UN METODO DE ANALISIS PARA VARIABLES TIPO QUANTAL (PROBIT)

Por J.P.López de la Fuente(*), C.Martínez Abad(**), E.Viñuela Sandoval(***)

1.- Introducción

Muchos de los problemas que se plantean en el campo de la investigación biológica y tecnológica, tratan de la relación existente entre un estímulo (insecticida, droga, tiempo que un animal se somete a un aprendizaje, etc.) y una respuesta (efecto que produce en el animal, el insecticida o la droga; medida de los conocimientos prácticos adquiridos, que muestra el animal como consecuencia del aprendizaje, etc.).

Cuando la respuesta no está exactamente determinada por el conocimiento del estímulo, y repeticiones de experimentos u observaciones para valores fijos de las variables independientes, no dan siempre la misma magnitud de respuesta, es necesario recurrir a la Estadística para la interpretación de los resultados. Con este fin, se han desarrollado entre otras técnicas el Análisis de Regresión en sus muchas manifestaciones, una de las cuales estudia un tipo especial de respuesta que se conoce como respuesta quantal o de todo o nada, por que no permite una gradación y sólo puede expresarse como "ocurriendo" o "no ocurriendo". El ejemplo más típico es la muerte: en muchos estudios de insecticidas el único objetivo del ensayo es ver si el insecto está o no muerto. Otros ejemplos de este tipo de respuesta son, el fallo en la germinación de las esporas en estudios de fungicidas; o en los estudios con drogas puede ser la curación de cierta situación sin considerar curas parciales, etc.

En cada situación dosis-respuesta hay que considerar dos componentes: el estímulo (por ejemplo un producto químico) y el sujeto (por ejemplo un insecto). El estímulo se aplicará al sujeto con cierta intensidad (expresada en unidades de concentración, tiempo, etc.) y bajo condiciones ambientales controladas; como consecuencia, el sujeto manifiesta una respuesta (cambio de color, crecimiento, muerte, etc.).

Si la respuesta es de tipo quantal, el que se dé o no, dependerá de la intensidad del estímulo. Para cada sujeto bajo condiciones controladas existirá un cierto nivel de intensidad, por debajo del cual la respuesta no se da y por encima del cual, sí se da; a este valor se le llama tolerancia y variará de un sujeto a otro dentro de la población utilizada.

2.- Descripción teórica

Para estudiar los datos de respuesta quantal, hay que conocer la distribución de frecuencias de la tolerancia en la población. Esta distribución medida en una escala natural raramente es simétrica, pero realizando una simple transformación de escala consistente en expresar en forma logarítmica las concentraciones de las dosis aplicadas (estímulos), y con la salvedad de no ser nunca negativos dichos logaritmos, se convierte en una distribución de tipo normal N (μ, σ):

(*) Becario Centro de Cálculo de la Universidad Complutense
(**) Analista del Centro de Cálculo de la Universidad Complutense
(***) I.Agrónomo, Dpto. Entomología Agrícola. E.T.S.I. Agrónomos

$$dP = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[- \frac{(x-\mu)^2}{2\sigma^2} \right] dx \qquad (1)$$

siendo μ = media de la distribución
σ = varianza de la distribución

DETERMINACION DE LA DOSIS MEDIA

Para caracterizar la efectividad de un estímulo en relación a la respuesta, se utiliza "la dosis letal media" o para el caso más general en que se incluyan otras respuestas diferentes a la muerte "la dosis efectiva media", que es la dosis que origina una respuesta en la mitad de la población. Se representan respectivamente por DL_{50} y DE_{50}.

