UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

UNA MODIFICACION A LA ESTADISTICA DE SHAPIRO - WILK

T E S I S

PARA OBTENER EL TITULO DE LICENCIADO EN ACTUARIA

P R E S E N T A

PEDRO MIGUEL QUIBRERA MATIENZO

MEXICO, D. F. ABRIL DE 1981

INDICE

INTRODUCCION.

El presente trabajo, es un breve estudio de las propiedades de potencia que presenta una modificación a la estadística de Shapiro y Wilk, usada en la prueba para decidir si una población tiene una distribución que se ajusta a la distribución normal. La investigación se realizó simulando tanto la distribución de la nueva estadística, como su potencia, puesto que a la fecha se desconoce su distribución teórica.

Las motivaciones que llevaron al estudio de esta modificación, son las siguientes ventajas que presenta sobre la estadística original: 1) Es posible calcularla para cualquier tamaño de muestra, pues no se requieren tablas en el proceso de cálculo. Así mismo, este último se realiza con gran rapidez, debido a la existencia de un método recursivo que facilita las operaciones.
2) Se conoce su distribución asintótica, lo que permite determinar aproximadamente su comportamiento, cuando el tamaño de muestra es suficientemente grande.

El contenido está distribuido como a continuación se indica: El primer capítulo trata de la estructura y propiedades analíticas de la estadística de Shapiro-Wilk. Toca enseguida, al turno a la modificación propuesta, discutiéndose con lujo de detalle su derivación teórica y sus propiedades asintóticas. El capítulo tercero describe los procedimientos usados en las simulaciones de su distribución y potencia. Finalmente, a través de un análisis comparativo de la efectividad de ambas estadísticas, se concluye respecto a la utilidad de la nueva estadística.

Nota aclaratoria: Este trabajo padece de apendicitis, debido al terco deseo del autor.

I.- LA PRUEBA DE SHAPIRO-WILK.

S.S Shapiro y M.B. Wilk, en su artículo: "An analysis of variance test for normality" (1965) derivan un nuevo procedimiento estadístico, para determinar si una muestra aleatoria proviene o no de una distribución normal.

Esencialmente el procedimiento consiste en comparar a través de un cociente, a dos estimadores de la varianza. El numerador es el cuadrado del estimador de Lloyd, muy sensibles a fallas en la normalidad; y el denominador es el estimador usual S^2, robusto ante no normalidad. Al efectuarse el cociente se logra que la nueva estadística sea no paramétrica, es decir, su distribución no depende de los parámetros de localización y escala, haciéndola apropiada para probar la hipótesis de normalidad.

Con el propósito de que el lector se forme una idea clara respecto a ésta estadística, se describirán brevemente la derivación y principales propiedades estadísticas de su numerador.

a) ESTIMADORES DE LLOYD.

En 1952, E.H. Lloyd, usando las estadísticas de orden, desarrolló un procedimiento para estimar los parámetros de localización y escala de una familia continua de distribuciones. El propósito de esta sección, es exponer escuetamente tal procedimiento.

Denotamos con las letras μ y σ a los parámetros de localización y escala respectivamente, nótese que no necesariamente son la media y varianza de la familia.

Sea X una variable aleatoria cuya distribución pertenece a la familia. Entonces si consideramos la estandarización de dicha variable:

$$Y = \frac{X-\mu}{\sigma}$$

la función de densidad de esta nueva variable, no depende de los parámetros. Por tal, motivo, tomando una muestra aleatoria de tamaño n, $x_1,\ldots,x_n$ de una distribución perteneciente a la familia, si estandarizamos cada variable y consideramos las n estadísticas de orden correspondientes $Y_{(1)}^{(n)}, Y_{(2)}^{(n)},\ldots,Y_{(n)}^{(n)}$, el vector de medias y la matriz de varianzas y covarianzas, son independientes de μ y σ, es decir:

$$\text{si} \quad \forall_{i\in\{1,\ldots,n\}} \quad E[\,Y_{(i)}^{(n)}\,] = m_i^{(n)}$$

$$\text{y} \quad \forall_{i,j\in\{1,\ldots,n\}} \quad \operatorname{cov}(\,Y_{(i)}^{(n)}, Y_{(j)}^{(n)}\,) = v_{ij}^{(n)}$$

tales cantidades dependen sólo de la función de densidad de Y.

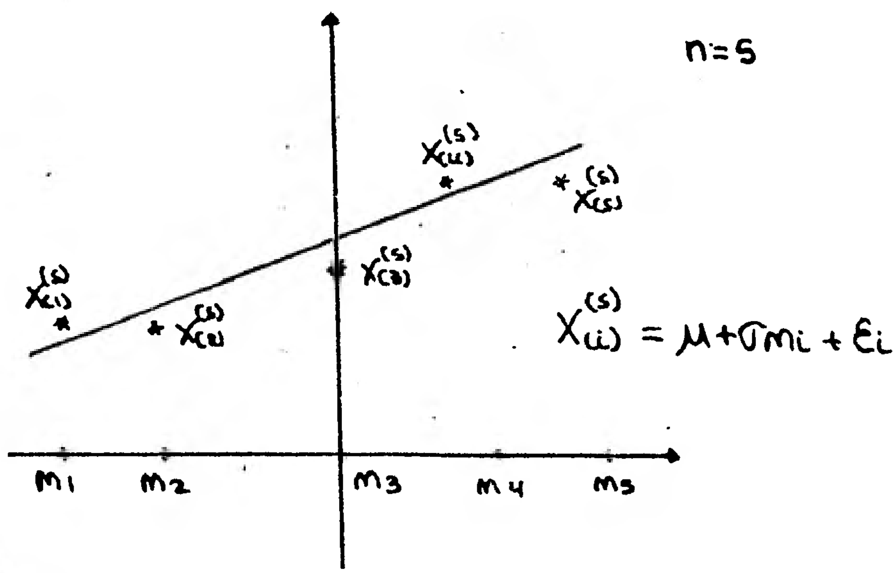
Ahora bien, como:

$$\forall_{i\in\{1,\dots,n\}} \quad Y_{(i)}^{(n)} = \frac{X_{(i)}^{(n)}-\mu}{\sigma}$$

se puede pensar que:

$$\forall_{i\in\{1,\dots,n\}} \quad E\left[X_{(i)}^{(n)} \right] = \mu+\sigma m_i^{(n)}$$

es una recta de regresión:



Podemos estimar a μ y σ, aplicando la teoría de mínimos cuadrados generalizados, obteniendo así, los estimadores insesgados de mínima varianza dentro de la clase de estimadores insesgados lineales en $X_{(1)}^{(n)},\dots,X_{(n)}^{(n)}$. Para visualizar la situación mas claramente, consideremos la notación matricial que comunmente se maneja en regresión.

Tomando $\underset{\sim}{X}^{(n)} = \begin{pmatrix} X_{(1)}^{(n)} \\ \vdots \\ X_{(n)}^{(n)} \end{pmatrix}$

$$\underset{\sim}{Y}^{(n)} = \begin{pmatrix} Y_{(1)}^{(n)} \\ \vdots \\ Y_{(n)}^{(n)} \end{pmatrix} , \quad \underset{\sim}{m}^{(n)} = \begin{pmatrix} m_{(1)}^{(n)} \\ \vdots \\ m_{(n)}^{(n)} \end{pmatrix}$$

y $\quad V^{(n)} = \left(v_{ij}^{(n)}\right)_{i,j \in \{1,\dots,n\}}$

las relaciones anteriores se expresan de la manera siguiente:

$$E\left[\underset{\sim}{Y}^{(n)}\right] = \underset{\sim}{m}^{(n)}$$

$$Var\left[\underset{\sim}{Y}^{(n)}\right] = V^{(n)}$$

y $\quad E\left[\underset{\sim}{X}^{(n)}\right] = \mu \underset{\sim}{1}^{(n)} + \sigma \underset{\sim}{m}^{(n)}$

$$= \left(\underset{\sim}{1}^{(n)}, \underset{\sim}{m}^{(n)}\right)\begin{pmatrix} \mu \\ \sigma \end{pmatrix}$$

y $\quad Var\left(\underset{\sim}{X}^{(n)}\right) = \sigma^2 V^{(n)}$

La estimación por mínimos cuadrados generalizados del vector $(\mu, \sigma)'$, se obtiene minimizando la forma cuadrática:

$$\left[\underset{\sim}{X}^{(n)} - \left(\underset{\sim}{1}^{(n)}, \underset{\sim}{m}^{(n)}\right)\begin{pmatrix} \mu \\ \sigma \end{pmatrix}\right]' V^{(n)^{-1}} \left[\underset{\sim}{X}^{(n)} - \left(\underset{\sim}{1}^{(n)}, \underset{\sim}{m}^{(n)}\right)\begin{pmatrix} \mu \\ \sigma \end{pmatrix}\right]$$

Al minimizar obtenemos:

$$\begin{pmatrix} \hat{\mu} \\ \hat{\sigma} \end{pmatrix} = \left[\left(\underset{\sim}{1}^{(n)}, \underset{\sim}{m}^{(n)}\right)' V^{(n)^{-1}} \left(\underset{\sim}{1}^{(n)}, \underset{\sim}{m}^{(n)}\right)\right]^{-1} \left(\underset{\sim}{1}^{(n)}, \underset{\sim}{m}^{(n)}\right)' V^{(n)^{-1}} \underset{\sim}{X}^{(n)}$$

cuya matriz de varianzas y covarianzas es:

$$\left[(\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)})' V^{(n)-1} (\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)})\right]^{-1} (\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)})' V^{(n)-1} \sigma^2 V^{(n)} \times$$

$$V^{(n)} (\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)}) \left[(\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)})' V^{(n)-1} (\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)})\right]^{-1}$$

$$= \sigma^2 \left((\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)})' V^{(n)-1} (\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)})\right)^{-1}$$

Ahora:

$$\left[(\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)})' V^{(n)-1} (\mathbf{1}^{(n)} \,\vdots\, \mathbf{m}^{(n)})\right]^{-1} = \begin{pmatrix} \mathbf{1}^{(n)'} V^{(n)-1} \mathbf{1}^{(n)} & \mathbf{1}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} \\ \mathbf{1}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} & \mathbf{m}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} \end{pmatrix}^{-1}$$

$$= \frac{1}{\Delta^{(n)}} \begin{pmatrix} \mathbf{m}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} & -\mathbf{1}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} \\ -\mathbf{1}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} & \mathbf{1}^{(n)'} V^{(n)-1} \mathbf{1}^{(n)} \end{pmatrix}$$

donde

$$\Delta^{(n)} = (\mathbf{1}^{(n)'} V^{(n)-1} \mathbf{1}^{(n)})(\mathbf{m}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)}) - (\mathbf{1}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)})^2$$

Usando esta reducción, las expresiones de $\hat{\mu}$ y $\hat{\sigma}$ se reducen a:

$$\begin{pmatrix} \hat{\mu} \\ \hat{\sigma} \end{pmatrix} = \frac{1}{\Delta^{(n)}} \begin{pmatrix} \mathbf{m}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} & -\mathbf{1}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} \\ -\mathbf{1}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} & \mathbf{1}^{(n)'} V^{(n)-1} \mathbf{1}^{(n)} \end{pmatrix} \begin{pmatrix} \mathbf{1}^{(n)'} V^{(n)-1} \\ \mathbf{m}^{(n)'} V^{(n)-1} \end{pmatrix} \mathbf{X}^{(n)}$$

$$= \frac{1}{\Delta^{(n)}} \begin{pmatrix} \mathbf{m}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} \mathbf{1}^{(n)'} V^{(n)-1} - \mathbf{1}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} \mathbf{1}^{(n)'} V^{(n)-1} \\ -\mathbf{1}^{(n)'} V^{(n)-1} \mathbf{m}^{(n)} \mathbf{1}^{(n)'} V^{(n)-1} + \mathbf{1}^{(n)'} V^{(n)-1} \mathbf{1}^{(n)} \mathbf{m}^{(n)'} V^{(n)-1} \end{pmatrix} \mathbf{X}^{(n)}$$

ie

$$\hat{\mu} = \frac{1}{\Delta^{(n)}} \left(-\mathbf{m}^{(n)'} V^{(n)-1} (\mathbf{1}^{(n)} \mathbf{m}^{(n)'} - \mathbf{m}^{(n)'} \mathbf{1}^{(n)'}) V^{(n)-1} \mathbf{X}^{(n)} \right)$$

$$\text{y} \quad \hat{\sigma} = \frac{1}{\Delta^{(n)}} \left(\mathbf{1}^{(n)'} V^{(n)-1} (\mathbf{1}^{(n)} \mathbf{m}^{(n)'} - \mathbf{m}^{(n)'} \mathbf{1}^{(n)'}) V^{(n)-1} \mathbf{X}^{(n)} \right)$$

De igual forma:

$$Var(\hat{\mu}) = \frac{\sigma^2}{\Delta} \underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}$$

$$Var(\hat{\sigma}) = \frac{\sigma^2}{\Delta} \underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{1}^{(n)}$$

y

$$cov(\hat{\mu}, \hat{\sigma}) = -\frac{\sigma^2}{\Delta} \underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}$$

Teniendo en cuenta que para nuestros propósitos las variables $X_1, \ldots, X_n$ se distribuyen normalmente bajo la hipótesis nula, podemos aplicar al cálculo de los estimadores de Lloyd, ciertos resultados algebraicos concernientes al vector de medias y a la matriz de varianzas y covarianzas de las estadísticas de orden estandarizadas, con el propósito de obtener una expresión más simple de la estadística de Shapiro - Wilk.

Tales resultados son:

(a) $\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)} = \underset{\sim}{0}^{(1)}$

(b) $\underset{\sim}{V}_{(n)}^{-1} \underset{\sim}{1}^{(n)} = \underset{\sim}{1}^{(n)}$

(la demostración se encuentra en el apéndice A).

Usando el resultado (a), la expresión denotada por Δ, se simplifica de:

$$\Delta^{(n)} = (\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{1}^{(n)})(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}) - (\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)})^2$$

a

$$\Delta^{*(n)} = (\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{1}^{(n)})(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)})$$

por lo que $\hat{\mu}$ se reduce de:

$$\hat{\mu} = -\frac{1}{\Delta^{(n)}}\, \underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \left(\underset{\sim}{1}^{(n)} \underset{\sim}{m}^{(n)'} - \underset{\sim}{m}^{(n)} \underset{\sim}{1}^{(n)'} \right) V_{(n)}^{-1} \underset{\sim}{X}^{(n)}$$

a

$$\hat{\mu} = \frac{\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)} \underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{X}^{(n)}}{\left(\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{1}^{(n)}\right)\left(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}\right)} = \frac{\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{X}^{(n)}}{\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{1}^{(n)}}$$

usando ahora el resultado (6)

$$V^{(n)} \underset{\sim}{1}^{(n)} = \underset{\sim}{1}^{(n)} \quad \text{ie} \quad \underset{\sim}{1}^{(n)'} V_{(n)}^{-1} = \underset{\sim}{1}^{(n)'}$$

$$\hat{\mu} = \frac{\underset{\sim}{1}^{(n)'} \underset{\sim}{X}^{(n)}}{\underset{\sim}{1}^{(n)'} \underset{\sim}{1}^{(n)}} = \overline{X}_{(n)} = \overline{X}$$ el promedio de la muestra sin ordenar.

La reducción de $\hat{\sigma}$ es de:

$$\hat{\sigma} = \frac{1}{\Delta^{(n)}}\, \underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \left(\underset{\sim}{1}^{(n)} \underset{\sim}{m}^{(n)'} - \underset{\sim}{m}^{(n)} \underset{\sim}{1}^{(n)'} \right) V_{(n)}^{-1} \underset{\sim}{X}^{(n)}$$ a

$$\hat{\sigma} = \frac{\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{1}^{(n)} \underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{X}^{(n)}}{\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{1}^{(n)} \underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}} = \frac{\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{X}^{(n)}}{\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}}$$

y en lo concerniente a las varianzas de los estimadores y a su covarianza, se obtienen:

$$\operatorname{Var}(\hat{\mu}) = \frac{\sigma^2}{\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{1}^{(n)}}$$

$$\operatorname{Var}(\hat{\sigma}) = \frac{\sigma^2}{\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}}$$ y $\operatorname{cov}(\hat{\mu}, \hat{\sigma}) = 0$

Como comentario final a esta sección, cabe notar la fuerte dependencia del estimador de Lloyd $\hat{\sigma}$ de la desviación estandar, de la hipótesis de normalidad, puesto que sus coeficientes se encuentran en función de $m^{(n)}$ y $V^{(n)}$, valores que a su vez dependen totalmente del tamaño de muestra y de la familia de donde se muestrea.

Este mismo hecho, ocaciona el uso inevitable de tablas para el cálculo de la estadística de Shapiro-Wilk, más lo peor del caso, es que tan sólo se conocen $m^{(n)}$ para $n \leq 50$ y $V^{(n)}$ hasta $n = 20$, teniéndo que recurrirse a la aproximación para tamaños de muestra mayores. Tal situación junto con el total desconocimiento de propiedades asintóticas, se traducen en grandes limitaciones de la estadística de Shapiro y Wilk.

b) LA ESTADISTICA DE PRUEBA.

