son necesarios de ingresar para proceder con los análisis estadísticos, a continuación se muestra en la Tabla 5.

Tabla 5. Administración de un Proyecto		
	DETALLES	ACCIONES
Tipo de proyecto	Se registran los tipos de proyectos a realizar; Ejm: BANARED.	CREAR, MODIFICAR, ELIMINAR Y CONSULTAR
Tipo de Análisis	Se registran los tipos de análisis a realizar en el proyecto; Ejm: ANÁLISIS DE SENSIBILIDAD.	
Tipo de Muestra	Se registran los tipos de muestras utilizados para el análisis, o sea el nombre de las muestras utilizadas para el análisis en laboratorio; Ejm: ASCOSPORAS,	
Tipo de Producto Químico	Se registran los tipos de químicos utilizados en los cultivos, estos deben estar relacionados con el tipo de muestra analizada previamente; Ejm: BENOMIL, PROPICONAZOLE.	
Tipo de Tratamiento	Se registran los tipos de tratamientos, esto son los tipos de dosis química que fueron suministrados sobre las muestras y su medida que será en Partes por Millón (PPM) ó en Porcentaje (%); Ejm: 0.0001ppm, 0.001ppm, 10%, 5%, 3%.	

Módulo de Proyectos.

Dentro de la aplicación informática se prepara la creación de un proyecto, que es el paso inicial para efectuar el registro de un proyecto de investigación que requiera hacer uso de esta herramienta informática; el mismo que se encuentra relacionado con los datos a ingresar y los análisis estadísticos a efectuarse.

En el módulo siguiente se procede a ingresar los detalles correspondiente al proyecto de investigación solicitado por los clientes, el CIBE realiza dicha tarea de manera manual, el sistema permitirá respaldar el tipo de análisis a realizar, el tipo de productos químicos o muestras, ubicación, nombre de las haciendas o fincas, fechas y responsables, procediendo finalmente al almacenamiento dentro de la base de datos.



NUEVO PROYECTO	
ID PROYECTO:	11
TIPO DE PROYECTO:	[Dropdown]
NOMBRE PROYECTO:	[Text]
TIPO DE ANÁLISIS:	[Dropdown]
TIPO DE MUESTRA:	1
LUGAR DE MUESTREO:	[Dropdown]
TEMPERATURA:	[Text]
PAIS:	[Text]
LOCALIDAD:	[Text]
NOMBRE DE FINCA:	[Text]
POBLACIÓN:	[Text]
FECHA DE MUESTREO:	27/07/2009
FECHA DE ANÁLISIS:	27/07/2009
ENCARGADO:	1.-Lindao Díaz-Sánchez Mario Gonzalo
GUARDAR	

FIGURA 2. MÓDULO DE PROYECTOS.

Módulo de Ingreso de Datos.

En el siguiente módulo se procede al almacenamiento de la información recolectada del campo o invernadero, en esta sección el encargado del proyecto procede a ingresar los datos numéricos obtenidos durante del proceso de evaluación.

Durante esta etapa se ingresa información adicional necesaria para el respaldo de la información, tales como el tipo de producto químico a analizar, la cantidad de tratamientos realizados a este producto, posteriormente se agregan las diferentes dosis utilizadas en laboratorio, con la finalidad de construir la matriz de datos que ayudará para los análisis estadísticos.

Posteriormente se procede a verificar los datos ingresados mediante el uso de técnicas de control, en donde se define como medida de control el promedio \pm desviación estándar en el cual indica si algún valor fue mal ingresado, caso contrario se procederá con el almacenamiento en la base de datos.

INFORMACIÓN DEL PROYECTO	
Id Proyecto:	País:
Tipo Proyecto:	Localidad:
Proyecto:	Finca:
Tipo de Análisis:	Población:
Tipo de Muestra:	Fecha de muestreo:
Lugar de Muestreo:	Fecha de análisis:
Temperatura:	Responsable:

DATOS DEL PROYECTO	
Tipo de Fungicida:	
# de Tratamientos: 2	Tipo de Tratamientos: + -

COD.	No.
1	1
2	2
3	3
4	4

FIGURA 3. MÓDULO DE INGRESO DE DATOS.

➤ Componente Analítico

En el siguiente componente se establecen los módulos de informes de los análisis estadísticos realizados para la herramienta, en el que se explica la funcionalidad de las técnicas estadísticas, los que sirven para realizar el reporte del proyecto de investigación.

Módulo de Análisis Descriptivo.

Es una aplicación desarrollada para el análisis exploratorio de los tratamientos evaluados, en el cual se detalla las medidas de tendencia central: suma, promedio, mínimo y máximo; las medidas de dispersión: varianza, desviación estándar y cuartiles; y los tipos de distribución de datos: curtosis y sesgo; adicional se muestra el gráfico de barras en el que se detalla de manera conjunta el aporte del promedio y la desviación estándar a las dosis expuestas.