Para su deteminación vamos a realizar una serie de transformaciones en la ecuación (1). La probabilidad, por tanto, se expresará como:

$$P = \int_{-\infty}^{X} \frac{1}{\sigma \sqrt{2\pi}} \exp\left[- \frac{(x-\mu)^2}{2\sigma^2} \right] dx \qquad (2)$$

Que mediante la transformación $x = \mu + \sigma u$, se convierte en:

$$P = \int_{-\infty}^{(X-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} \exp\left[- 1/2\, u^2 \right] du \qquad (3)$$

Gaddum propuso medir la probabilidad de las respuestas, en una escala transformada para lo cual definió la N.E.D. (Desviación equivalente normal) de cualquier valor P entre 0 y 1 como la abscisa que corresponde a una probabilidad P en una distribución normal con media 0 y varianza 1, es decir, la Y tal que:

$$P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Y} \exp\left[-1/2\, u^2 \right] du \qquad (4)$$

Considerando las ecuaciones (3) y (4), obtenemos:

$$Y = (X - \mu) / \sigma \qquad (5)$$

Con lo que la relación entre la dosis y la N.E.D. de la probabilidad de respuesta a esa dosis es una linea recta.

Bliss sugirió una nueva transformación, para lo cual definió el probit de P respecto de Y (PROBIT = probability unit), como el límite superior de la integral que verifica:

$$P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Y-5} \exp\left[- 1/2\, u^2 \right] du \qquad (6)$$

Para cualquier P el probit es el N.E.D. aumentado en 5 unidades. Esta transformación tiene la ventaja de que los probits son siempre positivos a menos que P sea excesivamente pequeño y la N.E.D. es negativa si P es inferior al 50%.

De las ecuaciones (5) y (6) obtenemos:

$$Y - 5 = (x - \mu) / \sigma \qquad\qquad Y = 5 + (1/\sigma)(x - \mu)$$

Que expresa la relación entre el probit de la proporción de respuesta esperada y la dosis.

Normalmente se escribe de la forma:

$$Y = \alpha + \beta x \qquad \text{siendo} \qquad \sigma = 1/\beta$$

$$\mu = (5 - \alpha)/\beta$$

La dosis media efectiva se estima como el valor de x que hace Y=5, luego el problema estriba en estimar la ecuación de regresión $Y = \alpha + \beta x$ siendo

x = dosis aplicada en logaritmos;
Y = respuesta calculada en unidades probits

En todos los problemas de experimentación insecticida se utiliza un control, que no es más que un conjunto de individuos a los que se aplica una concentración de la dosis igual a cero y cuya respuesta influirá en el cálculo de la línea de regresión.

Para hallar los valores de α y β, hay que maximizar la función:

$$\sum_{i-1}^{n} r_i \log P_i + (n_i - r_i) \log Q_i \qquad (7)$$

Donde $Q_i = 1 - P_i$

$$P_i = C + (1-C) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha + \beta x_i - 5} e^{-u/2} \, du$$

Con:

i = 1 ..... n° de dosis diferentes aplicadas.
C = respuesta dada por el control.
x_i= valor del nivel i-ésimo del producto aplicado (estímulo)
r_i= respuestas (número de individuos muertos) que corresponden al nivel i-ésimo del producto aplicado (estímulo).
n_i= número de individuos sometidos a ensayo en el nivel i-ésimo de producto aplicado (estímulo).
P_i= probabilidad de respuesta al nivel i-ésimo de producto aplicado (estímulo).
Q_i= probabilidad de no respuesta a ese mismo nivel.

3.- Descripción de la Biblioteca de programas en A.P.L.

La función principal, y las funciones auxiliares pueden encontrarse en la Biblioteca 11, en el workspace de nombre PROBITANAL. La descripción del workspace puede conseguirse ejecutando la función PROBITDES.

La función principal comienza resumiendo los datos suministrados:

- Número de dosis diferentes
- Valores de las dosis aplicadas
- Respuesta
- Número de individuos tratados
- Probabilidad entre las respuestas y los sujetos observados
- Valores del probit empírico para las dosis

A continuación, utilizando el método de los mínimos cuadrados, se ajusta una recta a la nube de puntos que constituyen los logaritmos de las dosis, y los probits de las respuestas para esas dosis. Se obtienen así, unos valores de α y β que sirven como punto inicial al método Newton-Raphson utilizado para maximizar la función (7).

Se calcula el test de la X^2, dando el valor de la función de densidad y el valor de la probabilidad acumulada dentro del intervalo.

