

# Instrucciones para usar Infostat®

## Descripción de Distribuciones de frecuencias

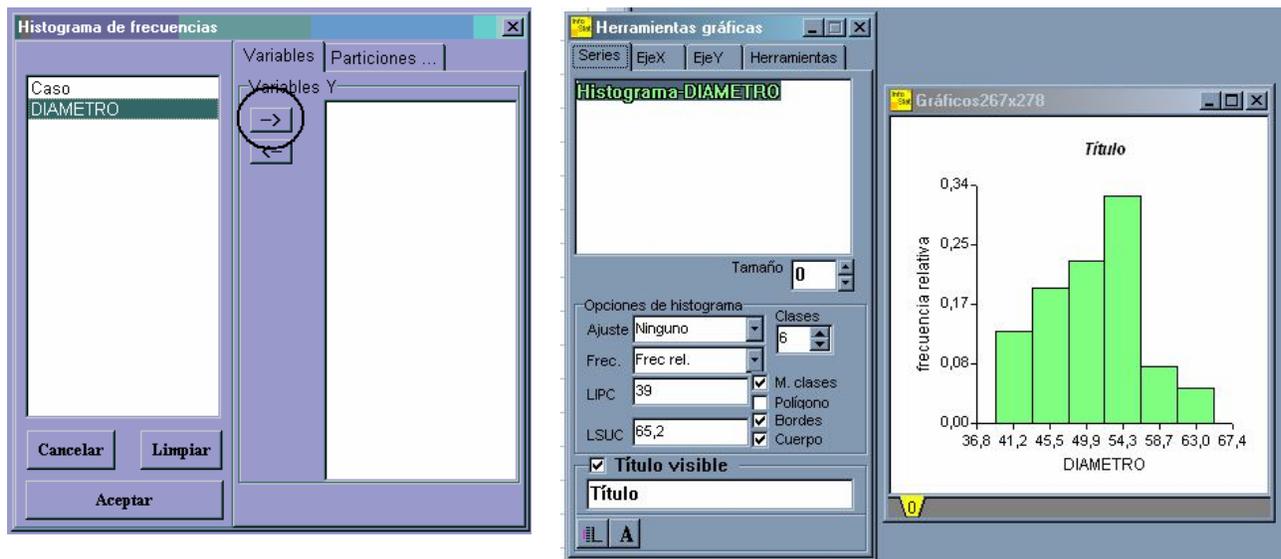
### Ingreso de datos

Ir a *Archivo* → *Nueva Tabla*; se deben ingresar los datos en una columna. Haciendo click en el encabezamiento de la columna se puede escribir el nombre de la variable

### Histogramas y Polígonos de frecuencias

Ir a *Gráficos* → *Histograma*. Aparece un cuadro de diálogo donde nos pide que indiquemos a cuál de las variables queremos construirle un Histograma. En nuestra planilla hay sólo una variable, la elegimos mediante la flecha y hacemos click en Aceptar.

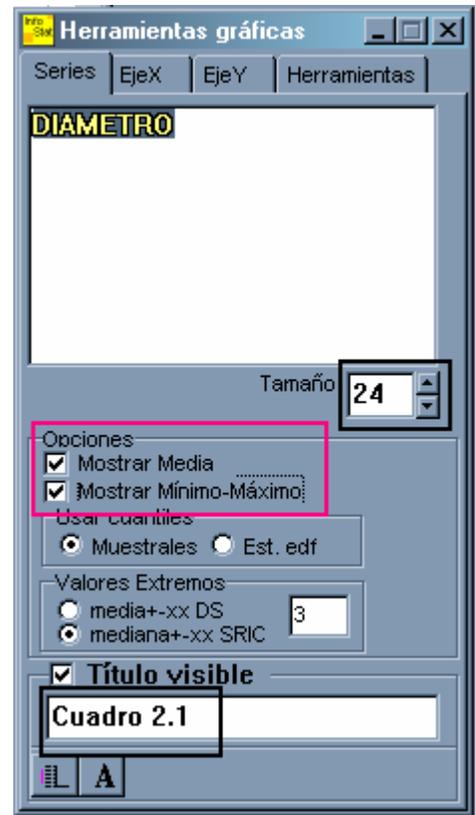
Aparece un histograma *default* con un menú de Herramientas Gráficas que permite seleccionar los diferentes tipos de gráficos y ajustar sus detalles.