Como se mencionó al inicio de este capítulo, el procedimiento estadístico de Shapiro y Wilk para la detección de normalidad en una muestra aleatoria, se basa en la comparación de dos estimadores del parámetro de escala de la distribución a través de un cociente: el cuadrado del estimador de Lloyd y el conocido estimador S^2.

Con el propósito de que la estadística de prueba tenga ciertas propiedades analíticas, los autores modificaron ligeramente al estimador de Lloyd, multiplicándolo por determinadas constantes. Tal modificación es:

$$\hat{\sigma}^* = \frac{\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}}{(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} V_{(n)}^{-1} \underset{\sim}{m}^{(n)})^{1/2}} \hat{\sigma}$$

Substituyendo el valor de $\hat{\sigma}$:

$$\hat{\sigma}^* = \frac{\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}}{(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} V_{(n)}^{-1} \underset{\sim}{m}^{(n)})^{1/2}} \cdot \frac{\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{x}^{(n)}}{\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}}$$

$$= \frac{\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{x}^{(n)}}{(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} V_{(n)}^{-1} \underset{\sim}{m}^{(n)})^{1/2}}$$

Tomando $\quad \underset{\sim}{a}^{(n)} = \dfrac{V_{(n)}^{-1} \underset{\sim}{m}^{(n)}}{(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} V_{(n)}^{-1} \underset{\sim}{m}^{(n)})^{1/2}}$

$$\hat{\sigma}^* = \frac{\underset{\sim}{a}^{(n)'} \underset{\sim}{x}^{(n)}}{1} = \underset{\sim}{a}^{(n)'} \underset{\sim}{x}^{(n)}$$

La expresión de la estadística de prueba, que de ahora en adelante denotaremos por $W^{(n)}$, es por tanto:

$$W^{(n)} = \frac{(\hat{\sigma}^*)^2}{(n-1)S^2}$$

$$= \frac{(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} \underset{\sim}{x}^{(n)})^2}{(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} V_{(n)}^{-1} \underset{\sim}{m}^{(n)})(n-1)S^2}$$

$$\text{i.e. } W^{(n)} = \frac{(\underset{\sim}{a}^{(n)\prime} \underset{\sim}{x}^{(n)})^2}{(n-1)S^2} = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n} (x_i-\bar{x})^2}$$

El estimador S^2 siempre es un estimador insesgado de la varianza poblacional, sea cual sea la familia de distribuciones de donde provenga la muestra, es decir S^2 es un estimador robusto ante no normalidad. Este hecho, aunado a la gran sensibilidad del estimador de la varianza de Lloyd a fallas de normalidad, provoca un comportamiento peculiar de $W^{(n)}$: cuando la hipótesis nula es cierta, es decir, cuando la muestra proviene efectivamente de una distribución normal, ambos estimadores estiman (salvo por una constante), a la varianza poblacional; mientras que en ausencia de normalidad, aún cuando el denominador sigue estimando a un múltiplo de la varianza, la enorme sensibilidad del numerador, ocaciona grandes distorciones en la estimación de dicho parámetro. Tal situación ha sido estudiada con lujo de detalle por los autores Shapiro y Wilk, quienes a través de la simulación han encontrado que para ciertas familias de distribución distintas de la normal, los valores medios de $W^{(n)}$, tienden a cargarse a la izquierda de los correspondientes a la familia normal. Así mismo, la varianza bajo normalidad; la varianza simulada de $W^{(n)}$ es en general más pequeña que la respectiva a las poblaciones no normales estudiadas.

Nótese que la modificación al estimador de Lloyd $\hat{\sigma}$, ya no es insesgado:

$$E(\hat{\sigma}^*) = E(\underset{\sim}{a}^{(n)\prime} \underset{\sim}{X}^{(n)})$$

$$= \frac{\underset{\sim}{m}^{(n)\prime} V^{(n)-1}}{(\underset{\sim}{m}^{(n)\prime} V^{(n)-1} V^{(n)-1} \underset{\sim}{m}^{(n)})^{1/2}} E(\mu \underset{\sim}{1}^{(n)} + \sigma \underset{\sim}{Y}^{(n)})$$

$$= \frac{\underset{\sim}{m}^{(n)\prime} V^{(n)-1}}{(\underset{\sim}{m}^{(n)\prime} V^{(n)-1} V^{(n)-1} \underset{\sim}{m}^{(n)})^{1/2}} (\mu \underset{\sim}{1}^{(n)} + \sigma \underset{\sim}{m}^{(n)})$$

$$= \sigma \frac{\underset{\sim}{m}^{(n)\prime} V^{(n)-1} \underset{\sim}{m}^{(n)}}{(\underset{\sim}{m}^{(n)\prime} V^{(n)-1} V^{(n)-1} \underset{\sim}{m}^{(n)})^{1/2}}$$

Las principales propiedades analíticas de $W^{(n)}$ son:

1) No depende de cambios en los parámetros de localización y escala.

2) En presencia de normalidad, la distribución de $W^{(n)}$, sólo depende del tamaño de muestra, no de los parámetros.

3) En presencia de normalidad, $W^{(n)}$ es independiente de S^2 y $\bar{X}$.

4)
$$\forall_{r \in \mathbb{N}} \quad E(W^{(n)r}) = \frac{E(\hat{\sigma}^{*2r})}{(n-1)\, E(S^{2r})}$$

5) $$E\left[W^{(n)}\right] = \frac{\underset{\sim}{m}^{(n)\prime} V^{(n)\prime}_{-1} \underset{\sim}{m}^{(n)} \left[\underset{\sim}{m}^{(n)\prime} V^{(n)\prime}_{-1} \underset{\sim}{m}^{(n)} + 1 \right]}{(n-1)\, \underset{\sim}{m}^{(n)\prime} V^{(n)\prime}_{-1} V^{(n)\prime}_{-1} \underset{\sim}{m}^{(n)}}$$

6) Los valores máximo y mínimo de $W^{(n)}$ son respectivamente: 1 y $\frac{n a_1^{(n)2}}{n-1}$ donde $a_1^{(n)}$ es la primera entrada del vector $\underset{\sim}{a}^{(n)}$.

En lo que concierne a la distribución de $W^{(n)}$, todavía no se ha podido determinar explícitamente para cualesquiera que sea el tamaño de la muestra, debido a la enorme complejidad que involucra el uso de los momentos de primer y segundo orden de las estadísticas de orden correspondientes a una muestra normal estandarizada. Tales circunstancias han conducido al uso de la simulación como único camino para un conocimiento aproximado de los principales porcentiles de $W^{(n)}$.

Así pues, muestreando de una distribución normal, los autores calcularon repetidos valores de $W^{(n)}$ y, ajustando los porcentiles obtenidos por el sistema de curvas de frecuencia de N.L. Johnson, se obtuvo entonces una aproximación empírica a la distribución nula de $W^{(n)}$.

Tal procedimiento lo efectuaron desde $n=3$ hasta $n=20$, con un tamaño de simulación de 5000 y desde $n=21$ hasta $n=50$ usaron: tamaño de la simulación igual a $100\,000/n$.

Recordará el lector que se conoce la matriz de varianzas y covarianzas solo para tamaños de muestra menores o iguales a veinte, por tanto, para valores mayores, los autores han aproximado el valor del vector:

$$\underset{\sim}{a}^{(n)} = \frac{m^{(n)'} V_n^{-1}}{\left(\underset{\sim}{m}^{(n)'} V_{(n)}^{-1} V_{(n)}^{-1} \underset{\sim}{m}^{(n)}\right)^{1/2}} \quad \text{por:}$$

$$\hat{a}_i^{(n)} = \frac{2 m_i^{(n)}}{\hat{C}^{2(n)}} \quad \text{para } i = 2, \dots, n-1$$

donde:

$$\hat{C}^{2(n)} = -2.722 + 4.083n$$

y para $i = 1, n$

$$\hat{a}_1^{2(n)} = \hat{a}_n^{2(n)} = \frac{\Gamma\left(\tfrac{1}{2}(n+1)\right)}{\left(2\Gamma\left(\tfrac{1}{2}n+1\right)\right)^{1/2}}$$

Una vez obtenidos los percentiles, la potencia de la prueba se simuló tomando 200 muestras aleatorias de diversas distribuciones, para distintos tamaños de muestra. En el artículo, se compara la efectividad de la prueba con otros procedimientos estadísticos como la prueba χ^2, la prueba de Kolmogorov-Smirnov y algunas mas. En el capítulo tercero se describirán con mas detalle los resultados del artículo, agregando los respectivos de la modificación que se propone en este trabajo.

II.- MODIFICACION PROPUESTA.

La modificación a la estadística de Shapiro - Wilk, consiste en substituir el estimador de Lloyd $\hat{\sigma}^{(n)}$, por otro cuyo comportamiento estadístico es muy parecido y que además presenta ciertas ventajas de índole asintótico y computacional. Las principales propiedades de este nuevo estimador, al cual denotaremos por $\hat{\sigma}_U^{(n)}$ son:

1) Es una combinación lineal de las estadísticas de orden (propiedad compartida por $\hat{\sigma}^{(n)}$).

2) Es un estimador insesgado del parámetro de escala σ. (propiedad compartida por $\hat{\sigma}^{(n)}$).

3) Debido a que su derivación descansa básicamente en la teoría de estadísticas U de Hoeffding, posee una distribución asintótica.

4) Existe un procedimiento recursivo por medio del cual es posible calcularlo para cualquier tamaño de muestra, sin depender del uso de tablas.

En las dos primeras secciones de éste capítulo se discutirán la derivación de $\hat{\sigma}_U^{(n)}$ y sus características estadísticas. Seguidamente serán expuestas las simulaciones respectivas, de la distribución nula y de la potencia.

a) DERIVACIÓN DEL ESTIMADOR $\hat{\sigma}_U^{(n)}$:

Comenzamos recordando el concepto clásico de: "estadística-U" de Hoeffding:

Sea D un conjunto de funciones de distribución, definidas en $\mathbb{R}^k$.

Sea $\Theta: D \to \mathbb{R}$, $F \to \Theta(F)$ una funcional definida en D.

Decimos que Θ es una funcional regular en D, si y solo si:

$$\exists\, n \in \mathbb{N} \quad \exists\, \Phi: \underbrace{\mathbb{R}^k \times \cdots \times \mathbb{R}^k}_{n \text{ veces}} \to \mathbb{R} \ \ni\ \forall F \in D \ \forall \underbrace{\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n}_{\text{muestra aleatoria de } F} \quad E[\Phi(\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n)] = \Theta(F)$$

A la función Φ, se la denomina el kernel de Θ, si además se cumple que n es el grado de Θ en D, es decir si n es el mínimo natural para el cual se satisface la proposición.

Por otro lado, si $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ son n vectores aleatorios de dimensión k y

$$\Phi: \underbrace{\mathbb{R}^k \times \cdots \times \mathbb{R}^k}_{m \text{ veces, con } m \leq n} \to \mathbb{R} \quad \text{una función,}$$

podemos definir una estadística de la siguiente forma:

$$U: \mathbb{R}^k \times \cdots \times \mathbb{R}^k \to \mathbb{R}$$

$$\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n \to \frac{1}{\binom{n}{m} m!} \sum{}'' \Phi(\underset{\sim}{X}_{i_1}, \ldots, \underset{\sim}{X}_{i_m})$$

donde $\sum''$, es la suma sobre todas las permutaciones posibles de m índices.

La estadística U, es simétrica en $\underset{\sim}{X_1}, \dots, \underset{\sim}{X_n}$ y se conoce como una estadística-U de Hoeffding. Cuando Φ es el Kernel de una funcional regular definida en un conjunto D de funciones de distribución k-dimensionales, claramente U es un estimador insesgado de Θ en D, ya que de la definición se tiene:

$$\forall_{F\in D} \quad \Theta(F) = \underbrace{\int_{\mathbb{R}^k}\cdots\int_{\mathbb{R}^k}}_{n \text{ veces}} U(\underset{\sim}{x_1}, \dots, \underset{\sim}{x_n})\, dF(\underset{\sim}{x_1})\cdots dF(\underset{\sim}{x_n})$$

Así pues, cuando $r=1$ y $D=\{N(\mu,\sigma^2),\ \mu\in\mathbb{R}, \sigma\in\mathbb{R}^+\}$, la funcional:

$$\begin{aligned} \Theta : D &\longrightarrow \mathbb{R} \\ N(\mu,\sigma^2) &\longrightarrow \sigma \end{aligned}$$

es regular en D, puesto que al tomar $n=2$ y

$$\begin{aligned} \Phi : \mathbb{R}\times\mathbb{R} &\longrightarrow \mathbb{R} \\ (x_1,x_2) &\longrightarrow k\,|x_1-x_2| \end{aligned}$$

con $k = \frac{\sqrt{\pi}}{2}$, tenemos que si $F\in D$ y X_1, X_2 es una muestra aleatoria de F, entonces:

$$\begin{aligned} E[\Phi(X_1,X_2)] &= E[k|X_1-X_2|] \\ &= \sigma, \end{aligned}$$

ya que si

$$Y_i = \frac{X_i-\mu}{\sigma} \qquad i=1,2$$

$Y_i \sim N(0,1)$ y por tanto:

$$E[k|X_1-X_2|] = E[k|(\mu+\sigma Y_1)-(\mu+\sigma Y_2)|]$$
$$= E[k|\sigma Y_1-\sigma Y_2|]$$
$$= k\sigma E[|Y_1-Y_2|]$$

Tomando $Z=Y_1-Y_2$ y notando que $Z\sim N(0,2)$ tenemos:

$$E[|Y_1-Y_2|] = E[|Z|] = \int_{-\infty}^{\infty} |Z| \frac{1}{\sqrt{2}\sqrt{2\pi}} e^{-\frac{1}{2(2)}Z^2} dZ$$

$$= \frac{2}{\sqrt{\pi}} = k^{-1}$$

$$\therefore E[\phi(X_1,X_2)] = k\sigma k^{-1} = \sigma$$

Como además con una muestra de tamaño uno, no es posible estimar la desviación estandar σ, de una distribución normal, ϕ es el Kernel de θ con grado dos. Por tal motivo, si $X_1,\dots,X_n$ es una muestra aleatoria de una $N(\mu,\sigma^2)$, la estadística-U basada en éste Kernel es:

$$U=U(X_1,\dots,X_n) = \frac{1}{\binom{n}{2}2!} \sum_{i\neq j} k|X_i-X_j|$$

Ésta estadística-U es precisamente el estimador $\hat{\sigma}_U^{(n)}$ propuesto. A continuación veremos que es una combinación lineal de las estadísticas de orden, y el método recursivo que facilita su cálculo.

Recordaré al lector que si $X_1,\dots,X_n$ es una muestra aleatoria de $F_X(\cdot)$, función de distribución, y $\Phi:\mathbb{R}^m \longrightarrow \mathbb{R}$ con $m\leq n$ una función $\ni$. $\Phi(X_1,\dots,X_m)$ sea variable aleatoria, entonces:

$$E\left[\Phi(X_1,\dots,X_m)\,\middle|\, X_{(1)}^{(n)},\dots,X_{(n)}^{(n)}\right] = \frac{1}{\binom{n}{m}m!}\sum{}''\Phi(X_{i_1},\dots,X_{i_m})$$

donde $\sum''$ es la suma sobre todas las permutaciones de m elementos $(i_1,\dots,i_m)$ del conjunto de índices $\{1,\dots,n\}$. Por tanto,

$$U(X_1,\dots,X_n) = \frac{1}{\binom{n}{2}2!}\sum_{i\neq j} k|X_i-X_j|$$
$$= E\left[k|X_1-X_2|\,\middle|\, X_{(1)}^{(n)},\dots,X_{(n)}^{(n)}\right].$$

es decir:

$$\hat{\sigma}_U^{(n)} = E\left[k|X_1-X_2|\,\middle|\, X_{(1)}^{(n)},\dots,X_{(n)}^{(n)}\right]$$
$$\text{con } k=\sqrt{\pi}/2.$$

Si observamos ahora que las n-estadísticas de orden $X_{(1)}^{(n)},\dots,X_{(n)}^{(n)}$, quedan completamente determinadas por el conocimiento de $X_{(1)}^{(n-1)},\dots,X_{(n-1)}^{(n-1)}$ y X_n, es decir que:

$$\sigma(X_{(1)}^{(n)},\dots,X_{(n)}^{(n)}) \subset \sigma(X_{(1)}^{(n-1)},\dots,X_{(n-1)}^{(n-1)},X_n)$$

donde $\sigma\{X_i\}_{i\in I}$ es el mínimo σ-álgebra $\ni$. $\forall_{i\in I}$ X_i es medible, obtenemos:

$$\mathbb{E}\left[k|X_1-X_2| \mid X_{(1)}^{(n)},\dots, X_{(n)}^{(n)} \right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[k|X_1-X_2| \mid X_{(1)}^{(n-1)},\dots, X_{(n-1)}^{(n-1)}, X_n \right] \middle| X_{(1)}^{(n)},\dots, X_{(n)}^{(n)} \right]$$

Ahora dado que $X_{(1)}^{(n-1)},\dots, X_{(n-1)}^{(n-1)}$ es independiente de X_n y $k|X_1-X_2|$, $X_{(1)}^{(n-1)},\dots, X_{(n-1)}^{(n-1)}$ es independiente de X_n, se tiene:

$$\mathbb{E}\left[k|X_1-X_2| \mid X_{(1)}^{(n-1)},\dots, X_{(n-1)}^{(n-1)}, X_n \right] = \mathbb{E}\left[k|X_1-X_2| \mid X_{(1)}^{(n-1)},\dots, X_{(n-1)}^{(n-1)} \right]$$

Esta última sigualdad es de vital importancia para la derivación del método recursivo que permite el cálculo fácil de $\hat{\sigma}_U^{(n)}$, por lo que se incluye en el Apéndice B una demostración completa de la misma.