El programa se detiene para que el usuario, consultando tablas de la distribución de la X^2, dé respuesta positiva o negativa a este test (SI o NO).
Una vez obtenida una solución, el programa continúa obteniendo los valores de $\overline{x}$, $\overline{y}$, Sx^2 y la pendiente de la recta, para pasar al cálculo de X^2.

Dependiendo de la respuesta, el programa continúa por un camino u otro.

Para X^2 no significativa:

- Cálculo del intervalo de confianza para la DL_{50}, dando el límite superior e inferior.

- Produce un cuadro con la DL_{50}, DL_{90}, DL_{95}, DL_{99}, expresándolas en logaritmos y en las unidades de experimentación.

- Dibujo de la recta. Se solicita del usuario el intervalo en la x en el que desea el dibujo, y se le dan instrucciones de manejo del terminal para conseguirlo.

- Cálculo de errores. Se calculan los errores de la pendiente, de de y, del log DL_{50}, log DL_{90}, ... y de la DL_{50}, DL_{90}, ...

Para X^2 significativa:

- Produce un cuadro con la DL_{50}, DL_{90}, DL_{95}, DL_{99}, expresándola en logaritmos y en las unidades de experimentación.

- Cálculo del factor de heterogeneidad.

- Dibujo de la recta y de las bandas de confianza. Se le pide al usuario el intervalo en el que quiere el dibujo, y el valor de la T de Student para los grados de libertad que se indican y el 95% de probabilidad, para el cálculo de las bandas, y se le dan instrucciones de manejo del terminal.

- Cálculo del intervalo de confianza para la DL_{50}.

- Cálculo de errores.

El programa acaba dando el tiempo invertido en el análisis.

Es importante resaltar que todos los cálculos se hacen utilizando logaritmos decimales, y sólo se obtiene el antilogaritmo una vez que se obtiene la respuesta en logaritmos. Este dato es fundamental, sobre todo en el cálculo de errores.

4.- Ejemplos

El primer ejemplo, cuyo nombre es CALANDRA, muestra un análisis con χ^2 significativa.

Se trata de un experimento para estudiar la toxicidad del óxido de etileno en la Calandra granaria.

Las dosis se midieron en mg/100 ml. El cuadro construido por el experimentador sobre 10 tratamientos fue:

										CONTROL
Muertos	23	30	29	22	23	7	12	17	10	0	0
Tratados	30	30	31	30	26	27	31	30	31	24	50
Lg.Dosis	.394	.391	.362	.322	.314	.260	.225	.199	.167	.033	.000

La salida de ordenador se obtuvo ejecutando

```
        CONTROL  PROBIT  M

donde CONTROL ← 0
y     M [1; ] ← 23 30 29 22 23 7 12 17 10 0
      M [2; ] ← 30 30 31 30 26 27 31 30 31 25
      M [3; ] ← .394 .391 .362 .322 .314 .260 .225 .199 .161 .033
```

```
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
                    ANALISIS PROBIT
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

NOMBRE DEL EXPERIMENTO: CALANDRA

NUMERO DE DOSIS DIFERENTES SIN LA DE CONTROL: 10
VALORES DE LAS DOSIS APLICADAS
    .39400    .33100    .35200    .32200    .31400    .26000    .22500    .19900    .16700    .03300

RESPUESTAS EN DICHAS DOSIS
23 30 29 22 23 7 12 17 10 0

NUMERO DE INDIVIDUOS TRATADOS EN DICHAS DOSIS
30 30 31 30 26 27 31 30 31 24

PROBABILIDAD ENTRE LAS RESPUESTAS Y LOS SUJETOS OBSERVADOS
    .76667   1.00000    .33548    .73333    .88462    .25926    .38710    .56667    .32258    .00000

VALORES DEL PROBIT EMPIRICO PARA DICHAS DOSIS
   5.727651004.00000   6.51824   5.62259   6.19850   4.35469   4.71354   5.16756   4.53993-994.00000

VALORES DE ALFA,BETA Y CONTROL EN LA 1 APROXIMACION
ALFA= 3.253123     BETA= 7.497863     CONTROL=  .000000
PROBITS ESPERADOS
   6.20728   6.18479   5.96735   5.66744   5.60745   5.20257   4.94014   4.74520   4.50527   3.50055



-----------------------------------------------------------------
 SOLUCION:    ALFA=2.937760143
              BETA=8.637353774
              CONTROL=0
PROBITS ESPERADOS
   6.34088   6.31497   6.06448   5.71899   5.64989   5.18347   4.88116   4.65659   4.38020   3.22279
-----------------------------------------------------------------
```

