ANALISIS RELACIONADOS CON LA NO-NORMALIDAD EN INVESTIGACIONES AGRONOMICAS.

ROBERTO CANALES CRUZ

T E S I S

PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS
EN ESTADISTICA EXPERIMENTAL

Universidad Autonoma Agraria
Antonio Narro

PROGRAMA DE GRADUADOS

Buenavista, Saltillo, Coah.
Septiembre de 1987.

Tesis elaborada bajo la supervisión del Comité Particular de asesoría y aprobada como requisito parcial, para optar el grado de:

MAESTRO EN CIENCIAS
ESPECIALIDAD EN ESTADISTICA EXPERIMENTAL

C O M I T E   P A R T I C U L A R

Asesor Principal: ______________________________
MC REGINO MORONES REZA

Asesor: ______________________________
MC EMILIO PADRON CORRAL

Asesor: ______________________________
MC VICTOR CANTU HERNANDEZ

Asesor: ______________________________
MC JUAN MANUEL CEPEDA DOVALA

____
DR. ELEUTERIO LOPEZ PEREZ
SUBDIRECTOR DE ASUNTOS DE POSTGRADO

BUENAVISTA, SALTILLO, COAHUILA. SEPTIEMBRE DE 1987

AGRADECIMIENTOS

A la Universidad Autónoma Agraria "Antonio Narro", por la oportunidad ofrecida para mi superación.

Al Consejo Nacional de Ciencia y Tecnología, por el apoyo brindado durante mis estudios de Postgrado.

Al Instituto Nacional de Investigaciones Agropecuarias y Forestales, por el estímulo brindado en la realización de mis estudios.

Al MC Regino Morones Reza, a quien con su profundo interés hizo posible la realización de este trabajo de tesis.

Al MC Emilio Padrón Corral, al MC Víctor Cantú Hernández y al MC Juan Manuel Cepeda Dovala por sus asesorías brindadas.

Al Ing. J. Concepción Loredo Osti por la ayuda desinteresada que brindó en el presente trabajo.

DEDICATORIA

A mis maestros que han consagrado parte de su vida al cumplimiento de su deber y que son un digno ejemplo a seguir.

A mi esposa e hijo:

Maricela

y

Ulises

A mis padres:

Apolinar

y

Timotea

A mis hermanos, como un testimonio a base de perseverancia.

A mis compañeros, Osti y Octavio por todo lo que compartimos juntos.

COMPENDIO

Análisis Relacionados con la No-Normalidad en Investigaciones Agronómicas

POR

ROBERTO CANALES CRUZ

MAESTRIA

ESTADISTICA EXPERIMENTAL

UNIVERSIDAD AUTONOMA AGRARIA ANTONIO NARRO

BUENAVISTA, SALTILLO, COAHUILA, AGOSTO 1987.

M.C. Regino Morones Reza - Asesor -

Palabras claves: Distribución, no-normalidad, desviación estándar, características agronómicas.

El presente estudio tuvo como objetivo evaluar la propiedad de distribución normal de las variables agronómicas que comúnmente se toman en trabajos de investigación que se realizan en la Universidad Autónoma Agraria "Antonio Narro".

Para llevar a cabo la evaluación, se obtuvieron aproximaciones satisfactorias de los valores esperados de los estadísticos de orden $\{m_{n-k+1}\}$ y coeficientes $\{a_{n-k+1}\}$ para la prueba de normalidad de Shapiro-Wilk.

Los cultivos estudiados fueron: maíz, frijol, sorgo, tomate, trigo, triticali, cebada, arroz, papa, zacate Estrella de Africa, durazno y nogal; las características agronómicas que se evaluaron fueron las que tienen relación con: componentes de rendimiento, propiedades bromatológicas, propiedades hídricas y propiedades físico-químicas del suelo.

En los análisis se encontró que en las propiedades físico-químicas y propiedades bromatológicas que se evalúan en meq, mg, ppm, por ciento y gr, por lo general, en más de un 30 por ciento de los casos presentaron no-normalidad.

ABSTRACT

Analysis Related With No-Normality in Agronomic Characteristics

BY

ROBERTO CANALES CRUZ

MASTER OF SCIENCE

EXPERIMENTAL STATISTICS

UNIVERSIDAD AUTONOMA AGRARIA ANTONIO NARRO

BUENAVISTA, SALTILLO, COAHUILA, AUGUST 1987.

M.C. Regino Morones Reza - Advisor -

Key words: Distribution, no-normality, standar desviation, agronomic characteristics.

This study had as objective to evaluate the property of normal distribution of the agronomic variables that commonly take in research works that are realize at Universidad Autónoma Agraria "Antonio Narro".

To carry out the evaluation, were obtained satisfactory approximations of expected values of order statistics $\{m_{n-k + 1}\}$ and coefficients $\{\hat{a}_{n-k + 1}\}$ for Shapiro-Wilk test for.

Cultivation studied were: maize, bean, sorghum, tomato, wheat, triticali, barley, rice, potato, Africa star grass, peach and walnut tree; the agronomic characteristics that were evaluated are: yielding components, bromatologic properties, hydric properties and physical-chemical soil properties.

In analysis was found that physical-chemical properties and bromatological properties that were evaluated in meq, mg, ppm, percent and gr, in general, more of 30 percent of the cases presented no-normality.

INDICE DE CONTENIDO

INDICE DE CUADROS

Página

1. INTRODUCCION

Los trabajos de investigación sólo tienen sentido cuando su objetivo es contribuír en la solución de un problema planteado por la sociedad. Una vez que se define el problema, viene la etapa de discernir que factores variables o características son relevantes para la explicación y/o interpretar el fenómeno relacionado al problema. La concepción del - problema por parte del investigador mucho ayudará en identificar las variables con posibilidades de evaluarse y que deberán considerarse en el problema que es objeto de estudio.

Dependiendo del conocimiento que sobre las propiedades distribucionales se tenga de las variables consideradas en el trabajo de investigación, será la metodología estadística a utilizar para analizar, interpretar y concluír sobre la información que arrojen los datos obtenidos.

En el análisis de una variable relevante a veces se utiliza un - procedimiento estadístico que requiere que la metodología sea adecuada o que la variable tenga una distribución normal. Muchas veces el análisis de dicha variable se realiza sin haberse tomado la precaución de que se satisfaga dicho requerimiento. Esto trae como consecuencia que pueda incurrir en problemas serios en la conclusión de sus resultados.

El presente trabajo tiene como objetivo evaluar, en lo que se refiere a su propiedad de distribución normal, las variables agronómicas - que comúnmente son evaluadas en trabajos de investigación que se realizan en nuestra Universidad Autónoma Agraria "Antonio Narro".

La hipótesis que se plantea es que ciertas variables que se eva--lúan en los trabajos no satisfacen la condición de distribuirse normal--mente.

2. REVISION DE LITERATURA

Generalidades Sobre el Análisis de Varianza

Steel y Torrie (1985) clasifican a los trabajos experimentales en tres clases: Preliminares, críticos y demostrativos. En el experimento preliminar, el investigador prueba un número grande de tratamientos con el objeto de obtener indicios para futuros trabajos; en este caso, los tratamientos aparecen una vez. En un experimento crítico el investigador compara las respuestas a diferentes tratamientos usando un número suficiente de observaciones de las respuestas para tener seguridad razonable de detectar diferencias significantes. Los experimentos demostrativos se llevan a cabo cuando los trabajadores de extensión comparan uno o más tratamientos nuevos con un patrón.

Kempthorne (1975) dice que una investigación diseñada estadísticamente, consiste de las etapas siguientes:

1. Definición del problema.
2. Formación de hipótesis.
3. Sugerencias de la técnica experimental y el diseño.
4. Examen de los posibles resultados y referencias detrás de las razones, para la indagación que asegure que el experimento proporcionará la información requerida y en la extensión adecuada.
5. Consideración de los posibles resultados desde el punto de

vista de los procedimientos estadísticos, los cuales les serán aplicados para asegurar que las condiciones necesarias estén satisfechas para que estos procedimientos sean válidos.

6. Ejecución del experimento.
7. Aplicación de las técnicas estadísticas a los resultados experimentales.
8. Extracción de conclusiones con medidas de confiabilidad de las estimaciones de cualquiera de las cantidades que son evaluadas, empezando por dar cuidadosa consideración a la validez de las conclusiones para la población de objetivos o eventos a la cual se van a aplicar.
9. Evaluación de la investigación completa particularmente con otras investigaciones del mismo o similares problemas.

Scheffé (1959) define al análisis de varianza como una técnica estadística para analizar mediciones dependientes de varias clases de efectos que operan simultáneamente, para decidir cuales clases de efectos son importantes y estimar sus consecuencias. Las mediciones u observaciones pueden ser de una ciencia experimental como la genética o no experimental como la astronomía. En una teoría para analizar mediciones, naturalmente tiene implicación el como fue planteado o el como fueron tomadas las observaciones, o sea el diseño experimental. Históricamente, la presente técnica de análisis de varianza ha sido desarrollada principalmente en relación con problemas de experimentación agrícola.

Eisenhart (1947) considera que el análisis de varianza es usado para proporcionar soluciones a dos tipos de problemas: Clase I, localización y estimación de relaciones fijas (constantes) entre las medias de subconjuntos del universo de objetos considerados; Clase II, localización y estimación de componentes de variación (aleatoria) asociados a una población compuesta.

Para problemas de la Clase I, Eisenhart (1947) da el Teorema I: La condición necesaria y suficiente para la estricta validez de los procedimientos de análisis de varianza, para resolver problemas de Clase I con respecto a los datos ordenados como en la Tabla 1 (la tabla de referencia, es una que contiene r hileras y c columnas, la cual no se consigna aquí), son que:

$$x_{ij} = m_{..} + (m_{i.} - m) + (m_{.j} - m_{..}) + z_{ij} \qquad 2.1.1.$$

$$i = 1,2,\ldots, r \text{ (hileras)}$$

$$j = 1,2,\ldots, c \text{ (columnas)}$$

donde las $m_{i.}$, $m_{.j}$ y $m_{..}$ son constantes con:

$$m_{..} = \sum_{i=1}^{r} m_{i.} / r = \sum_{j=1}^{c} m_{.j} / c \qquad 2.1.2.$$

y las z_{ij} se distribuyen normal e independientemente con media cero y varianza común σ^2 [otros autores resumen el teorema anterior, diciendo: El modelo es aditivo y los errores están distribuídos normal e independientemente con media cero y varianza constante σ^2].

Para problemas de la Clase II, Eisenhart (1947) hace la suposición de que las medias de $(m_{i} - m_{..})$, $(m_{.j} - m_{..})$ y las z_{ij} son todas -

cero. Supone además que las variables aleatorias ($m_{i.} - m_{..}$), -
($m_{.j} - m_{..}$) y z_{ij} se distribuyen con covarianzas entre ellas cero. Asume además que las distribuciones son normales.

Little y Hill (1981) comentan que si los datos no concuerdan con las suposiciones del análisis de varianza, el análisis de los datos pueden dar lugar a que el investigador llegue a conclusiones que no tienen justificación. Así mismo, el investigador puede descuidar conclusiones - importantes que se alcanzarían si los datos fuesen analizados adecuada-- mente.

Cochran (1947) en su disertación sobre las consecuencias cuando - las suposiciones del análisis de varianza no son satisfechas llega entre otras conclusiones a las siguientes: Que la falta de cualquiera de las - suposiciones alterará en alguna extensión las propiedades estándar entre las cuales se aplica la técnica dependiente; que en general los factores que causan disturbios más severos son sesgos extremos, presencia de errores robustos, comportamientos anómalos de ciertos tratamientos o partes del experimento, marcado alejamiento a la relación con la media o cier-- tos tratamientos o parte del experimento; que los métodos principales para un mejor análisis son la omisión de ciertas observaciones, tratamientos, repeticiones, subdivisión de la varianza del error y transformación de otra escala antes del análisis.

Sobre Pruebas para los Supuestos del Análisis de Varianza

Snedecor y Cochran (1980) consideran que la prueba de aditividad de Tukey es útil, por que: (i) ayuda a decidir si la transformación es - necesaria; (ii) sugiere la transformación apropiada; (iii) informa si la transformación ha tenido éxito para producir aditividad.

Infante (1972) considera que para el modelo:

$$Y_{ij} = \mu + \tau_i + \beta_j + \xi_{ij} \qquad 2.2.1.$$

$i = 1,2,\ldots, t$

$j = 1,2,\ldots, b$

$\xi_{ij} \sim NI(0,\sigma^2)$

si se presenta no-aditividad, se tendría:

$$Y_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \xi_{ij} \qquad 2.2.2.$$

$i = 1,2,\ldots, t$

$j = 1,2,\ldots, b$

$\xi_{ij} \sim NI(0,\sigma^2)$

Entonces el método de Tukey para probar no-aditividad consiste en probar Ho : $(\tau\beta)_{ij} = 0$, $\forall_i$, $\forall_j$.

Para el primer modelo se tendría que la SC del error está dada - por:

$$\text{SC Error} = \sum_{i=1}^{t} \sum_{j=1}^{t} Y_{ij}^2 - \frac{Y_{..}^2}{rt} - \left(\frac{\sum_{1=1}^{t} Y_{i.}^2}{b} - \frac{Y_{..}^2}{tb}\right) - \left(\frac{\sum_{j=1}^{b} Y_{.j}^2}{t} - \frac{Y_{..}^2}{tb}\right)$$

$$= \sum_{i=1}^{t} \sum_{j=1}^{b} (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 \qquad 2.2.3.$$

El método de Tukey consiste en particionar esta suma de cuadra--- dos, en dos componentes:

1. $$SCNA = \frac{\left[\sum_{i=1}^{t}\sum_{j=1}^{b} Y_{ij}\,(\bar{Y}_{i.} - \bar{Y}_{..})\,(\bar{Y}_{.j} - \bar{Y}_{..})\right]^2}{\sum_{i=1}^{t}(\bar{Y}_{i.} - Y_{..})^2\;\sum_{j=1}^{b}(\bar{Y}_{.j} - \bar{Y}_{..})^2} \qquad 2.2.4.$$

que es la suma de cuadrados de no-aditividad con un grado de libertad.