Usando ambas afirmaciones obtenemos:

$$\hat{\sigma}_U^{(n)} = \mathbb{E}\left[k|X_1-X_2| \mid X_{(1)}^{(n)},\dots, X_{(n)}^{(n)} \right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[k|X_1-X_2| \mid X_{(1)}^{(n-1)},\dots, X_{(n-1)}^{(n-1)}, X_n \right] \middle| X_{(1)}^{(n)},\dots, X_{(n)}^{(n)} \right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[k|X_1-X_2| \mid X_{(1)}^{(n-1)},\dots, X_{(n-1)}^{(n-1)} \right] \middle| X_{(1)}^{(n)},\dots, X_{(n)}^{(n)} \right]$$

$$= \mathbb{E}\left[\hat{\sigma}_U^{(n-1)} \mid X_{(1)}^{(n)},\dots, X_{(n)}^{(n)} \right]$$

es decir: $\hat{\sigma}_U^{(n)} = \mathbb{E}\left[\hat{\sigma}_U^{(n-1)} \mid X_{(1)}^{(n)},\dots, X_{(n)}^{(n)} \right]$.
Esta es, finalmente, la fórmula recursiva.
Así pues, para $n=2$:

$$\hat{\sigma}_U^{(2)} = E\left[k|X_1 - X_2| \;\middle|\; X_{(1)}^{(2)}, X_{(2)}^{(2)} \right]$$

$$= E\left[k\left(X_{(2)}^{(2)} - X_{(1)}^{(2)} \right) \;\middle|\; X_{(1)}^{(2)}, X_{(2)}^{(2)} \right]$$

$$= k\, X_{(2)}^{(2)} - k\, X_{(1)}^{(2)}$$

$$= (-k, k) \begin{pmatrix} X_{(1)}^{(2)} \\ X_{(2)}^{(2)} \end{pmatrix} = \underset{\sim}{b}^{(2)'} \underset{\sim}{X}^{(2)}$$

Para $n = 3$:

$$\hat{\sigma}_U^{(3)} = E\left[k|X_1 - X_2| \;\middle|\; X_{(1)}^{(3)}, X_{(2)}^{(3)}, X_{(3)}^{(3)} \right]$$

$$= E\left[\hat{\sigma}_U^{(2)} \;\middle|\; X_{(1)}^{(3)}, X_{(2)}^{(3)}, X_{(3)}^{(3)} \right]$$

$$= E\left[\underset{\sim}{b}^{(2)'} \underset{\sim}{X}^{(2)} \;\middle|\; X_{(1)}^{(3)}, X_{(2)}^{(3)}, X_{(3)}^{(3)} \right]$$

Dados $X_{(1)}^{(3)}, X_{(2)}^{(3)}; X_{(3)}^{(3)}$

$$X_{(1)}^{(2)} \text{ pudo ser } \begin{cases} X_{(1)}^{(3)} & \text{con prob. } \frac{2}{3} \\ & \text{ó} \\ X_{(2)}^{(3)} & \text{con prob. } \frac{1}{3} \end{cases}$$

De igual forma:

$$X_{(2)}^{(2)} \text{ pudo ser } \begin{cases} X_{(2)}^{(3)} & \text{con prob } \frac{1}{3} \\ & \text{ó} \\ X_{(3)}^{(3)} & \text{con prob } \frac{2}{3} \end{cases}$$

de donde

$$\hat{\sigma}_U^{(3)} = E\left[kX_{(2)}^{(2)} - k\,X_{(1)}^{(2)} \;\middle|\; X_{(1)}^{(3)}, X_{(2)}^{(3)}, X_{(3)}^{(3)} \right]$$

$$= k\left(\tfrac{1}{3}X_{(2)}^{(3)} + \tfrac{2}{3}X_{(3)}^{(3)}\right)$$

$$-k\left(\tfrac{2}{3}X_{(1)}^{(3)} + \tfrac{1}{3}X_{(2)}^{(3)}\right)$$

$$= -\tfrac{2}{3}k\,X_{(1)}^{(3)} + \tfrac{2}{3}k\,X_{(3)}^{(3)}$$

$$= \left(-\tfrac{2}{3}k, 0, \tfrac{2}{3}k\right)\begin{pmatrix} X_{(1)}^{(3)} \\ X_{(2)}^{(3)} \\ X_{(3)}^{(3)} \end{pmatrix}$$

$$= \left[\begin{pmatrix} \frac{2}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} \end{pmatrix}\begin{pmatrix} -k \\ k \end{pmatrix}\right]'\begin{pmatrix} X_{(1)}^{(3)} \\ X_{(2)}^{(3)} \\ X_{(3)}^{(3} \end{pmatrix}$$

$$= \underset{\sim}{b}^{(3)'}\,\underset{\sim}{X}^{(3)}$$

donde $\underset{\sim}{b}^{(3)} = \begin{pmatrix} 2/3 & 0 \\ 1/3 & 1/3 \\ 0 & 2/3 \end{pmatrix}\underset{\sim}{b}^{(2)}$

$$= A_3\,\underset{\sim}{b}^{(2)}$$

es decir: $\hat{\sigma}_U^{(3)} = (A_3\,\underset{\sim}{b}^{(2)})'\,\underset{\sim}{X}^{(3)}$

Para $n=4$

$$\hat{\sigma}_U^{(4)} = E\left[\hat{\sigma}_U^{(3)} \,\middle|\, X_{(1)}^{(4)}, X_{(2)}^{(4)}, X_{(3)}^{(4)}, X_{(4)}^{(4)}\right]$$

$$\hat{\sigma}_v^{(4)} = E\left[\underset{\sim}{b_3}' \underset{\sim}{X}_{(3)} \,\middle|\, X_{(1)}^{(4)}, X_{(2)}^{(4)}, X_{(3)}^{(4)}, X_{(4)}^{(4)} \right]$$

$$= E\left[-\tfrac{2}{3}kX_{(1)}^{(3)} + \tfrac{2}{3}k\, X_{(3)}^{(3)} \,\middle|\, X_{(1)}^{(4)}, X_{(2)}^{(4)}, X_{(3)}^{(4)}, X_{(4)}^{(4)} \right]$$

Ahora

$$X_{(1)}^{(3)} \text{ pudo ser } \begin{cases} X_{(1)}^{(4)} & \text{con } \frac{3}{4} \text{ de prob.} \\ \text{ó} & \\ X_{(2)}^{(4)} & \text{con } \frac{1}{4} \text{ de prob.} \end{cases}$$

$$X_{(2)}^{(3)} \text{ pudo ser } \begin{cases} X_{(2)}^{(4)} & \text{con } \frac{2}{4} \text{ de prob.} \\ \text{ó} & \\ X_{(3)}^{(4)} & \text{con } \frac{2}{4} \text{ de prob.} \end{cases}$$

$$X_{(3)}^{(3)} \text{ pudo ser } \begin{cases} X_{(3)}^{(4)} & \text{con } \frac{1}{4} \text{ de prob.} \\ \text{ó} & \\ X_{(4)}^{(4)} & \text{con } \frac{3}{4} \text{ de prob.} \end{cases}$$

$$\therefore\ E\left[X_{(1)}^{(3)} \mid \mathcal{O}_4 \right] = \frac{3}{4} X_{(1)}^{(4)} + \frac{1}{4} X_{(2)}^{(4)}$$

$$E\left[X_{(3)}^{(3)} \mid \mathcal{O}_4 \right] = \frac{1}{4} X_{(3)}^{(4)} + \frac{3}{4} X_{(4)}^{(4)}$$

$$\therefore\ \hat{\sigma}_v^{(4)} = -\frac{2}{3}k\left(\frac{3}{4} X_{(1)}^{(4)} + \frac{1}{4} X_{(2)}^{(4)} \right) + \frac{2}{3}k\left(\frac{1}{4} X_{(3)}^{(4)} + \frac{3}{4} X_{(4)}^{(4)} \right)$$

$$= \left((-k)\tfrac{2}{3}\cdot\tfrac{3}{4},\ (-k)\tfrac{2}{3}\cdot\tfrac{1}{4},\ (k)\tfrac{2}{3}\cdot\tfrac{1}{4},\ (k)\tfrac{2}{3}\cdot\tfrac{3}{4} \right) \underset{\sim}{X}_{(4)}$$

$$= \underset{\sim}{b_4}' \underset{\sim}{X}_{(4)}$$

Tomando $A_4 = \begin{pmatrix} 3/4 & 0 & 0 \\ 1/4 & 2/4 & 0 \\ 0 & 2/4 & 1/4 \\ 0 & 0 & 3/4 \end{pmatrix}$

$$A_4 A_3 \underset{\sim}{b}_{(2)} = \begin{pmatrix} 3/4 & 0 & 0 \\ 1/4 & 2/4 & 0 \\ 0 & 2/4 & 1/4 \\ 0 & 0 & 3/4 \end{pmatrix} \begin{pmatrix} 2/3 & 0 \\ 1/3 & 1/3 \\ 0 & 2/3 \end{pmatrix} \begin{pmatrix} -k \\ k \end{pmatrix}$$

$$= \begin{pmatrix} \frac{2}{3}\cdot\frac{3}{4} & 0 \\ \frac{1}{4}\cdot\frac{2}{3}+\frac{2}{4}\cdot\frac{1}{3} & \frac{2}{4}\cdot\frac{1}{3} \\ \frac{2}{4}\cdot\frac{1}{3} & \frac{2}{4}\cdot\frac{1}{3}+\frac{1}{4}\cdot\frac{2}{3} \\ 0 & \frac{2}{3}\cdot\frac{2}{4} \end{pmatrix} \begin{pmatrix} -k \\ k \end{pmatrix}$$

$$= \underset{\sim}{b}_4$$

ie $\underset{\sim}{b}_4 = A_4 \underset{\sim}{b}_3 = A_4 A_3 \underset{\sim}{b}_2$

$$\therefore \hat{T}_U^{(4)} = (A_4 A_3 \underset{\sim}{b}_2)' \underset{\sim}{X}_{(4)}$$

Para el caso general:

Tomando $X_{(1)}^{(n-1)}, \ldots, X_{(n-1)}^{(n-1)}$; la nueva observación X_n cae en intervalos de la forma

$$X_{(j-1)}^{(n-1)} < X_n \leq X_{(j)}^{(n-1)} \qquad j=1,\ldots,n$$

entonces $X_{(1)}^{(n)}, \ldots, X_{(n)}^{(n)}$ se transforma en:

$$X_{(i)}^{(n)} = \begin{cases} X_{(i)}^{(n-1)} & \text{si } i<j \\ X_n & \text{si } i=j \\ X_{(i-1)}^{(n-1)} & \text{si } i>j \end{cases}$$

Por tanto, dados $X_{(1)}^{(n)}, \dots, X_{(n)}^{(n)}$

$X_{(i)}^{(n-1)}$ pudo ser $X_{(i)}^{(n)}$ si $X_n = X_{(j)}^{(n)}$ con $j > i$

Hay $n-i$ de estos casos

ó

$X_{(i)}^{(n-1)}$ pudo ser $X_{(i+1)}^{(n)}$ si $X_n = X_{(j)}^{(n)}$ con $j \leq i$

Hay i de estos casos

$\therefore$

$$X_{(i)}^{(n-1)} = \begin{cases} X_{(i)}^{(n)} & \text{con prob } \frac{n-i}{n} \\ X_{(i+1)}^{(n)} & \text{con prob } \frac{i}{n} \end{cases}$$

Así pues

$$E\left[X_{(i)}^{(n-1)} \mid X_{(1)}^{(n)}, \dots, X_{(n)}^{(n)} \right]$$

$$= \frac{n-i}{n} X_{(i)}^{(n)} + \frac{i}{n} X_{(i+1)}^{(n)}$$

$$\therefore E\left[\underline{b}'_{n-1} \underset{\sim}{X}_{(n-1)} \mid Q_n \right] = E\left[\sum_{i=1}^{n-1} b_i X_{(i)}^{(n-1)} \mid X_{(1)}^{(n)}, \dots, X_{(n)}^{(n)} \right]$$

$$= \sum_{i=1}^{n-1} b_i \left(\frac{n-i}{n} X_{(i)}^{(n)} + \frac{i}{n} X_{(i+1)}^{(n)} \right)$$

$$= \sum_{i=1}^{n-1} \left(b_i \frac{n-i}{n} X_{(i)}^{(n)} + b_i \frac{i}{n} X_{(i+1)}^{(n)} \right)$$

$$= b_1 \frac{n-1}{n} X_{(1)}^{(n)} + \sum_{i=2}^{n-1} \left(b_i \frac{n-i}{n} + b_{i-1} \frac{i-1}{n} \right) X_{(i)}^{(n)} +$$

$$b_{n-1} \frac{n-1}{n} X_{(n)}^{(n)}$$

$$\begin{bmatrix} b_1 \frac{n-1}{n} \\ b_2 \frac{n-2}{n} + b_1 \frac{1}{n} \\ b_3 \frac{n-3}{n} + b_2 \frac{2}{n} \\ \vdots \\ b_{n-2} \frac{2}{n} + b_{n-3} \frac{n-3}{n} \\ b_{n-1} \frac{1}{n} + b_{n-2} \frac{n-2}{n} \\ b_{n-1} \frac{n-1}{n} \end{bmatrix} \qquad \underset{\sim}{X}^{(n)} = \underset{\sim}{b}^{(n)} \underset{\sim}{X}^{(n)}$$

Notando ahora que $\underset{\sim}{b}^{(n)}$ se puede expresar como:

$$\begin{bmatrix} n-1/n & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1/n & n-2/n & 0 & \cdots & 0 & 0 & 0 \\ 0 & 2/n & n-3/n & \cdots & 0 & 0 & 0 \\ 0 & 0 & 3/n & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 3/n & 0 & 0 \\ 0 & 0 & 0 & \cdots & n-3/n & 2/n & 0 \\ 0 & 0 & 0 & \cdots & 0 & n-2/n & 1/n \\ 0 & 0 & 0 & \cdots & 0 & 0 & n-1/n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_{n-2} \\ b_{n-1} \end{bmatrix} = \underset{\sim}{b}^{(n)}$$

$$= A_n \underset{\sim}{b}^{(n-1)}$$

Tenemos que:

$$E\left[\underset{\sim}{b}'_{n-1}\, \underset{\sim}{X}_{(n-1)} \mid \mathcal{O}_n \right] = \left(A_n\, \underset{\sim}{b}_{n-1} \right)' X_{(n)}$$

∴ continuando con el cálculo recursivo de $\hat{\sigma}_u$:

Para $n=5$

$$\begin{aligned} \hat{\sigma}_u^{(5)} &= E\left[\hat{\sigma}_u^{(4)} \mid \mathcal{O}_5 \right] \\ &= E\left[\underset{\sim}{b}'_4\, \underset{\sim}{X}_{(4)} \mid \mathcal{O}_5 \right] \\ &= \left(A_5\, \underset{\sim}{b}'_4 \right)' \underset{\sim}{X}_{(5)} \end{aligned}$$

pero como

$$\underset{\sim}{b}'_4 = \left(A_4 A_3\, \underset{\sim}{b}_2 \right)'$$

$$\hat{\sigma}_u^{(5)} = \left(A_5 A_4 A_3\, \underset{\sim}{b}_2 \right)' \underset{\sim}{X}_{(5)} = \underset{\sim}{b}'_{(5)}\, \underset{\sim}{X}_{(5)}$$

Hasta que finalmente

$$\begin{aligned} \hat{\sigma}_u^{(n)} &= \underset{\sim}{b}'_n\, \underset{\sim}{X}_{(n)} \\ &= \left(A_n \times A_{n-1} \times \cdots \times A_3 \times \underset{\sim}{b}_2 \right)' \underset{\sim}{X}_{(n)} \end{aligned}$$

con $\underset{\sim}{b}_2 = \begin{pmatrix} -k \\ k \end{pmatrix}$ $\quad k = \frac{\sqrt{\pi}}{2}$

y A_i, matriz de $i \times (i-1)$

$$= \begin{pmatrix} \frac{i-1}{i} & 0 & 0 & & 0 & 0 & 0 \\ \frac{1}{i} & \frac{i-2}{i} & 0 & \cdots & 0 & 0 & 0 \\ 0 & \frac{2}{i} & \frac{i-3}{i} & & 0 & 0 & 0 \\ & \vdots & & & & \vdots & \\ 0 & 0 & 0 & & \frac{i-3}{i} & \frac{2}{i} & 0 \\ 0 & 0 & 0 & \cdots & 0 & \frac{i-2}{i} & \frac{1}{i} \\ 0 & 0 & 0 & & 0 & 0 & \frac{i-1}{i} \end{pmatrix}$$

