UNA RECOMENDACIÓN PARA LA DEMOSTRACIÓN DE LA INDEPENDENCIA DE LA MEDIA Y LA VARIANZA MUESTRALES.

E. Menéndez[1] -Dpto. Matemática Aplicada. Facultad de Matemática y Computación. Universidad de La Habana.Cuba,
J. Villagómez[2] y M. A. Herrera[3] -Unidad Académica Facultad de Matemática. Sede Acapulco. Universidad Autónoma de Guerrero. México

RESUMEN

Con este trabajo se desea divulgar una demostración de la independencia entre la media y la varianza cuando son calculadas a partir de una muestra aleatoria de la distribución Normal. Esta demostración no necesita de una avanzada herramienta matemática que en la mayoría de los casos el estudiante no posee cuando se estudia este aspecto. Esta demostración está basada en el uso del álgebra matricial, en particular en las transformaciones ortogonales (Kalbfleisch, 1985). Otros resultados que se obtienen con este enfoque son las distribuciones de la media muestral, la varianza muestral y la de una combinación lineal de las variables aleatorias.

Palabras claves: distribución t-Student; independencia; media muestral; varianza muestral;

ABSTRACT

With this work it is wanted to disclose a demonstration of the independence between the mean and the variance when they are calculated from a random sample of the Normal distribution. This demonstration does not need of an advanced mathematical tool that in most of the cases the student does not possess when this aspect is studied. This demonstration is based on the use of the matrix algebra, in particular the orthogonal transformations (Kalbfleisch, 1985). Other results that are obtained with this focus are the probability distributions of: the sample mean, the sample variance, and a lineal combination of the random variables.

Keywords: independence; sample mean; sample variance; t-Student distribution

1.- INTRODUCCIÓN.

Sea $X_1, X_2, \ldots, X_n$ un conjunto de n variables aleatorias independientes e idénticamente distribuidas (IID) según la distribución Normal con media μ y varianza σ^2. Si se representa por $\bar{X}$ y S^2 la media y la varianza muestrales respectivamente, la función aleatoria

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \qquad (1.1)$$

es ampliamente usada como "pivote" en la determinación de un intervalo de confianza para μ o como "estadístico de prueba" en una prueba de hipótesis sobre μ, en ambos casos con el supuesto de que la varianza σ^2 es desconocida.

En cualquiera de los dos procesos de la inferencia estadística citados antes se utiliza el hecho de que la función aleatoria T definida en (1.1) posee una distribución t-Student con n-1 grados de libertad. ["Student" (W. S. Gosset), 1908], [Casella y Berger, 2002]. En aras de abreviar este hecho se utilizará la

E-mail: [1]ema@matcom.uh.cu
[2]juanvillagomez2006@yahoo.com
[3] herrerapolo@hotmail.com

notación T~ $t^{(n-1)}$. Este resultado es considerado sin mucho más detalle en cursos elementales de Probabilidad y/o de Estadística para estudiantes universitarios. En cursos menos elementales se suele sistematizar el estudio de la distribución Normal y aquellas que de una manera más o menos directa se obtienen a partir de la distribución Normal, como son las distribuciones Chi-Cuadrado χ^2 y la t-Student. Incluso en una carrera de Matemática y alguna otra con un nivel matemático más elevado, donde se incluya el estudio más riguroso de la teoría de las Probabilidades y la Estadística, suele ocurrir que en el momento de abordarse este aspecto el estudiante no posea todas las herramientas matemáticas que permita el estudio más profundo de estas distribuciones de probabilidad; en particular la demostración de la independencia entre la media y la varianza muestrales, lo cual se necesita en la demostración de que la distribución de probabilidad de la función aleatoria T dada en (1.1) es la t-Student con n-1 grados de libertad.

2.- LA DISTRIBUCIÓN T-STUDENT CON ν GRADOS DE LIBERTAD.

Aún cuando se suele definir la distribución t-Student y todas las demás mediante la función de densidad correspondiente, para fines prácticos resulta más interesante destacar cómo se obtiene, cuando esto sea posible, cada distribución. En el caso de la distribución t-Student, la misma se puede obtener como la distribución del cociente de las variables aleatorias X e Y, siempre que éstas sean independientes y con distribuciones Normal estándar y Chi-Cuadrado con ν grados de libertad respectivamente, mediante la siguiente relación. ["Student" (W. S. Gosset), 1908]

$$\frac{X}{\sqrt{\frac{Y}{\nu}}} \quad . \tag{2.1}$$

Como quiera que la expresión (1.1) mediante transformaciones muy elementales puede escribirse como

$$T = \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} = \frac{\frac{\frac{\bar{X}-\mu}{\sigma}}{\sqrt{n}}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2}} \quad , \tag{2.2}$$

para la demostración de que T ~ $t^{(n-1)}$ bastará demostrar a partir de (2.1) que el numerador de (2.2) se distribuye Normal estándar y que $\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2_{n-1}$, siendo ambos términos independientes.

Como se aprecia de lo expresado anteriormente, el establecimiento de cómo surge una distribución t-Student, requiere de otra distribución adicional a la Normal, siendo ésta la Chi-Cuadrado.

Una variable con distribución Chi-Cuadrado con n grados de libertad se obtiene cuando se suman los cuadrados de n variables aleatorias IID según la Normal estándar. En resumen, Y ~ χ^2_n si

$$Y = \sum_{i=1}^{n} Z_i^2 \, , \tag{2.3}$$

donde $Z_1, Z_2, \ldots Z_n$ son IID~ N (0,1). [Hogg y Tanis, 2006, pág. 282]

3.- DEMOSTRACIÓN DE QUE LA FUNCIÓN ALEATORIA T POSEE UNA DISTRIBUCIÓN T-STUDENT.