### Gráfico de Caja y Bigotes (Box plot)

Ir a *Gráficos* → *Gráfico de Cajas (Box Plot)*. Aparece un cuadro de diálogo donde nos pide que indiquemos a cuál de las variables queremos construirle un Gráfico de Caja y Bigotes. En nuestra planilla hay sólo una variable, la elegimos mediante la flecha y hacemos click en Aceptar.

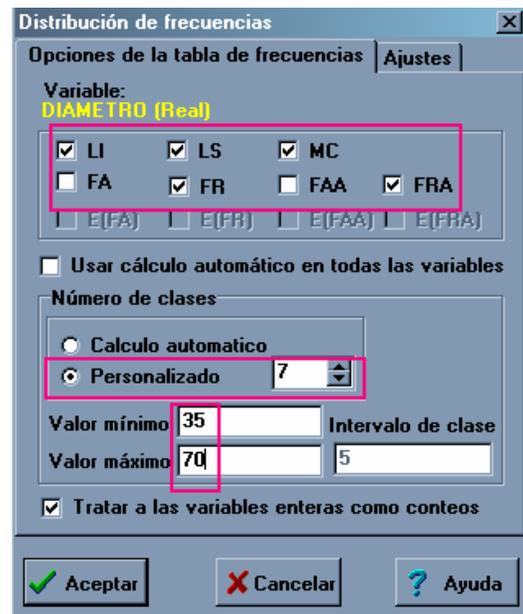
Aparece un gráfico *default* con un menú de Herramientas Gráficas que permite seleccionar los diferentes tipos de gráficos y ajustar sus detalles.



## Tablas de Frecuencias

Ir a *Estadísticas* → *Tablas de Frecuencias*. Aparece un cuadro de diálogo donde nos pide que indiquemos a cuál de las variables queremos construirle una Tabla de Frecuencias. Elegimos la variable mediante la flecha y hacemos click en Aceptar.

Aparece un menú de Distribución de frecuencias que permite seleccionar diferentes opciones para construir la Tabla.



## Medidas de posición y de dispersión

Ir a *Estadísticas* → *Medidas resumen*. Aparece un cuadro de diálogo donde nos pide que indiquemos cuál de las variables queremos resumir. Elegimos la variable mediante la flecha y hacemos click en Aceptar.

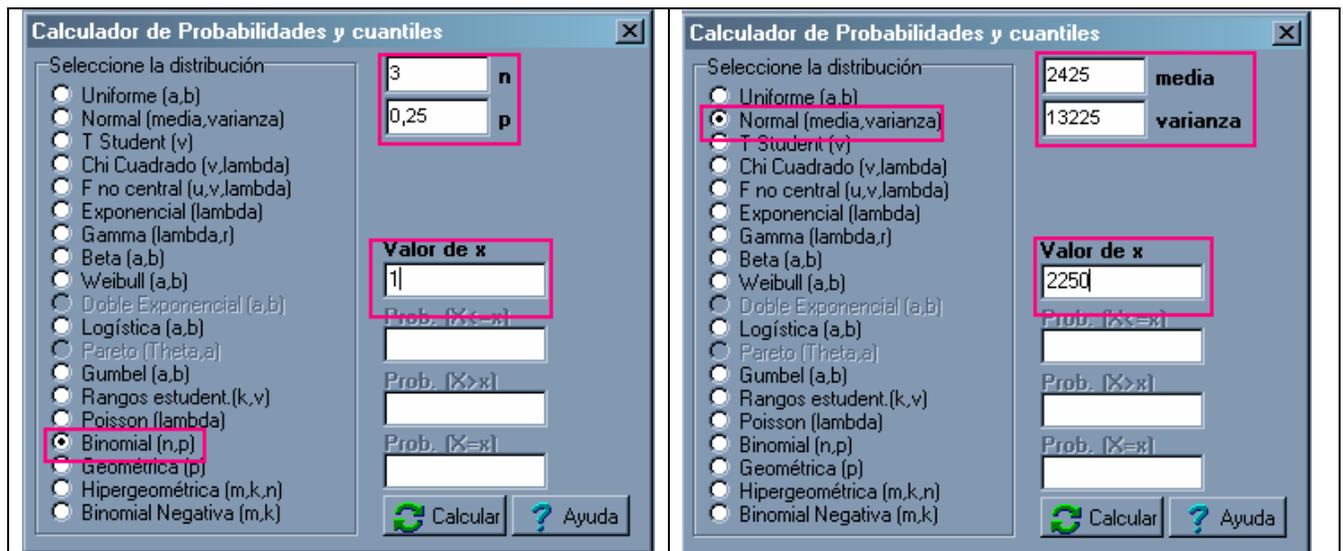
Aparece un menú de Estadística Descriptiva que permite seleccionar las diferentes medidas resumen.