Ahora bien, podemos encontrar la forma explícita de b_n, ya que:

$$\forall_{n\geq 3} \quad A_n A_{n-1} \dots A_4 A_3 = \frac{2}{n(n-1)} \begin{pmatrix} n-1 & 0 \\ n-2 & 1 \\ n-3 & 2 \\ \vdots & \vdots \\ 2 & n-3 \\ 1 & n-2 \\ 0 & n-1 \end{pmatrix}$$

Veamos la demostración por inducción:

Para $n=3$
$$A_3 = \begin{pmatrix} 2/3 & 0 \\ 1/3 & 1/3 \\ 0 & 2/3 \end{pmatrix} = \frac{2}{3\cdot 2} \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix}$$

Suponemos que es válido para $n-1$, i.e.

$$A_{n-1} A_{n-2} \dots A_4 A_3 = \frac{2}{(n-1)(n-2)} \begin{pmatrix} n-2 & 0 \\ n-3 & 1 \\ \vdots & \vdots \\ 1 & n-3 \\ 0 & n-2 \end{pmatrix}$$

Para n:

$$A_n A_{n-1} \dots A_4 A_3 = \frac{1}{n} \begin{pmatrix} n-1 & & & & & 0 \\ 1 & n-2 & & & & \\ & 2 & n-3 & & & \\ & & 3 & \ddots & & \\ & & & \ddots & 3 & \\ & 0 & & & n-3 & 2 \\ & & & & & n-2 & 1 \\ & & & & & & n-1 \end{pmatrix} A_{n-1} \dots A_3$$

$$= \frac{2}{n(n-1)(n-2)} \begin{pmatrix} n-1 & & & 0 \\ 1 & n-2 & & \\ & 2 & \ddots & \\ & & \ddots & 2 \\ & 0 & & n-2 & 1 \\ & & & & n-1 \end{pmatrix} \begin{pmatrix} n-2 & 0 \\ n-3 & 1 \\ \vdots & \vdots \\ 1 & n-3 \\ 0 & n-2 \end{pmatrix}$$

$$= \frac{2}{n(n-1)(n-2)} A_n^* \Pi_{n-1} \qquad \text{donde:}$$

$$(A_n^*)_{ik} = \begin{cases} n-i & \text{Para } k=i \text{ con } i=1,\dots,n-1 \\ i-1 & \text{Para } k=i-1 \text{ con } i=2,\dots,n \\ 0 & \text{en otro caso con } i=1,\dots,n \text{ y } \\ & k=1,\dots,n-1 \end{cases}$$

y

$$(\Pi_{n-1})_{ki} = \begin{cases} (n-1)-k & \text{para } i=1,\ k=1,\dots,n-1 \\ k-1 & \text{para } i=2,\ k=1,\dots,n-1 \end{cases}$$

Por tanto; para $i=2,\dots,n-1$

$$(A_n^* \Pi_{n-1})_{i1} = (A_n^*)_{i,i-1}(\Pi_{n-1})_{i-1,1} + (A_n^*)_{i,i}(\Pi_{n-1})_{i,1}$$

$$= (i-1)(n-i) + (n-i)(n-1-i)$$
$$= (n-i)(i-1+n-i-i)$$
$$= (n-i)(n-2)$$

$$(A_n^* \Pi_{n-1})_{i2} = (A_n^*)_{i,i-1}(\Pi_{n-1})_{i-1,2} + (A_n^*)_{i,i}(\Pi_{n-1})_{i,2}$$

$$= (i-1)((i-1)-1) + (n-i)(i-1)$$
$$= (i-1)(i-2+n-i) =$$
$$= (i-1)(n-2)$$

y

$$(A_n^* \Pi_{n-1})_{11} = (n-1)(n-2) \qquad (A_n^* \Pi_{n-1})_{12} = 0$$

$$(A_n^* \Pi_{n-1})_{n1} = 0 \qquad (A_n^* \Pi_{n-1})_{n2} = (n-1)(n-2)$$

ie

$$\left(A^*_n \Pi_{n-1} \right) = \begin{pmatrix} (n-1)(n-2) & 0 \\ (n-2)(n-2) & 1(n-2) \\ (n-3)(n-2) & 2(n-2) \\ \vdots & \vdots \\ 2(n-2) & (n-3)(n-2) \\ 1(n-2) & (n-2)(n-2) \\ 0(n-2) & (n-1)(n-2) \end{pmatrix}$$

ie

$$\left(A^*_n \Pi_{n-1} \right) = \begin{pmatrix} n-1 & 0 \\ n-2 & 1 \\ n-3 & 2 \\ \vdots & \\ 2 & n-3 \\ 1 & n-2 \\ 0 & n-1 \end{pmatrix}$$

Así pués:

$$\underset{\sim}{b}^{(n)} = \frac{2}{n(n-1)} \begin{pmatrix} n-1 & 0 \\ n-2 & 1 \\ \vdots & \\ 1 & n-2 \\ 0 & n-1 \end{pmatrix} \begin{pmatrix} -k \\ k \end{pmatrix}$$

ie

$$\underset{\sim}{b}^{(n)} = \frac{2k}{n(n-1)} \begin{pmatrix} 1-n \\ 3-n \\ 5-n \\ \vdots \\ n-5 \\ n-3 \\ n-1 \end{pmatrix} \quad \text{con } k = \frac{\sqrt{n}}{2}$$

Por tanto, la expresión de $\hat{\mathcal{G}}^{(n)}$ como combinación lineal de las n estadísticas de orden, es:

$$\hat{\sigma}_U^{(n)} = \underset{\sim}{b}^{(n)\prime}\underset{\sim}{X}^{(n)} = \frac{\sqrt{\pi}}{n(n-1)}\begin{pmatrix} 1-n \\ 3-n \\ 5-n \\ \vdots \\ n-5 \\ n-3 \\ n-1 \end{pmatrix} \underset{\sim}{X}^{(n)}$$

$$= \frac{\sqrt{\pi}}{n(n-1)} \sum_{i=1}^{n} (2i-(n+1))\, X_{(i)}$$

Comparando estas expresiones con las correspondientes a $\hat{\sigma}^{(n)}$, podemos percatarnos de la simplicidad de $\hat{\sigma}_U^{(n)}$; su cálculo es muy sencillo y fácil de efectuar, lo que aunado a sus propiedades asintóticas, concede a la modificación ciertas ventajas sobre otros métodos estadísticos.

b) PROPIEDADES ANALÍTICAS Y ASINTÓTICAS DEL ESTIMADOR $\hat{\sigma}_U^{(n)}$.

Dando por asentada la naturaleza estadística del estimador $\hat{\sigma}_U^{(n)}$, como una estadística-U, podemos determinar su varianza y distribución asintótica a través de la teoría de Hoeffding.

Según dicha teoría, la varianza de una estadística-U:

$$U = \binom{n}{m}^{-1} \sum' \Phi(X_{\alpha_1}, \dots, X_{\alpha_m})$$

basada en el Kernel simétrico Φ de una funcional lineal Θ, con $X_1, \dots, X_n$ muestra aleatoria, esta dada por:

$$\operatorname{Var}(U) = \binom{n}{m}^{-1} \sum_{c=1}^{m} \binom{m}{c} \binom{n-m}{m-c} \zeta_c$$

donde:

$$\zeta_c = E[\Phi_c^2(X_1, \dots, X_c)] - \Theta^2$$

$$\text{y} \quad \Phi_c(x_1, \dots, x_c) = E[\Phi(x_1, \dots, x_c, X_{c+1}, \dots, X_m)].$$

Por tanto:

$$\operatorname{Var}(\hat{\sigma}_U^{(n)}) = \frac{2}{n(n-1)} \left[2(n-2) \zeta_1^{(1)} - \zeta_2^{(1)} \right]$$

$$\text{con} \quad \zeta_1^{(1)} = E[\Phi_1^{(1)}(X_1)]^2 - \sigma^2$$

$$\text{y} \quad \zeta_2^{(1)} = E[\Phi_2^{(1)}(X_1, X_2)]^2 - \sigma^2$$

Ahora como $\Phi^{(1)}(x_1, x_2) = k|x_1 - x_2|$

$$\Phi_1^{(1)}(x_1) = E[\Phi(x_1, X_2)] = E[k|x_1 - X_2|]$$

$$= \int_{-\infty}^{\infty} k|x_1 - x_2| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_2-\mu)^2} dx_2$$

puesto que $X_2 \sim N(\mu, \sigma^2)$

de donde:

$$E[\varphi_1^{(1)}(X_1)]^2 =$$

$$\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} k|x_1-x_2| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_2-\mu)^2} dx_2\right)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_1-\mu)^2} dx_1$$

$$= \left(\frac{\pi}{12} + \frac{\sqrt{3}}{2}\right)\sigma^2 \qquad \text{(Ver anexo D, (1))}$$

$$\therefore \quad \zeta_1^{(1)} = \left(\frac{\pi}{12} + \frac{\sqrt{3}}{2} - 1\right)\sigma^2$$

$$\cong .127824792\ \sigma^2$$

De igual forma:

$$\varphi_1^{(2)}(x_1, x_2) = k|x_1 - x_2|$$

$$\therefore \quad E[\varphi_1^{(2)}(X_1,X_2)]^2 = E[k^2(X_1-X_2)^2]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k^2(x_1-x_2)^2 \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}[(x_1-\mu)^2+(x_2-\mu)^2]} dx_1 dx_2$$

$$= \frac{\pi}{2}\sigma^2 \qquad \therefore \quad \zeta_2^{(1)} = \left(\frac{\pi}{2} - 1\right)\sigma^2$$

$\therefore$ finalmente:

$$\operatorname{Var}(\hat{\sigma}_U^{(n)}) = \frac{2}{n(n-1)}\left[2(n-2)\left(\frac{\pi}{12} + \frac{\sqrt{3}}{2} - 1\right)\sigma^2 - \left(\frac{\pi}{2} - 1\right)\sigma^2\right]$$

$$= \left[\frac{4(n-2)}{n(n-1)}\left(\frac{\pi}{12} + \frac{\sqrt{3}}{2} - 1\right) - \frac{2}{n(n-1)}\left(\frac{\pi}{2} - 1\right)\right]\sigma^2$$

De aquí que:

$$\lim_{n\to\infty} n\mathrm{Var}\left(\hat{\sigma}_U^{(n)}\right) = 4\left(\frac{\pi}{12}+\frac{\sqrt{3}}{2}-1\right)\sigma^2$$
$$= m^2 \xi_1^{(1)} = 4\xi_1^{(1)}$$

Siguiendo en la misma línea, si aplicamos ahora el teorema 7.2 del apéndice C, obtenemos la distribución asintótica de $\hat{\sigma}_U^{(n)}$:

7.2 nos dice que:

$$\lim_{n\to\infty} \sqrt{n}\left(\hat{\sigma}_U^{(n)}-\sigma\right) \xrightarrow{\mathcal{D}} N(0, 4\xi_1^{(1)})$$

esto es, para n suficientemente grande:

$$\hat{\sigma}_U^{(n)} \dot{\sim} N\left(\sigma, \frac{4\xi_1^{(1)}}{n}\right)$$

i.e $$\hat{\sigma}_U^{(n)} \dot{\sim} N\left(\sigma, \frac{.5113}{n}\sigma^2\right)$$

Por lo que $\hat{\sigma}_U^{(n)}$, es un estimador insesgado, consistente y con distribución asintótica normal, lo cual le da ciertas ventajas sobre otros estimadores de la desviación estandar σ.

c) ESTRUCTURA Y CARACTERISTICAS DE LA ESTADÍSTICA PROPUESTA.

Una vez establecida la naturaleza estadística del estimador $\hat{\sigma}_U^{(n)}$, entraremos de lleno a la discusión de la modificación -

propuesta, valiéndonos para ello de toda la información contenida en las secciones anteriores.

Se mencionó en el principio de este capítulo que la modificación propuesta, consiste en cambiar el numerador de la estadística de Shapiro-Wilk, el cuadrado del estimador de Lloyd del parámetro de escala σ, por el cuadrado de la estadística-U $\hat{\sigma}_U^{(n)}$, la cual es también una combinación lineal de las estadísticas de orden y un estimador insesgado de σ. Así pues, la modificación propuesta, denotada por el símbolo $W_U^{(n)}$, es:

$$W_U^{(n)} := \frac{(\hat{\sigma}_U^{(n)})^2}{S^2} = \frac{(\underset{\sim}{b}^{(n)'} \underset{\sim}{X}_{(n)})^2}{S^2}$$

$$= \frac{n \left[\sum_{i=1}^{n} (2i-(n+1) X_{(i)}) \right]^2}{[n(n-1)]^2 \cdot \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}}$$

Es de esencial importancia el notar que el estimador robusto de la varianza, S^2, es una estadística-U basada en el Kernel simétrico de tamaño dos:

$$\phi^{(2)}(x_1, x_2) = \frac{(x_1 - x_2)^2}{2}$$

de la funcional regular de grado dos: $\theta^{(2)}(F) = Var(F) = \sigma^2$, definida en el conjunto de todas las familias de distribución unidimensionales con varianza finita.

Es decir:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$
$$= \binom{n}{2}^{-1}\sum_{i<j}\frac{(X_i-X_j)^2}{2}$$

y $\forall$ F función de distribución unidimensional con varianza finita $\forall$ $X_1, X_2, \dots X_n$ muestra aleatoria de F $E(S^2)=\sigma^2$

Haciendo uso nuevamente de la teoría de Hoeffding, tenemos:

$$Var(S^2) = \frac{2}{n(n-1)}\left[2(n-2)\zeta_1^{(2)} + \zeta_2^{(2)}\right]$$

donde

$$\zeta_1^{(2)} = E[\phi_1^{(2)}(X_1)]^2 - (\sigma^2)^2$$

y

$$\zeta_2^{(2)} = E[\phi_2^{(2)}(X_1,X_2)]^2 - (\sigma^2)^2$$

Como $\phi_1^{(2)}(x_1) = E[\phi^{(2)}(x_1,X_2)]$

$$= E\left[\frac{(x_1-X_2)^2}{2}\right]$$

$$\zeta_1^{(2)} = \frac{1}{2}\sigma^4 \quad (\text{Ver apéndice D (4)})$$

De manera similar, como

$$\phi_2^{(2)}(x_1,x_2) = \phi^{(2)}(x_1,x_2) = \frac{(x_1-x_2)^2}{2}$$

$$\zeta_2^{(2)} = 2\sigma^4 \quad (\text{Ver apéndice D (5)})$$

$$\therefore \quad Var(s^2) = \frac{2}{n(n-1)}\left[2(n-2)\tfrac{1}{2}\sigma^4 + 2\sigma^4\right]$$
$$= \frac{2}{n-1}\sigma^4$$

Podemos constatar nuevamente que:

$$\lim_{n\to\infty} n\,Var(s^2) = m^2 \zeta_1^{(2)}$$

pués $\lim_{n\to\infty} n\,Var(s^2) = \lim_{n\to\infty} n\left(\frac{2}{n-1}\sigma^4\right)$
$$= 2^2\,\tfrac{1}{2}\sigma^4 = m^2\zeta_1^{(2)}$$

Aplicando ahora el teorema 7.2 del apéndice C, tenemos:

$$\lim_{n\to\infty} \sqrt{n}\,(s^2-\sigma^2) \xrightarrow{\mathcal{D}} N(0,2\sigma^4)$$

es decir, para n suficientemente grande

$$s^2 \;\dot{\sim}\; N\left(\sigma^2, \tfrac{2}{n}\sigma^4\right)$$

Lo que nos interesa ahora, es determinar si la modificación propuesta posee o no una distribución asintótica, y si existe, encontrar explícitamente su expresión. Para tales fines, volvamos a hechar mano de la teoría de estadísticas-U. En primera instancia, el teorema 7.2 del apéndice C nos indica la existencia de una distribución conjunta asintótica, para la función de distribución conjunta de $\sqrt{n}\,(\hat{\sigma}_U^{(n)}-\sigma)$ y

$\sqrt{n}(S^2-\sigma^2)$. Tal afirmación nos indica que:

$$\sqrt{n}\begin{pmatrix}\hat{\sigma}_U^{(n)}-\sigma\\ S^2-\sigma^2\end{pmatrix}\xrightarrow[n\to\infty]{\mathcal{D}} N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}4\zeta_1^{(1)}, & 4\zeta_1^{(1,2)}\\ 4\zeta_1^{(1,2)}, & 4\zeta_2^{(1)}\end{pmatrix}\right)$$

Es decir:

$$\sqrt{n}\begin{pmatrix}\hat{\sigma}_U^{(n)}-\sigma\\ S^2-\sigma^2\end{pmatrix}\xrightarrow[n\to\infty]{\mathcal{D}} N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}4\left(\frac{\pi}{12}+\frac{\sqrt{3}}{2}-1\right)\sigma^2 & \sigma^3\\ \sigma^3 & 2\sigma^4\end{pmatrix}\right)$$

ie

$$\sqrt{n}\begin{pmatrix}\hat{\sigma}_U^{(n)}-\sigma\\ S^2-\sigma^2\end{pmatrix}\dot{\sim} N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}.5112\sigma^2 & \sigma^3\\ \sigma^3 & 2\sigma^4\end{pmatrix}\right)$$

para n "grande"

(Los cálculos de $\zeta_1^{(1)}$, $\zeta_1^{(1,2)}$ y $\zeta_2^{(1)}$ se encuentran en el apéndice D (1), (3) y (4))

Es posible determinar la existencia de una distribución asintótica de $W_U^{(n)}$, en base al teorema 7.5 mencionado en el apéndice C. Tal teorema establece las condiciones suficientes para la existencia de una distribución asintótica, de una función $Z = h(U^{(1)},\dots,U^{(9)})$ donde $U^{(1)},\dots,U^{(9)}$ son 9-estadísticas-U.