2. $$SCB = \sum_{i=1}^{t}\sum_{j=1}^{b}(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 - SCNA \qquad 2.2.5.$$

que se conoce como suma de cuadrados de balance.

Las propiedades distribucionales de estas sumas de cuadrados son:

Bajo $H_o : (\tau\beta)_{ij} = 0 \ \forall_i, \forall_j$

(i) $SCNA/\sigma^2 \sim X_1^2$ g.l.

(ii) $SCB/\ \sigma^2 \sim X^2_{[(t-1)(b-1) - 1]}$ g.l.

(iii) Las dos variables $X^2_{,s}$ son independientes.

Considerando lo anterior se tiene:

$$Fc = \frac{(SCNA)\ [(t-1)(b-1)-1]}{SCB} \sim F_{1,[(t-1)(b-1)-1]} \qquad 2.2.6.$$

si y sólo si $H_o : (\tau\beta)_{ij} = 0$ es cierta.

De manera que si $Fc \geqq F_{\alpha,\ 1,\ [(t-1)(b-1)-1]}$ se rechaza la hipótesis $H_o : (\tau\beta)_{ij} = 0$.

Para el supuesto de homogeneidad de varianza la prueba de Bartlett es de las más conocidas para evaluar la validez de dicho supuesto. Snedecor y Cochran (1980) presentan el procedimiento para esta prueba:

Si se tiene t estimaciones S_1^2, S_2^2, ..., S_t^2 con f_1, f_2, ..., f_t - grados de libertad respectivamente y si S^2 es el estimador de varianza - ponderada por grados de libertad, es decir:

$$S^2 = \sum_{i=1}^{t} S_i^2 \Big/ \sum_{i=1}^{t} f_i \tag{2.2.7.}$$

por otra parte sean:

$$M = 2.3026[(\sum_{i=1}^{t} f_i)(\log_{10} S^2) - \sum_{i=1} f_i \log_{10} S_i^2] \tag{2.2.8.}$$

$$C = 1 + 1/3(t-1)[\sum_{i=1}^{t} 1/f_i - 1/\sum_{i=1}^{t} f_i] \tag{2.2.9.}$$

Bartlett (1947) demostró que $X_c^2 = M/C$ tiende a distribuirse como una X^2 con t-1 grados de libertad si y solamente si $H_o : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_t^2$ es cierta.

La regla de decisión es:

Si $X_c^2 \geqq X^2_{\alpha,\ t-1\ g.l.}$ se rechaza la hipótesis de homogeneidad de - varianzas (es decir, se rechaza la hipótesis $H_o : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_t^2$).

Sobre la independencia de los errores Snedecor y Cochran (1980) - consideran que si no se toma precaución puede suceder que en el experi-- mento se induzcan correlaciones positivas entre errores para diferentes repeticiones del mismo tratamiento. En ausencia de efectos reales de tratamientos el cuadrado medio de éstos, estimaría a $\sigma^2[1-(r-1)\rho_I]$, donde r es el número de repeticiones y ρ_I es el coeficiente de correlación interclase; el cuadrado medio del error estimaría a $\sigma^2(1-\rho_I)$. De manera que - la razón F_c sería estimado por:

$$F_c = 1-(r-1)\rho_I/(1-\rho_I) \qquad 2.2.10.$$

Por lo anterior si ρ_I es positiva la razón anterior puede ser mucho más grande que 1; así que correlaciones positivas entre los errores dentro de un tratamiento restan validez a la prueba de F, dando demasiados resultados significativos. La precaución más efectiva es el uso cuidadoso de asignación aleatoria de los tratamientos a las unidades experimentales.

Hay varios procedimientos estadísticos para probar la normalidad de los resultados de variables aleatorias. Aquí presentamos la prueba debida a Shapiro y Wilk (1965) la cual consiste en probar en una muestra - completa su normalidad. El estadístico de prueba es obtenido dividiendo el cuadrado de una apropiada combinación lineal de los estadísticos de - orden muestral por el usual simétrico estimador de varianza. Este cociente es invariante en ambas escalas y origen y por lo tanto la estadística es apropiada para una prueba de hipótesis compuestas de normalidad.

Para derivar la estadística W los autores consideran lo siguiente:

Si $m' = (m_1, m_2, \ldots, m_n)$ denota el vector ordenado de valores esperados de los estadísticos de orden de la normal estándar y $V = (V_{ij})$ - es la correspondiente matriz de covarianzas de dimensión n x n. Esto es, si $X_1 \leqq X_2 \leqq \ldots \leqq x_n$ denota una muestra aleatoria ordenada de tamaño n - de una distribución normal con media cero y varianza $\sigma^2 = 1$, entonces:

$$E(X_i) = m_i \qquad i = 1, 2, \ldots, n \qquad 2.2.11.$$

y

$$Cov(X_i, X_j) = V_{ij} \qquad i, j = 1, 2, \ldots, n \qquad 2.2.12.$$

Por otro lado, sea $Y' = (Y_1, Y_2, \ldots, Y_n)$ un vector ordenado de observaciones aleatorias. El objetivo es derivar una prueba para la hipótesis de que ésta es una muestra de una distribución normal con media μ desconocida y varianza σ^2 desconocida.

Si el $\{Y_i\}$ es una muestra de una normal entonces Y_i puede expresarse como:

$$Y_i = \mu + \sigma X_i, \qquad i = 1, 2, \ldots, n \qquad 2.2.13.$$

Por el teorema de mínimos cuadrados generalizado se sigue que los mejores estimadores lineales insesgados para μ y σ se obtienen minimizando la forma cuadrática $(Y-\mu 1-\sigma m)'v^{-1}(Y-\mu 1-\sigma m)$, donde $1' = (1,1, \ldots, 1)$. Estos estimadores son respectivamente:

$$\hat{\mu} = [m'v^{-1}m(m1'1m')v^{-1}y]/[1'v^{-1}1m'v^{-1}m-(1'v^{-1}m)^2] \qquad 2.2.14.$$

y

$$\hat{\sigma} = [1'v^{-1}(1m'-m1')v^{-1}y]/[1'v^{-1}1m'v^{-1}m-(1'v^{-1}m)^2] \qquad 2.2.15.$$

Para distribuciones simétricas, $1'v^{-1}m = 0$ y entonces:

$$\hat{\mu} = \sum_{i=1}^{n} Y_i/n = \bar{Y}, \text{ y } \hat{\sigma} = (m'v^{-1}y)/(m'v^{-1}m) \qquad 2.2.16.$$

Sea $S^2 = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$ el cual denota el usual simétrico estimador insesgado de $(n-1)\sigma^2$.

La W de la prueba estadística para normalidad está definida por:

$$W = (R^4\hat{\sigma}^2)/(C^2S^2) = b^2/S^2 = (a'Y)^2/S^2 = (\sum_{i=1}^{n} a_iY_i)^2/\sum_{i=1}^{n}(Y_i-\bar{Y})^2 \qquad 2.2.17.$$

donde: $R^2 = m'V^{-1}m$ 2.2.18.

$C^2 = m'V^{-1}V^{-1}m$ 2.2.19.

$a' = (a_1, a_2, \ldots, a_n) = (m'V^{-1})/(m'V^{-1}V^{-1}m)^{\frac{1}{2}}$ 2.2.20.

$b = R^2\sigma/C$ 2.2.21.

Así, b es el mejor estimador insesgado de la pendiente de una regresión lineal de las observaciones ordenadas, Y_i, sobre los valores esperados, m_i, de los estadísticos de orden de la normal estándar. La constante C es definida para que los coeficientes lineales estén normalizados.

Puede notarse que si la muestra es verdaderamente de una población normal, entonces el numerador b^2 y el denominador S^2 de W son ambos estimadores de la misma cantidad σ^2. Para poblaciones no-normales esas cantidades en general no estiman la misma cosa.

Algunas propiedades analíticas de W son:

Lema 1. W es invariante en escala y origen.

Corolario 1. W tiene una distribución la cual depende solamente del tamaño de la muestra n, para muestras de una distribución normal.

Corolario 2. W es estadísticamente independiente de S^2 y de $\bar{Y}$, para muestras de una distribución normal.

Corolario 3. $E(W^r) = E(b^{2^r})/E(S^{2^r})$:

para cualquier r. 2.2.22.

Lema 2. El máximo valor de W es 1.

Lema 3. El mínimo valor de W es $na_1^2(n-1)$.

Lema 4. El medio y primer momento de W están dados por:

$$E(W^{\frac{1}{2}}) = \frac{R^2\sqrt{\frac{n-1}{2}}}{C\sqrt{\frac{n}{2}}\sqrt{2}} \qquad 2.2.23.$$

$$E(W) = \frac{R^2\ (R^2+1)}{C^2\ (n-1)} \qquad 2.2.24.$$

Lema 5. Una distribución conjunta involucrando a W está definida por:

$$h(W,\Theta_2,\ldots,\Theta_{n-2}) = KW^{\frac{1}{2}}\ (1-W)^{\frac{1}{2}(n-4)}\mathrm{Cos}^{n-4}\Theta_2\ldots\mathrm{Cos}\Theta_{n-3} \qquad 2.2.25.$$

sobre una región T en la cual las $\Theta_{i,s}$ y W no son independientes y donde K es una constante.

Corolario 4. Para n = 3, la densidad de W es:

$$(3/\pi)(1-W)^{-\frac{1}{2}}W^{-\frac{1}{2}},\ 3/4 \leqq W \leqq 1 \qquad 2.2.26.$$

El procedimiento para probar la hipótesis nula compuesta de normalidad de una muestra aleatoria completa de tamaño n, X_1, X_2 ..., X_n, se da a continuación:

(i) Ordenar las observaciones para obtener una muestra ordenada $Y_1 \leqq Y_2 \leqq \ldots \leqq Y_n$.

(ii) Calcular:

$$S^2 = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 \qquad 2.2.27.$$

(iii)(a) Si n es par, n = 2k, calcule:

$$b = \sum_{i=1}^{k} a_{n-i+1}(Y_{n-i+1}-Y_i) \qquad 2.2.28.$$

donde los de a_{n-i+1} son los valores tabulados para esta - prueba [Apêndice nueve de Anderson y McLean (1974)].

(b) Si n es impar, $n = 2k + 1$, el cálculo es hecho como en (iii)(a), siendo $a_{k+1} = 0$ cuando $n = 2k + 1$.

Asî encontramos

$$b = a_n(Y_n-Y_1) + \ldots + a_{k+2}(Y_{k+2}-Y_k) \qquad 2.2.29.$$

donde el valor de Y_{k+1}, la mediana muestral no entra en el cálculo de b.

(iv) Calcule: $W = b^2/S^2$ 2.2.30.

(v) uno, cinco, 10, 50, 90, 95 y 99 por ciento para la distribu ción de W de tabla [Apêndice 10 de Anderson y McLean - (1974)]. Valores pequeños de W de tabla son significativos, o sea que indican no-normalidad.

Anderson y McLean (1974) comentan que hay varias pruebas para nor malidad; sin embargo, ellos recomiendan la prueba de Shapiro-Wilk porque consideran que es una prueba apropiada para hipôtesis compuestas de normalidad, dado que no incluye a la media, la varianza como parte de la hi pôtesis, como es comûn en algunas otras pruebas tales como la de - Kolmogorov-Smirnov y la Chi-cuadrada.

Shapiro <u>et al</u>. (1968) realizaron un estudio para comparar nueve - pruebas para normalidad; las estadîsticas consideradas en dicho trabajo fueron W (de Shapiro-Wilk 1965), $\sqrt{b_1}$ (tercer momento estándar), b_2 - (cuarto momento estándar), K_s (Kolmogorov-Smirnov), CM (Cramer-Von - Mises), WCW (CM ponderada), D(KS modificada), CS (Chi-cuadrada) y μ - (rango estudentizado).

Los autores encontraron que: (i) la estadística W provee en general un medio superior para medir la no-normalidad; (ii) las pruebas contrastantes (KS, CM, WCM, D) son típicamente insensitivas; (iii) la estadística μ es excelente, controla la simetría, especialmente de cola corta, en distribuciones pero tiene virtualmente no sensibilidad para la - asimetría; (iv) una combinación de ambas $\sqrt{b_1}$ y b_2 usualmente provee un juicio sensible pero a la par con su ejecución combinada es generalmente dominada por W; (v) con un procedimiento sensible buen indicador de ex-- trema no-normalidad (por ejemplo, la distribución exponencial) puede ser llevada a cabo con muestras de tamaño menor que veinte.

Shapiro y Francia (1972) presentan una modificación a la estadística W de Shapiro-Wilk para la prueba de normalidad: La prueba original de los coeficientes y porcentajes de puntos para muestras de tamaño me-- nor de 100. La modificación consiste en usar coeficientes que sólo dependen de los valores esperados de las estadísticas de orden de la nor----- mal. Argumentan que es deseable que la prueba use coeficientes que dependen solamente de los valores esperados de las estadísticas de orden de - la normal. Consecuentemente, una prueba W' aproximada la definen como:

$$W' = (\sum_{i=1}^{n} b_i Y_i)^2 / \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \qquad 2.2.31.$$

donde

$$b' = (b_1, b_2, \ldots, b_n) = m'/(m'm)^{\frac{1}{2}} \qquad 2.2.32.$$

los valores de m están dados por Harter (1961).