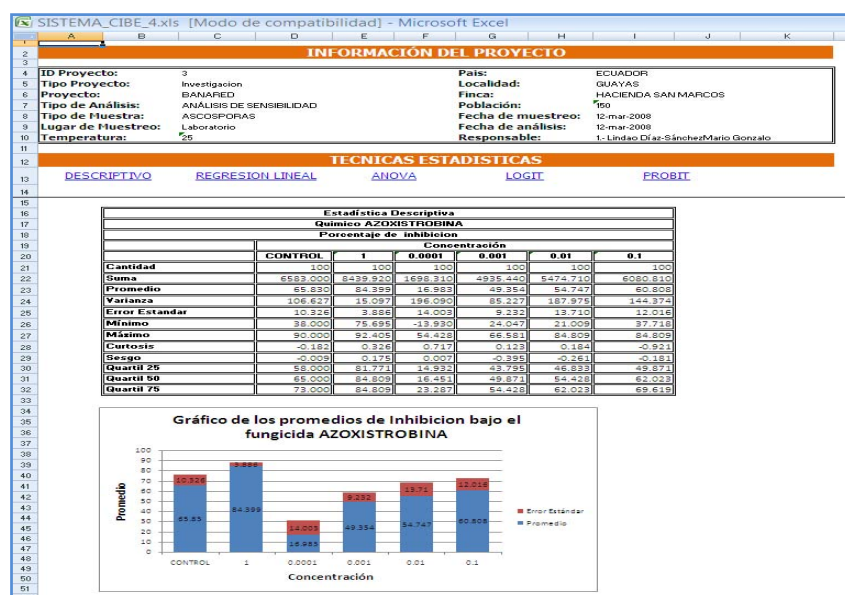


FIGURA 4. MÓDULO DE ANÁLISIS DESCRIPTIVO.

Módulo de Regresión Lineal.

El siguiente módulo fue desarrollado utilizando el método de regresión lineal simple, para el análisis del cálculo de la dosis letal media (DL50), el que se determina mediante la ecuación $50 = x * m + b$; donde x es la variable a encontrar (DL₅₀), m la pendiente y b la intersección de la ecuación, el punto de intersección y la pendiente es el mejor ajuste de la línea de regresión entre el promedio de inhibición de cada tratamiento y la transformación logarítmica natural de cada dosis, que se muestra en el módulo, adicional se muestran los gráficos lineales en el que se detalla la línea de tendencia lineal y logarítmica como ajuste de regresión.

Cabe recalcar que este módulo es una mejora del modelo que usaba el CIBE en su anterior aplicación.

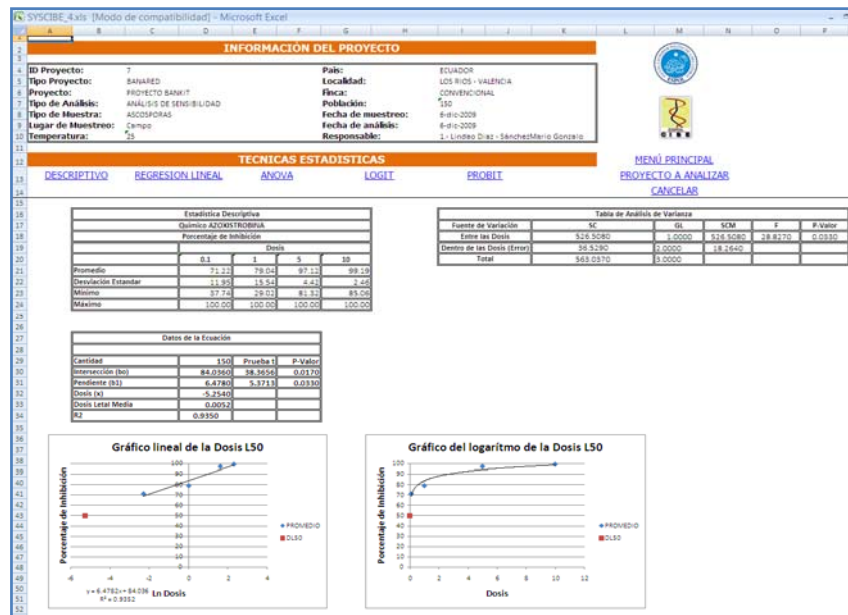


FIGURA 5. MÓDULO DE REGRESIÓN LINEAL.

Módulo de Análisis de Varianza (ANOVA).

El Análisis de Varianza es el método que compara dos o más tratamientos, se construye el módulo de ANOVA que es aquel encargado de mostrar las diferencias significativas existentes en los tratamientos, permitiendo al usuario determinar cuál es la dosis más adecuado a suministrar, adicional se construye el grafico de las medias de cada dosis.

El proceso consiste en el cálculo de las sumas y medias cuadráticas, con los que se determinará el *P- Valor* que será utilizado como medida de decisión, para esto se desarrollará el test de homogeneidad para determinar si existen varianzas iguales o no; el cual permite construir la prueba de LCD de Fisher que nos indicaran cual de los tratamientos internamente son diferentes.

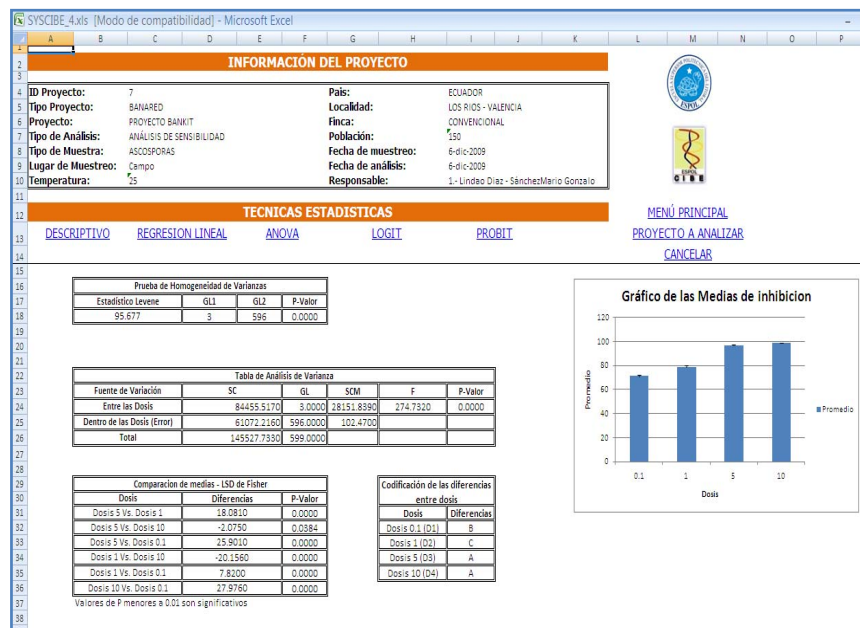


FIGURA 6. MÓDULO DE ANOVA.