X MEDIA =0.2689663745
Y MEDIA = 5.260917873
SX2 = 1.039975704
PENDIENTE DE LA RECTA = 8.637353774

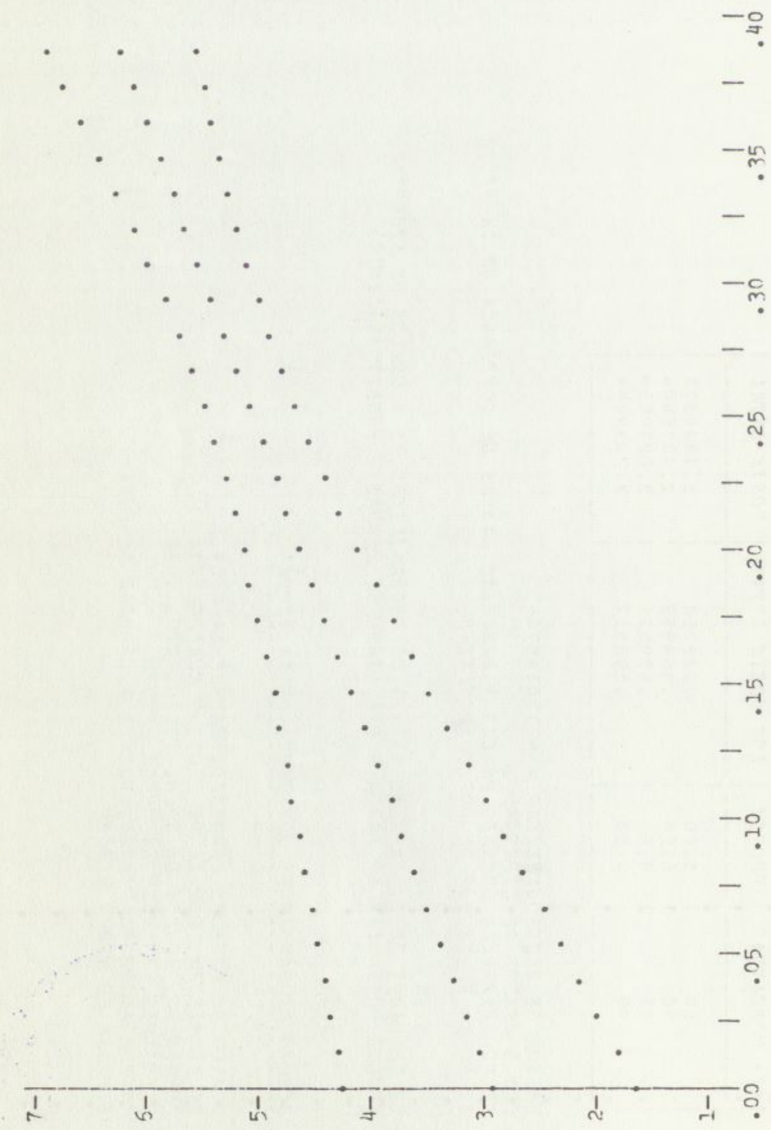
VALOR DE LA CHI2=33.20439014
CON 8 GRADOS DE LIBERTAD
CALCULO DE LA DISTRIBUCION PARA LA CHI2
MEDIANTE METODOS APROXIMATIVOS
EL VALOR DE LA FUNCION DE DENSIDAD ES :2.350030964E¯5
EL VALOR DE LA PROBABILIDAD ACUMULADA ES ;0.9999443815
ES SIGNIFICATIVA PARA USTED LA CHI2: SI

°/° MORTANDAD	PROBIT	LOG DOSIS LETAL	DOSIS LETAL
50	5.00	.238758	1.7328393
90	6.28	.386952	2.4375404
95	6.64	.428631	2.6830654
99	7.33	.508517	3.2249044