Weisberg y Bingham (1975), hacen una aproximación a la estadística de Shapiro-Francia con la estadística $\tilde{W}'$ que es más conveniente para máquinas computadoras. Ellos muestran que $\tilde{W}'$ es equivalente a W' y discuten el porcentaje de puntos.

Los autores definen el vector:

$$\tilde{m}' = (\tilde{m}_1, \tilde{m}_2, \ldots, \tilde{m}_n) \qquad 2.2.33.$$

por

$$\tilde{m}_i = \Phi^{-1}[(i-3/8)/(n+1/4)] \qquad 2.2.34.$$

$$i = 1, 2, \ldots, n,$$

donde $\Phi^{-1}(\rho)$ es el inverso de la función de distribución acumulada de la normal estándar evaluada en ρ. Como prueba de normalidad proponen el uso de:

$$\tilde{W}' = [(\tilde{m}'X)^2/(\tilde{m}'\tilde{m})]/\sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad 2.2.35.$$

siendo X un vector aleatorio consistiendo de las estadísticas de orden de una muestra aleatoria de n observaciones.

Sobre la Transformación de Datos a Otra Escala

Infante (1972) considera que debido a fallas en alguno de los supuestos del análisis de varianza, puede sugerirse la conveniencia de transformar los datos originales. Las razones para hacer una transformación son entre otras las siguientes:

(i) Heterogeneidad de varianzas. Esto frecuentemente debido a una relación entre la media y la varianza.

(ii) No aditividad en el modelo. También es muy frecuente que esta falta de aditividad ocasione el problema (i).

(iii) Puesto que algunas transformaciones, al mismo tiempo que homogeinizan las varianzas son útiles también para aproximar datos no-normales a la normal.

Respecto a cual transformación es la adecuada, la lógica de cada cambio de variable puede sugerirla. Tukey ha sugerido un método para seleccionar una transformación de acuerdo con los resultados de su prueba de aditividad.

Para el caso de heterogeneidad de varianzas, la teoría para seleccionar la transformación Graybill (1961) considera lo siguiente: Sea Y una variable aleatoria y supongamos que conocemos a la media de Y por la función f(t). Es decir, si $\mu = E(Y)$ y $\sigma^2 = V(Y)$, entonces $\sigma^2 = f(\mu)$. Por ejemplo, si Y es una variable que se distribuye como una poisson con parámetro m, entonces $\mu = m$ y $\sigma^2 = m$; así $\mu = \sigma^2$. Si Y es una variable binomial con parámetro ρ, $\mu = p$ y $\sigma^2 = p(1-p)$; así $\sigma^2 = \mu(1-\mu)$. Ahora bien si las condiciones son satisfechas, la variable Y es una variable normal y en general esperamos que la media y la varianza no estarán correlacionadas. Por esta razón, si sabemos que $\sigma^2 = f(\mu)$, deseamos encontrar una transformación h(Y) tal que V[h(Y)] no esté relacionada con E[h(Y)]. Supongamos que X = h(Y) es la variable transformada. El problema es encontrar una función h(Y) tal que la varianza de X sea una constante no relacionada con E(X). Supongamos que h(Y) puede expandirse alrededor de μ por una serie de Taylor. Así tenemos:

$$X = h(Y) = h(\mu) + (Y-\mu)h'(\mu) + R \qquad 2.3.1.$$

si se considera R despreciable, entonces:

$$X = h(\mu) + (Y-\mu)h'(\mu) \qquad 2.3.2.$$

donde $h'(\mu)$ es la derivada de $h(Y)$ evaluada en $Y = \mu$.

Ahora

$$E(X) = E[h(\mu) + (Y-\mu)h'(\mu)] = h(\mu) \qquad 2.3.3.$$

dado que :

$$E(Y) = \mu$$

también:

$$\begin{aligned} V(X) &= E[X-E(X)]^2 = E[X - h(\mu)]^2 \\ &= E[h(\mu) + (Y-\mu)h'(\mu) - h(\mu)]^2 \\ &= E[(Y -\mu)h'(\mu)]^2 = [h'(\mu)]^2 E[Y -\mu]^2 \\ &= \sigma^2[h'(\mu)]^2 \qquad 2.3.4. \end{aligned}$$

esto debido a que:

$$\sigma^2 = E[Y -\mu]^2 \qquad 2.3.5.$$

como se asumió que $\sigma^2 = f(\mu)$, donde $f(\mu)$ es conocida, entonces:

$$V(X) = f(\mu)[h'(\mu)]^2 \qquad 2.3.6.$$

por otra parte, se desea que $V(X)$ sea independiente de μ se establece - que $V(X)$ es igual a una constante C^2. Entonces:

$$C^2 = f(\mu)[h'(\mu)]^2 \Longrightarrow h'(\mu) = C/\sqrt{f(\mu)} \qquad 2.3.7.$$

entonces:

$$h(\mu) = C \int dt/\sqrt{f}(t) = CG(\mu) + K \qquad 2.3.8.$$

donde $G(\mu)$ es la integral indefinida de la función:

$$1/\sqrt{f(\mu)} \qquad 2.3.9.$$

y K es la constante de integración. De hecho, las constantes C^2 y K no - son de importancia, dado que ellas no dependen de μ. A ellas puede asignárseles un valor conveniente (con tal de que no estén relacionadas con μ). Puede ser que convenga hacer $K = 0$ y $C^2 = 1$.

Según Infante (1972) la transformación raíz cuadrada es obtenida cuando la función que relaciona a la varianza con la media es - $\sigma^2 = f(\mu) = \mu$. Esta función recuerda a la distribución poisson donde media y varianza son iguales. La transformación raíz cuadrada se usa en casos donde los resultados son números enteros que se refieren a conteos - de eventos raros tales como artículos defectuosos, accidentes, número de plantas en una área; insectos sobrevivientes después de una aplicación, etcétera. Si los números son menores que diez, Bartlett (1947) sugiere - usar $\sqrt{Y+0.5}$, $\sqrt{Y+1}$ o transformación logarítmica es obtenida cuando la función que relaciona a varianza con la media es de la forma:

$$\sigma^2 = f(\mu) = w^2\mu^2 \qquad 2.3.10.$$

es decir, cuando la desviación estándar es proporcional a la media. Esta transformación puede usarse cuando el rango es proporcional a la media. La transformación se usa para proporciones pero más a menudo para con--- teos. Bartlett (1947) sugiere que si los datos contienen ceros se utilice $\log_e(Y + 0.5)$ o $\log_e (Y + 1)$.

La transformación Arco Seno o Seno Inverso, se sugiere cuando la varianza es proporcional a $\mu(1-\mu)/n$; es decir, la función que relaciona a la varianza con la media es de la forma:

$$\sigma^2 = f(\mu) = \mu(1-\mu)/n \qquad 2.3.11.$$

el uso de esta transformación es para cuando se tienen proporciones o porcentajes, pero no si se tienen variables contínuas y cada observación se divide por un valor constante para formar proporciones. Bartlett (1947) ha sugerido que con $n < 50$, una proporción cero debe tomarse como $(1/4)n$ y una proporción 100 porciente como $(n - 1/4)/n$.

Ostle (1983) presenta una tabla donde resume cuatro transformaciones, indicando en que situaciones deben de usarse:

Tabla 1.

TRANSFORMACION	ECUACION	CONDICIONES QUE CONDUCEN A SU APLICACION
Logarítmica	$Y' = \log_e Y$	1. Los efectos verdaderos son multiplicativos. 2. La desviación estándar es proporcional a la media.
Raíz cuadrada	$Y' = \sqrt{Y}$ $Y' = \sqrt{Y+1}$	La varianza es proporcional a la media (por ejemplo cuando los datos originales son muestras de una distribución Poisson.
Arc Sen	$Y' = \text{arc se}\sqrt{\rho}$	La varianza es proporcional a $\mu(1-\mu)$ como por ejemplo cuando los datos originales son muestras (expresados como proporciones o frecuencias relativas) de poblaciones binomiales.
Recíproca	$Y' = 1/Y$	La desviación estándar es proporcional al cuadrado de la media.

Algunos Aspectos de los Estadísticos de Orden

Ostle (1983) dice que las observaciones aleatorias en algunos casos se encuentran ordenadas de acuerdo a su magnitud. Ello ocurre de dos maneras:

1. Inicialmente las observaciones se obtienen al azar, pero posteriormente se ordenan de acuerdo a su magnitud y

2. Las observaciones que se obtienen por naturaleza según su magnitud.

Como ejemplo del segundo caso, considera el ensayo de duración respecto a la vida o duración de pipetas. La primera observación que surge es la asociada al tubo más débil (o sea el tubo de vida más corto), la segunda observación está asociada al siguiente tubo débil y así sucesivamente.

Mendenhall *et al.* (1986) dicen que muchas funciones de variables aleatorias que son de interés en la práctica, dependen de las magnitudes relativas de las variables observadas. Por ejemplo, nos puede interesar el tiempo de máxima velocidad en una carrera automovilista o el ratón de mayor peso entre los ratones alimentados con cierta dieta.

Así ordenamos muchas veces las variables aleatorias según sus magnitudes. Las variables ordenadas resultantes se conocen como estadísticos de orden.

Brunk (1979) considera que generalmente para analizar la información de una muestra depende de una u otra de dos suposiciones básicas:

1. El tamaño de muestra es suficientemente grande, de tal modo que ciertas funciones de las variables aleatorias muestrales (por ejemplo la media muestral) tienen distribuciones que son aproximadas por distribuciones límite;

2. La población principal es normal. Dice el autor que aunque en la práctica es muy frecuente que ocurra una de estas suposiciones, no es de ninguna manera siempre de ese modo. Presenta algunos métodos que dependen sólo de la suposición.

3. La muestra aleatoria viene de una población cuya función de distribución de probabilidad es continua. Considera que para este caso, los métodos no son mejores para encontrar intervalos de confianza para la media de la población (un parámetro) o para probar la hipótesis de que la varianza de la población (un parámetro) es un número específico, cuando la forma de la distribución se especifica; estos métodos son conocidos como métodos no paramétricos. Estos métodos no dependen de suposiciones sobre las distribuciones como en uno y dos (excepto cuando la función sea continua); estos métodos son conocidos también como métodos de distribución libre. Estos métodos frecuentemente requieren el empleo de estadísticos de orden. El mismo autor de la siguiente definición: Definición (estadísticos de orden). Dejemos que Y_1 denote la más pequeña de las variables aleatorias muestrales, Y_2 la segunda más pequeña y así sucesivamente, de tal modo que Y_n sea la mayor. Las variables aleatorias $Y_1, Y_2, \ldots, Y_n$ son las estadísticas de orden de la muestra.

Lindgren (1970) indica que para las variables aleatorias continuas e independientes $Y_1, Y_2, \ldots, Y_n$ con función de densidad f(g) y denotando a las variables aleatorias, ordenadas Y_i por $Y_1 \leqq Y_2 \leqq \ldots \leqq Y_n$ - se sigue que:

(i) La función de densidad conjunta para $Y_1, Y_2, \ldots, Y_n$ es:

$$f y_1, y_2, \ldots, y_n(y_1, y_2, \ldots y_n) = n! \prod_{i=1}^{n} f(Y_i), \quad Y_1 \leqq Y_2 \leqq \ldots \leqq Y_n$$

$$= 0 \text{ d.o.f.} \qquad 2.4.1.$$

(ii) La función de distribución y la función de densidad para la K-ésima observación más pequeña son:

$$F y_K(Y) = P[Y_K \leqq Y] = P[K \text{ o más de las n observaciones son } \geqq Y]$$

$$= \sum_{j=K}^{n} C_j^n \, [F(Y)]^j [1-F(Y)]^{n-J} \qquad 2.4.2.$$

$$F y_K(Y) = nC_{K-1}^{n-1} \, [F(Y)]^{K-1} [1-F(Y)]^{n-K} \qquad 2.4.3.$$

(iii) Si K = n, se obtiene la función de distribución y función de densidad para la observación mayor Y_n:

$$F y_n(Y) = [F(Y)]^n, \quad f y_n(Y) = n[F(Y)]^{n-1} f(Y) \qquad 2.4.4.$$

(iv) Si K = 1, se obtiene la función de distribución y la función de densidad para la observación más pequeña Y_1 :

$$F y_1(Y) = 1-[1-F(Y)]^n, \quad f y_1(Y) = n[1-F(Y)]^{n-1} f(Y) \qquad 2.4.5.$$

(v) La función de distribución y la función de densidad conjunta para Y_1 y Y_n son:

$$F_{Y_1 Y_n}(\mu,\nu) = [F(\nu)]^n - [F(\mu)]^n, \text{ si } \mu \leqq \nu$$

$$= [F(\nu)]^n, \qquad \text{si } \mu > \nu \qquad 2.4.6.$$

$$f_{Y_1 Y_n}(\mu,\nu) = n(n-1)[F(\nu)-F(\mu)]^{n-2} f(\nu)f(\mu) \text{ si } \mu \leqq \nu$$

$$= 0, \text{ si } \mu > \nu \qquad 2.4.7.$$

(vi) La función de densidad para el rango muestral, $R = Y_n - Y_1$, es:

$$f_R(r) = n(n-1) \int_{-\alpha}^{\alpha} [F(\mu+r)-F(\mu)]^{n-2} f(\mu)f(\mu+r)d\mu \qquad 2.4.8.$$

el cual es válido para $r > 0$ y es cero si $r < 0$

Harter (1961) discute la obtención de los valores esperados de los estadísticos de orden de la normal. El valor esperado para la K-ésima mayor observación de una muestra de tamaño n de una población estándar está dada por la ecuación:

$$E(X_{k/n}) = \frac{n!}{(n-K)!(K-1)!} \int_{-\alpha}^{\alpha} X[\tfrac{1}{2}-\Phi(X)]^{n-K} \phi(X)dX \qquad 2.4.9.$$

$$\text{donde} \quad \phi(X) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}X^2} \qquad 2.4.10.$$

$$\text{y} \quad \Phi(X) = \int_0^X \phi(X)dX \qquad 2.4.11.$$

Para lo anterior hace uso de la aproximación de Blom para estimar el i-ésimo estadístico de orden para una muestra de tamaño n por medio - de la relación:

$$E(S_i) = \phi^{-1}\left(\frac{i - \alpha}{n-2\alpha+1}\right) , \qquad 2.4.12.$$

donde

$$\Phi(X) = \int_{-\alpha}^{\alpha} \phi(X)dX \quad y \quad \phi(X) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}X^2} \qquad 2.4.13.$$