Módulo de Regresión Logística (LOGIT).

El módulo de regresión logística es una aplicación construida en base a los algoritmos de la minería de datos, el cual utiliza como medida de predicción o de discriminación el logaritmo natural obteniendo la función logística, con el cual se clasifica los datos en dos grupos, mediante el uso de la probabilidad al evento 0 = no existe crecimiento del tubo germinativo y 1 = existe crecimiento del tubo germinativo.

El módulo presenta un resumen, el cual muestra el número de iteraciones efectuadas, la variabilidad o error del modelo con el que se mide la precisión y el R cuadrado que es la robustez del modelo.

Se realiza la tabla de variables de la ecuación Logit, en el cual se calcula los betas de cada dosis para el modelo, además se calcula el test de WALD y su *P-Valor* el que nos permitirá decidir si una variable se debe o no usar en el modelo, también se realizan los cálculos del factor de riesgo (Odds Ratio) para cada dosis, con este factor se puede determinar si una dosis tuvo un mayor efecto de inhibición.

Adicional se muestra la matriz de correlaciones de las dosis evaluadas, que indica cuales son las variables con mayor relación lineal, y el grafico de la función logística que muestra las probabilidades de cada observación que se ajusta al modelo encontrado.

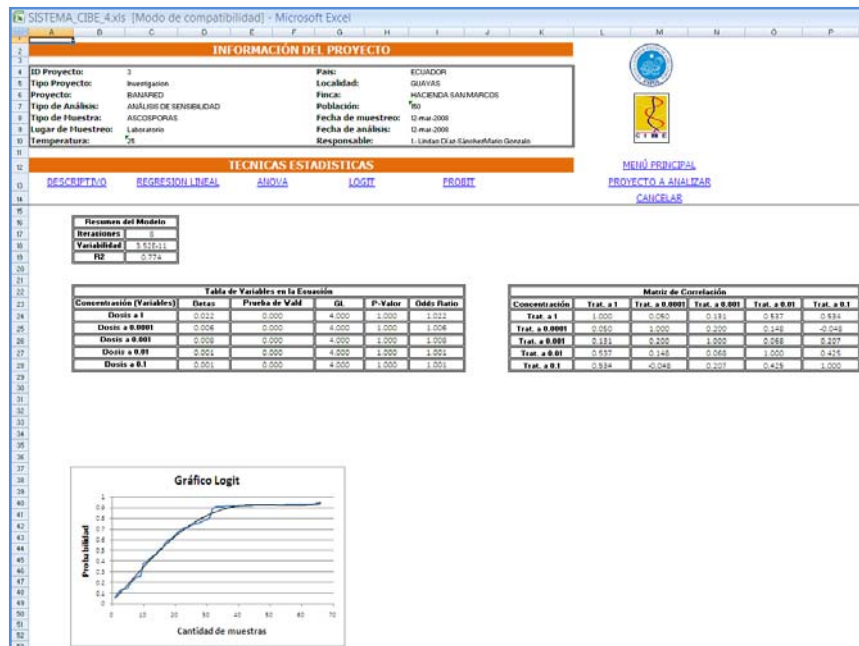


FIGURA 7. MÓDULO DE REGRESIÓN LOGIT.

Módulo de Regresión Probabilística (PROBIT).

El módulo de regresión probabilística PROBIT utiliza como medida de predicción o de discriminación la función de Distribución Normal Estándar obteniendo de esta forma la función probabilística, con el cual se clasifica los datos en dos grupos, mediante el uso de la probabilidad al evento 0 = no existe crecimiento del tubo germinativo y 1 = existe crecimiento del tubo germinativo.

El módulo presenta un resumen, el cual muestra el número de iteraciones efectuadas, la variabilidad o error del modelo con el que se mide la precisión y el R cuadrado que es la robustez del modelo.

Se realiza la tabla de variables de la ecuación Probit, en el cual se calcula los betas de cada dosis para el modelo, además se calcula el

test de WALD y su *P-Valor* el que permite decidir si una variable se debe o no usar en el modelo, también se realizan los cálculos del factor de riesgo (Odds Ratio) para cada dosis, con este factor se puede determinar si una dosis tuvo un mayor efecto de inhibición.

Adicional se muestra el gráfico de la función Probit que muestra las probabilidades de cada observación que se ajusta al modelo encontrado.