Así pues, tomando la fución:

$$h(y_1,y_2)=\frac{y_1^2}{y_2} \quad \text{tenemos que:}$$

$$h(\hat{\sigma}_U^{(n)}, S^2) = \frac{(\hat{\sigma}_U^{(n)})^2}{S^2} = W_U^{(n)}$$

Esta función cumple con todas las condiciores del teorema 7.5, por lo que la variable aleatoria

$$\begin{aligned} Z &= \sqrt{n}\left(h\left(\hat{\sigma}_U^{(n)}, s^2\right) - h(\sigma, \sigma^2)\right) \\ &= \sqrt{n}\left(W_U^{(n)} - 1 \right) \end{aligned}$$

tiene por distribución asintótica una $N(0, \lambda^2)$ con

$$\lambda^2 = 16\left(\frac{\pi}{12} + \frac{\sqrt{3}}{2} - 1\right) - 2 \simeq .0452$$

Esto es, para n suficientemente grande

$$W_U^{(n)} \dot{\sim} N\left(1, \frac{.0452}{n}\right)$$

Finalmente, podemos observar que los valores generados por simulación de $W_U^{(n)}$ para $n = 40$ y 50 se adecuan en buena medida a la aproximación anterior.

III.- SIMULACIONES DE LA DISTRIBUCION NULA DE $W_U^{(n)}$ Y DE SU POTENCIA.

Como en el caso de la estadística de Shapiro-Wilk, la complejidad estadística de la modificación propuesta, no ha permitido hasta ahora, el determinar exactamente su distribución nula teórica, por lo que se ha tenido que recurrir a la técnica de la simulación, para obtener una aproximación empírica.

Una vez obtenidas las aproximaciones de los principales porcentiles, se simuló su potencia para cuatro tamaños de muestra, muestreando de varias familias de distribución. En las dos primeras secciones de este capítulo, se exponen tales simulaciones, y en la sección tercera se discuten las concluciones respecto a la efectividad de la modificación propuesta.

a) UNA APROXIMACIÓN A LA DISTRIBUCIÓN NULA DE $W_U^{(n)}$.

El procedimiento usado en la simulación de la distribución nula de $W_U^{(n)}$ a grandes rasgos fue el siguiente:

Paso 1) A través de una subrutina generadora de valores aleatorios independientes de una normal estandarizada, se obtuvieron $M(n)$ muestras aleatorias de tamaño n.

Paso 2) Ordenando las muestras y aplicando la expresión algebraica derivada en el capítulo pasado, se calculó el valor correspondiente $W_U^{(n)}$ de todas y cada una de las muestras, generándose así, una muestra aleatoria de tamaño $M(n)$ de la distribución nula de $W_U^{(n)}$.

Paso 3) Finalmente, por medio de la distribución empírica de la muestra anterior, se obtuvo una aproximación a los principales porcentiles de la distribución nula de $W_U^{(n)}$.

En la tabla I, aparecen las cantidades derivadas por el procedimiento anterior, para tamaños de muestra de 10, 20, 40 y 50. Asímismo, usando la aproximación asintótica a la distribución nula de $W_U^{(n)}$, se encontraron los respectivos porcentiles para cada n. Tales valores, se listan también en la tabla I.

Los porcentiles simulados difieren muy poco de los correspondientes valores que resultan al usar la aproximación normal mencionada. En la tabla podemos observar que la discrepancia mayor es del orden de seis centécimas; acentuándose el grado de adecuación a medida que aumenta el tamaño de muestra n.

Dicha similitud se refleja también entre la media y varianza de los porcentiles simulados ($\bar{x}(n)_s$ y $s^2(n)_s$) y la media y varianza de la aproximación normal ($\mu(n)_a$ y $\sigma^2(n)_a$), resultando aún mucho menor la diferencia en cada caso.

TABLA I. APROXIMACIONES A LA DISTRIBUCIÓN NULA DE $W_U^{(n)}$:

n	10	20	40	50
$M(n)$	4000	4000	3000	3000
.01 s	.82084	.86296	.90890	.92328
.01 a	.84342	.88933	.92191	.92987
.05 s	.92897	.93367	.94907	.95312
.05 a	.88912	.92163	.94456	.95034
.10 s	.97191	.96343	.96669	.97078
.10 a	.91398	.93920	.95699	.96147
.90 s	1.12256	1.07608	1.04977	1.04413
.90 a	1.08602	1.06080	1.04301	1.03853
.95 s	1.12893	1.08109	1.05508	1.04938
.95 a	1.11088	1.07838	1.05544	1.04967
.99 s	1.13771	1.08764	1.06198	1.05622
.99 a	1.15658	1.11068	1.07829	1.07013
$\bar{x}(n)$ s	1.01848	1.0081	.99858	.99949
$\bar{x}(n)$ a	1	1	1	1
$s^2(n)$ s	.01729	.00891	.00427	.00330
$s^2(n)$ a	.01968	.00884	.00442	.00355
$\mu(n)$ a	1	1	1	1
$\sigma^2(n)$ a	.00451	.00225	.00112	.00090

s – Aproximaciones simuladas

a – Aproximaciones asintóticas

b) POTENCIA EMPÍRICA DE $W_U^{(n)}$.

Contando con las aproximaciones a los principales percentiles de la distribución nula de $W_U^{(n)}$ que determinan las regiones de rechazo de la prueba, valores pequeños de la estadística de prueba, es posible simular su potencia tomando diversas muestras provenientes de varias poblaciones y contando el número de rechazos que ocurran en cada caso. Realizando tal procedimiento un número suficientemente grande de veces, podemos esperar que el porcentaje de rechazos, sea una buena aproximación a la potencia de la prueba.

El tamaño de la simulación fue de 600 para los cuatro tamaños de muestra: $n=10, 20, 40$ y 50. Las poblaciones muestreadas fueron: $N(0,1)$, $Exp(10)$, lognormal (1), $\chi^2(1)$, $\chi^2(2)$, $\chi^2(4)$, $\chi^2(10)$, Binomial $(4, 1/2)$ y Poisson (1).

Shapiro y Wilk, de manera similar, simularon también la potencia de su prueba al 5%, para un tamaño de muestra igual a 20, tomando 200 muestras de las poblaciones mencionadas y algunas otras. Como se refirió en la última sección del primer capítulo, compararon la efectividad de su prueba respecto a otros métodos ya conocidos, como la prueba χ^2, la Kolmogorov-Smirnov, la Cramér-Von Mises y varias más, simulando para ello la potencia de cada una con el mismo procedimiento.

De forma similar, con el propósito de establecer un criterio de comparación entre la prueba original de Shapiro y Wilk y la modificación propuesta, se simuló la potencia de la primera conjuntamente a la de $W_U^{(n)}$, usando los coeficientes y percentiles correspondientes a cada tamaño de muestra, de acuerdo a las tablas que se encuentran en el artículo de los autores.

En la tabla II, se encuentran las potencias al 5% de W y $W_U^{(n)}$ simuladas y las respectivas derivadas por Shapiro y Wilk para un tamaño de muestra igual a 20. Podemos observar que la potencia simulada de la estadística W, concuerda en buen grado con la derivada por sus autores, resultando en ambos casos sensible a la asimetría, a la dispersión por presencia de colas largas y, en algunos casos a la de colas cortas.

La potencia simulada de $W_U^{(n)}$ se encuentra siempre por debajo de la respectiva a W, lo cual era de esperarse debido a que ésta "usa" más información del comportamiento normal que $W_U^{(n)}$. Sin embargo, comparándola con los métodos b_2, KS, CVM, WCVM, D y M, la potencia simulada de $W_U^{(n)}$ es mayor para casi todas las familias. En el caso de la χ^2, $W_U^{(n)}$ detecta mejor la no normalidad para aquellas fami-

TABLA II. POTENCIAS SIMULADAS PARA UNA PRUEBA DE TAMAÑO 5% PARA $n = 20$

$W_U^{(n)}$ - Prueba Propuesta
W_1 - Prueba de Shapiro-Wilk (simulada)
W_2 - Prueba de Shapiro-Wilk (simulada actores)
χ^2 - Prueba de la Ji-cuadrada
$\sqrt{b_1}$ - Prueba del 3er momento estandarizado
b_2 - Prueba del 4to momento estandarizado

KS - Prueba de Kolmogorov-Smirnov
CVM - Prueba de Cramer-Von Mises
WCVM - Prueba de Cramer-Von Mises ponderada por $F/1-F$ (fun dist)
D - Prueba de Durbin
$\mathcal{U}$ - Prueba de David (rango/desv. estandar)

	$W_U^{(n)}$	W_1	W_2	χ^2	$\sqrt{b_1}$	b_2	KS	CVM	WCVM	D	$\mathcal{U}$
$\chi^2(1)$	.88	.98	.98	.94	.89	.53	.44	.44	.54	.87	.10
$\chi^2(2)$	.59	.80	.84	.33	.74	.34	.27	.23	.27	.42	.08
$\chi^2(4)$	.33	.54	.50	.13	.49	.27	.18	.13	.16	.15	.06
$\chi^2(10)$	.16	.24	.29	.07	.29	.19	.11	.10	.11	.07	.06
Lognor.	.81	.94	.93	.95	.89	.58	.44	.48	.62	.82	.06
Pois(1)	.60	.99	.99	1.0	.26	.11	.55	.22	.31	1.0	.35
Bin(4,½)	.33	.75	.71	1.0	.02	.03	.38	.15	.17	1.0	.20

lias que se asemejan mucho a la normal, mientras que para aquellas que son muy distintas, como la binomial y Poisson simuladas, es mucho más efectiva la prueba χ^2. Una situación contraria se presenta con el método $\sqrt{b_1}$, el cual tiene potencias más altas que $W_U^{(n)}$ para las familias parecidas a la normal y menores en las muy disimiles.

En la tabla III aparecen las potencias simuladas para $n=10, 40$ y 50 de $W_U^{(n)}$ y W, haciéndose evidente que, aunque sigue siendo superior la efectividad de W, va disminuyendo la diferencia entre los dos métodos conforme aumenta el tamaño de muestra n, manteniéndose lijeramente mayor para las familias: $\chi^2(4)$, $\chi^2(10)$ y $Bin(4, 1/2)$.

TABLA III. POTENCIAS SIMULADAS DE $W_U^{(n)}$ y W CON UN TAMAÑO DE PRUEBA DE 5%.

Tamaño de muestra	n=10		n=40		n=50	
Familias ↓	$W_U^{(n)}$	W	$W_U^{(n)}$	W	$W_U^{(n)}$	W
N (0,1)	.06	.05	.04	.03	.03	.03
Exp (10)	.37	.44	.86	.99	.91	.99
Lognor.	.55	.62	.99	1.0	.99	1.0
$\chi^2(1)$	.58	.71	.98	1.0	.99	1.0
$\chi^2(2)$	.37	.44	.85	.99	.91	1.0
$\chi^2(4)$	.23	.24	.55	.89	.63	.95
$\chi^2(10)$	.10	.10	.23	.50	.31	.37
B(4,1/2)	.25	.48	.50	1.0	.56	1.0
Poiss (1)	.46	.75	.91	1.0	.96	1.0

c) CONCLUSIONES.

La estadística W compara la configuración estadística de la muestra en objeto, con la configuración de una muestra normal, a través de los momentos de primer y segundo orden de las estadísticas de orden normales. Similarmente la estadística $W_U^{(n)}$ realiza tal comparación, usando tan sólo la información contenida en los momentos de primer orden de las estadísticas de orden, de una muestra normal de tamaño dos. Lo notable es que con tan poca información se haya obtenido una potencia que, aunque siempre es menor que la respectiva de W, no lo es terriblemente.

A la fecha, existen proyectos para mejorar tal situación, cambiando el kernel de $\hat{\sigma}_U^{(n)}$ por otros que incorporen un conocimiento semejante al de W.

Así pués, la idea de modificar $\hat{\sigma}_L$ por un estimador que sea estadística-U, es la parte medular de la modificación propuesta, y como dijimos antes, si se explota adecuadamente, puede llevarnos al descubrimiento de una prueba con potencia muy semejante a la ideada por Shapiro y Wilk, que aparte tenga todas las ventajas ya mencionadas de $W_U^{(n)}$.

APENDICE A.

Demostración de las expresiones:

(a) $\underset{\sim}{1}^{(n)'} V_{(n)}^{-1} \underset{\sim}{m}^{(n)} = 0.$

(b) $V_{(n)}^{-1} \underset{\sim}{1}^{(n)} = \underset{\sim}{1}^{(n)}$

Afirmación 1.

Sea X va continua cuya función de densidad $f_X(\cdot)$ es simétrica respecto a la recta $x=\mu$, entonces:

$$\forall_{i\in\{1,\dots,n\}} \quad \forall_{x\in\mathbb{R}} \quad f_{X_{(i)}}(\mu+x) = f_{X_{(n-i+1)}}(\mu-x)$$

dem

Recordamos que:

$$f_{X_{(i)}}(x) = B(i,n-i+1)^{-1} F_X^{i-1}(x)\left[1-F_X(x)\right]^{n-i} f_X(x)$$

Ahora dado que $f_X(\cdot)$ es simétrica respecto a $x=\mu$, se tiene: $\forall_{x\in\mathbb{R}}\ F_X(\mu-x) = 1-F_X(x+\mu)$

Por lo que:

$$f_{X_{(n-i+1)}}(\mu-x) = B(n-i+1, n-(n-i+1)+1)^{-1} F_X^{n-i+1-1}(\mu-x) \times \left[1-F_X(\mu-x)\right]^{n-(n-i+1)} f_X(\mu-x)$$

$$= B(n-i+1,i)^{-1} F_X^{n-i}(\mu-x)\left[1-F_X(\mu-x)\right]^{i-1} f_X(\mu-x)$$

$$= B(i,n-i+1)^{-1}\left[1-F_X(\mu+x)\right]^{n-i} F_X(\mu+x)^{i-1} \cdot f_X(\mu+x) = f_{X_{(i)}}(\mu+x) \quad \blacksquare$$

Afirmación 2.