El autor para encontrar un valor α requerido para evaluar $E(X_i)$, hace uso de las siguientes expresiones de acuerdo al tamaño de mues----- tra. Cuando ésta es par, n = 2K, y si es impar, entonces: n = 2K +1.

donde:

$$E(X_i) = -E(X_{n-i+1}) \text{ si } i = 1,2,\ldots K \qquad 2.4.14.$$

$$E(X_{K+1}) = 0 \qquad \text{si } n = 2K + 1 \qquad 2.4.15.$$

$$E(X_{n-i+1}) = \Phi^{-1}\left(1- \frac{i-\alpha_{i,n}}{n-2\alpha_{i,n}+1}\right) \qquad 2.4.16.$$

o

$$E(X_i) = \Phi^{-1} \left(\frac{i -\alpha_{i,n}}{n-2\alpha_{i,n}+1}\right) \qquad 2.4.17.$$

y despejando $\alpha_{i,n}$ de ésta

$$\alpha_{i,n} = \frac{i-(n+1)\phi[E(X_i)]}{1-2\Phi[E(X_i)]}; \; E(X_i) \neq 0 \qquad 2.4.18.$$

$$\alpha_{i,n} = \alpha_{n-i+1,\ n};\ i = 1,2,\ldots,K \qquad 2.4.19.$$

$$\alpha_{K+1,n} = \tfrac{1}{2} \qquad \text{si } n - 2K+1 \qquad 2.4.20.$$

Ahora bien, para la estimación de los valores $\alpha_{1,n}$, $\alpha_{2,n}$ y α_n , se usan las siguientes ecuaciones:

$$\hat{\alpha}_{i,n} = \alpha_{1,n} \cong .315065+0.057974\text{Loge}(n)-0.009776\text{Loge}(n)^2 \qquad 2.4.21.$$

$$\hat{\alpha}_{2,n} = \alpha_{2,n} \cong 0.327511+0.058212\text{Loge}(n)-0.007909\text{Loge}(n)^2 \qquad 2.4.22.$$

$$\hat{\alpha}_n = \alpha_n \cong 0.314195+0.063336\text{Loge}(n)-0.010895\text{Loge}(n)^2 \qquad 2.4.23.$$

3. MATERIALES Y METODOS

De la Fuente de Información

Para realizar el estudio se dispuso de la información de los trabajos de investigación utilizados como tesis para obtener el grado de licenciatura o de maestría en la División de Ingeniería de la Universidad Autónoma Agraria "Antonio Narro".

De los trabajos de investigación revisados se seleccionaron aquellos que presentaron los resultados originales de las características evaluadas. En total fueron 105 ensayos sobre las que se realizaron 395 pruebas de normalidad en muestras de tamaño diferente.

El efecto del diseño experimental y de tratamiento, fue eliminado de manera que la característica fuese evaluada, en cuanto a su normalidad, libre del efecto mencionado.

Estimación de los Valores Esperados de los Estadísticos de Orden de una Distribución Normal Estándar

Dado que la prueba de normalidad de Shapiro-Francia (1972) requiere de los valores esperados de los estadísticos de orden para obtener la Estadística W', es necesario conocer dichos valores; sin embargo, su comparación no es sencilla y las tablas donde éstos se presentan no son fáciles de conseguir.

Por otra parte, si se utiliza $m_K = \Phi^{-1}\left(\frac{K-3/8}{n+1/4}\right)$ como una -

aproximación para el valor esperado del K-ésimo estadístico de orden y con ellos se deriva el estadístico W' como sugiere Weisberg y Bringham (1975) se considera que los resultados que se obtienen se alejan de los de la prueba de Shapiro-Wilk. Por lo anterior, se analizó la conveniencia de derivar una aproximación más adecuada para valores esperados de los estadísticos de orden.

Algunas Propiedades de los Valores Esperados de los Estadísticos de Orden

Como se estableció en el capítulo de revisión de literatura, los valores esperados de los estadísticos de orden de una muestra aleatoria de tamaño n , tomada de una población normal estándar, satisfacen la siguiente relación:

$$m_K = -m_{n-K+1}; \; K=1,2,\ldots,n; \text{ para todo } n \in N \qquad 3.2.1.1.$$

$$m_{r+1} = 0 \qquad \text{Si} \qquad n = 2r + 1 \; ; \; r = 0,1,2\ldots$$

donde m_i es el valor esperado del i-ésimo estadístico de orden.

También se señaló que m_K puede expresarse como:

$$m_K = \Phi^{-1}\left(\frac{K-\alpha_{K,n}}{n-2\alpha_{K,n}+1}\right); \; K = 1,2,\ldots,n \qquad 3.2.1.2.$$

para alguna constante apropiada $\alpha_{K,n}$; donde $\Phi(.)$ denota la función de distribución normal estándar.

Es fácil ver que debido a la relación 3.2.1.1. y a 3.2.1.2., se tiene que:

$$m_{n-K+1} = \Phi^{-1}\left(1 - \frac{K-\alpha_{K,n}}{n-2\alpha_{K,n}+1}\right); \; K = 1,2,\ldots,n \qquad 3.2.1.3.$$

de modo que al igualar los términos dentro del paréntesis, se hace evi-- dente que:

$$\alpha_{K,n} = \alpha_{n-K+1,n}; \; K = 1,2,\ldots,n; \text{ para todo } n \in N \qquad 3.2.1.4.$$

$$\alpha_{r+1} = 1/2 \quad \text{si} \quad n = 2_r+1, \quad r = 0,1,2,\ldots$$

y por otro lado, que:

$$\Phi(m_K) = \frac{K-\alpha_{K,n}}{n-2\alpha_{K,n}+1} \qquad K = 1,2,\ldots n$$

lo cual se obtiene de 3.2.1.2.

Obtención de la Aproximación para los Valores Esperados de los Estadísticos de Orden

De la discusión anterior se desprende que si para alguna muestra dada de tamaño n , conocemos el conjunto $\{\alpha_{K,n}\}$ de constantes que satisfacen 3.2.1.2., utilizando algún método numérico, podemos conocer el conjunto $\{m_K\}$ de valores esperados de los estadísticos de orden. Es claro - entonces que el problema de encontrar una aproximación para los valores del conjunto $\{m_K\}$, es equivalente al de encontrar una aproximación para el conjunto $\{\alpha_{K,n}\}$. Es decir, si denotamos por $\bar{m}_K$; la aproximación para el valor del K-ésimo estadístico de orden m_K , y por $\bar{\alpha}_{K,n}$ la aproxima--- ción para la constante $\alpha_{K,n}$ asociada a m_K, entonces,

$$\bar{m}_K = \Phi^{-1}\left(\frac{K-\alpha_{K,n}}{n-2\bar{\alpha}_{K,n}+1}\right); \; K = 1,2,\ldots n \qquad 3.2.2.1.$$

Es evidente que la bondad de la aproximación para m_K, depende en general, de que tan "buena" es la aproximación para $\alpha_{K,n}$.

En base a una exploración numérica se encontró que:

$$\bar{\alpha}_{1,n} = 0.315065+0.057944\mathrm{Log}_{10}(n)-0.010895\mathrm{Log}_{10}(n)^2 \qquad 3.2.2.2.$$

$$= 0.315065+0.0251777882\mathrm{9log}_e(n) - 0.001843861\mathrm{og}_e(n)^2$$

$$\bar{\alpha}_{K,n} = 0.327511+0.058212\log_{10}(n)-0.007909\log_{10}(n)^2 \qquad 3.2.2.3.$$

$$= 0.327511+0.025811503\mathrm{8log}_e(n)-0.001491731\mathrm{og}_e(n)^2$$

y

$$\bar{\alpha}_{K,n} = \frac{(K-1) - (n + \frac{1471}{1470})\,\Phi\,(\bar{m}_{K-1})}{1 - 2\Phi(\bar{m}_{K-1})}; \; K>2 \qquad 3.2.2.4.$$

proporcionan una aproximación aceptable para el conjunto $\{\alpha_{K,n}\}$.

La discrepancia entre los valores de esta aproximación para $\bar{m}_K$ y los valores tabulados por Harter (1961), es del orden del tercer dígito después del punto decimal, considerándose que de esta manera Φ^{-1} (.) se obtiene con una precisión razonable; además, la discrepancia disminuye - cuando n aumenta.

Procedimiento de Evaluación para la Aproximación al Valor Esperado de los Estadísticos de Orden

Obsérvese que de acuerdo a la relación de simetría descrita en 3.2.1.1., solamente es necesario evaluar la aproximación para una parte de los valores esperados y automáticamente queda efectuada la evaluación de la parte complementaria.

También obsérvese que para las evaluaciones de $\{\bar{m}_K\}$ y $\{\bar{\alpha}_{K,n}\}$, es necesario tener un procedimiento para probar $\Phi(X)$ y $\Phi^{-1}(P)$. En el presente trabajo se utilizaron los siguientes:

(i) Procedimiento de Hasting para evaluar $\Phi(X)$:

$$\Phi(X) = 1-\phi(X)\ (b_1t+b_2t^2+b_3t^3+b_4t^4+b_5t^5) \qquad 3.2.3.1.$$

donde

$$t = \frac{1}{1+0.2316419X}$$

b_1 =-0.319381530

b_2 = 0.356563782

b_3 = 1.781477937

b_4 = -1.821255978

b_5 = 1.330274429

(ii) Un Procedimiento para evaluar $\Phi^{-1}(P)$, es decir el valor de X para el cual $P = \Phi(X)$, pudiera ser el método iterativo de Newton-Raphson, que a continuación se presenta. Sea X_i el valor obtenido para $\Phi^{-1}(P)$ en la i-ésima iteración, $i = 0,1,2,\ldots,$

entonces:

$$X_{i+1} = X_i - \frac{[\Phi(X_i)-P]}{P(X_i)} \qquad 3.2.3.2.$$

donde $p(.)$ es la función de densidad de la distribución normal estándar.

El procedimiento iterativo corre como sigue:

Etapa 1. Se propone un valor inicial X_0, el cual es dado por la aproximación racional.

Etapa 2. Se obtiene X_{i+1} usando 3.2.3.2.

Etapa 3. Se verifica si se ha alcanzado la precisión deseada (en nuestro caso si $/\Phi(P)-P/ < 10^{-8}$); si se alcanzó, el procedimiento termina asignando a X el último valor obtenido en la etapa 2, ésto es, el valor obtenido en la iteración i+1. Si no se alcanzó la precisión debe pasarse nuevamente a la etapa 2.

El método de aproximación racional para evaluar X_0 es:

$$X_0 = B - \frac{a_0 + a_1 B + a_2 B^2}{1 + C_1 B + C_2 B^2 + C_3 B^3} \qquad 3.2.3.3.$$

$$B = \sqrt{-2 \log_e (1-P)}\ , \ 0<P<1/2$$

donde:

$a_0 = 2.515517$ $c_1 = 1.432788$

$a_1 = 0.802853$ $c_2 = 0.0189269$

$a_2 = 0.016328$ $c_3 = .001308$

El procedimiento de evaluación para la aproximación de cada uno de los valores esperados se hace utilizando lo descrito en (i) e (ii) para cada caso y usando las fórmulas 3.2.2.1, 3.2.2.2, 3.2.2.3 y 3.2.2.4.

Obtención de la Aproximación para los valores de los Coeficientes $\{\hat{a}_{n-K+1}\}$ en la Prueba de Normalidad

Para estimar la aproximación a los coeficientes $\{\hat{a}_{n-K+1}\}$ por la ecuación:

$$b' = (b_1, b_2, \ldots, b_n) = \frac{m'}{(m'm)^{\frac{1}{2}}} \qquad 3.3.1.$$

Estos valores del coeficiente {bn-K+1} difieren como aproximación de los $\{a_{n-K+1}\}$, por lo que, mediante un proceso numérico se encontró que:

$$\hat{a}_{1,n} = .413090-1.220316(b_{1,n})+4.326634(b_{1,n})^2-2.847795(b_{1,n})^3 \qquad 3.3.2.$$

$$\hat{a}_{K,n} = Bo + EXP\ B_1\ (bn-K+1)\ ,\ K = 2,3,\ldots,n \qquad 3.3.3.$$

donde:

$Bo = (-4.684563E-03+1.834622E-04(n)-2.210639E-06(n^2)$

$B_1 = [-0.465886+0.306314(Log_e n)-0.069049(log_e n^2)+4.714998E-03(log_e n^3)]$

encontrándose así, la aproximación razonable para los coeficientes $\{a_{n-K+1}\}$. Dado que con este procedimiento se obtiene la misma relación simétrica observada anteriormente para los estadísticos de orden en la ecuación 3.2.1.1., sólo se evalúa la primera parte correspondiente a [int(n/2)].