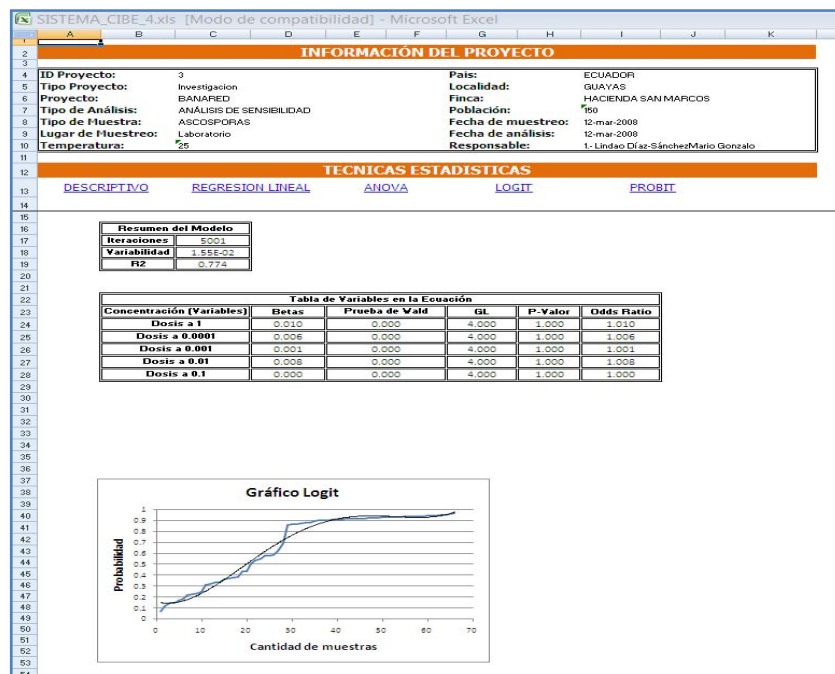


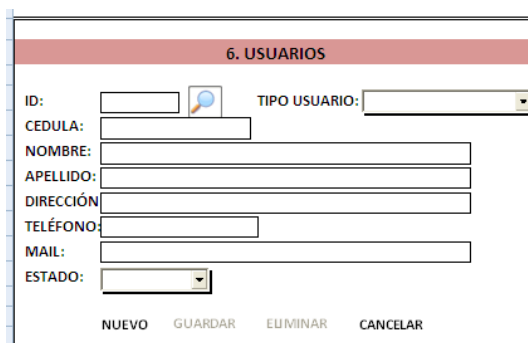
FIGURA 8. MÓDULO DE REGRESIÓN PROBIT.

➤ Componente Seguridad

En el siguiente componente se establece los módulos de seguridad que permite que la aplicación sea utilizada de manera correcta y la información sea menos sensible y más confidencial.

Módulo de Usuarios.

Es una aplicación desarrollada para administrar los niveles de acceso a las diferentes actividades que realicen las personas que utilicen el sistema, en el cual se ingresa información acerca del usuario y su rol como encargado de los proyectos del centro.



The image shows a web form titled "6. USUARIOS". It contains the following fields and controls:

- ID:
- TIPO USUARIO:
- CEDULA:
- NOMBRE:
- APELLIDO:
- DIRECCIÓN:
- TELÉFONO:
- MAIL:
- ESTADO:

At the bottom of the form, there are four buttons: NUEVO, GUARDAR, ELIMINAR, and CANCELAR.

FIGURA 9. MÓDULO DE USUARIOS.

Iniciación de la herramienta informática.

A continuación se detalla paso a paso el funcionamiento de la herramienta.

1. Para abrir la aplicación dar un click con el mouse en el archivo de Excel llamado SYSCIBE.
2. Se presenta a continuación el ingreso del primer usuario, el que será el usuario administrador de la aplicación. Ingresar los datos solicitados en la ventana y luego guardar la información, este emitirá un mensaje "Proceso realizado con éxito" y le indicará que su usuario y contraseña serán su correo electrónico y su número de cedula respectivamente.

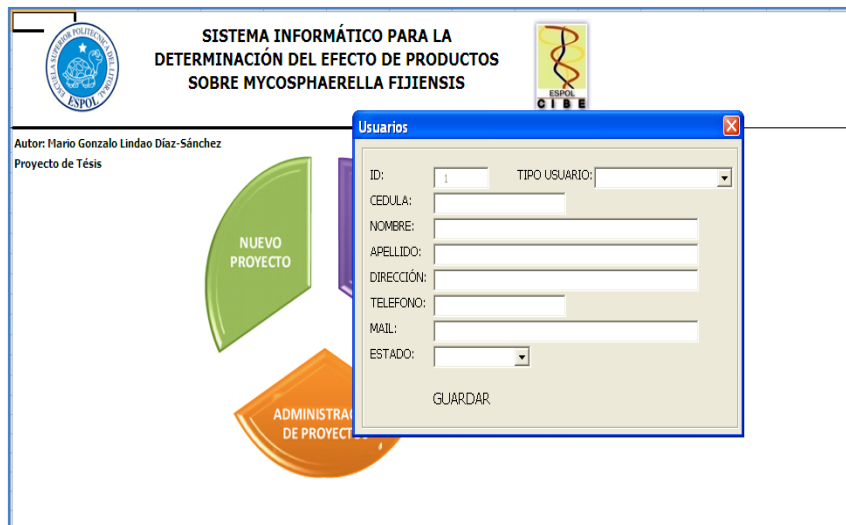


FIGURA 10. INGRESO DE USUARIOS.

3. Debe ingresar el usuario y clave para el uso de la aplicación.

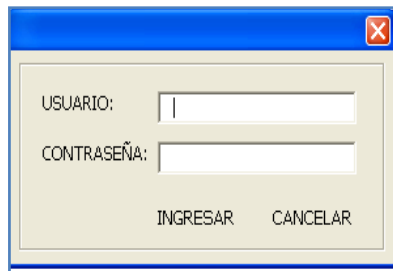


FIGURA 11. VALIDACIÓN DE USUARIO.

4. La pantalla inicial de la aplicación muestra las actividades a realizar durante este en el sistema, antes de iniciar un proyecto debemos ingresar cierta información que nos ayudara, damos click en Administración de Proyectos.