FACTOR DE HETEROGENEIDAD = 4.150548768
CON 8 GRADOS DE LIBERTAD
DE EL INTERVALO DESEADO EN LAS X PARA LAS BANDAS DE CONFIANZA DE LA RECTA
EN FORMA DE VECTOR DE DOS COMPONENTES
⎕:
0 .4
DE EL VALOR DE T, PARA EL 95 °/° DE PROBABILIDAD, Y 8 GRADOS DE LIBERTAD
(PUEDE CONSULTAR LA TABLA 8 DEL LIBRO DE FINNEY : PROBIT ANALYSIS)
⎕:
2.31

DIBUJO DE BANDAS DE CONFIANZA DE LA RECTA

CUANDO SE PARE EL TERMINAL, CAMBIE LA BOLA,
DE ESCRITURA, POR LA DE PLOT, Y CUANDO HAYA
TERMINADO, AJUSTE LA HOJA, FOLIO, ETC... AL
PRINCIPIO, Y PULSE LA TECLA RETURN.
CUANDO ACABE EL DIBUJO, VUELVA A CAMBIAR LA BOLA
Y PULSE DE NUEVO RETURN.


.00
.05
.10
.15
.20
.25
.30
.35
.40
1
2
3
4
5
6
7

```
LIMITE INFERIOR=1.470758004
DOSISLETAL AL 50°/°=1.732839323
LIMITE SUPERIOR=1.931252397


CALCULO DE ERRORES
------------------

ERROR DE LA PENDIENTE=1.997750113
ERROR DE LA Y MEDIA=0.1727830709
ERROR DEL LOGARITMO DE LA DOSIS LETAL AL 50°/° =0.02118922536
ERROR DE LA DOSIS LETAL AL 50°/° =0.08454522094
ERROR DEL LOGARITMO DE LA DOSIS AL 90°/° =0.03383579069
ERROR DE LA DOSIS AL 90°/° =0.1899082577
ERROR DEL LOGARITMO DE LA DOSIS AL 95°/° =0.0419991951
ERROR DE LA DOSIS AL 95°/° =0.2594704584
ERROR DEL LOGARITMO DE LA DOSIS AL 99°/° =0.05890671568
ERROR DE LA DOSIS AL 99°/° =0.4374186938
SIENDO LA CHI2 SIGNIFICATIVA

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
 TIEMPO INVERTIDO:      12.40   SEGUNDOS
 FINAL DEL ANALISIS PROBIT PARA EL EXPERIMENTO  CALANDRA
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
```

El segundo ejemplo, cuyo nombre es CERATITIS, muestra un análisis con X^2 no significativa.

El experimento es sobre el estudio de la toxicidad del DIPTEREX, en la Ceratitis capitata.

Las dosis se han medido en μg/gr de peso vivo.

El cuadro construido fue:

						CONTROL
Muertos	40	27	9	1	1	0
Tratados	80	80	80	80	80	80
Lg. Dosis	.82	.74	.66	.58	.50	.00

La salida de ordenador se obtuvo ejecutando

```
CONTROL  PROBIT  M
```

```
donde   CONTROL  ← 0
y       M [1;]   ← 40 27 9 1 1
        M [2;]   ← 80 80 80 80 80
        M [3;]   ← .82 .74 .66 .58 .50
```

```
ANALISIS PROBIT

NOMBRE DEL EXPERIMENTO: CERATITIS

NUMERO DE DOSIS DIFERENTES SIN LA DE CONTROL: 5
VALORES DE LAS DOSIS APLICADAS
   .92000    .74000    .66000    .58000    .50000

RESPUESTAS EN DICHAS DOSIS
40 27 9 1 1

NUMERO DE INDIVIDUOS TRATADOS EN DICHAS DOSIS
80 80 80 80 80

PROBABILIDAD ENTRE LAS RESPUESTAS Y LOS SUJETOS OBSERVADOS
   .50000    .33750    .11250    .01250    .01250

VALORES DEL PROBIT EMPIRICO PARA DICHAS DOSIS
  5.00000   4.58114   3.78653   2.75816   2.75816


VALORES DE ALFA,BETA Y CONTROL EN LA 1 APROXIMACION
ALFA=¯1.426212    BETA= 7.883346    CONTROL=  .000000
PROBITS ESPERADOS
  5.03813   4.40746   3.77680   3.14613   2.51546


-----------------------------------------------------------
 SOLUCION:   ALFA=¯1.467802891
             BETA=7.980742152
             CONTROL=0
PROBITS ESPERADOS
  5.07641   4.43795   3.79949   3.16103   2.52257
-----------------------------------------------------------
```