Programa de Computación para la Prueba de Normalidad de Shapiro-Wilk

Se elaboró un programa en lenguaje Basic Cuadro 1A para el micro-computador CROMEMCO-CS100, que contempla:

(i) Elimina el efecto del diseño experimental empleando en el ensayo, para evaluar su normalidad.

(ii) Cálculo de las aproximaciones de los estadísticos de orden $\{\hat{m}_i\}$.

(iii) Obtención de la primera aproximación a los coeficientes $\{a_{n-K+1}\}$ de Shapiro-Wilk por la ecuación 3.3.1.

(iv) Estimación aproximada de los coeficientes $\{\hat{a}_{n-K+1}\}$ de Shapiro-Wilk para la prueba de normalidad.

(v) Cálculo de la estadística W' para la prueba de normalidad de Shapiro-Wilk.

Análisis de la Información Obtenida Mediante la Prueba de Normalidad de Shapiro-Wilk

Se formaron Cuadros de doble entrada con los resultados de los análisis y donde fue posible, se probó la hipótesis de igual proporción de no-normalidad para las características correspondientes a una clasificación agronómica considerada.

4. RESULTADOS

Características Agronómicas que se les Realizó Prueba de Normalidad de Shapiro-Wilk

Los cultivos agrícolas a que hicieron referencia los trabajos de investigación de donde se obtuvo la información para el presente estudio fueron: nogal, frijol, sorgo, tomate, maíz, zacate, durazno, trigo, cebada, arroz, ajo y papa.

Las características detectadas se presentan en el Cuadro 4.1.1., el cual hace referencia a la unidad de medición, frecuencia de aparición, rango de tamaño de muestra y rango de las unidades, en el estudio donde se detectó la normalidad y la no-normalidad, se clasifica la característica desde el punto de vista agronómico.

Resultados de la Prueba de Normalidad Según la Clasificación Agronómica

Con el objeto de resumir los resultados obtenidos, se hizo uso de Cuadros de doble entrada, donde por un lado se consideran características que pertenecen a la clasificación agronómica y por el otro frecuencias de normalidad y no-normalidad.

Cuadro 4.1.1. Concentración de los valores y características obtenidas con la prueba de normalidad de Shapiro-Wilk en el estudio realizado. U.A.A.A.N., Buenavista, Saltillo, Coah. 1987.

No. de Orden	Características y Unidades	Frec. de Aparición	Clasif. Agr.	Rango de tamaño muestra	Rango de unidad de normalidad	Rango de unidad de no-normalidad
				28		7.4 - 7.7
1	PH^{+}	24	Cationes	7 - 24	7.2 - 9.0	
				7 - 24		7.0 - 8.32
				12	23.9 - 31.9	
2	$C.I.C^{+}$	17	Cationes	7 - 15	12.0 - 54.6	
	meq/100 gr.			14 - 15		11.3 - 48.72
				28		1.12- 25
3	$C.E.^{+}$	23	Cationes	7 - 24	.10 - 90	
	$mmhos/cm^3$			7 - 21		.305-31.57
4	N_2^{+} %	10	Cationes	9 - 24	0 - 13.82	
				7 - 20	1.5 - 151.0	
5	Ca^{+} meq/lt.	14	Cationes	7 - 20		1.7 - 63
				7 - 20	.35 - 1.81	
6	Mg^{++} meq/lt.	14	Cationes	7 - 20		.02 - 10.0
				7 - 15	.26 - 13.36	
7	Na^{++} meq/lt.	14	Cationes	7 - 20		.26 - 5.56

Cuadro 4.1.1. ....continuación

No. de Orden	Características y Unidades	Frec. de Aparición	Clasif. Agr.	Rango de tamaño muestra	Rango de unidad de normalidad	Rango de unidad de no-normalidad
8	K^{++} meq/lt	2	Cationes	15	.096 - 13.26	
				20		.15 - 1.64
9	P_2O_5 PPM	8	Aniones	7 - 12	2.4 - 33.0	
				7		2.0 - 33.0
10	CO_3 meq/lt	5	Aniones	7	27.5 - 59.4	
				20		.3 - .6
11	CO_3 %	14	Aniones	49	82.1 - 83.55	
				28	88.63 -109.03	
				7 - 20	12.88 - 71.6	
				14 - 15		29.53 - 55.24
12	HCO_3 meq/lt	2	Aniones	15	4.0 - 6.0	
				20		2.0 - 4.0
13	CL meq/lt	2	Aniones	15	44.55 -249.55	
				20		3.5 -118.4
14	SO_4 meq/lt	2	Aniones	15 - 20	0 - 96.1	

Cuadro 4.1.1. ....continuación

No. de Orden	Características y Unidades	Frec. de Aparición	Clasif. Agr.	Rango de tamaño muestra	Rango de unidad de normalidad	Rango de unidad de no-normalidad
15	N Kg/ha	5	Elementos mayores	12	19.1 - 45.6	
				7	2.1 - 18.6	
				7		2 - 34.1
16	P_2O_5 Kg/ha	14	Elementos mayores	12		710.8 - 1568.4
				28	87.65 - 268.7	
				7 - 20	3 - 414.0	
				9 - 24		12.15- 180.0
17	K Kg/ha	11	Elementos mayores	28		710.8 - 1568.4
				15 - 24	0 -1125.0	
				9 - 15		260 - 900
18	Ca^{+} mg/100 gr.	4	Elementos menores	12	747 -1032	
				12		277 - 947
19	Mg^{+} mg/100 gr.	4	Elementos menores	12	14 - 73	32 - 69
				12		
20	Mn mg/100 gr.	5	Elementos menores	12	3.7 - 43	
21	Fe^{+} PPM	5	Elementos menores	12	5 - 150	
22	Cu^{+} PPM	5	Elementos menores	12		.60 - 28

Cuadro 4.1.1. ....continuación

No. de Orden	Características y Unidades	Frec. de Aparición	Clasif. Agr.	Rango de tamaño muestra	Rango de unidad de normalidad	Rango de unidad de no-normalidad
23	Zn^{+} PPM	5	Elementos menores	12		1.0 - 315
24	Arena %	6	Propiedades físicas	12 - 15 20	9.2 - 39.2	12.4 - 64.4
25	Limo %	6	Propiedades físicas	12 - 20	7.2 - 45.6	
26	Arcilla %	6	Propiedades físicas	12 - 20	21.2 - 61.6	
27	C.C %	4	Propiedades físicas	12 15 14 - 15	24.47 - 28.07 31.38 - 36.77	33.75 - 46.2
28	P.M.P %	4	Propiedades físicas	12 15 14	13.22 - 15.26 17.01 - 25.11	18.34 - 25.11
29	Da gr/cm^{3}	4	Propiedades físicas	12 14 - 15	1.22 - 1.85	1.0 - 1.2

Cuadro 4.1.1. ....continuación

No. de Orden	Características y Unidades	Frec. de Aparición	Clasif. Agr.	Rango de tamaño muestra	Rango de unidad de normalidad	Rango de unidad de no-normalidad
30	Ds gr/cm^3	3	Propiedades físicas	12	2.45 - 2.77	
				15	2.48 - 2.54	
				14		1.83 - 2.59
31	M.O. %	25	Propiedades Bromatológica.	12	1.2 - 1.9	
				49		.62 - 1.5
				7 - 20	.25 - 93.8	
				7 - 16		.93 - 96.8
32	M.S.T. %	8	Propiedades Bromatológica.	60		.409- 5.6
				16	93.73 - 98.21	
				16		95.6 - 97.95
33	P.C %	5	Propiedades Bromatológica.	16	5.08 - 12.5	
34	F.C. %	5	Propiedades Bromatológica.	16	26.68 - 37.84	
						31.81 - 35.6
35	E.E. %	5	Propiedades Bromatológica.	16	6.48 - 8.57	
				16		3.17 - 11.1
36	Cenizas %	5	Propiedades Bromatológica.	16	6.48 - 8.57	
				16		3.17 - 11.1

Cuadro 4.1.1. ....continuación

No. de Orden	Características y Unidades	Frec. de Aparición	Clasif. Agr.	Rango de tamaño muestra	Rango de unidad de normalidad	Rango de unidad de no-normalidad
37	E.L.N. %	5	Propiedades Bromatológicas.	16 16	48.7 - 59.6	51.2 - 63.3
38	Diámetro de Fruto. (cm)	3	Componente de Rend.	22 22	.46 - 1.06	4.7 - 5.1
39	Diámetro de Tallos (cm)	1	Componente de Rend.	72	.85 - 1.11	
40	Altura (cm)	12	Componente de Rend.	70 15 4 - 72	8.6 - 88.8 13 - 82.7	.985 - 1.18
41	No. de Espigas.	2	Componente de Rend.	20 24 4	576 - 780 304 - 648 72 - 142	
42	No. de granos en espigas o vaina.	8	Componente de Rend.	12 20 24 4 - 72 4	4 - 6 31.4 - 56.3 12 - 51	29.1 - 40.5 32 - 39
43	Peso 100 semillas (gr)	7	Componente de Rend.	12 24 20 4	25.15 - 28.6 24.7 - 38.1 27.7 - 38.1	24.3 - 47.6

Cuadro 4.1.1. ....continuación

No. de Orden	Características y Unidades	Frec. de Aparición	Clasif. Agr.	Rango de tamaño muestra	Rango de unidad de normalidad	Rango de unidad de no-normalidad
44	No. de Tallos	4	Componente de Rend.	18 24 36	394 - 570 147 - 260 4 - 7	
45	No. de Hojas/Tallo	1	Componente de Rend.	72	9 - 14	
46	No. de Plantas	1	Componente de Rend.	15	54.7 - 78.3	
47	Longitud de Espiga (cm)	10	Componente de Rend.	20 20 24 4 4	8.1 - 10.8 21.2 - 25.5 4.8 - 6.8 6.89 - 9.12	 7.43 - 7.82
48	Longitud de Raíz (cm)	1	Componente de Rend.			35.3 - 61.7
49	Longitud de Hoja (m)	1	Componente de Rend.	72	24 - 30	
50	Acame %	3	Componente de Rend.	20 - 21 20 - 21	0 - 30	 0 - 40

Cuadro 4.1.1. ....continuación

No. de Orden	Características y Unidades	Frec. de Aparición	Clasif. Agr.	Rango de tamaño muestra	Rango de unidad de normalidad	Rango de unidad de no-normalidad
51	Infección	4	Componente de Rend.	16 - 21		.06 - 40.0
52	Humedad en grano. %	1	Componente de Rend.	22	10.96 - 19.06	
53	Rendimiento %	4	Componente de Rend.	72	8.275 - 16.8	52.15 - 66.5
54	Indice de cosecha %	5	Componente de Rend.	24	.129 - .337	
				4	.239 - .417	
				4		.261 - .424
55	Rend. Biológico. kg/ha	1	Componente de Rend.	49	425 - 1075	
56	Rend. Agronómico. kg/ha	22	Componente de Rend.	12	854.5 - 1358.4	
				20	4216 - 7866	
				24	1494 - 4520	
				40	1534 - 4145	
				60	1375 - 8983	
				15	19340 - 28640	
				20	3888 - 7883	

Cuadro 4.1.1. ....continuación

No. de Orden	Características y Unidades	Frec. de Aparición	Clasif. Agr.	Rango de tamaño muestra	Rango de unidad de normalidad	Rango de unidad de no-normalidad
				80	3000 - 7500	
				96	0 - 10805	
				96		700 - 5425
				16	6220 - 22453	
				18	330 - 705	
				35	5.15 - 35.03	
				14	20.33 - 52.23	
				24	3.94 - 5.72	
				27	.18 - 9.45	
57	Longitud escurrida (cm)	2	C. Hidrológica	90		0 - 90.0
58	R.A.S	2	C. Hidrológica	12 - 21		.105 - 19.3

Del Cuadro 4.1.2. al 4.1.11 se presentan las características que corresponden a la clasificación agronómica considerada y la frecuencia - del resultado en la prueba de normalidad. Además, al pie del Cuadro se - da el valor de X_i^2 , $X^2_{0.05,\ K-1}$ y la decisión con respecto a la prueba de hipótesis de igualdad de proporción de no-normalidad para las K características que contiene la clasificación de referencia.

En el Cuadro 4.1.12 se muestra la frecuencia de los casos de no-normalidad y normalidad para las diferentes categorías de la clasificación agronómica y la decisión sobre la hipótesis de igual proporción de casos.

Las Aproximaciones de los Valores Esperados de los Estadísticos de Orden

En el Cuadro 4.2.1. para diferente tamaño de muestra, se presentan los valores esperados de los estadísticos de orden $\{m_{n-K+1}\}$ reportados por Harter (1961) y las aproximaciones obtenidas para los mismos valores esperados de los estadísticos de orden $\{\hat{m}_{n-K+1}\}$ y la correspondiente desviación que da idea del grado de precisión alcanzada.

Cuadro 4.1.2. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para las características correspondientes a cationes, en el estudio realizado. U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
PH^{+}	15 (62.5)	9 (37.5)
$C.I.C^{+}$ meq/100 gr.	15 (88.24)	2 (11.765)
$C.E^{+}$ mmhos/cm^{3}	12 (52.17)	11 (47.83)
N_2^{+} %	10 (100.0)	0 (0)
Ca^{++} meq/lt	10 (71.43)	4 (28.57)
Mg^{++} meq/lt	9 (64.29)	5 (35.71)
Na^{++} meq/lt	8 (57.14)	6 (42.86)
K^{++} meq/lt	1 (50.0)	1 (50.0)
TOTAL	80 (67.80)	38 (32.20)

$X_c^2 = 24.631$, $X^2_{0.05,7} = 14.067$,

Se rechaza la hipótesis de que hay igual proporción de no-normalidad entre las características correspondientes a cationes.