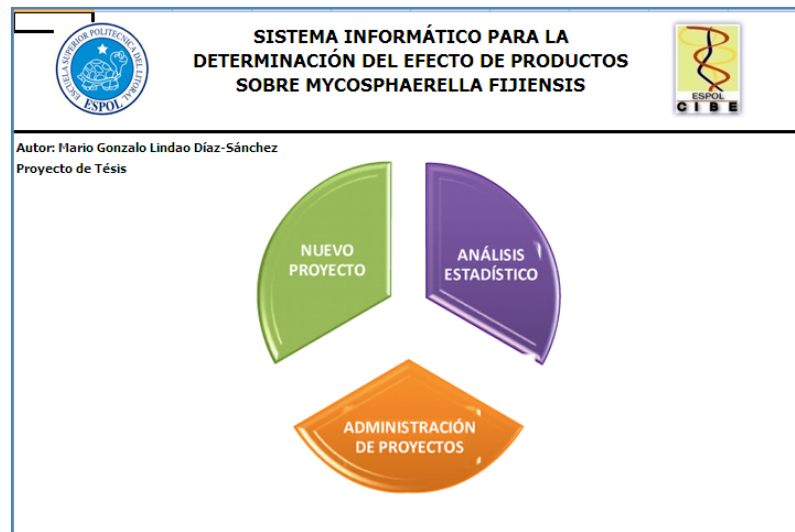




FIGURA 12. PANTALLA INICIAL.

5. En la parte superior se encuentra el menú principal de administración, en donde se seleccionará la opción “Tipo de Proyecto” y dando click en “Nuevo” se activarán los campos en donde escribiremos el “Detalle” del tipo de proyecto y el estado en que se encuentra dicho registro Activo o Inactivo (Ver Ayuda) y luego se selecciona la opción de “Guardar”.

6. Para modificar un registro, primero buscamos el registro presionando el icono de buscar  el cual mostrará una ventana con las lista de los registros existentes Grafico 14, damos doble click sobre el registro a ser modificado, para luego escribir el nuevo valor en el campo “Detalle”, finalmente guardamos el registro.

7. Para eliminar un registro, primero buscamos el registro presionando el icono de buscar  el cual mostrará una ventana con las lista de los registros existentes Grafico 14, damos doble click sobre el registro a ser eliminado y

finalmente seleccionamos la opción “Eliminar”.

The screenshot shows a web application interface for project management. At the top, there is a header with the logo of the National Institute of Health (INSPI) and the word 'ADMINISTRACIÓN'. Below the header, there are navigation tabs: TIPO DE PROYECTO, TIPO DE ANÁLISIS, TIPO DE MUESTRA, TIPO DE PRODUCTO QUÍMICO, TIPO DE DOSIS, and USUARIOS. The main content area displays a form titled '1. TIPO DE PROYECTO'. The form has three input fields: 'ID' with the value '1', 'DETALLE' with the value 'BANAREO', and 'ESTADO' with a dropdown menu showing 'Activo'. Below the form are four buttons: 'NUEVO', 'GUARDAR', 'ELIMINAR', and 'CANCELAR'.

FIGURA 13. ADMINISTRACIÓN DE PROYECTOS.

id	detalle	estado
1	BANAREO	Activo
3	Investigacion	Activo

FIGURA 14. LISTA DE REGISTROS.

8. Continuamos ingresando registros de igual manera para las opciones en el menú de administración “Tipo de Análisis”, “Tipo de Muestra”, “Tipo de Producto Químico” y “Tipo de Dosis”, los que son necesarios de ingresar para poder construir un proyecto.

The screenshot displays the 'ADMINISTRACIÓN' web application interface. At the top, there is a header with the application title 'ADMINISTRACIÓN' and a logo on the left. On the right, there are navigation links: 'MENU PRINCIPAL' and 'PROYECTOS'. Below the header, a horizontal menu contains several options: 'TIPO DE PROYECTO', 'TIPO DE ANÁLISIS', 'TIPO DE MUESTRA', 'TIPO DE PRODUCTO QUÍMICO', 'TIPO DE DOSIS', and 'USUARIOS'. The main content area features three distinct forms, each with a colored header:

- 1. TIPO DE PROYECTO** (light blue header): Includes fields for 'ID:', 'DETALLE:', and 'ESTADO:'. Below the fields are buttons for 'NUEVO', 'GUARDAR', 'ELIMINAR', and 'CANCELAR'.
- 2. TIPO DE ANÁLISIS** (light red header): Includes fields for 'ID:', 'DETALLE:', and 'ESTADO:'. Below the fields are buttons for 'NUEVO', 'GUARDAR', 'ELIMINAR', and 'CANCELAR'.
- 3. TIPO DE MUESTRA** (light orange header): Includes fields for 'ID:', 'DETALLE:', 'DESCRIPCION:', and 'ESTADO:'. Below the fields are buttons for 'NUEVO', 'GUARDAR', 'ELIMINAR', and 'CANCELAR'.

FIGURA 15. ADMINISTRACIÓN DE PROYECTOS.

9. Regresamos al menú principal y seleccionamos la opción de “Nuevo Proyecto”.

10. Se procede a construir un nuevo proyecto, para esto se necesita llenar los siguientes campos: Tipo de proyecto, Nombre del Proyecto, Tipo de Análisis, Tipo de Muestra, Lugar de Muestreo (Campo, Invernadero, Laboratorio), País, Localidad, Nombre de la Finca, Población, Fecha de Muestreo, Fecha de Análisis y Responsable (por defecto es la persona que ingreso a la aplicación) y luego selecciona la opción “Guardar”.


NUEVO PROYECTO


ID PROYECTO: 8 PAIS: LOCALIDAD: NOMBRE PROYECTO: NOMBRE DE FINCA: POBLACIÓN: TIPO DE ANÁLISIS: TIPO DE MUESTRA: FECHA DE MUESTREO: 15/01/2010 LUGAR DE MUESTREO: FECHA DE ANÁLISIS: 15/01/2010 TEMPERATURA: RESPONSABLE: 1.-Lindao Diaz - Sánchez Mario Gonzalo

GUARDAR

FIGURA 16. NUEVO PROYECTO.