## Distribuciones de Probabilidad

### Cálculo de probabilidades con diferentes modelo de distribución

Ir a *Estadísticas* → *Probabilidades y cuantiles*. Aparece un menú donde seleccionamos el modelo de distribución de probabilidad, escribimos los valores de los parámetros y hacemos click en Calcular.



## ***Muestras Aleatorias y Medias Muestrales***

### **Generación de todas las muestras posibles a partir de un conjunto de unidades**

Permite obtener todas las muestras aleatorias posibles de cierto tamaño a partir de un conjunto de observaciones (población). Para cada muestra generada permite calcular la media y otros estadísticos muestrales. El programa sólo informa los valores de los estadísticos muestrales, no los datos que pertenecen a cada muestra generada. Este módulo puede ser usado con fines pedagógicos para la visualización de la distribución muestral de medias.

Ir a *Archivo* → *Nueva Tabla*; se deben ingresar en una columna los datos de la población de la cual se desean extraer las muestras.

#### **Todas las muestras posibles**

Ir a *Aplicaciones* → *Didácticas* → *todas las muestras posibles*; aparece un cuadro de diálogo donde se debe seleccionar la variable sobre la cual se desea generar muestras aleatorias → *Aceptar*.

Aparece un nuevo cuadro de diálogo donde se debe indicar el tamaño muestral deseado y, allí mismo, el programa hace un cálculo combinatorio de cuántas muestras aleatorias es posible generar para esa población. Tildando sobre el botón M.A.S. (muestreo aleatorio simple) y clickeando sobre *Aceptar* se obtienen las muestras deseadas.

### **Generación de muestras aleatorias a partir de una distribución empírica**

Permite obtener la cantidad solicitada de muestras aleatorias de cierto tamaño a partir de un conjunto de observaciones (población). A diferencia de la aplicación anterior, el programa informa los datos que pertenecen a cada muestra generada. Estas muestras pueden ser almacenadas en varias columnas (**Muestras en distintas columnas**) o en una sola columna (**Muestras en una única columna**). Este módulo puede ser usado con fines pedagógicos para visualización de la distribución muestral de estadísticos obtenidos a partir de la distribución empírica.

#### **Ingreso de datos**

Ir a *Archivo* → *Nueva Tabla*; se deben ingresar en una columna los datos de la población de la cual se desean extraer las muestras.

#### **Muestrear Desde La Distribución Empírica**

Ir a *Aplicaciones* → *Didácticas* → *Muestreo desde la distribución empírica*; aparece un cuadro de diálogo donde se debe seleccionar la variable sobre la cual se desea generar muestras aleatorias → *Aceptar*. Aparece un nuevo cuadro de diálogo donde se puede indicar cuántas muestras y de qué tamaño cada una, se desean obtener → *Aceptar*.

## ***Estimación de la Media y de la Varianza Poblacionales***

### **Intervalo de confianza para la media, con varianza poblacional desconocida**

#### **Ingreso de datos**

Ir a *Archivo* → *Nueva Tabla*; los datos de la muestra se deben ingresar en una columna. Atención: Infostat efectúa el análisis sólo si se ingresan las observaciones individuales; si sólo se cuenta con medidas resumen (media y desvío estándar muestrales) no es posible construir el IC utilizando el programa.