Cuadro 4.2.3. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para las características correspondientes a aniones, en el estudio realizado. U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
P_2O_5 PPM	6 (75.0)	2 (25.0)
CO_3 meq/lt	2 (40.0)	3 (60.0)
CO_3 %	12 (85.71)	2 (14.29)
HCO_3 meq/lt	1 (50.0)	1 (50.0)
CL meq/lt	1 (50.0)	1 (50.0)
SO_4 meq/lt	2 (100.0)	9 (0)
TOTAL	24 (72.73)	9 (27.27)

$X^2_c = 3.666$, $X^2_{0.05,5} = 11.070$

No se rechaza la hipótesis de que hay igual proporción de no-normalidad entre las características correspondientes a aniones.

Cuadro 4.1.4. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para las características correspondientes a elementos mayores en el estudio realizado. - U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
N Kg/ha	3 (60.0)	2 (40.0)
P_2O_5 Kg/ha	10 (71.43)	4 (28.57)
K Kg/ha	7 (63.64)	4 (36.36)
TOTAL	20 (66.67)	10 (33.33)

$X_c^2 = 0.80$, $X^2_{0.05,3} = 7.815$

No se rechaza la hipótesis de igualdad de proporción de no-normalidad.

Cuadro 4.1.5. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para las características correspondientes o elementos menores en el estudio realizado. - U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
Ca^+ mg/100 gr	3 (75.0)	1 (25.0)
Mg^+ mg/100 gr	3 (75.0)	1 (25.0)
Mn^+ mg/100	5 (100.0)	0 (0)
Fe^+ PPM	5 (100.0)	0 (0)
Cu^+ PPM	0 (0)	5 (100.0)
Zn^+ PPM	0 (0)	5 (100.0)
TOTAL	16 (57.14)	12 (42.86)

$X_c^2 = 14.0$, $X^2_{.05,5} = 10.07$; se rechaza la hipótesis de igualdad de proporción de no-normalidad para las características correspondientes a elementos menores.

Cuadro 4.1.6. Frecuencias de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para las características correspondientes a las propiedades físicas del suelo, en el estudio realizado. U.A.A.A.N., Buenavista, Saltillo, Coah., - 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
Arena %	5 (83.33)	1 (16.67)
Limo %	6 (100.0)	0 (0)
Arcilla %	6 (100.0)	0 (0)
C.C. %	2 (50.0)	2 (50.0)
P.M.P. %	3 (75.0)	1 (25.0)
Da. gr/cm^3	1 (25.0)	3 (75.0)
Ds gr/cm^3	2 (66.67)	1 (33.33)
TOTAL	25 (75.76)	8 (24.24)

$X^2_c = 6.0$, $X^2_{0.05,6} = 12.592$; no se rechaza la hipótesis de igual proporción de no-normalidad para las características correspondientes a propiedades físicas de suelo.

Cuadro 4.1.7. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para las características correspondientes a las propiedades bromatológicas en el estudio realizado. U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
M.O. %	17 (68.0)	8 (32.0)
M.S.T %	5 (71.43)	2 (28.57)
P.C. %	5 (100.0)	0 (0)
F.C. %	4 (80.0)	1 (20.0)
E.E. 5	3 (60.0)	2 (40.0)
CENIZAS %	1 (20.0)	4 (80.0)
E.L.N. %	4 (80.0)	1 (20.0)
TOTAL	38 (67.86)	18 (32.14)

$X^2_c = 17.00$, $X^2_{0.05,6} = 12.592$; se rechaza la hipótesis de igual proporción de no-normalidad para las características correspondientes a propiedades bromatológicas.

Cuadro 4.1.8. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para las características correspondientes a las propiedades de componentes de rendimiento en el estudio realizado. U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
Diámetro de fruto. (cms)	2 (66.67)	1 (33.33)
Diámetro de tallos. (cms)	1 (100.0)	0 (0)
Altura (cms)	11 (91.67)	1 (8.33)
No. de Espigas, Vainas.	2 (100.0)	0 (0)
No. de granos en espiga o vaina.	6 (75.0)	2 (25.0)
Peso de 100 semillas (gr)	6 (85.78)	1 (14.29)
No. de Tallos	4 (100.0)	0 (0)
No. de Hojas	1 (100.0)	0 (0)
No. de Plantas	1 (100.0)	0 (0)
Long. de Espiga (cm)	8 (80.0)	2 (20.0)
Long. de Raíz (cm)	0 (0)	1 (100.0)
Long. de Hoja (cm)	1 (100.0)	0 (0)
TOTAL	43 (84.31)	8 (15.69)

$X^2_c = 10.0$, $X^2_{0.05,11} = 19.675$; no se rechaza la hipótesis de igual proporción de no-normalidad para las características correspondientes a componentes de rendimiento.

Cuadro 4.1.9. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para las características agroquímicas, en el estudio realizado. U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
Acame %	2 (66.67)	1 (33.33)
Infección %	0 (0)	4 (100.0)
Humedad en grano %	1 (100.0)	0 (0)
Rendimiento %	2 (50.0)	2 (50.0)
Indice de cosecha %	3 (60.01)	2 (40.0)
TOTAL	8 (47.06)	9 (52.94)

$X^2_c = 4.889$, $X^2_{0.05,4} = 9.488$; no se rechaza la hipótesis de igual proporción de no-normalidad para las características agronómicas evaluadas.

Cuadro 4.1.10. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para la característica rendimiento en el estudio realizado. U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
Rendimiento Biológico kg/ha	1 (100.0)	0 (0)
Rendimiento Agronómico kg/ha	19 (86.36)	3 (13.14)
TOTAL	20 (86.96)	3 (13.04)

Cuadro 4.1.11. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para las características hidrológicas, en el estudio realizado. U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características y Unidades	Normalidad (%)	No-normalidad (%)
Longitud de lámina escurrido (m)	0 (0)	2 (100.0)
R.A.S.	0 (0)	2 (100.0)
TOTAL	0 (0)	4 (100.0)

$X_c^2 = 0.0$, $X_c^2 = 3.84$; no se rechaza la hipótesis de igual proporción - de no-normalidad para las características evaluadas.

Cuadro 4.1.12. Frecuencia de casos de normalidad y no-normalidad con la prueba de Shapiro-Wilk para la clasificación agronómica en el estudio. U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Categorías de Clasificación	Normalidad	No-normalidad
Rendimiento agronómico	22	3
Hidrológicas	0	4
Propiedades físicas	25	8
Componentes del rendimiento	43	8
Aniones	24	8
Características agronómicas en porcentaje	8	9
Elementos mayores	20	10
Elementos menores	16	12
Propiedades bromatológicas	38	18
Cationes	80	38
TOTAL	276	119

$X_c^2 = 74.67$, $X^2_{0.05,9} = 16.912$; se rechaza la hipótesis de igual proporción de no-normalidad en las categorías de clasificación agronómica.

Las Aproximaciones de los Valores $\{\hat{a}_{n-K+1}\}$ de la Tabla para la Prueba de Normalidad de Shapiro-Wilk

En el Cuadro 4.3.1. para diferente tamaño de muestra, se dan los valores de tabla reportados por Shapiro-Wilk, las aproximaciones que, para ellos se obtienen $\{\hat{a}_{n-K+1}\}$ y la correspondiente discrepancia entre ellas.

La Aproximación de la Estadística W de Shapiro-Wilk

Para diferentes tamaños de muestra, se tomaron datos reales y fueron evaluados mediante los valores de tabla de Shapiro-Wilk y con las aproximaciones $\{\hat{a}_{n-K+1}\}$ en la prueba de normalidad. Los resultados obtenidos se presentan en el Cuadro 4.4.1., y se observa la desviación de la estadística W.

Cuadro 4.2.1. Desviación de valores esperados de los estadísticos de orden $\{m_{m-K+1}\}$ y sus correspondientes aproximaciones $\{\hat{m}_{m-K+1}\}$.

n = 4

i	m_{n-K+1}	m_{n-K+1}	d
1	1.02938	1.03249	-.0031
2	0.29701	0.30254	-.0055

Cuadro 4.2.1. ....continuación

n = 10

i	m_{n-K+1}	$\hat{m}_{n-K+1}$	d
1	1.53875	1.54104	-.0023
2	1.00136	1.00482	-.0035
3	0.65606	0.66019	-.0041
4	0.37576	0.38085	-.0051
5	0.12267	0.12875	-.0061

Cuadro 4.2.1. ....continuación

n = 16

i	m_{n-K+1}	$\hat{m}_{n-K+1}$	d
1	1.76599	1.76828	-.0023
2	1.28474	1.28773	-.0030
3	0.99027	0.99363	-.0034
4	0.76317	0.76719	-.0040
5	0.57001	0.57463	-.0046
6	0.39622	0.40140	-.0052
7	0.23375	0.23948	-.0057
8	0.07729	0.08361	-.0063

Cuadro 4.2.1. ....continuación

n = 25

i	m_{n-K+1}	$\hat{m}_{n-K+1}$	d
1	1.96531	1.96768	-.0024
2	1.52430	1.52696	-.0027
3	1.26275	1.26560	-.0028
4	1.06679	1.07017	-.0034
5	0.90501	0.90886	-.0038
6	0.76405	0.76830	-.0042
7	0.63690	0.64150	-.0046
8	0.51935	0.52427	-.0049
9	0.40860	0.41384	-.0052
10	0.30268	0.30824	-.0056
11	0.20006	0.20596	-.0059
12	0.09953	0.10578	-.0063

Cuadro 4.2.1. ....continuación

n = 38

i	m_{n-K+1}	$\hat{m}_{n-K+1}$	d
1	2.14009	2.14251	-.0024
2	1.72914	1.73157	-.0024
3	1.49061	1.49306	-.0025
4	1.31514	1.318044	-.0029
5	1.17280	1.17610	-.0033
6	1.05095	1.05458	-.0036
7	0.94300	0.94691	-.0039
8	0.84508	0.84924	-.0042
9	0.75468	0.75906	-.0044
10	0.67009	0.67467	-.0046
11	0.59005	0.59484	-.0048
12	0.51363	0.51860	-.0050
13	0.44012	0.4530	-.0052
14	0.36892	0.37430	-.0054
15	0.29954	0.30514	-.0056
16	0.23159	0.23740	-.0058
17	0.16469	0.17072	-.0060
18	0.09853	0.10479	-.0063
19	0.03280	0.03931	-.0065

Cuadro 3.2.1. ....continuación

i	m_{n-K+1}	$\hat{m}_{n-K+1}$	d
28	0.40111	0.406380	-.0053
29	0.36746	0.37283	-.0054
30	0.33423	0.33969	-. 006
31	0.30136	0.30692	-. 006
32	0.26881	0.27448	-.0057
33	0.23655	0.24232	-.0058
34	0.20453	0.210403	-.0059
35	0.17272	0.17870	-. 006
36	0.14108	0.147176	-.0061
37	0.10959	0.115795	-.0062
38	0.07820	0.084526	-.0063
39	0.04689	0.053338	-.0064
40	0.01562	0.022199	-.0066

Cuadro 4.2.1. ....continuación

i	m_{n-K+1}	$\hat{m}_{n-K+1}$	d
28	0.40111	0.406380	-.0053
29	0.36746	0.37283	-.0054
30	0.33423	0.33969	-. 006
31	0.30136	0.30692	-. 006
32	0.26881	0.27448	-.0057
33	0.23655	0.24232	-.0058
34	0.20453	0.210403	-.0059
35	0.17272	0.17870	-. 006
36	0.14108	0.147176	-.0061
37	0.10959	0.115795	-.0062
38	0.07820	0.084526	-.0063
39	0.04689	0.053338	-.0064
40	0.01562	0.022199	-.0066

Cuadro 4.2.1. ....continuación

$n \doteq 80$

i	m_{n-K+1}	$\hat{m}_{n-K+1}$	d
1	2.42677	2.42912	-.0024
2	2.05714	2.05924	-.0021
3	1.84832	1.85021	-.0019
4	1.69798	1.70019	-.0022
5	1.57836	1.58089	-.0025
6	1.57836	1.58089	-.0025
7	1.57836	1.58089	-.0025
8	1.31236	1.31558	-.0032
9	1.24165	1.24509	-.0034
10	1.17666	1.180167	-.0035
11	1.11631	1.119941	-.0036
12	1.05978	1.063524	-.0037
13	1.00644	1.010304	-.0039
14	0.95584	0.959802	-. 004
15	0.90757	0.911637	-.0041
16	0.86134	0.865498	-.0042
17	0.81687	0.821132	-.0043
18	0.77398	0.778326	-.0043
19	0.73246	0.736899	-.0044
20	0.69217	0.696701	-. 005
21	0.65297	0.657597	-. 005
22	0.61476	0.619473	-. 005
23	0.57742	0.58223	-.0048
24	0.54088	0.545734	-.0049
25	0.50504	0.51003	-. 005
26	0.48985	0.47492	-.0051
27	0.43522	0.440394	-.0052

Cuadro 4.2.1. ....continuación

n = 50

i	m_{n-K+1}	$\hat{m}_{n-K+1}$	d
1	2.24907	2.25149	-.0024
2	1.85487	1.85717	-.0023
3	1.62863	1.63086	-.0022
4	1.46374	1.46637	-.0026
5	1.33109	1.33409	-.0030
6	1.21846	1.22175	-.0030
7	1.11948	1.12302	-.0030
8	1.03042	1.03418	-.0038
9	0.94887	0.95283	-.0039
10	0.87321	0.87734	-.0041
11	0.80225	0.80655	-.0042
12	0.73513	0.73958	-.0044
13	0.67117	0.67578	-.0046
14	0.60986	0.61461	-.0048
15	0.55077	0.55566	-.0049
16	0.49354	0.49858	-.0050
17	0.43789	0.44308	-.0052
18	0.38357	0.38891	-.0053
19	0.33036	0.33585	-.0055
20	0.27807	0.28372	-.0057
21	0.22653	0.23235	-.0058
22	0.17559	0.18158	-.0060
23	0.12511	0.13127	-.0062
24	0.07494	0.08129	-.0064
25	0.02496	0.03150	-.0065