Sea X va continua cuya función de densidad $f_X(\cdot)$ es simétrica respecto a la recta $x=0$, con esperanza finita, entonces:

$$\forall_{i \in \{1,\dots,n\}} \quad E[X_{(i)}] = -E[X_{(n-i+1)}]$$

dem

Sea $i \in \{1 \dots n\}$

Por la proposición 1.

$$\forall_{x \in \mathbb{R}} \quad f_{X_{(n-i+1)}}(-x) = f_{X_{(i)}}(x)$$

$\therefore$

$$E[X_{(n-i+1)}] = \int_{-\infty}^{\infty} x f_{X_{(n-i+1)}}(x)\,dx$$

Haciendo $x = -y$

$$E[X_{(n-i+1)}] = \int_{\infty}^{-\infty} (-y) f_{X_{(n-i+1)}}(-y)(-dy)$$

$$= -\int_{-\infty}^{\infty} y f_{X_{(n-i+1)}}(-y)\,dy$$

$$= -\int_{-\infty}^{\infty} y f_{X_{(i)}}(y)\,dy$$

$$= -E[X_{(i)}]$$

■

Afirmación 3.
Sea X una variable aleatoria continua cuya función de densidad sea simétrica respecto a la recta $x = \mu$, entonces:

$$\forall_{\substack{i,j\in\{1\ldots n\}\\ i\le j}} \quad \forall_{\substack{x,y\in\mathbb{R}\\ y\le x}} \; f_{X_{(n-j+1)},X_{(n-i+1)}}(\mu-x,\mu-y)$$

$$= f_{X_{(i)},X_{(j)}}(\mu+y,\mu+x)$$

dem

Sean $i,j\in\{1,\ldots,n\}$ con $i\le j$
Recordamos que la función de densidad conjunta de $X_{(i)}$ y $X_{(j)}$ está dada por:

$$f_{X_{(i)},X_{(j)}}(x,y) = \begin{cases} \dfrac{n!}{(i-1)!(j-i-1)!(n-j)!} F_X^{i-1}(x) f_X(x) \left[F_X(y)-F_X(x)\right]^{j-i-1} \\ \qquad f_X(y)\left[1-F_X(y)\right]^{n-j} & \text{para } x\le y \\ 0 & \text{en otro caso} \end{cases}$$

Sean $x,y\in\mathbb{R}$ con $y\le x$, entonces

$$f_{X_{(n-j+1)},X_{(n-i+1)}}(\mu-x,\mu-y) =$$

$$n!/(n-j+1-1)!\,(n-i+1-(n-j+1)-1)!\,(n-(n-i+1))!$$

$$\times F_X^{n-j+1-1}(\mu-x) f_X(\mu-x) \left[F_X(\mu-y)-F_X(\mu-x)\right]^{n-i+1-(n-j+1)-1}$$

$$\times f_X(\mu-y)\left[1-F_X(\mu-y)\right]^{n-(n-i+1)}$$

$$= \frac{n!}{(n-j)!(j-i-1)!(i-1)!} F_X^{n-j}(\mu-x) f_X(\mu-x)$$

$$\times \left[F_X(\mu-y)-F_X(\mu-x)\right]^{j-i-1} f_X(\mu-y)\left[1-F_X(\mu-y)\right]^{i-1}$$

Usando nuevamente que:

$$\forall_{x\in\mathbb{R}} \quad F_X(\mu-x) = 1 - F_X(\mu+x) \quad y \quad f(\mu-x) = f(\mu+x)$$

se tiene:

$$f_{X_{(n-j+1)},X_{(n-i+1)}}(\mu-x,\mu-y) = \frac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!}$$

$$\cdot \left[1-F_X(\mu+x)\right]^{n-j} f_X(\mu+x) \left[1-F_X(\mu+y)-(1-F_X(\mu+x))\right]^{j-i-1}$$

$$f_X(\mu+y)\left[1-(1-F_X(\mu+y))\right]^{i-1}$$

$$= \frac{n!}{(i-1)!\,(j-i-1)!\,(n-j)!}\left[1-F_X(\mu+x)\right]^{n-j} f_X(\mu+x)$$

$$\cdot \left[F_X(\mu+x)-F_X(\mu+y)\right]^{j-i-1} f_X(\mu+y)\, F_X^{i-1}(\mu+y$$

$$= f_{X_{(i)},X_{(j)}}(\mu+x,\mu+y) \quad \blacksquare$$

Afirmación cuatro.
Sea X una variable aleatoria continua cuya función de densidad sea simétrica respecto a $x=0$, con media y varianza finitas, entonces:

$$\forall_{\substack{i,j\in\{1,\dots,n\}\\ i<j}} \quad \operatorname{cov}(X_{(i)},X_{(j)}) = \operatorname{Cov}(X_{(n-i+1)},X_{(n-j+1)})$$

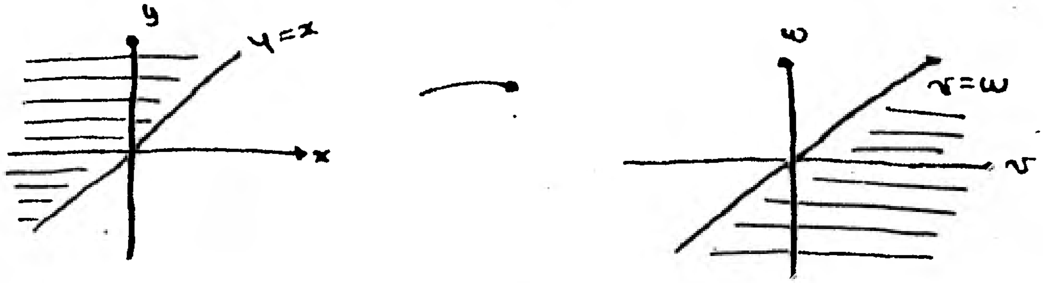
dem.

$\operatorname{cov}(X_{(n-j+1)}, X_{(n-i+1)})$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{y} (x - E[X_{(n-j+1)}])(y - E[X_{(n-i+1)}])\, f_{X_{(n-j+1)},X_{(n-i+1)}}(x,y)\, dx\, dy$$

Haciendo el cambio de variables

$$x = -v$$
$$y = -w$$



$$\left| \det \frac{D(x,y)}{D(v,w)} \right| = \left| \det \begin{pmatrix} +0 & -1 \\ -1 & 0 \end{pmatrix} \right|$$
$$= |1| = 1$$

Usando la afirmación 3

$\operatorname{cov}(X_{(n-j+1)}, X_{(n-i+1)})$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{v} (-v + E(X_{(j)}))(-w + E(X_{(i)}))\, f_{X_{(n-j+1)}X_{(n-i+1)}}(-v,-w)\, dw\, dv$$

y ahora por la afirmación 4 con $\mu = 0$, la integral anterior es igual a:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{v} (v - E(X_{(i)}))(w - E(X_{(j)}))\, f_{X_{(i)},X_{(j)}}(w,v)\, dw\, dv$$
$$= \operatorname{Cov}(X_{(i)}, X_{(j)}) \blacksquare$$

Afirmación 5.

Sea X una variable aleatoria continua con función de densidad $f_X(\cdot)$ continua y con media y varianza finitas, entonces:

a) $\sum_{i=1}^{n} E(X_{(i)}) = n\,E(X)$

b) $\sum_{i=1}^{n} E(X_{(i)}^2) = n\,E(X^2)$

c) $\sum_{i=1}^{n} \sum_{j=1}^{n} E(X_{(i)}, X_{(j)}) = n\,E(X^2) + n(n-1)\,E^2(X)$

d) $\sum_{i=1}^{n} \sum_{j=1}^{n} Cov(X_{(i)}, X_{(j)}) = n\,Var(X)$

dem

Teniendo en cuenta que:

$$\forall_{k,m \in \mathbb{N}} \quad \left(\sum_{i=1}^{n} X_{(i)}^k \right)^m = \left(\sum_{i=1}^{n} X_i^k \right)^m$$

tendremos: para toda $k \in \mathbb{N}$ la siguiente expresión:

$$E\left[\sum_{i=1}^{n} X_{(i)}^k \right] = E\left[\sum_{i=1}^{n} X_i^k \right]$$

$$\therefore \quad \sum_{i=1}^{n} E[X_{(i)}^k] = \sum_{i=1}^{n} E[X_i^k] = n\,E[X^k].$$

Con $k=1$ se obtiene a) y con $k=2$ b).

De igual forma con $k=1$ y $m=2$

$$E\left[\sum_{i=1}^{n} X_{(i)} \right]^2 = E\left[\sum_{i=1}^{n} X_i \right]^2$$

ie

$$\sum_{i=1}^{n} \sum_{j=1}^{n} E[X_{(i)} X_{(j)}] = \sum_{i=1}^{n} \sum_{j=1}^{n} E[X_i X_j]$$

y El miembro derecho de esta igualdad a su vez es igual a:

$$\sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} E[X_i X_j] + \sum_{i=1}^{n} E[X_i^2]$$

$$= \sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} E[X_i] E[X_j] + n E[X^2]$$

$$= n(n-1) E^2[X] + n E[X^2]$$

$$\therefore \sum_{i=1}^{n} \sum_{j=1}^{n} E[X_{(i)} X_{(j)}] = n E[X^2] + n(n-1) E^2[X]$$

lo cual demuestra c)

Por último usando: a) y el mismo argumento:

$$\sum_{i=1}^{n} (X_{(i)} - E(X_{(i)})) \text{ es igual a}$$

$$\sum_{i=1}^{n} (X_i - E(X_i))$$

de donde:

$$E\left[\left(\sum_{i=1}^{n} (X_{(i)} - E(X_{(i)}))\right)^2\right] = E\left[\left(\sum_{i=1}^{n} (X_i - E(X_i))\right)^2\right]$$

ie

$$E\left[\sum_{i=1}^{n} \sum_{j=1}^{n} (X_{(i)} - E(X_{(i)}))(X_{(j)} - E(X_{(j)}))\right] =$$

$$E\left[\sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - E(X_i))(X_j - E(X_j))\right]$$

ie $$\sum_{i=1}^{n}\sum_{j=1}^{n} cov(X_{(i)}, X_{(j)}) = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j)$$

y dado que $\forall_{i,j \in \{1 \dots n\}} \; i \neq j \Leftrightarrow cov(X_i, X_j) = 0$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j) = \sum_{i=1}^{n} Var(X_i)$$
$$= n\,Var(X)$$

$$\therefore \sum_{i=1}^{n}\sum_{j=1}^{n} cov(X_{(i)}, X_{(j)}) = n\,Var(X)$$

por lo que d) ha quedado demostrado y con ello la proposición cinco. ■

Afirmación 6.
Sea X una variable aleatoria que se distribuye según una $N(0,1)$, entonces:

$$V\underline{1} = \underline{1}$$

donde V es la matriz de varianzas y covarianzas de $X_{(1)}, \dots, X_{(n)}$, las n-estadísticas de orden, correspondientes a una muestra aleatoria $X_1, \dots, X_n$ de la misma distribución.
dem.

Aplíquese en orden ascendente las afirmaciones 2, 4 y 5. ■

Afirmación 7.
Sea X una variable aleatoria que se distribuye

según una $N(0,1)$, entonces:

$$\underline{1}' V^{-1} \underline{m} = 0$$

donde V es la misma matriz definida en la afirmación anterior y $\underline{m}$ es el vector de medias respectivo.

dem

Dado que la familia normal es simétrica respecto a la media, en este caso igual a cero; la función de densidad conjunta de $X_{(1)}, \ldots, X_{(n)}$ es la misma que la del vector $-X_{(n)}, -X_{(n-1)}, \ldots, -X_{(2)}, -X_{(1)}$, es-decir

$$\forall_{(x_1,\ldots x_n) \in \mathbb{R}^n} \quad f_{X_{(1)}, X_{(2)}, \ldots, X_{(n)}}(x_1,\ldots,x_n) = f_{-X_{(n)},-X_{(n-1)},\ldots,-X_{(2)},-X_{(1)}}(x_1,\ldots x_n)$$

Ahora tomando

$$J = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & & & \vdots & \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

se tiene

$$\underline{X}^* := \begin{pmatrix} -X_{(n)} \\ -X_{(n-1)} \\ \vdots \\ -X_{(2)} \\ -X_{(1)} \end{pmatrix} = -J \begin{pmatrix} X_{(1)} \\ X_{(2)} \\ \vdots \\ X_{(n-1)} \\ X_{(n)} \end{pmatrix} = -J\underline{X}$$

Dado que ambos vectores tienen la misma función de densidad, ambos tienen por tanto, igual vector de medias e idéntica matriz de varianzas y covarianzas. En notación matricial:

$$E[\underline{X}^*] = E[\underline{X}]$$

$$\text{ie} \quad E[-J\underline{x}] = E[\underline{x}]$$
$$\text{ie} \quad -J\,E[\underline{x}] = E[\underline{x}]$$
$$\text{ie} \quad -J\,\underline{m} = \underline{m}$$

De igual forma:

$$\operatorname{Var}(\underline{x}^*) = \operatorname{Var}(\underline{x})$$
$$\text{ie} \quad \operatorname{Var}(-J\underline{x}) = \operatorname{Var}(\underline{x})$$
$$\text{ie} \quad (-J)'\operatorname{Var}(\underline{x})(-J) = \operatorname{Var}(\underline{x})$$

Ahora por definición de la matriz J se tiene que:

$$J = J'$$

$\therefore$ la última igualdad se transforma en:

$$J\operatorname{Var}(\underline{x})J = \operatorname{Var}(\underline{x})$$
$$\text{ie} \quad JVJ = V$$

$$\therefore \quad V^{-1} = (JVJ)^{-1}$$
$$= J^{-1}V^{-1}J^{-1}$$

Nuevamente por definición de J

$$J^{-1} = J$$

$$\therefore \quad V^{-1} = JV^{-1}J$$

De aquí y del resultado anterior

$$\underline{1}'V^{-1}\underline{m} = \underline{1}'(JV^{-1}J)\,\underline{m}$$
$$= \underline{1}'(JV^{-1}J)(-J\underline{m})$$
$$= -(\underline{1}'J)\,V^{-1}(J\cdot J)\,\underline{m}$$

Pero, por definición de J:

$$J^{-1} = J \quad y \quad \underset{\sim}{1}' J = \underset{\sim}{1}'$$

$$\therefore \quad \underset{\sim}{1}' V^{-1} \underset{\sim}{m} = -\underset{\sim}{1}' V^{-1} \underset{\sim}{m}$$

$$\text{ie} \quad \underset{\sim}{1}' V^{-1} \underset{\sim}{m} = 0$$

APÉNDICE B.

Demostración de la proposición:

$$E\left[k|X_1-X_2|\,\middle|\,X_{(1)}^{(n-1)},\dots,X_{(n-1)}^{(n-1)},X_n\right]$$

$$= E\left[k|X_1-X_2|\,\middle|\,X_{(1)}^{(n-1)},\dots,X_{(n-1)}^{(n-1)}\right]$$

debido a la independencia de

$X_{(1)}^{(n-1)},\dots,X_{(n-1)}^{(n-1)}$ y X_n y entre
$k|X_1-X_2|, X_{(1)}^{(n-1)},\dots,X_{(n-1)}^{(n-1)}$ y X_n.

dem:

Denotamos por $\mathcal{O}_{n-1}$ a $X_{(1)}^{(n-1)},\dots,X_{(n-1)}^{(n-1)}$
Como $E[k|X_1-X_2|\,\mathcal{O}_{n-1}]$ es $\sigma(\mathcal{O}_{n-1})$ medible y $\sigma(\mathcal{O}_{n-1})\subset\sigma(\mathcal{O}_{n-1},X_n)$, también es $\sigma(\mathcal{O}_{n-1},X_n)$ medible, por lo que si demostramos que:

$$\forall B\in\sigma(\mathcal{O}_{n-1},X_n)\quad \int_B E[k|X_1-X_2|\,|\,\mathcal{O}_{n-1}]\,dP=\int_B k|X_1-X_2|\,dP$$

por la unicidad de la esperanza condicional, tendríamos la igualdad deseada:

$$E[k|X_1-X_2|\,|\,\mathcal{O}_{n-1},X_n]=E[k|X_1-X_2|\,|\,\mathcal{O}_{n-1}]$$

Sean pués, $U_1,U_2,\dots,U_{n-1},V$ abiertos de $\mathbb{R}$, denotemos por A al conjunto:

$$A:=\left(X_{(1)}^{(n-1)}\right)^{-1}(U_1)\cap\left(X_{(2)}^{(n-1)}\right)^{-1}(U_2)\dots\cap\left(X_{(n-1)}^{(n-1)}\right)^{-1}(U_{n-1})$$

claramente:

$$A\in\sigma(\mathcal{O}_{n-1})\ \text{y}\ A\cap X_n^{-1}(V)\in\sigma(\mathcal{O}_{n-1},X_n).$$

Sea $\left\{ S_m = \sum_{i=1}^{r_m} a_{im} \chi_{A_{im}} \right\}_{m \in \mathbb{N}}$ una sucesión de funciones simples crecientes que converjan a $E[k|X_1-X_2| \mid \mathcal{Q}_{n-1}]$. Dado que $\mathcal{Q}_{n-1}$ y X_n son independientes, también lo son $E[k|X_1-X_2| \mid \mathcal{Q}_{n-1}]$ y X_n, por lo que es posible tomar las A_{im}-s independientes de $X_n^{-1}(V)$, i.e.:

$$\forall_{m \in \mathbb{N}} \ \forall_{i \in \{1 \dots r_m\}} \quad A_{im} \text{ es independiente de } X_n^{-1}(V),$$

y dado que $A \in \mathcal{Q}_{n-1}$, tenemos que:

$$\forall_{m \in \mathbb{N}} \ \forall_{i \in \{1 \dots r_m\}} \quad A_{im} \cap A \text{ es independiente de } X_n^{-1}(V).$$

Por tanto:

$$\forall_{m \in \mathbb{N}} \int_{A \cap X_n^{-1}(V)} S_m \, dP = \sum_{i=1}^{r_m} a_{im} P(A \cap X_n^{-1}(V) \cap A_{im})$$

$$= P(X_n^{-1}(V)) \int_A S_m \, dP$$

$$\therefore \int_{A \cap X_n^{-1}(V)} E[k|X_1-X_2| \mid \mathcal{Q}_{n-1}] \, dP = \lim_{m \to \infty} \int_{A \cap X_n^{-1}(V)} S_m \, dP$$

$$= P(X_n^{-1}(V)) \lim_{m \to \infty} \int_A S_m \, dP$$

$$= P(X_n^{-1}(V)) \int_A E[k|X_1-X_2| \mid \mathcal{Q}_{n-1}] \, dP$$

i.e.

$$\int_{A \cap X_n^{-1}(V)} E[k|X_1-X_2| \mid \mathcal{Q}_{n-1}] \, dP = P(X_n^{-1}(V)) \int_A E[k|X_1-X_2| \mid \mathcal{Q}_{n-1}] \, dP$$

Sea $\left\{ t_m = \sum_{i=1}^{q_m} b_{im} \chi_{B_{im}} \right\}_{m\in\mathbb{N}}$ una sucesión de funciones simples

crecientes, que converge a $k|X_1-X_2|$. Ahora dado que $k|X_1-X_2|$, $\mathcal{O}_{n-1}$ es indep. de X_n, podemos tomar las $B_{im}s$ $\ni$.