11. Al cerrar la ventana de Nuevo Proyecto, esta se direcciona a la hoja de ingreso de datos, en donde aparecerá en la parte superior la información del proyecto, luego se procede a construir la matriz de datos, para esto necesitamos llenar los campos “Tipo de Fungicida”, “Cantidad De Dosis” y “Tipo de Dosis”.

12. Para ingresar los datos debemos ingresar las dosis utilizadas durante el proceso de sensibilidad, para esto elegimos un número determinado de dosis por ejemplo 5 esto significa que abra un máximo de 5 dosis para ingresar, luego seleccionamos los tipos de dosis por ejemplo Dosis al 5% y presionamos el icono de ingresar  , esto creara una columna adicional en la cual procederemos a ingresar los datos; y así con las demás dosis, finalmente digitamos los datos en sus respectiva dosis.

13. Para eliminar una columna de dosis, fijamos la fila de “COD.” seleccionando la celda referente a la Dosis a eliminar y procedemos a presionar el icono de eliminar .

INFORMACIÓN DEL PROYECTO

Id Proyecto: 7	País: ECUADOR
Tipo Proyecto: BANARED	Localidad: LOS RIOS - VALENCIA
Proyecto: PROYECTO BANKIT	Finca: CONVENCIONAL
Tipo de Análisis: ANÁLISIS DE SENSIBILIDAD	Población: 150
Tipo de Muestra: ASCOSPORAS	Fecha de muestreo: 6-dic-2009
Lugar de Muestreo: Campo	Fecha de análisis: 6-dic-2009
Temperatura: 25	Responsable: 1.- Lindao Díaz - SánchezMario Gord

DATOS DEL PROYECTO

Tipo de Fungicida:

Cantidad de Dosis:
Tipo de Dosis:
+ X

[MENÚ PRINCIPAL](#)
[GUARDAR DATOS](#)
[CANCELAR](#)



COD.	0	10	11	13	17
No.	CONTROL	5	1	10	0.1
1	85	5	25	0	24
2	80	5	15	0	10
3	66	15	22	0	25
4	96	5	26	0	35
5	53	5	15	0	14
6	74	10	5	0	15
7	90	10	16	2	16
8	75	10	18	0	20
9	74	10	15	0	24
10	66	5	20	0	26
11	74	5	20	0	25
12	75	5	27	5	25
13	75	5	10	3	20
14	89	7	8	0	23
15	66	0	25	0	36
16	66	0	11	0	16
17	60	0	12	0	10
18	98	0	20	0	20
19	92	0	15	0	30
20	93	0	10	0	25
21	58	0	18	0	26
22	67	0	5	0	25
23	92	0	8	0	15
24	54	0	10	0	21
25	93	0	7	0	37

FIGURA 17. INGRESO DE DATOS.

14. Finalmente guardamos los datos digitados en “Guardar Datos”, este emitirá un primer mensaje “Posibles datos mal ingresados, desea revisar?. ACEPTAR para revisar y CANCELAR para continuar” Grafico 18, si elige la opción de revisar ciertas celdas se encontraran sombreadas lo cual indica un posible mal ingreso, caso contrario continuamos con el proceso de almacenamiento.

INFORMACIÓN DEL PROYECTO

Id Proyecto: 7	Pais: ECUADOR
Tipo Proyecto: BANARED	Localidad: LOS RIOS - VALENCIA
Proyecto: PROYECTO BANKIT	Finca: CONVENCIONAL
Tipo de Análisis: ANÁLISIS DE SENSIBILIDAD	Población: 150
Tipo de Muestra: ASCOSPORAS	Fecha de muestreo: 6-dic-2009
Lugar de Muestreo: Campo	Fecha de análisis: 6-dic-2009
Temperatura: 25	Responsable: 1.- Lindao Diaz - SánchezMario Gor

DATOS DEL PROYECTO

Tipo de Fungicida: 3.- AZOXISTROBINA + X

Cantidad de Dosis: 5 **Tipo de Dosis:** MENÚ PRINCIPAL


[GUARDAR DATOS](#)
[CANCELAR](#)

COD.	0	10	11	13	17
No.	CONTROL	5	1	10	0.1
1	85				
2	80				
3	66				
4	96				
5	53				
6	74				
7	90				
8	75	10	18	0	20
9	74	10	15	0	24
10	66	5	20	0	26
11	74	5	20	0	25
12	75	5	27	5	25
13	75	5	10	3	20
14	89	7	8	0	23
15	66	0	25	0	36
16	66	0	11	0	16
17	60	0	12	0	10
18	98	0	20	0	20
19	92	0	15	0	30
20	93	0	10	0	25
21	58	0	18	0	26
22	67	0	5	0	25
23	92	0	8	0	15
24	54	0	10	0	21
25	93	0	7	0	37

Mensaje

Posibles datos mal ingresados, desea revisar?. ACEPTAR para revisar y CANCELAR para continuar

FIGURA 18. VERIFICACIÓN DE DATOS.

15. Para consultar los datos ingresados, seleccionamos el icono de buscar , y dando doble click sobre el registro deseado aparecerá en la hoja los datos ingresados.

16. La ultima de etapa de la aplicación es la presentación de los resultados, en el menú principal seleccionamos la opción de “Análisis Estadístico”.

17. Para iniciar el análisis estadístico de los datos damos un click en la opción “PROYECTO A ANALIZAR”, el cual mostrara una ventana con la lista de proyectos ingresados, seleccionamos y damos doble click sobre el registro a analizar, este nos presentará la información sobre el proyecto.

18. En la parte inferior se encuentra el menú de las Técnicas o Métodos Estadísticos desarrollados en la aplicación que son: “Descriptivo”, “Regresión Lineal”, “ANOVA”, “LOGIT”, “PROBIT”, dando un click sobre cualquiera de las opciones este genera el análisis automáticamente.