Cuadro 4.3.1. Desviaciones de los coeficientes $\{a_{n-K+1}\}$ y sus correspondientes aproximaciones $\{\hat{a}_{n-K+1}\}$, considerando diferentes tamaños de muestra.

n = 4

i	a_{n-K+1}	$\hat{a}_{n-K+1}$	b
1	.6872	.68745	-.00025
2	.1677	.16522	.0025

Cuadro 4.3.1. ....continuación

n = 10

i	a_{n-K+1}	$\hat{a}_{n-K+1}$	b
1	0.5739	0.5732	.0007
2	0.3291	0.3292	-.0001
3	0.2141	0.2152	-.00011
4	0.1224	0.1229	-.0005
5	0.0399	0.0395	.0004

Cuadro 4.3.1. ....continuación

n = 16

i	a_{n-K+1}	$\hat{a}_{n-K+1}$	b
1	0.5056	0.50485	-.00075
2	0.3290	0.32805	-.00095
3	0.2521	0.25250	-.00050
4	0.1939	0.19451	-.00061
5	0.1447	0.14511	-.00041
6	0.1005	0.10066	-.00016
7	0.0593	0.05912	-.00018
8	0.0196	0.01913	-.00047

Çuadro 4.3.1. ....continuación

n = 25

i	a_{n-K+1}	$\hat{a}_{n-K+1}$	b
1	0.4450	0.44610	-.0011
2	0.3059	0.30670	.0002
3	0.2543	0.25397	.00033
4	0.2148	0.21450	.00030
5	0.1822	0.18196	.00024
6	0.1539	0.15359	.00031
7	0.1283	0.12800	.00030
8	0.1046	0.10430	.00030
9	0.0823	0.08210	.00020
10	0.0610	0.06070	.00030
11	0.0403	0.04010	.00020
12	0.0200	0.019990	.00010

Cuadro 4.3.1. ....continuación

n = 38

i	a_{n-K+1}	$\hat{a}_{n-K+1}$	b
1	0.4015	0.40051	.00099
2	0.2774	0.27798	-.00058
3	0.2391	0.23957	-.00047
4	0.2110	0.21138	-.00038
5	0.1881	0.18852	-.00042
6	0.1586	0.16894	-.00034
7	0.1513	0.15160	-.00030
8	0.1356	0.12587	-.00027
9	0.1211	0.12135	-.00025
10	0.1075	0.10776	-.00026
11	0.0947	0.09450	-.00020
12	0.0824	0.08262	-.00022
13	0.0706	0.07081	-.00021
14	0.0592	0.05938	-.00018
15	0.0481	0.04824	-.00014
16	0.0372	0.03733	-.00013
17	0.0264	0.02659	-.00019
18	0.0158	0.01597	-.00017
19	0.0053	0.00543	-.00013

Cuadro 4.3.1. ....continuación

n = 50

i	a_{n-K+1}	$\hat{a}_{n-K+1}$	b
1	0.3751	0.3764	-.0013
2	0.2574	0.2568	.0006
3	0.2260	0.2254	.0006
4	0.2032	0.2025	.0007
5	0.1847	0.1842	.0005
6	0.1691	0.1686	.0005
7	0.1554	0.1549	.0005
8	0.1430	0.1425	.0005
9	0.1317	0.1212	.0005
10	0.1212	0.1208	.0004
11	0.1113	0.1109	.0004
12	0.1020	0.1010	.0004
13	0.0932	0.0928	.0004
14	0.0846	0.0843	.0003
15	0.0764	0.0761	.0003
16	0.0685	0.0682	.0003
17	0.0608	0.0605	.0003
18	0.0532	0.0530	.0002
19	0.0459	0.0456	.0003
20	0.0386	0.0384	.0002
21	0.0314	0.0312	.0002
22	0.0244	0.0242	.0002
23	0.0174	0.0172	.0002
24	0.0104	0.0102	.0002
25	0.0035	0.0033	.0002

Cuadro 4.4.1. Desviación de la estadística W_c considerando los valores a_{n-K+1} de tabla (tomados de Shapiro-Wilk 1965) y la $\hat{W}_c$ obtenida con las aproximaciones $\hat{a}_{n-K+1}$, de casos reales en muestras de tamaño diferente que se les realizó a la prueba de normalidad.

n	W_c	$\hat{W}_c$	d
6	.9345	.9365	-.0020
6	.8869	.8889	-.0020
9	.9685	.9692	-.0007
12	.9658	.9644	.0014
20	.8617	.8617	.0000
20	.9460	.9440	.0020
24	.9263	.9269	.0006
24	.9543	.9539	.0005
32	.9083	.9095	-.0012
40	.8482	.8499	-.0017
45	.9780	.9779	-.0001
48	.9875	.9859	.0016
50	.7679	.7658	.0021

5. DISCUSIONES

De las Características Agronómicas Consideradas en el Estudio

Del Cuadro 4.1.2. se desprende que de las 395 pruebas de normalidad realizadas, se obtuvieron 119, que representan el 30.13 por ciento - en las que se rechazó la hipótesis de que los datos pertenecían a una población normal. Lo anterior nos indica que la no-normalidad es un evento que tiene probabilidad alta de ocurrir en características que se evalúan en experimentos agronómicos.

Considerando la clasificación agronómica del Cuadro 5.1.1. se obtiene que la no-normalidad ocurre desuniformemente en las diferentes categorías y de acuerdo al porcentaje observado, podemos decir que el más alto corresponde a la categoría hidrológica y la menor al rendimiento - agronómico. Lo anterior significa que debe ponerse mayor atención a los datos provenientes de las categorías agronómicas reportados en porcentaje, elementos menores, elementos mayores, cationes, propiedades bromatológicas, aniones, propiedades físicas y componentes del rendimiento. Esto concuerda como lo señalado por muchos autores como Eisenhart (1947), Little y Hill (1981), Steel y Torrie (1985) y otros.

En las características evaluadas en porcentaje cuyos resultados - se encuentran en 30 y 70 generalmente se tiene distribución normal, tal y como se reporta en la literatura.

La Aproximación de los Estadísticos de Orden

Considerando de que no es fácil tener a la mano los valores esperados de los estadísticos de orden $\{m_{n-K+1}\}$, resulta recomendable utilizar las aproximaciones que se obtienen en este estudio (Cuadro 4.2.1.), ya que éstos presentan prácticamente los mismos valores ya que su diferencia es en el tercer dígito con respecto al valor de tabla.

De los Valores de Aproximación a Shapiro-Wilk

Los resultados que se obtuvieron en el Cuadro 4.3.1. para los coeficientes $\{a_{n-K+1}\}$ de la tabla de Shapiro-Wilk por medio de la aproximación $\{\hat{a}_{n-K+1}\}$ de este estudio resultan satisfactorios, ya que las pruebas de normalidad hechas con ellos son prácticamente las mismas (Cuadro 4.4.1.).

Cuadro 5.1.1. Frecuencia de casos de no-normalidad con la prueba de Shapiro-Wilk para la clasificación agronómica en el estudio. U.A.A.A.N., Buenavista, Saltillo, Coah., 1987.

Características de Clasificación	normalidad	no-normalidad	% no-normal
Hidrológicas	0	4	100.00
Características Agronómicas en Porcentaje	8	9	52.94
Elementos Menores	16	12	42.86
Elementos Mayores	20	10	33.33
Cationes	80	38	32.20
Propiedades Bromatológicas	38	18	32.14
Aniones	24	9	27.29
Propiedades Físicas	25	8	24.24
Componentes de Rendimiento	43	8	15.69
Rendimiento Agronómico	22	3	12.00
TOTAL	276	119	30.13

6. CONCLUSIONES

1. Hasta donde sea posible, hacer la prueba de normalidad de los datos resultantes de los trabajos de investigación que se realicen en la U.A.A.A.N.

2. Debe ponerse esmerado cuidado a datos que son evaluados en meq, mg, PPM, porciento y gr, particularmente a cationes, aniones, propiedades físicas, bromatológicas, elementos mayores y menores, que por lo general no se distribuyen en forma normal.

3. Si se detecta no-normalidad en los datos, la literatura reporta que puede tenerse heterogeneidad de varianza y una transformación al mismo tiempo que corrige la heterogeneidad, consigue que éstos tengan un comportamiento más próximo a la normal.

4. El programa de cómputo desarrollado para el presente estudio, puede usarse con toda confianza para realizar la prueba de normalidad de Shapiro-Wilk y también para obtener aproximaciones satisfactorias de los valores esperados de los estadísticos de orden.

7. R E S U M E N

Se realizó un estudio para conocer el comportamiento de características agronómicas con respecto a la distribución normal. Para lo anterior, se utilizaron resultados de investigación obtenidos en la División de Ingeniería de la Universidad Autónoma Agraria "Antonio Narro".

Se consideraron 58 características que dieron lugar a 395 pruebas de normalidad utilizando la prueba de Shapiro-Wilk. Para la ejecución de la prueba se desarrolló un programa de computación que contempla la obtención de las aproximaciones a los valores de tabla de los estadísticos de orden de Harter y a los coeficientes de Shapiro-Wilk a partir de conocer el tamaño de muestra.

Se encontró que un 30 porciento de las pruebas realizadas resultaron con distribución no-normal; la mayor proporción se refirió a características que están ligadas a propiedades físicas y químicas del suelo; las características que más relación tienen con el rendimiento de cultivos agrícolas tienden a comportarse, en un alto porcentaje, como variables con distribución normal.

8. LITERATURA CITADA

Abramowitz, Milton, and I.A. Stegun. 1964. Handbook of mathematical functions with formulas, graphs, and mathematical tables. National Bureau of Standards Applied Mathematics series No. 55. Washington, D.C. pp. 927-963.

Anderson, V.L. and R.A. Mclean. 1974, Design of experiments. A Realistic Approach. Marcel Dekker. New York. 450 p.

Bartlett, M.S. 1947. The use of transformations. Biometrics. 3:39-52. United States of America.

Brunk, H.D. 1979. Introducción a la estadística matemática. 3 ed. Trillas. México 597 p.

Cochran, W.G. 1947. Some consequences when the assumptions for the analysis of variance are not satisfied. Biometrics. 3:22-38. United States of America.

________________ y G.M. Cox. 1980. Diseños Experimentales. Trillas. México. 661 p.

Eisenhart, C. 1947. The assumption underlying the analysis of variance. Biometrics. 3:1-21. United State of America.

Graybill, F.A. 1961. An Introduction to linear statistical models. McGraw-Hill. New York. 463 p.

Harter, H.L. 1961. Expected values of normal order statistics. Biometrika. 48:151-185. Great Britain.

Infante, G.,S. 1972. Apuntes del curso de diseños experimentales I. Clave Est. 621. Colegio de Postgraduados de Chapingo, México, 620 p.

Kempthorne, O. 1975. The disign and analysis of experiments. 6a. ed. - Krieger. New York. U.S.A. 631 p.

Lindgren, B.W. 1970. Statistical Theory. 2a. ed. McMillan. New York. - 521 p.

Little, T.M. y F.J. Hill 1981. Métodos estadísticos para la investiga--- ción en la agricultura. Trillas. México. 270 p.

Mendenhall, W.; R.L. Sheaffer y D.D. Wa Cherly. 1986. Estadística matemá tica con aplicaciones. 3 ed. Iberoamérica. México. 751 p.

Ostle, B. 1983. Estadística aplicada. Técnicas de estadística moderna, - cuando y donde aplicarlas. Limusa. México. 629 p.

Scheffé, H. 1959. The analysis of variance, Wiley, New York U.S.A. - 477 p.

Shapiro, S.S. and R.S. Francia. 1972. An approximate analysis of varian- ce test for normality. J. of the Amer. Stat. Assoc. - 67(337):215-216. New York. U.S.A.

________________ and M.B. Wilk. 1965. An analysis of variance test - for normality. Complete Samples. Biometrika. 52:591-611. - Great Britain.

______________________________ M.B. Wilk and H.J. Chen, 1968. A comparative study of various test for normality. J. of the Amer. Stat. Assoc. 63:1343-72. New York. U.S.A.

Snedecor, G.W. y W.G. Cochran. 1980. Métodos estadísticos. C.E.C.S.A. Mé xico. 703 p.

Steel, R.G.D. y J.H. Torrie. 1985. Bioestadística. Principios y procedi- mientos. 2a. Ed. McGraw-Hill Colombia. 622 p.

Weisberg, S. and. C. Bingham. 1975. An aproximate analysis of variance - test for no-normality suitable for machine calculation. - Technometrics. 17(1): 133-134.