## Intervalo de confianza

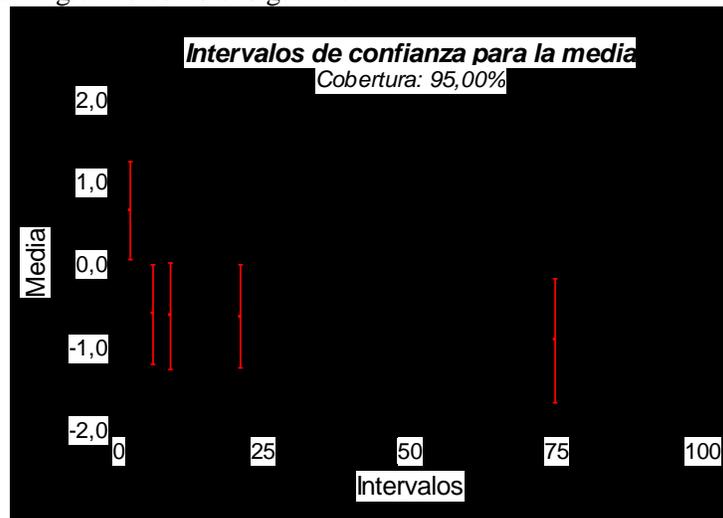
Ir a *Estadísticas* → *Estimación de características poblacionales* → *Muestreo aleatorio simple*. Aparece el cuadro de diálogo para elegir la variable para la cual se desea obtener el IC. Una vez elegida, se le da *Aceptar* y obtiene un nuevo cuadro de diálogo en el cual se debe elegir el botón *Promedio*, tildar la casilla donde dice *Intervalo de confianza para el parámetro poblacional* e indicar el nivel de confianza deseado. Luego se le da *Aceptar* y el programa devuelve los valores de la media muestral, el error de la estimación y los límites inferior (LI) y superior (LS) del IC

## Simulación de Intervalos de confianza para la media, con varianza poblacional conocida

Permite obtener por simulación un conjunto de IC para la media de una variable aleatoria con distribución normal cuyos parámetros son indicados. Esto facilita la interpretación de los conceptos subyacentes en el proceso de estimación, como precisión y nivel de confianza.

Ir a *Aplicaciones* → *Didácticas* → *Intervalos de Confianza* aparece un cuadro de diálogo donde se debe indicar la media poblacional y la variancia poblacional de la distribución normal de la cual se extraerán las muestras, el número de IC deseados, el nivel de confianza y el tamaño muestral → *Aceptar*.

Por ejemplo, se solicita que, para una variable aleatoria con distribución normal, con un promedio de 0 y variancia de 1, se extraigan 100 muestras aleatorias de tamaño 10 y que para cada una de ellas se calcule el IC al 95%. Se obtiene un gráfico como el siguiente:



La línea horizontal indica el valor del parámetro  $\mu$ . En el título del gráfico se indica el porcentaje de intervalos, de los obtenidos, que contienen al verdadero valor del parámetro (*Cobertura*). En *Herramientas gráficas* se puede cambiar la cantidad de intervalos (*Intervalos*), el nivel de confianza (*Confianza*), el tamaño de la muestra (*Tamaño muestral*) y el tamaño de los puntos que señalan el centro de cada intervalo (*Tamaño*). Además puede solicitarse al programa que genere un nuevo conjunto de IC (*Generar otro conjunto de intervalos*).

# ***Pruebas de Hipótesis acerca de la Media Poblacional***

## **Prueba de hipótesis acerca de una media poblacional**

### **Ingreso de datos**

Ir a *Archivo* → *Nueva Tabla*; los datos de la muestra se deben ingresar en una columna. Atención: Infostat efectúa el análisis sólo si se ingresan las observaciones individuales; si sólo se cuenta con medidas resumen (media y desvío estándar muestrales) no es posible resolver la prueba utilizando el programa.

### **Prueba *t* para una muestra**

Ir a *Estadísticas* → *Inferencia basada en una muestra* → *Prueba *t* para una parámetro* aparece un cuadro de diálogo donde se debe indicar la variable que se desea para la cual se desea hacer la prueba → *Aceptar*. Aparece otro cuadro de diálogo donde se puede indicar de qué tipo de prueba de hipótesis se trata (unilateral o bilateral) y qué información se desea obtener → *Aceptar*.

## **Prueba de hipótesis acerca de dos medias poblacionales, observaciones independientes**

### **Ingreso de datos**

Ir a *Archivo* → *Nueva Tabla*; los datos se deben ingresar en dos columnas, una para la variable aleatoria de interés y otra para la variable que se utilizará para identificar las dos muestras. El orden de las columnas es indistinto.

### **Prueba *t* para observaciones independientes**

Ir a *Estadísticas* → *Inferencia basada en dos muestras* → *Prueba *t**; aparece un cuadro de diálogo donde en *Variables* se debe indicar la variable aleatoria y en *Criterio de clasificación*, la variable (columna) donde está indicado a qué muestra pertenece cada observación → *Aceptar*. Aparece un cuadro de diálogo donde se debe indicar el tipo de prueba de hipótesis (unilateral o bilateral) y la información que se desea obtener → *Aceptar*.