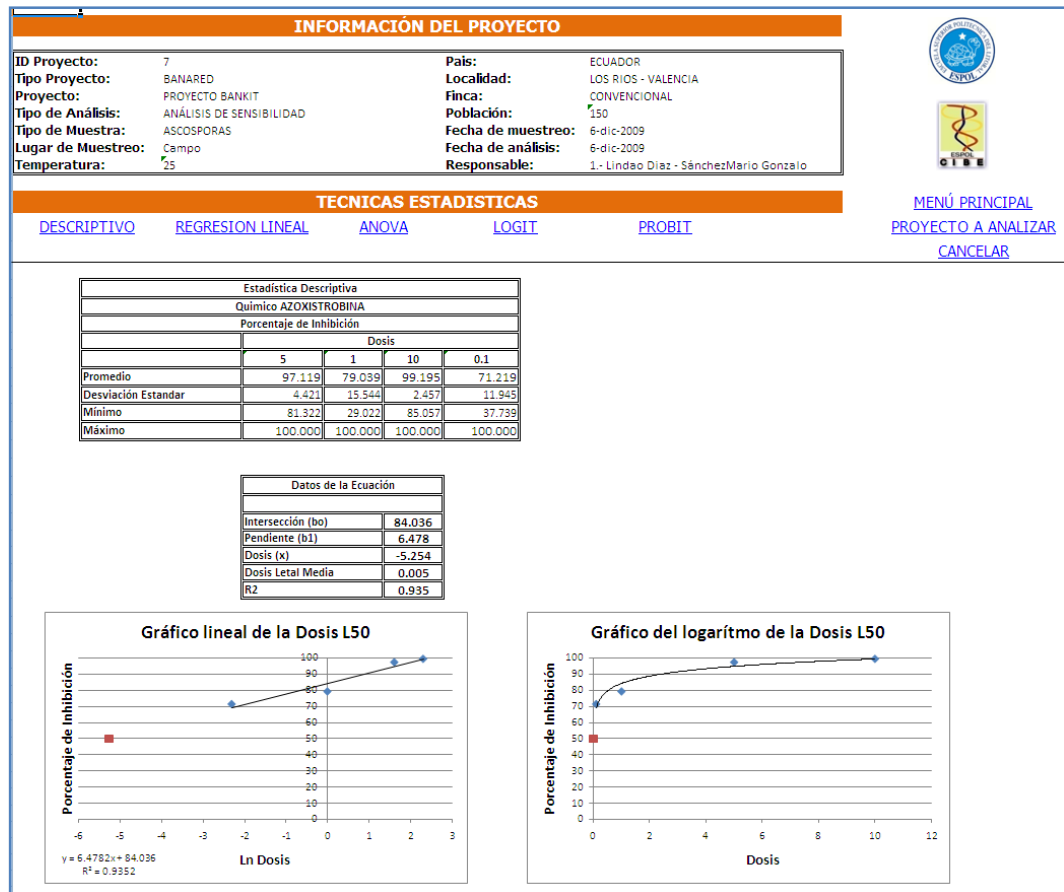



FIGURA 19. ANÁLISIS ESTADÍSTICOS E INFORMES.

19. Cierre la aplicación como cualquier otra, el botón en la esquina superior izquierda, no es necesario utilizar la opción guardar .

Ayuda de la herramienta informática.

1. El estado de Activo nos indica que el registro estará en uso, caso contrario el registro no será utilizado durante la etapa de resultados.
2. Es necesario ingresar todos los campos para poder hacer un buen registro de la información, el cual emitirá un mensaje “Proceso fue realizado con éxito”, caso contrario mostrará el mensaje de “Ingrese la información requerida”.
3. Si existe un registro utilizado en varias proyectos y se desea eliminar este emitirá “El elemento está siendo utilizado, imposible eliminar”, debido a que está relacionado con resultados anteriores.
4. Para una mayor agilidad sobre la aplicación puede usar también los accesos directos presionando las hojas o fichas de Excel “INICIO”, “INGRESO DE DATOS”, “ANALISIS DE DATOS” y “ADMINISTRACIÓN”.
5. En la hoja de Administración se encuentra la opción de Usuarios en donde se podrá ingresar algún otro usuario que también maneje la aplicación.
6. Con el fin de evitar errores en los ingresos las hojas y celdas están bloqueadas, solo en la hoja de Análisis de Datos se pueden copiar los resultados para poder emitir los reportes respectivos.
7. Al cerrar el archivo de Excel SYSCIBE emitirá un mensaje “Desea guardar el archivo” el cual el usuario tomará la opción “No”, debido a que no es necesario que utilice dicha acción, porque todo está almacenado en la base de datos.

BIBLIOGRAFÍA

- [1] BAQUERO M., FERNÁNDEZ G. y GARZÓN P. (2004) *El banano en el Ecuador: Estructura de mercados y formación de precios*, DIRECCIÓN DE INVESTIGACIONES ECONÓMICAS DEL BANCO CENTRAL DEL ECUADOR, Guayaquil-Ecuador.
- [2] BERRY M., y LINOFF G., (2004) *Data Mining Techniques*, Segunda Edición, Publicación por Wiley, Indianápolis – EEUU.
- [3] CABALLERO J. (2008), *Modelos de regresión logística incondicional*, Página Web: <http://saei.org/hemero/epidemiol/nota5.html#1>, Octubre 2009.
- [4] CALLE C., (2004) *Estimación de un modelo Probit y Logit para la evolución viral de los pacientes con VIH caso hospital de infectología José Daniel Rodríguez de Guayaquil*, Tesis de Grado de ESPOL, Guayaquil-Ecuador.
- [5] CARLIER, J. y DE WAELE, D. (2003) *Evaluación global de la resistencia de los bananos al marchitamiento por fusarium, enfermedades de las manchas foliares causados por Mycosphaerella y Nematodos*, Impreso Arceaux 49, Montpellier – Francia.