A P E N D I C E

Cuadro 1A. Programa para estimar los valores esperados de los estadísticos de orden $\{\hat{m}_{n-K+1}\}$, los coeficientes $\{\hat{a}_{n-K+1}\}$ y la Prueba de Normalidad por el método Shapiro-Wilk. (Lenguaje Basic).

```
1Ø   Clear
2Ø   Input "Num. de tratamientos= "; M
3Ø   Input "Num. de repeticiones= "; N
4Ø   O=M*N
5Ø   Dim A(M,N):  DIM B(M): Dim C(N): Dim D(M,N)
6Ø   Dim F(N): Dim Y(O): Dim E(M): Dim H(O): G(O)
7Ø   Print "Teclee un 4=(C.A.), 5=(B.A) o un 6=(si es un
     muestreo)=";:Input G
8Ø   IF G=4 then 29Ø
9Ø   IF G=6 then 11Ø
1ØØ  IF G=5 then 12Ø
11Ø  N=1
12Ø  For I=1 to M: For j=1 to N: Print "Y(";I;",";J;")=";:
     Input A(I,J)
13Ø  S=S+A(I,J):  NEXT J: NEXT I
14Ø  G=S/O
15Ø  For I=1 to M:S=Ø; K=Ø: Z=Ø: For J=1 to N: S=S+A(I,J):
     F(J)=F(J)+A(I,J)
16Ø  Next J
17Ø  B(I)=S/N:NEXT I
18Ø  For J=1 to N: C(J)/M
19Ø  Next J
2ØØ  IF N=1 then 22Ø
21Ø  Go to 25Ø
22Ø  For I=1 to M: For J=1 to N: D(I,J) = A(I,J)-(C(J):
     A=A+D(I,J)↑2
23Ø  Next J: Next I
24Ø  Go to 42Ø
25Ø  For I=1 to M: For J=1 to N
26Ø  D(I,J)=A(I,J)-(B(J))-(C(J))+G
27Ø  A=A+D(I,J)↑2:  Next J: Next I
28Ø  Go to 42Ø
29Ø  O=Ø
```

```
300   For I=1 to M: Print "R";I;: Input (EI):O=O+E(I)
320   For J=1 to E(I): Print "Y(";I;";";J;")=";:Input A(I,J):
      S=S+A(I,J): Next J
340   B(I)=S/E(I): S=Ø: Next I
360   For I=1 to M: For J=1 to E(I):D(I,J)=A(I,J)-(B(J):
      A=A+D(I,J)↑2
370   Next J: Next I
380   Print: Print: Print: Print "Espere estoy calculando":
      Print: Print: Print
390   For I=1 to M: For J=1 to (EI): Y=Y+1
400   H(Y)=D(I,J): Next J: Next I
410   Go to 72Ø
420   Print: Print: Print:
460   Print "Espere estoy calculando"; Print: Print:
470   For I=1 to M
500   For J=1 to N: Y=Y+1
600   H(Y)=D(I,J)
700   Next J: Nest I
720   Z=Ø
1000  For Y=1 to O-1: For Z=Y+1 to O
1050  IF H(Y) <H(Z) then 12ØØ
1100  T=H(Y): H(Y)=H(Z)
1150  H(Z)=T
1200  Next Z: Next Y
1250  W=Ø: Z=Ø: T=Ø: P=Ø
1300  Go to 15ØØ
1350  W=2.718281828<(Q<2/2)
1360  Z=(1/2.5Ø6628275)/W
1370  T=1+Ø.2316419*Q
1380  P=Z*(Ø.31938153/T-Ø.356563782/T↑2+1.7881477937/T↑3-
      1.821255978/T↑4+1.33Ø274429/T↑5)
1400  Return
1500  R=Int(O/2)
1550  Print "Aproximaciones de los estadísticos de orden"
1570  Print: Print
1600  AØ=Ø.315Ø65+Ø.Ø2517778829*Log_e(O)-1.84386E-Ø3*Log_e(O)↑2
1620  For k=1 to R
```

```
1700  L=(k-AØ)(/(O-2*AØ+1):B= Sqr(-2*Loge(L))
1800  Q=(B-(2.515517+Ø.8Ø2853*B+Ø.Ø16328*B*B)/(1+1.432788*
      B+Ø.Ø189269*B*B+1.3Ø8E-Ø3*B↑3))
1900  Gosub 135Ø
1920  E=L-P:Q=Q-(E/Z):IF Abs(E)>1E-Ø8 then 19ØØ
2000  AØ=(k-(O+1471/1470)*P)/(1-2*P)
2050  IF k=1 then AØ=Ø.327511+Ø.Ø258115Ø38*Loge(O)-1.49173E-
      Ø3*Loge(O)↑2
2100  Print "Y(";O-k+1;")=";Q
2120  SS-SS+Q*Q:Y(k)=-Q:Y(O-k+1)=Q: Next k
2140  k=k-(O-2*R=O)
2160  For I=k to 1 step -1
2170  Print "Y(";I;")=";Y(I)
2180  Next I
2200  Print: Print
2240  Print "Aproximaciones de An-k+1 de Shapiro-Wilk"
2250  Print: Print
2260  SS=2*SS
2270  U=SS↑Ø.5
2280  For Z=1 to O:Y(Z)=Y(Z)/U
2350  Next Z
2410  C1=Ø.413Ø9Ø-1.22Ø316*Y(O)+4.326634*Y(O)↑2-2.847795*
      Y(O)↑3
2420  G(1)=-C1:G(O)=C1
2430  IF O<4 then 25ØØ
2440  For Z=2 to R
2450  E1=(-4.684563E-Ø3+1.834622E-Ø4*O-2.210639E-Ø6*O↑2)
2460  E2=(-Ø.465886+Ø.3Ø6314*Loge(O)-Ø.Ø69Ø49*Loge(O)↑2+
      4.714998E-Ø3*Loge(O)↑3)
2470  E3=2.718281828↑E2:E4=E3*Y(O-Z+1):I(Z)=E1+E4
2480  G(Z)=-I(Z):G(O-Z+1)=I(Z)
2490  Next Z
2500  For Z=1 to O: Print "A(";Z;")=";G(O-Z+1): Next Z:B=O
2550  For Z = 1 to R
2600  B=B+G(O-Z+1)*(H(O-Z+1)-(H(Z)))
2700  Next Z
2750  Print: Print
```

```
2800  Print "b↑2=";B↑2: Print "S↑2=";A:Print "W cal=b↑2/S↑2
2810  Print: Print:
2820  Print "W cal=";B↑2; "1";A;"=";(B)↑2/A
2850  Print: Print
2900  Input "Alfa=", C:Print "WT(";C;",";O:")=";
3000  Input D: IF W≤D then 3150
3100  IF W>D then 3250
3150  Print: Print
3200  Print "RECHAZAMOS-HIPOT DE NORMALIDAD":END
3250  Print: Print
3300  Print "NO RECHAZAMOS-HIPOT DE NORMALIDAD: END:
```

Tabla 2A. Porcentaje de puntos de la prueba W para n=3(1)100 (tomada de Shapiro-Francia, 1972)

	Nivel						
n	.01	.05	.10	.50	.90	.95	.99
3	.7530	.7670	.7890	.9590	.9980	.9990	1.0000
4	.6870	.7480	.7920	.9350	.9870	.9920	.9970
5	.6860	.7620	.8060	.9270	.9790	.9860	.9930
6	.7130	.7880	.8260	.9270	.9740	.9810	.9890
7	.7300	.8030	.8380	.9280	.9720	.9790	.9880
8	.7490	.8180	.8510	.9320	.9720	.9780	.9870
9	.7640	.8290	.8590	.9350	.9720	.9780	.9860
10	.7810	.8420	.8690	.9380	.9720	.9780	.9860
11	.7920	.8500	.8760	.9400	.9730	.9790	.9860
12	.8050	.8590	.8830	.9430	.9730	.9790	.9860
13	.8140	.8660	.8890	.9450	.9740	.9790	.9860
14	.8250	.8740	.8950	.9470	.9750	.9800	.9860
15	.8350	.8810	.9010	.9500	.9750	.9800	.9870
16	.8440	.8870	.9060	.9520	.9760	.9810	.9870
17	.8510	.8920	.9100	.9540	.9770	.9810	.9870
18	.8580	.8970	.9140	.9560	.9780	.8920	.9880
19	.8630	.9010	.9170	.9570	.9780	.9820	.9880
20	.8680	.9050	.9200	.9590	.9790	.9830	.9880
21	.8730	.9080	.9230	.9600	.9800	.9830	.9890
22	.8780	.9110	.9260	.9610	.9800	.9840	.9890
23	.8810	.9140	.9280	.9620	.9810	.9840	.9890
24	.8840	.9160	.9300	.9630	.9810	.9840	.9890
25	.8880	.9180	.9310	.9640	.9810	.9850	.9890
26	.8910	.9200	.9330	.9650	.9820	.9850	.9890
27	.8940	.9230	.9350	.9850	.9820	.9850	.9900
28	.8960	.9240	.9360	.9660	.9820	.9850	.9900
29	.8980	.9260	.9370	.9660	.9820	.9850	.9900
30	.9000	.9270	.9370	.9670	.9830	.9850	.9900
31	.9020	.9290	.9400	.9670	.9830	.9860	.9900
32	.9040	.9300	.9410	.9680	.9830	.9860	.9900
33	.9060	.9310	.9420	.9680	.9830	.9860	.9900
34	.9080	.9330	.9430	.9690	.9830	.9860	.9900
35	.9100	.9340	.9440	.9690	.9840	.9860	.9900
36	.9120	.9350	.9450	.9700	.9840	.9860	.9900
37	.9140	.9360	.9460	.9700	.9840	.9870	.9900
38	.9160	.9380	.9470	.9710	.9840	.9870	.9900
39	.9170	.9390	.9480	.9710	.9840	.9870	.9910
40	.9190	.9400	.9490	.9720	.9850	.9870	.9910
41	.9200	.9410	.9500	.9720	.9850	.9870	.9910
42	.9220	.9420	.9510	.9720	.9850	.9870	.9910
43	.9230	.9430	.9510	.9730	.9850	.9870	.9910
44	.9240	.9440	.9520	.9730	.9850	.9870	.9910
45	.9260	.9450	.9530	.9730	.9850	.9880	.9910
46	.9270	.9450	.9530	.9740	.9850	.9880	.9910
47	.9280	.9460	.9540	.9850	.9880	.9880	.9910
48	.9290	.9470	.9540	.9740	.9850	.9880	.9910
49	.9290	.9470	.9550	.9740	.9850	.9880	.9910
50	.9300	.9470	.9550	.9740	.9850	.9880	.9910

Cuadro 2A..............continuación

n	.01	.05	.10	.50	.90	.95	.99
51	.9350	.9540	.9640	.9810	.9900	.9920	.9940
52	.9371	.9570	.9640	.9815	.9900	.9920	.9940
53	.9380	.9570	.9640	.9820	.9900	.9920	.9940
54	.9386	.9575	.9645	.9825	.9905	.9920	.9940
55	.9400	.9580	.9650	.9830	.9910	.9920	.9940
56	.9420	.9590	.9655	.9830	.9910	.9920	.9940
57	.9440	.9610	.9660	.9830	.9910	.9920	.9940
58	.9449	.9619	.9665	.9830	.9910	.9920	.9940
59	.9450	.9620	.9670	.9830	.9910	.9920	.9940
60	.9459	.9623	.9675	.9835	.9910	.9920	.9940
61	.9470	.9630	.9680	.9840	.9910	.9920	.9940
62	.9470	.9635	.9690	.9840	.9915	.9925	.9940
63	.9470	.0640	.9700	.9840	.9920	.9930	.9940
64	.9470	.9645	.9705	.9845	.9920	.9930	.9945
65	.9480	.9650	.9710	.9850	.9920	.9930	.9950
66	.9490	.9655	.9710	.9850	.9920	.9930	.9950
67	.9500	.9660	.9710	.9850	.9920	.9930	.9950
68	.9505	.9662	.9715	.9855	.9920	.9930	.9950
69	.9510	.9663	.9720	.9860	.9920	.9930	.9950
70	.9520	.9663	.9720	.9860	.9920	.9935	.9950
71	.9530	.9670	.9720	.9860	.9920	.9940	.9950
72	.9543	.9675	.9725	.9860	.9925	.9940	.9950
73	.9560	.9680	.9730	.9860	.9930	.9940	.9950
74	.9560	.9685	.9730	.9860	.9930	.9940	.9950
75	.9560	.9690	.9730	.9860	.9930	.9940	.9950
76	.9565	.9690	.9735	.9865	.9930	.9940	.9950
77	.9570	.9690	.9740	.9870	.9930	.9940	.9950
78	.9572	.9693	.9745	.9870	.9930	.9940	.9955
79	.9573	.9700	.9750	.9870	.9930	.9940	.9960
80	.9575	.9700	.9750	.9870	.9940	.9940	.9960
81	.9580	.9700	.9750	.9870	.9930	.9940	.9960
82	.9588	.9705	.9755	.9875	.9930	.9940	.9960
83	.9597	.9710	.9760	.9880	.9930	.9940	.9960
84	.9607	.9715	.9765	.9880	.9930	.9940	.9960
85	.9610	.9720	.9770	.9880	.9930	.9940	.9960
86	.9611	.9720	.9770	.9880	.9935	.9940	.9960
87	.9612	.9720	.9770	.9880	.9940	.9940	.9960
88	.9612	.9720	.9770	.9880	.9940	.9945	.9960
89	.9612	.9720	.9770	.9880	.9940	.9950	.9960
90	.9613	.9723	.9775	.9885	.9940	.9950	.9960
91	.9620	.9730	.9780	.9890	.9940	.9950	.9960
92	.9626	.9730	.9785	.9940	.9940	.9950	.9960
93	.9630	.9730	.9790	.9890	.9940	.9950	.9960
94	.9638	.9735	.9790	.9890	.9940	.9950	.9960
95	.9646	.9740	.9790	.9890	.9940	.9950	.9960
96	.9654	.9745	.9790	.9890	.9940	.9950	.9960
97	.9657	.9750	.9790	.9890	.9940	.9950	.9960
98	.9656	.9755	.9795	.9890	.9940	.9950	.9960
99	.9662	.9760	.9800	.9890	.9940	.9950	.9960
100	.9674	.9765	.9800	.9890	.9940	.9950	.9960