Y MEDIA =0.7236817373
Y MEDIA = 4.355598908
SX2 = 1.08840345
PENDIENTE DE LA RECTA = 7.980742152

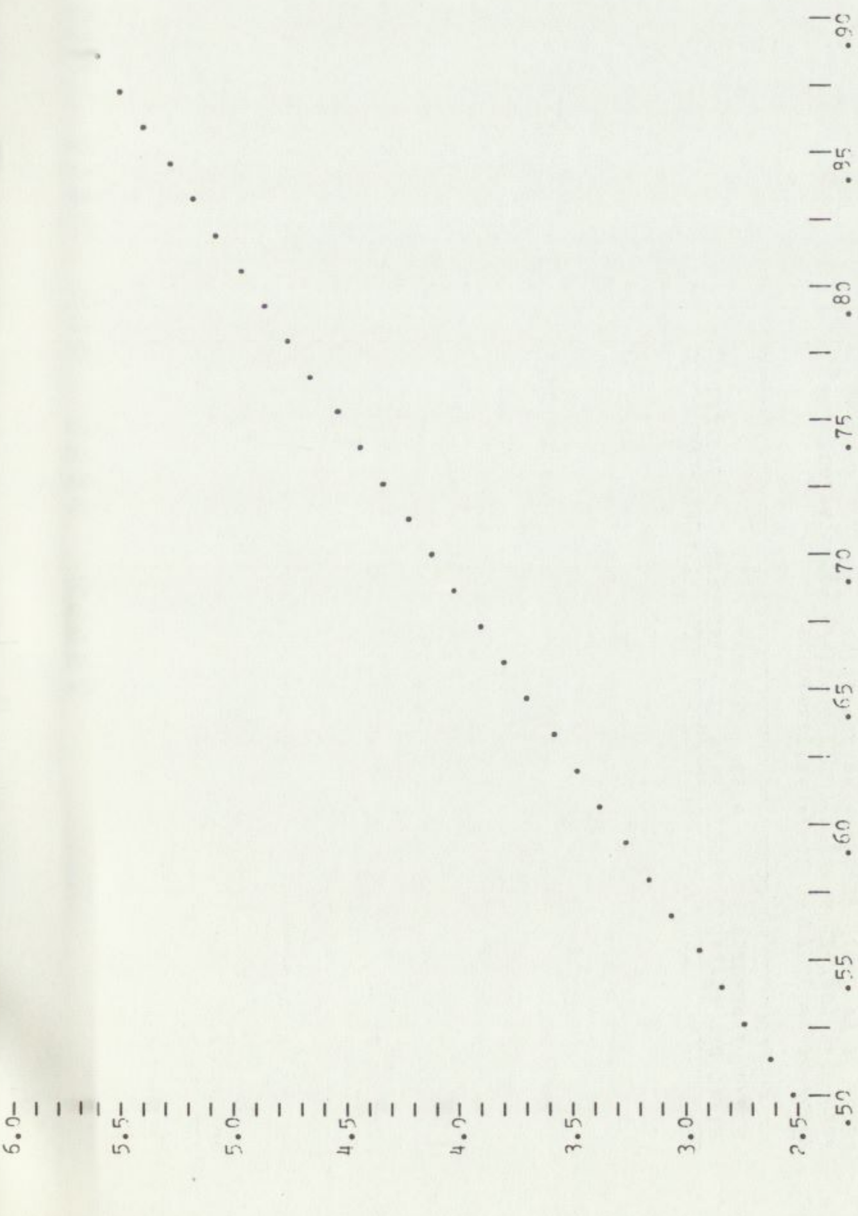
VALOR DE LA CHI2=2.78307514
CON 3 GRADOS DE LIBERTAD
CALCULO DE LA DISTRIBUCION PARA LA CHI2
MEDIANTE METODOS APROXIMATIVOS
EL VALOR DE LA FUNCION DE DENSIDAD ES :0.1655142173
EL VALOR DE LA PROBABILIDAD ACUMULADA ES ;0.573318712
ES SIGNIFICATIVA PARA USTED LA CHI2: NO
INTERVALO DE CONFIANZA PARA LA DOSIS AL 50°/° DE MORTANDAD
SIENDO LA CHI*2 NO SIGNIFICATIVA,CON α=0.05
LIMITE INFERIOR=6.10963807
DOSIS LETAL AL 50°/°=6.462882266
LIMITE SUPERIOR=6.987327644