- [6] CASTRO R. y ZAVALA A., (1998) *Programa en "basic" para el cálculo de $dI50$ por el método de probits*, Pagina Web: http://sisbib.unmsm.edu.pe/BvRevistas/Medicina_Experimental/v15_n1-2/progra_basic.htm, Octubre 2009.
- [7] FREITEZ J., (2007) *Desarrollo de un modelo predictivo del brote de la Sigatoka negra para las plantaciones de plátano al sur del lago de Maracaibo*, Tesis de Maestría de Universidad de los Andes, Mérida - Venezuela.
- [8] FREUD J., MILLER I., y MILLER M., (2000) *Estadística matemática con aplicaciones*, Sexta Edición, Publicación por Pearson Educación, México.
- [9] FUENTES E., SORIA J., y SORO G. (2003) *Proyecto de agroecoturístico de banano orgánico en la parroquia Juan Gómez Rendón*, Tesis de Grado de ESPOL, Guayaquil-Ecuador.
- [10] GALINDO, E., (1999) *Estadística para la Administración y la Ingeniería*, Primera Edición, Publicación Graficas Mediavilla Hnos., Ecuador.
- [11] GARCIA J., RODRIGUEZ J. y BRAZALES A., (1999) *Aprenda Visual Basic 6.0*, Manual de Programación, Publicación por San Sebastián, NAVARRA.
- [12] GIRON J., MARTINEZ M. y GARCIA J., (2003) *Modelo para la regresión robusta politómica*, Departamento de estadística e investigación operativa, Málaga-España.

- [13] GONZALEZ J, (2002) *Modelo Logit*, Página Web: www.docirs.cl/scoring_htm/Logit_function.htm, Agosto 2009.
- [14] GUTIÉRREZ J., GUTIÉRREZ O. y NIETO D. (2003), *Resistencia a Benomil y Tiabendazol en aislamientos de Colletotrichum gloeosporioides (Penz.) Penx. y Sacc. Obtenidos de mango en cinco regiones de México*, Revista Mexicana de Fitopatología, Oregon - México.
- [15] JELSOFT ENTERPRISES LTD., (2009) *Vb-mundo-programacion visual*, Pagina Web: <http://www.vb-mundo.com>, Octubre 2009.
- [16] JONES D.R., (1999) *Diseases of banana, abaca and enset*, CABI Publishing, Oxon - United Kingdom of Great Britain.
- [17] MICROSOFT, (2007) *Microsoft developer network*, Pagina Web: <http://msdn.microsoft.com/es-es/library/microsoft.office.tools.excel.aspx>, Octubre 2009.
- [18] MONTGOMERY D. (2003) *Diseño y análisis de experimentos*, Segunda Edición, Publicación por Limusa Wiley, Arizona–EEUU.
- [19] NÚÑEZ R. (1989) *El cultivo de banano*, Ministerio de Agricultura y Ganadería, Sección cooperativas, Ecuador.
- [20] ORJEDA, G. (1998) *Evaluación de la resistencia de los bananos a las enfermedades de Sigatoka y marchitamiento por fusarium*, Impreso por Louis – Jean.

- [21] OSORIO J., (2007) *Análisis del procedimiento para la determinación de la dl50 (dosis letal media) a través del método de Probit en un bioensayo*, Pagina Web: <http://www.monografias.com/trabajos14/dosis-letal/dosis-letal.shtml#RESUM>, Octubre 2009.
- [22] PELÁEZ J., VÁSQUEZ L., DÍAZ T., CASTAÑEDA D., RODRÍGUEZ E. y ARANGO R. (2006), *Use of a micro tittle plate dilution assay to measure activituse activity of antifungal compounds against mycosphaerella fijiensis morelet*, Revista Facultad Nacional de Agronomía, Medellín - Colombia.
- [23] RENCHER, A., (2002) *Methods of Multivariate Análisis*, Segunda Edición, Publicación por Wiley-Interscience, New York – EEUU.
- [24] RUIZ O., (2009) *El mejoramiento del proceso de análisis de sensibilidad en banano a productos químicos, como iniciativa de autogestión del CIBE, con miras a obtener una acreditación*, Tesis de Magíster En Gestión De La Productividad Y La Calidad, Guayaquil-Ecuador.
- [25] UCA, 7 de Agosto del 2005, *SPSS 10 Guía para análisis de Datos*, Pagina Web: www.uca.es/serv/ai/formacion/spss/Pantalla/verguia.pdf, Diciembre 2008.
- [26] VARGAS M., 2003, *Estimación de un modelo Probit Multivariante*, Tesis de Grado, MPRA.

- [27] VEGA G., (1997), *La Sigatoka Negra (Mycosphaerella fijiensis Morelet) en el plátano*, Pagina Web: <http://www.monografias.com/trabajos33/sigatoka-negra/sigatoka-negra.shtml>, Noviembre 2008.
- [28] VILLON R., 2006, *Proceso estadístico para discriminar datos aberrantes usando un Modelo de regresión logística*, CENACAD - CISE.
- [29] VIVERES J. y CASTALO J. (2006), *Evaluación In Vitro de extractos vegetales sobre Mycosphaerella fijiensis morelet*, Universidad de Caldas, Caldas-Colombia.
- [30] WITTEN I., y FRANK E., (2005) *Data mining practical machine learning tolos and techniques*, Segunda Edición, Publicación por Morgan Kaufmann, EEUU.
- [31] YONSON B., (2005) *Uso de regresión multinomial logística para determinar el perfil del estudiante de la ESPOL según preferencias alimenticias*, Tesis de Grado de ESPOL, Guayaquil-Ecuador.