## **Prueba de hipótesis acerca de dos medias, observaciones apareadas**

### **Ingreso de datos**

Ir a *Archivo* → *Nueva Tabla*; los datos se deben ingresar en dos columnas, una para las observaciones provenientes de la muestra 1 y otra para las provenientes de la muestra 2. El orden de las columnas es indistinto.

Para el ejemplo de Comparación de muestras apareadas desarrollado en el capítulo 4:

Caso	Técnica nueva	Técnica conocida
1	45,00	41,00
2	40,00	45,00
3	30,00	43,00
4	42,00	39,00
5	24,00	39,00
6	25,00	55,00
7	24,00	45,00
8	30,00	51,00
9	40,00	37,00
10	25,00	41,00

Real registros: 10\*2

### Prueba *t* para observaciones apareadas

Ir a *Estadísticas* → *Inferencia basada en dos muestras* → *Prueba t apareada* aparece un cuadro de diálogo donde se debe indicar las dos variables que contiene los datos correspondientes a cada tratamiento (deben estar en dos columnas separadas) → *Aceptar*. Aparece un cuadro de diálogo donde se debe indicar el tipo de prueba de hipótesis (unilateral o bilateral) y la información que se desea obtener → *Aceptar*.

En el ejemplo:

#### Prueba T (muestras apareadas)

Obs(1)	Obs(2)	N	media(dif)	DE(dif)	T	p(Unilateral I)
Técnica nueva	Técnica conocida	10	-11,10	11,83	-2,97	0,0079

Media(dif) y DE(dif) corresponden a la media y al desvío estándar de la variable diferencia. T y bilateral indican el valor del estadístico y del valor *p* de la prueba de hipótesis seleccionada.

## **Análisis de la Asociación entre dos Variables Cuantitativas: Regresión y Correlación**

### Análisis de Regresión

#### Ingreso de datos

Ir a *Archivo* → *Nueva Tabla*; los datos se deben ingresar en dos columnas, una para la variable predictora X y otra para la variable respuesta Y. El orden de las columnas es indistinto.

Ejemplo:

Caso	P	MS
1	0,00	10,00
2	20,00	10,60
3	40,00	14,40
4	60,00	15,80
5	80,00	15,30
6	100,00	20,20

### Diagrama de dispersión

Ir a *Gráficos* → *Diagrama de dispersión*; aparece un cuadro de diálogo donde se debe indicar la variable correspondiente al eje Y (variable dependiente, MS en este ejemplo) y la variable correspondiente al eje X (variable independiente o regresora, P en este ejemplo) → *Aceptar*.

### Regresión Lineal Simple

Ir a *Estadísticas* → *Regresión Lineal* aparece un cuadro de diálogo donde se debe indicar la variable dependiente y la variable independiente (regresora) → *Aceptar*. Aparece un cuadro con 5 solapas de las cuales solo utilizaremos en este curso, las dos primeras (*General* y *Diagnóstico*).

Dentro de la solapa *General* tienen que estar tildadas las casillas que dicen *Coefficientes de regresión y estadísticos asociados* y *Ordenada al origen*.

Dentro de la solapa *Diagnóstico*, tildar *Residuos*, *Predichos* e *Intervalos de Confianza*, y en *Graficar*, tildar *Ajuste* y *Bandas de confianza*.

En el ejemplo:

#### Análisis de regresión lineal

Variable	N	R <sup>2</sup>	R <sup>2</sup> Aj	ECMP	AIC	BIC
MS	6	0,90	0,87	4,15	23,98	23,36

#### Coefficientes de regresión y estadísticos asociados

Coef	Est.	E.E.	LI(95%)	LS(95%)	T	p-valor
const	9,63	0,96	6,97	12,30	10,04	0,0006
P	0,10	0,02	0,05	0,14	5,99	0,0039

En la tabla de Coeficientes de regresión y estadísticos asociados, se reportan para cada parámetro del modelo (Coef) el valor estimado (Est.), el error estándar de la estimación (E.E.), los límites inferior y superior del intervalo de confianza al (1- $\alpha$ )% (LI y LS), el valor del estadístico *t* para poner a prueba la hipótesis que el parámetro vale cero (T) y el valor *p* (p-valor) para dicha prueba de hipótesis. La primera fila (const) corresponde a la ordenada al origen, mientras que la segunda (P en este caso), corresponde a la pendiente.

## Análisis de Correlación

### Ingreso de datos

Ir a *Archivo* → *Nueva Tabla*; los datos se deben ingresar en dos columnas, una para cada variable aleatoria analizada. El orden de las columnas es indistinto.