$$\forall_{m\in\mathbb{N}} \ \forall_{i\in\{1,\dots,q_m\}} \quad B_{im}\cap A \text{ es indep. de } X_n^{-1}(V)$$

$$\therefore \ \forall_{m\in\mathbb{N}} \int_{A\cap X_n^{-1}(V)} t_m \, dP = \sum_{i=1}^{q_m} b_{im} P(A\cap X_n^{-1}(V)\cap B_{im})$$

$$= P(X_n^{-1}(V)) \int_A t_m \, dP$$

$$\therefore \int_{A\cap X_n^{-1}(V)} k|X_1-X_2| \, dP = \lim_{m\to\infty} \int_{A\cap X_n^{-1}(V)} t_m \, dP$$

$$= P(X_n^{-1}(V)) \int_A k|X_1-X_2| \, dP$$

i.e.

$$\int_{A\cap X_n^{-1}(V)} k|X_1-X_2| \, dP = P(X_n^{-1}(V)) \int_A k|X_1-X_2| \, dP$$

y como $A\in\sigma(\mathcal{O}_{n-1})$ $\int_A k|X_1-X_2| \, dP = \int_A \mathbb{E}[k|X_1-X_2| \mid \mathcal{O}_{n-1}] \, dP$

$$\therefore \int_{A\cap X_n^{-1}(V)} \mathbb{E}[k|X_1-X_2| \mid \mathcal{O}_{n-1}] \, dP = P(X_n^{-1}(V)) \int_A \mathbb{E}[k|X_1-X_2| \mid \mathcal{O}_{n-1}] \, dP$$

$$= P(X_n^{-1}(V)) \int_A k|X_1-X_2| \, dP = \int_{A\cap X_n^{-1}(V)} k|X_1-X_2| \, dP.$$

Con lo cual, hemos demostrado que:

$$\forall_{\substack{U_1,\dots,U_{n-1},V \\ \text{abiertos de } \mathbb{R}}} \int_{(X_{(1)}^{n-1})^{-1}(U_1)\cap\dots\cap(X_{(n-1)}^{n-1})^{-1}(U_{n-1})\cap X_n^{-1}(V)} E\left[k|X_1-X_2| \mid \mathcal{O}_{n-1}\right] dP$$

$$= \int_{(X_{(1)}^{n-1})^{-1}(U_1)\cap\dots\cap(X_{(n-1)}^{n-1})^{-1}(U_{n-1})\cap X_n^{-1}(V)} k|X_1-X_2|\, dP$$

$$= \int_{(X_{(1)}^{n-1})^{-1}(U_1)\cap\dots\cap(X_{(n-1)}^{n-1})^{-1}(U_{n-1})\cap X_n^{-1}(V)} E\left[k|X_1-X_2| \mid \mathcal{O}_{n-1}, X_n\right] dP$$

Con lo que ambas esperanzas condicionales poseen la misma integral sobre cualquier conjunto de la clase

$$\mathcal{C} = \left\{ (X_{(1)}^{n-1})^{-1}(U_1)\cap\dots\cap(X_{(n-1)}^{n-1})^{-1}(U_{n-1})\cap X_n^{-1}(V) \;\middle|\; U_1,\dots,U_{n-1},V \text{ son abiertos de } \mathbb{R} \right\}$$

Y dado que tal clase es cerrada bajo intersecciones, por el teorema de extensión:

$$\forall_{B\in\sigma(\mathcal{C})} \int_B E\left[k|X_1-X_2| \mid \mathcal{O}_{n-1}\right] dP = \int_B E\left[k|X_1-X_2| \mid \mathcal{O}_{n-1}, X_n\right] dP$$

pero $\sigma(\mathcal{C}) = \sigma(\mathcal{O}_{n-1}, X_n)$; lo cual demuestra nuestra afirmación:

$$E\left[k|X_1-X_2| \mid \mathcal{O}_{n-1}\right] \underset{c.s.}{=} E\left[k|X_1-X_2| \mid \mathcal{O}_{n-1}, X_n\right]$$

APÉNDICE C.

De los teoremas principales de la teoría de Estadísticas-U de Hoeffding, usados en la determinación de las distribuciones asintóticas de $\hat{T}_U^{(n)}$ y $W_U^{(n)}$.

TEOREMA 7.1

Sea $X_1, \dots, X_n$ muestra aleatoria de $F_X(\cdot)$.
Sean $\varphi^{(i)}(x_1, \dots, x_{m(i)})$ $i = 1, \dots, g$, g funciones reales que no dependan de n, simétricas en sus $m(i) \leq n$ argumentos.
Sean

$$\forall_{i \in \{1,\dots,g\}} \; U^{(i)} = \binom{n}{m(i)}^{-1} \sum' \varphi^{(i)}(X_{j_1}, \dots, X_{j_{m(i)}})$$

donde $\sum'$ es la suma sobre todas las permutaciones ordenadas del conjunto de índices $\{1, \dots, d(i)\}$.
Si existen $\forall_{i \in \{1,\dots,g\}}$ $E[\varphi^{(i)}(X_1, \dots, X_{m(i)})] = \Theta^{(i)}$ y $E[\varphi^{(i)}(X_1, \dots, X_{m(i)})]^2$

entonces:
la distribución conjunta de

$$\sqrt{n}\,(U^{(1)} - \Theta^{(1)}, U^{(2)} - \Theta^{(2)}, \dots, U^{(g)} - \Theta^{(g)})$$

tiende, cuando $n \to \infty$, a una normal g-variada con media cero y matriz de varianzas y covarianzas:

$$V = (m(i)m(j)\, \zeta_1^{(i,j)})_{i,j \in \{1,\dots,g\}}$$

donde

$$\zeta_1^{(i,j)} = E[(\varphi_1^{(i)}(X_1) - \Theta^{(i)})(\varphi_1^{(j)}(X_1) - \Theta^{(j)})]$$

con $\varphi_1^{(i)}(x_i) = E\left[\varphi^{(i)}(x_1, X_2, \dots, X_{m(i)})\right]$

para $i \in \{1, \dots, g\}$.

Además, si el determinante de la matriz $\left(\zeta_1^{(i,j)}\right)_{i,j \in \{1,\dots,g\}}$ es positivo, entonces la función de distribución límite es no-singular.

Nota: Para $g=1$, el teorema anterior nos indica, bajo ciertas condiciones, la distribución asintótica de una estadística-U.

TEOREMA 7.2

Bajo las mismas hipótesis del teorema 7.1, y si $\forall_{i \in \{1,\dots,g\}}$ $\zeta_1^{(i)} > 0$ entonces la distribución conjunta de:

$$\left(\frac{U^{(1)} - \Theta^{(1)}}{\sqrt{Var(U^{(1)})}}, \dots, \frac{U^{(g)} - \Theta^{(g)}}{\sqrt{Var(U^{(g)})}} \right)$$

tiende, cuando $n \to \infty$, a una normal g-variada con media cero y matriz de varianzas y covarianzas: $\left(\rho(i,j)\right)_{i,j \in \{1,\dots,g\}}$ donde

$$\rho^{(i,j)} = \lim_{n \to \infty} \frac{cov(U^{(i)}, U^{(j)})}{\left(Var(U^{(i)}) Var(U^{(j)})\right)^{1/2}}$$

$$= \frac{\zeta_1^{(i,j)}}{\left(\zeta_1^{(i)} \zeta_1^{(j)}\right)^{1/2}} \qquad i,j \in \{1,\dots,g\}$$

Este teorema es un corolario del anterior

TEOREMA 7.3

Sea $\forall_{i \in \{1,\dots,g\}}$ $\{Y_n^{(i)}\}_{n=1}^{\infty}$ una sucesión de variables aleatorias tal que:

$$\lim_{n \to \infty} E[Y_n^{(i)}]^2 = 0$$

Sea $\forall_{i \in \{1,\dots,g\}}$ $U^{(i)'} = U^{(i)} + \frac{Y_n^{(i)}}{\sqrt{n}}$

donde $U^{(i)}$ está dada por las mismas condiciones del teorema 7.1, entonces la función de distribución conjunta de

$$\sqrt{n}\,(U^{(1)'} - \theta^{(1)}, \dots, U^{(g)'} - \theta^{(g)})$$

tiende, cuando $n \to \infty$, a una normal g-variada con media cero y matriz de varianzas y covarianzas $(m(i)m(j)\,\zeta_1^{(i,j)})_{ij}$

TEOREMA 7.5.

Sean $U^{(i)'}$ $i=1,\dots,g$ como en el teorema anterior. Sea $h(y_1,\dots,y_g) \in C^2(B_\varepsilon(\underset{\sim}{\theta}))$ donde $\varepsilon > 0$ y $\underset{\sim}{\theta} = (\theta^{(1)},\dots,\theta^{(g)})$, entonces si h no involucra a n, la distribución de:

$$Z_n = \sqrt{n}\,(h(U^{(1)'},\dots,U^{(g)'}) - h(\underset{\sim}{\theta}))$$

tiende asintóticamente a una distribución normal con media cero y varianza:

$$\sum_{i=1}^{g} \sum_{j=1}^{g} m(i)m(j)\, h_{y_i}(\underset{\sim}{\theta})\, h_{y_j}(\underset{\sim}{\theta})\, \zeta_1^{(i,j)}.$$

APENDICE D

De los cálculos necesarios para la determinación de las distribuciones asintóticas de $\hat{\sigma}_U^{(n)}$ y $W_U^{(n)}$.

1) $E[\phi_1^{(1)2}(X_1)] = \left(\frac{\pi}{12} + \frac{\sqrt{3}}{2}\right)\sigma^2$

Como $\phi^{(1)}(x_1,x_2) = k|x_1-x_2|$ con $k = \sqrt{\pi}/2$

$$\phi_1^{(1)}(x_1) = E[\phi(x_1,X_2)] = E[k|x_1-X_2|]$$

$$= \int_{-\infty}^{\infty} k|x_1-x_2| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_2-\mu)^2} dx_2$$

Ahora tomando $\tau = \frac{x_2-\mu}{\sigma}$

$$\phi_1(x_1) = \int_{-\infty}^{\infty} \frac{k}{\sqrt{2\pi}} |x_1-\sigma\tau-\mu| e^{-\frac{1}{2}\tau^2} d\tau$$

$$= \frac{k}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x_1-\mu}{\sigma}} (x_1-\sigma\tau-\mu) e^{-\frac{1}{2}\tau^2} d\tau + \frac{k}{\sqrt{2\pi}} \int_{\frac{x_1-\mu}{\sigma}}^{\infty} (\sigma\tau+\mu-x_1) e^{-\frac{1}{2}\tau^2} d\tau$$

$$= \frac{k}{\sqrt{2\pi}}(x_1-\mu)\left[\int_{-\infty}^{\frac{x_1-\mu}{\sigma}} e^{-\frac{1}{2}\tau^2} d\tau - \int_{\frac{x_1-\mu}{\sigma}}^{\infty} e^{-\frac{1}{2}\tau^2} d\tau\right]$$

$$- \frac{\sigma k}{\sqrt{2\pi}}\left[-e^{-\frac{1}{2}\tau^2}\Big|_{-\infty}^{\frac{x_1-\mu}{\sigma}} + e^{-\frac{1}{2}\tau^2}\Big|_{\frac{x_1-\mu}{\sigma}}^{\infty}\right]$$

$$= \frac{k(x_1-\mu)}{\sqrt{2\pi}}\left[\int_{-\left(\frac{x_1-\mu}{\sigma}\right)}^{\infty} e^{-\frac{1}{2}\tau^2} d\tau - \int_{\frac{x_1-\mu}{\sigma}}^{\infty} e^{-\frac{1}{2}\tau^2} d\tau\right]$$

— continúa en la página siguiente.

$$-\frac{\sigma K}{\sqrt{2\pi}}\left[-e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2}-e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2}\right]$$

$$=\frac{K(x_1-\mu)}{\sqrt{2\pi}}\int_{-\left(\frac{x_1-\mu}{\sigma}\right)}^{\left(\frac{x_1-\mu}{\sigma}\right)} e^{-\frac{1}{2}\tau^2}\,d\tau+\frac{2\sigma K}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2}$$

$$\therefore E\left[\Phi_1^{(1)2}(X_1)\right]=$$

$$\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}K|x_1-x_2|\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x_2-\mu)^2}dx_2\right)^2\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x_1-\mu)^2}dx_1$$

$$=\int_{-\infty}^{\infty}\left[\frac{2\sigma K}{\sqrt{2\pi}}\left(\left(\frac{x_1-\mu}{\sigma}\right)\int_0^{\left(\frac{x_1-\mu}{\sigma}\right)}e^{-\frac{1}{2}\tau^2}d\tau+e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2}\right)\right]^2$$

$$\times\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2}dx_1$$

Tomando $\psi=\dfrac{x_1-\mu}{\sigma}$

$$E\left[\Phi_1^{(1)2}(X_1)\right]=$$

$$\int_{-\infty}^{\infty}\left[\frac{2\sigma K}{\sqrt{2\pi}}\left(\psi\int_0^{\psi}e^{-\frac{1}{2}\tau^2}d\tau+e^{-\frac{1}{2}\psi^2}\right)\right]^2\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\psi^2}\sigma\,d\psi$$

$$=\frac{4\sigma^2K^2}{(2\pi)^{3/2}}\int_{-\infty}^{\infty}\left(\psi\int_0^{\psi}e^{-\frac{1}{2}\tau^2}d\tau+e^{-\frac{1}{2}\psi^2}\right)^2e^{-\frac{1}{2}\psi^2}d\psi$$

Denotamos por Int a la integral anterior.
Sea
$$h(\psi)=\int_0^{\psi} e^{-\frac{1}{2}\tau^2}d\tau \quad \therefore\; h'(\psi)=e^{-\frac{1}{2}\psi^2}$$

Con esta notación Int queda como:
$$\mathrm{Int}=\int_{-\infty}^{\infty}(\psi h(\psi)+h'(\psi))^2 h'(\psi)\,d\psi$$
$$=\int_{-\infty}^{\infty}\left(\psi^2 h(\psi)^2 h'(\psi)+2\psi h(\psi)h'(\psi)^2+h'(\psi)^3\right)d\psi$$

Sean
$$(A)=\int_{-\infty}^{\infty}\psi^2 h(\psi)^2 h'(\psi)\,d\psi$$
$$(B)=\int_{-\infty}^{\infty}2\psi h(\psi)h'(\psi)^2\,d\psi$$
y
$$(C)=\int_{-\infty}^{\infty}h'(\psi)^3\,d\psi$$

$$(A)=\int_{-\infty}^{\infty}\psi^2 h(\psi)^2 h'(\psi)\,d\psi$$

$$dv=\psi^2 h'(\psi) \qquad \mu=h^2(\psi)$$
$$v=-\psi h'(\psi)+h(\psi) \qquad d\mu=2h(\psi)h'(\psi)$$

$$(A)=(-\psi h'\psi+h(\psi))h^2(\psi)\Big|_{-\infty}^{\infty}$$
$$-\int_{-\infty}^{\infty}(-\psi h'(\psi)+h(\psi))\,2h(\psi)h'(\psi)\,d\psi$$

Sean $(a) = \left(-\psi h'(\psi)+h(\psi)\right) h^2(\psi) \Big|_{-\infty}^{\infty}$

$$(b) = 2\int_{-\infty}^{\infty} \psi h(\psi) h'(\psi)^2 d\psi$$

y $(c) = -2\int_{-\infty}^{\infty} h(\psi)^2 h'(\psi) d\psi$

entonces $(A) = (a)+(b)+(c)$

Ahora $(a) = -\psi h'(\psi) h^2(\psi) \Big|_{-\infty}^{\infty}$

$$+ h^3(\psi)\Big|_{-\infty}^{\infty} = h^3(\infty) - h^3(-\infty)$$

$$= \left(\int_0^{\infty} e^{-\frac{1}{2}z^2} dz\right)^3 - \left(\int_0^{-\infty} e^{-\frac{1}{2}z^2} dz\right)^3$$

$$= 2\left(\frac{\sqrt{2\pi}}{2}\right)^3 = \frac{(2\pi)^{3/2}}{4}.$$

$$(b) = 2\int_{-\infty}^{\infty} \psi h(\psi) h'(\psi)^2 d\psi$$

$dv = \psi h'(\psi)^2 \qquad u = h(\psi)$

$v = -\frac{1}{2} h'(\psi)^2 \qquad du = h'(\psi)$

$$\therefore (b) = 2h(\psi)\left(-\frac{1}{2} h'(\psi)^2\right)\Big|_{-\infty}^{\infty}$$

$$-2\int_{-\infty}^{\infty} \left(-\frac{1}{2}\right) h'(\psi)^3 d\psi$$

i.e. $(b) = \int_{-\infty}^{\infty} h'(\psi)^3 d\psi = \int_{-\infty}^{\infty} (e^{-\frac{1}{2}\psi^2})^3 d\psi$

$$= \int_{-\infty}^{\infty} e^{-\frac{3}{2}\psi^2} d\psi = \frac{\sqrt{2\pi}}{\sqrt{3}}$$

$$(e) = -2\int_{-\infty}^{\infty} h(\psi)^2 h'(\psi)\, d\psi = -\frac{2}{3} h(\psi)^3 \Big|_{-\infty}^{\infty}$$

$$= -\frac{2}{3}\left[h^3(\infty) - h^3(-\infty) \right] = -\frac{2}{3}\left[\frac{(2\pi)^{3/2}}{4} \right]$$

$$= -\frac{(2\pi)^{3/2}}{6}$$

$$\therefore (A) = \frac{(2\pi)^{3/2}}{4} + \frac{(2\pi)^{1/2}}{3^{1/2}} - \frac{(2\pi)^{3/2}}{6}$$

$$= \frac{(2\pi)^{3/2}}{12} + \frac{(2\pi)^{1/2}}{3^{1/2}}$$

Observando que $(B) = (b)$, tenemos:

$$(B) = \frac{(2\pi)^{1/2}}{3^{1/2}}$$

De igual forma $(c) = (b)$

$$\therefore (c) = \frac{(2\pi)^{1/2}}{3^{1/2}}$$

por lo que, finalmente:

$$Int = A + B + C = \frac{(2\pi)^{3/2}}{12} + \frac{(2\pi)^{1/2}}{3^{1/2}} + \frac{(2\pi)^{1/2}}{3^{1/2}} + \frac{(2\pi)^{1/2}}{3^{1/2}}.$$

con $\psi = \dfrac{x_2-\mu}{\sigma}$, por lo que:

$$\zeta_1^{(1,2)} = \int_{-\infty}^{\infty} \frac{2\sigma k}{2\sqrt{2\pi}} \left[\psi \int_0^{\psi} e^{-\frac{1}{2}\tau^2} d\tau + e^{-\frac{1}{2}\psi^2}\right]$$

$$\times \left[\sigma^2\psi^2 + \sigma^2\right] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\psi^2} d\psi - \sigma^3$$

$$= \frac{\sigma^3 k}{(\sqrt{2\pi})^2} \int_{-\infty}^{\infty} (\psi h(\psi) + h'(\psi))(\psi^2+1) h'(\psi)\, d\psi - \sigma^3$$

con $h(\psi) = \int_0^{\psi} e^{-\frac{1}{2}\tau^2} d\tau$.