°/° MORTANDAD	PROBIT	LOG DOSIS LETAL	DOSIS LETAL
50	5.00	.810426	6.4628823
90	6.28	.970812	9 3500155
95	6.64	1.015921	10.3733951
99	7.33	1.102379	12.6584066

DE EL INTERVALO DESEADO EN LAS X
PARA LA RECTA, EN FORMA DE VECTOR DE
DOS COMPONENTES
□:
.5 .9

DIBUJO DE LA RECTA

CUANDO SE PARE EL TERMINAL, CAMBIE LA BOLA
DE ESCRITURA, POR LA DE PLOT, Y CUANDO HAYA
TERMINADO, AJUSTE LA HOJA, FOLIO, ETC... AL
PRINCIPIO Y PULSE LA TECLA RETURN
CUANDO ACABE EL DIBUJO, VUELVA A CAMBIAR LA
BOLA Y PULSE DE NUEVO RETURN.


6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
.50
.55
.60
.65
.70
.75
.80
.85
.90

```
CALCULO DE ERRORES
------------------

ERROR DE LA PENDIENTE=0.9585285326
ERROR DE LA Y MEDIA=0.0834434874?
ERROR DEL LOGARITMO DE LA DOSIS LETAL AL 50°/° =0.01426125469
ERROR DE LA DOSIS LETAL AL 50°/° =0.212226528
ERROR DEL LOGARITMO DE LA DOSIS AL 90°/° =0.0307908583
ERROR DE LA DOSIS AL 90°/° =0.6629027408
ERROR DEL LOGARITMO DE LA DOSIS AL 95°/° =0.0359338003
ERROR DE LA DOSIS AL 95°/° =0.8583012722
ERROR DEL LOGARITMO DE LA DOSIS AL 99°/° =0.04596793227
ERROR DE LA DOSIS AL 99°/° =1.339830006

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
 TIEMPO INVERTIDO          8.52   SEGUNDOS
 FINAL DEL ANALISIS PROBIT PARA EL EXPERIMENTO  CERATITIS
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
```

5.- Bibliografia

PROBIT ANALYSIS
D.J.FINNEY
THIRD EDITION
CAMBRIGDE AT THE UNIVERSITY PRESS 1971

EXPERIMENTATION INSECTICIDE
TRAVAUX PRACTIQUES
INSTITUT NATIONAL AGRONOMIQUE PARIS - GRIGNON

HANDBOOK OF MATHEMATICAL FUNCTIONS
V.S.DEPARTMENT OF COMMERGE
NATIONAL BUREAU OF STANDARDS APPLIED MATHEMATICS
SERIES 1966

APROXIMATIONS FOR DIGITAL COMPUTERS
C. HASTINGS
PRINCETON UNIV. PRESS, PRINCETON, N.J., 1955

NONLINEAR PROGRAMMING: SEQUENTIAL UNCONSTRAINED
MINIMIZATION TECHNIQUES
ANTHONY V. FIACCO AND GARTH P. McCORMICK
JOHN WILEY AND SONS, INC., 1968

COMPUTATIONAL METHODS IN OPTIMIZATION
A UNIFIED APPROACH
E. POLAK
MATHEMATICS IN SCIENCE AND ENGINEERING: VOL, 77
ACADEMIC PRESS, 1971