## Correlación

Ir a *Estadísticas* → *Análisis de correlación* → *Coefficientes de correlación*; aparece un cuadro de diálogo donde se debe indicar las dos variables Y → *Aceptar*. Aparece un cuadro donde se debe seleccionar el coeficiente de correlación (*Pearson*)

Para el ejemplo de Correlación desarrollado en el capítulo 5:

### Coeficientes de correlación

*Correlación de Pearson: coeficientes\probabilidades*

	CO	PI
CO	1,000	8,1E-05
PI	0,878	1,000

Los resultados se muestran como una matriz, donde los elementos de la diagonal son todos iguales a 1 ya que representan la correlación de una variable consigo misma; por debajo de la diagonal se encuentra el coeficiente de correlación y por encima, el valor *p* asociado a la prueba de hipótesis de correlación nula.

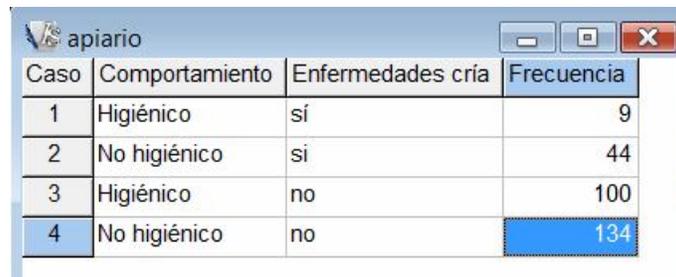
## Pruebas de Hipótesis acerca de Variables Categóricas

### Pruebas de independencia y de homogeneidad

#### Ingreso de datos

Ir a *Archivo* → *Nueva Tabla*; los datos se deben ingresar en tres columnas, una para cada variable cualitativa y la restante para las frecuencias observadas.

Para el ejemplo de Prueba de independencia desarrollado en el capítulo 6:



Caso	Comportamiento	Enfermedades cría	Frecuencia
1	Higiénico	sí	9
2	No higiénico	si	44
3	Higiénico	no	100
4	No higiénico	no	134

#### Prueba Chi cuadrado

Ir a *Estadísticas* → *Datos Categorizados* → *Tablas de Contingencia*. Aparece un cuadro de diálogo donde en *Criterios de clasificación* se deben ingresar las dos variables cualitativas (en nuestro ejemplo Comportamiento y Enfermedades cría) y en *Frecuencias*, la columna conteniendo las frecuencias observadas → *Aceptar*. Aparece un nuevo cuadro de diálogo donde se debe indicar cuál es la variable que representa las columnas de la tabla de contingencia, y cuál representa las filas. En nuestro caso, las columnas son Enfermedades cría y las filas, Comportamiento. En la solapa *Opciones*, se pueden solicitar las *Frecuencias esperadas bajo independencia*, para verificar si existe alguna violación a los

requerimientos de la prueba, *Frecuencias relativas por filas* y *Frecuencias relativas por columnas*, para la discusión de los resultados. → *Aceptar*.

En el ejemplo:

**Tablas de contingencia**

*Frecuencias: Frecuencia*

*Frecuencias absolutas*

*En columnas: Enfermedades cría*

Comportamiento	no	sí	Total
Higiénico	100	9	109
No higiénico	134	44	178
Total	234	53	287

Tabla de contingencia de frecuencias observadas

*Frecuencias relativas por filas*

*En columnas: Enfermedades cría*

Comportamiento	no	sí	Total
Higiénico	0,92	0,08	1,00
No higiénico	0,75	0,25	1,00
Total	0,82	0,18	1,00

Tabla de frecuencias relativas observadas por tipo de comportamiento

*Frecuencias esperadas*

*En columnas: Enfermedades cría*

Comportamiento	no	sí	Total
Higiénico	88,87	20,13	109,00
No higiénico	145,13	32,87	178,00
Total	234,00	53,00	287,00

Tabla de contingencia de frecuencias esperadas

Estadístico	Valor	gl	p
Chi Cuadrado Pearson	12,17	1	0,0005
Chi Cuadrado MV-G2	13,39	1	0,0003
Irwin-Fisher bilateral	0,16		0,0005
Coef. Conting. Cramer	0,15		
Coef. Conting. Pearson	0,20		
Coeficiente Phi	0,21		

⇒ Estadístico  $X^2$  y valor  $p$