Sean:

$$(A) = \int_{-\infty}^{\infty} \psi^3 h(\psi) h'(\psi)\, d\psi$$

$$(B) = \int_{-\infty}^{\infty} \psi h(\psi) h'(\psi)\, d\psi$$

$$(C) = \int_{-\infty}^{\infty} \psi^2 h'(\psi)^2\, d\psi$$

y $$(D) = \int_{-\infty}^{\infty} h'(\psi)^2\, d\psi$$

ie

$$\zeta_1^{(1;2)} = \frac{\sigma^3 k}{2\pi}\Big((A) + (B) + (C) + (D) \Big) - \sigma^3$$

$$(A) = \int_{-\infty}^{\infty} \psi^3 h(\psi) h'(\psi)\, d\psi$$

$$dv = \psi^3 h'(\psi) \qquad u = h(\psi)$$
$$v = -h'(\psi)(\psi^2+2) \qquad du = h'(\psi)$$

$$(A) = -h'(\psi)(\psi^2+2)h(\psi)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -h'(\psi)^2(\psi^2+2)\,d\psi$$

$$= \int_{-\infty}^{\infty} h'(\psi)^2(\psi^2+2)\,d\psi$$

$$= \int_{-\infty}^{\infty} \psi^2 h'(\psi)^2\,d\psi + \int_{-\infty}^{\infty} 2h'(\psi)^2\,d\psi$$

$$= -\frac{1}{2}\psi\, h'(\psi)^2\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -\frac{1}{2}h'(\psi)^2\,d\psi + 2\pi^{1/2}$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} h'(\psi)^2\,d\psi + 2\pi^{1/2} = \frac{\pi^{1/2}}{2} + 2\pi^{1/2}$$

$$(B) = \int_{-\infty}^{\infty} \psi h(\psi) h'(\psi)\,d\psi$$

$$dv = \psi h'(\psi) \qquad u = h(\psi)$$
$$v = -h'(\psi) \qquad du = h'(\psi)$$

$$(B) = -h'(\psi)h(\psi)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -h'(\psi)\,h'(\psi)\,d\psi$$

$$= \int_{-\infty}^{\infty} h'(\psi)^2\,d\psi = \pi^{1/2}$$

$$(C) = \int_{-\infty}^{\infty} \psi^2 h'(\psi)^2\,d\psi = \frac{\pi^{1/2}}{2}$$

$$(D) = \int_{-\infty}^{\infty} h'(\psi)^2\,d\psi = \pi^{1/2}$$

ie $I_{\Omega T} = \frac{(2\pi)^{3/2}}{12} + (6\pi)^{1/2}$

$\therefore\ E[\Phi^2(X_1)] = \frac{4\sigma^2 k^2}{(2\pi)^{3/2}} I_{\Omega T}$

$= \frac{4\sigma^2 k^2}{(2\pi)^{3/2}} \left[\frac{(2\pi)^{3/2}}{12} + (6\pi)^{1/2} \right]$

y como $k = \sqrt{\pi}/2$, entonces:

$$E[\Phi^{(1)2}(X_1)] = \frac{\pi\sigma^2}{(2\pi)^{3/2}} \left[\frac{(2\pi)^{3/2}}{12} + (6\pi)^{1/2} \right]$$

$$= \left(\frac{\pi}{12} + \frac{\sqrt{3}}{2} \right)\sigma^2$$

$$\simeq (1.1278247416)\sigma^2 \quad \blacksquare$$

2) $E[\Phi_2^{(1)2}(X_1, X_2)] = \frac{\pi}{2}\sigma^2$

$\Phi_2^{(1)}(x_1, x_2) = k|x_1 - x_2|$

$\therefore\ E[\Phi_2^{(1)2}(X_1, X_2)] = E[k^2(X_1 - X_2)^2]$

$= k^2 E[(X_1 - X_2)^2]$

Ahora como $X_1, X_2 \sim N(\mu, \sigma^2)$

$X_1 - X_2 \sim N(0, 2\sigma^2)$

$\therefore\ E(X_1 - X_2) = 0$ y $Var(X_1 - X_2) = 2\sigma^2$

ie $2\sigma^2 = E((X_1 - X_2)^2) + E^2(X_1 - X_2)$

$\therefore\ E((X_1 - X_2)^2) = 2\sigma^2$

$\therefore\ E[\Phi_2^{(1)2}(X_1, X_2)] = k^2 \cdot 2\sigma^2$

$= \frac{\pi}{2}\sigma^2 \quad \blacksquare$

3) $\zeta_1^{(1,2)} = \frac{1}{4}\sigma^3$

Por definición $\zeta_1^{(1,2)} = E\left[\psi_1^{(1)}(X_1)\psi_1^{(2)}(X_1)\right]$

donde $\psi_1^{(1)}(x_1) = E\left[\Phi^{(1)}(x_1,X_2) - \sigma\right]$

y $\psi_1^{(2)}(x_1) = E\left[\Phi^{(2)}(x_1,X_2) - \sigma^2\right]$

Ahora $\Phi^{(1)}(x_1,X_2) = k|x_1 - X_2|$

$$\therefore \psi_1^{(1)}(x_1) = \int_{-\infty}^{\infty} k|x_1 - x_2| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_2-\mu}{\sigma}\right)^2} dx_2 - \sigma$$

De igual forma: $\Phi^{(2)}(x_1,X_2) = \frac{(x_1 - X_2)^2}{2}$

$$\therefore \psi_1^{(2)}(x_1) = \frac{1}{2} E\left[x_1^2 - 2x_1X_2 + X_2^2\right] - \sigma^2$$

$$= \frac{1}{2}\left[(x_1-\mu)^2 + \sigma^2\right] - \sigma^2$$

De donde:

$$\zeta_1^{(1,2)} = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} k|x_1 - x_2| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_2-\mu}{\sigma}\right)^2} dx_2\right)$$

$$\times \frac{1}{2}\left[(x_1-\mu)^2 + \sigma^2\right]\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2} dx_1 - \sigma^3$$

Pero por la sección 1) de este apéndice

$$\int_{-\infty}^{\infty} k|x_1 - x_2| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_2-\mu}{\sigma}\right)^2} dx_2 =$$

$$\frac{2\sigma k}{\sqrt{2\pi}}\left[\psi\int_0^{\psi} e^{-\frac{1}{2}\tau^2} d\tau + e^{-\frac{1}{2}\psi^2}\right]$$

$$\therefore \ \zeta_1^{(1,2)} = \frac{\sigma^3 k}{2\pi}\left[\frac{\pi^{1/2}}{2} + 2\pi^{1/2} + \pi^{1/2} + \frac{\pi^{1/2}}{2} + \pi^{1/2}\right] - \sigma^3$$

$$= \frac{\sigma^3 k}{2\pi} \cdot 5\pi^{1/2} - \sigma^3 \qquad \text{como } k = \frac{\pi^{1/2}}{2}$$

$$= \frac{5}{4}\sigma^3 - \sigma^3 = \frac{1}{4}\sigma^3 \quad \blacksquare$$

4) $\zeta_1^{(2,2)} = \frac{1}{2}\sigma^4$

Por definición $\zeta_1^{(2,2)} = E\left[\varphi_1^{(2)}(x_1)\right]^2 - (\sigma^2)^2$

donde $\varphi_1^{(2)}(x_1) = E\left[\varphi^{(2)}(x_1, x_2)\right]$

$$= E\left[\frac{(x_1 - x_2)^2}{2}\right]$$

Ahora $E\left[\frac{1}{2}(x_1 - x_2)^2\right] = E\left[\frac{1}{2}(x_1^2 - 2x_1x_2 + x_2^2)\right]$

$$= \frac{1}{2}\left(x_1^2 - 2x_1\mu + \mu^2 + \sigma^2\right)$$

$$= \frac{1}{2}\left((x_1 - \mu)^2 + \sigma^2\right)$$

$$= \frac{1}{2}\left(\sigma^2\left(\frac{x_1 - \mu}{\sigma}\right)^2 + \sigma^2\right)$$

$$= \frac{\sigma^2}{2}\left(\left(\frac{x_1 - \mu}{\sigma}\right)^2 + 1\right)$$

$$\therefore \ E\left[\varphi_1^{(2)}(x_1)\right]^2 = E\left[\frac{\sigma^2}{2}\left(\left(\frac{x_1 - \mu}{\sigma}\right)^2 + 1\right)\right]^2$$

$$= \frac{\sigma^4}{4} E\left[\left(\frac{x_1 - \mu}{\sigma}\right)^2 + 1\right]^2$$

Como $X_1 \sim N(\mu, \sigma^2)$ entonces $\frac{X_1 - \mu}{\sigma} \sim N(0,1)$.

y $\therefore \left(\frac{X_1 - \mu}{\sigma}\right)^2 = Z^2 \sim \chi^2(1)$

$$\therefore \quad E[Z^2] = 1 \quad y \quad Var(Z^2) = 2$$

$$\therefore \quad E[(Z^2)^2] = Var(Z^2) + E^2(Z^2)$$

$$= 2+1 = 3$$

ie $$E\left[\left(\frac{X_1-\mu}{\sigma}\right)^2\right]^2 = 3$$

$$\therefore \quad E\left[\varphi_1^{(2)}(X_1)\right]^2 = \frac{\sigma^4}{4} E\left[(Z^2+1)^2\right]$$

$$= \frac{\sigma^4}{4} E\left[(Z^2)^2 + 2Z^2 + 1\right]$$

$$= \frac{\sigma^4}{4}\left[3+2+1\right] = \frac{3}{2}\sigma^4$$

$$\therefore \quad \zeta_1^{(2,2)} = \frac{3}{2}\sigma^4 - \sigma^4 = \frac{1}{2}\sigma^4 \quad \blacksquare$$

5) $\zeta_2^{(2,2)} = 2\sigma^4$

Por definición $\zeta_2^{(2,2)} = E\left[\varphi_2^{(2)}(X_1,X_2)\right]^2 - (\sigma^2)^2$

donde $\varphi_2^{(2)}(x_1,x_2) = \frac{(x_1-x_2)^2}{2}$

$$\therefore \quad E\left[\varphi_2^{(2)}(X_1,X_2)\right]^2 = E\left[\frac{(X_1-X_2)^4}{4}\right]$$

Dado que $X_1, X_2 \sim N(\mu,\sigma^2)$

$X_1 - X_2 \sim N(0, 2\sigma^2)$

$$\therefore \quad \frac{X_1-X_2}{\sqrt{2}\,\sigma} \sim N(0,1)$$

$$\therefore \quad \left(\frac{X_1-X_2}{\sqrt{2}\,\sigma}\right)^2 = \frac{(X_1-X_2)^2}{2\sigma^2} \sim \chi^2_{(1)}$$

$$\therefore E\left[\frac{(X_1-X_2)^2}{2\sigma^2}\right]=1 \quad y \quad Var\left[\frac{(X_1-X_2)^2}{2\sigma^2}\right]=2$$

$$\therefore E\left[\frac{(X_1-X_2)^2}{2\sigma^2}\right]^2 = 2+1=3$$

i.e.

$$E\left[\frac{(X_1-X_2)^4}{4}\right] = 3\sigma^4$$

$$\therefore \quad \zeta_2^{(2,2)} = 3\sigma^4 - \sigma^4 = 2\sigma^4 \quad \blacksquare$$

6) $$\sqrt{n}\left(W_U^{(n)}-1\right) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,\lambda^2)$$

donde $$\lambda^2 = 16\left(\frac{\pi}{12}+\frac{\sqrt{3}}{2}-1\right)-2$$

$$\simeq .045196669$$

dem.

Se aplica el teorema 7.5 del apéndice anterior, tomando: $g=2$

$$U(1)'=U(1)=\hat{\sigma}_U^{(n)} \qquad m(1)=2$$
$$U(2)'=U(2)=S^2 \qquad m(2)=2$$

$$y \quad h(y_1,y_2)=\frac{y_1^2}{y_2} \qquad \text{Veamos:}$$

Como $h_{y_1}=\frac{2y_1}{y_2}$, $h_{y_2}=\frac{-y_1^2}{y_2^2}$, $h_{y_1y_2}=-\frac{2y_1}{y_2^2}$,

$h_{y_2y_1}=\frac{-2y_1}{y_2^2}$, $h_{y_1y_1}=\frac{2}{y_2}$, $h_{y_2y_2}=\frac{2y_1^2}{y_2^3}$

y h son continuas en $\underset{\sim}{\theta}=(\sigma,\sigma^2)$, la variable aleatoria:

$\sqrt{n}\left(h(\hat{\sigma}_U^{(n)},S^2)-h(\sigma,\sigma^2)\right)$ tiene por distribución asintótica a una $N(0,\lambda^2)$, donde λ^2 es igual a:

$$\lambda^2 = m(1)^2 h_{y_1}^2(0,\sigma^2)\, \xi_1^{(1,1)} +$$
$$2\, m(1)\, m(2)\, h_{y_1}(0,\sigma^2)\, h_{y_2}(0,\sigma^2)\, \xi_1^{(1,2)}$$
$$+ (m(2))^2\, h_{y_2}^2(0,\sigma^2)\, \xi_1^{(2,2)}$$

Ahora, por las secciones anteriores de este apéndice

$$\xi_1^{(1,1)} = \left(\frac{\pi}{12} + \frac{\sqrt{3}}{2} - 1\right)\sigma^2$$

$$\xi_1^{(1,2)} = \frac{1}{4}\sigma^3 \quad y$$

$$\xi_1^{(2,2)} = \frac{1}{2}\sigma^4$$

$$\therefore\ \lambda^2 = 4\left(\frac{2}{\sigma}\right)^2\left(\frac{\pi}{12} + \frac{\sqrt{3}}{2} - 1\right)\sigma^2$$
$$+ 8\left(\frac{2}{\sigma}\right)\left(\frac{-1}{\sigma^2}\right)\frac{1}{4}\sigma^3$$
$$+ 4\left(\frac{-1}{\sigma^2}\right)^2\frac{1}{2}\sigma^4$$
$$= 16\left(\frac{\pi}{12} + \frac{\sqrt{3}}{2} - 1\right) - 4 + 2$$
$$= 16\left(\frac{\pi}{12} + \frac{\sqrt{3}}{2} - 1\right) - 2$$

ya que $\quad h_{y_1}(0,\sigma^2) = \dfrac{2\sigma}{\sigma^2} = \dfrac{2}{\sigma}$

y $\quad h_{y_2}(0,\sigma^2) = \dfrac{-\sigma^2}{(\sigma^2)^2} = -\dfrac{1}{\sigma^2}$

■

REFERENCIAS

[1].- SHAPIRO S.S. and WILK M.B (1965)
"An analysis of variance test for normality (complete samples)".
BIOMETRIKA 52, 591-611.

[2].- HOEFFDING W. (1942)
"A class of statistics with asymptotically normal distribution".
THE ANNALS OF MATHEMATICAL STAISTICS
19, 293-325.

[3].- DAVID H.A. (1970)
"Order statistics".
Wiley, New York.

[4].- KENDALL M. and STUART A. (1979)
"The advanced theory of statistics"
Volumen 2. "Inference and relationship"
Fourth Edition.
Griffin, London.

[5].- LOÉVE M. (1955)
"Probability theory"
Volumen 2.
Fourth Edition
Springer Verlarg, New York.