Dado que en la demostración que se pretende se utiliza el concepto de matriz ortogonal, es conveniente que se retome éste, así como algunas consecuencias que se derivan del mismo en caso de que se hayan estudiado con anterioridad en un curso de Álgebra Matricial. En caso contrario, éstos bien pudieran incluirse, dejándose como ejercicio la demostración de algunas de las consecuencias mencionadas antes. [Kurosh,1994]

a) La matriz cuadra C de orden n es ortogonal si CC' =C'C = **I**, donde **I** es la matriz identidad de orden n. De esta definición se obtiene inmediatamente que si C es ortogonal también lo es C'. C' denota la transpuesta de C.
b) Si la matriz C es ortogonal, entonces su determinante es ± 1.
La demostración de este resultado es muy sencilla.
1 = det(**I**) = det(C'C) = det(C').det(C) = (det(C))^2, de lo cual det(C) = ± 1
c) El resultado que ahora se expone no posee una demostración tan sencilla como las anteriores, por lo que en caso de no conocerse por los estudiantes debe asumirse su cumplimiento. En primer lugar debe destacarse que un vector está normalizado cuando su longitud, distancia al origen o producto escalar por si mismo es la unidad. Sean q vectores ortonormales de tamaño n , digamos U_1, U_2,...,U_q (1≤ q < n), entonces existe una matriz ortogonal C cuyas primera q filas son los vectores U_1, U_2,...,U_q y por tanto la matriz ortogonal C' tiene a estos mismos vectores como sus primeras q columnas. El método de ortogonalización de Gram-Schmidt permite realizar este proceso. [Kurosh,1994, Bru et al., 2004, Rao, 2002]
d) Si Y es un vector de dimensión n y C una matriz ortogonal de tamaño n, entonces el producto escalar de Y por si misma es invariante por una transformación ortogonal de éste.
La demostración consiste en definir U = CY. Entonces U'U= Y'C'CY, pero como C es ortogonal C'C = **I** y entonces U'U = Y'**I**Y = Y'Y.
Este resultado también puede expresarse como que la suma de los cuadrados de los componentes de un vector es invariante por una transformación ortogonal de éste. Este enunciado tiene un significado más vinculado con las Probabilidades y la Estadística.
e) Siempre que se produzca una transformación ortogonal de un vector, digamos U = CY, el Jacobiano de la transformación, entendido éste como del determinante de la matriz cuyos elementos $c_{i\,j}$ están dados por $\dfrac{\partial U_i}{\partial Y_j}$, para i,j = 1, 2,...,n, resulta ser el determinante de C, que como ya se señaló en (b) es ± 1. Y por tanto $|J| = 1$.

Con estos resultados establecidos puede ahora pasarse a la presentación de la demostración del aspecto sustantivo de este trabajo, mismo que puede consultarse en Kalbfleisch (1985).

TEOREMA 1. Si las variables aleatorias U_1, U_2,...,U_n son IID ~ N(0 , 1), entonces $Z_1, Z_2, \ldots Z_n$ son también IID ~ N(0 , 1), donde Z = ($Z_1, Z_2, \ldots Z_n$)' = CU y C es una matriz ortogonal.

DEMOSTRACIÓN.

Se basa en el hecho de que cuando se transforma una variable aleatoria continua, la densidad de la nueva variable aleatoria se obtiene mediante la densidad de la variable original evaluada en la nueva variable aleatoria multiplicada por el valor absoluto del Jacobiano de la transformación. En nuestro caso,

$$f_Z\left(z_1, \ldots, z_n\right) = f_U\left(z_1, \ldots, z_n\right)|J|.$$

Por (b) se tiene que $|J| = 1$ y además por (d) la suma de los cuadrados es invariante, entonces

$$f_Z\left(z_1, \ldots, z_n\right) = \left(\sqrt{2\pi}\right)^{-n} e^{-\frac{1}{2}\sum_{i=1}^{n} z_i^2},$$

de lo que resulta que las $Z_1, Z_2, \ldots Z_n$ son IID ~ N(0, 1).

TEOREMA 2. Sean $U_1, U_2,...,U_n$ son variables aleatorias IID ~ N(0 , 1) y $a_1,...,a_n$ constantes tales que $\sum_{i=1}^{n} a_i^2 = 1$. Definamos $Z_1 = \sum_{i=1}^{n} a_i U_i$ y $V = \sum_{i=1}^{n} U_i^2 - \left(\sum_{i=1}^{n} a_i U_i\right)^2$. Entonces Z_1 y V son independientes con distribuciones N(0, 1) y χ^2_{n-1} respectivamente.

DEMOSTRACIÓN.

Por (c) se puede asumir que existe una matriz C ortogonal, cuya primera fila es $a_1,...,a_n$, ya que $\sum_{i=1}^{n} a_i^2 = 1$. Sea entonces Z = CU y por teorema 1, $Z_1, Z_2, \ldots Z_n$ son IID ~ N(0, 1). Luego $Z_1 = \sum_{i=1}^{n} a_i U_i$ ~ N(0, 1). Pero además, ya que $\sum_{i=1}^{n} U_i^2 = \sum_{i=1}^{n} Z_i^2$ por (d) se tiene que

$$V = \sum_{i=1}^{n} U_i^2 - \left(\sum_{i=1}^{n} a_i U_i\right)^2 = \sum_{i=1}^{n} U_i^2 - z_1^2 = \sum_{i=2}^{n} Z_i^2$$

y por (2.3) V ~ χ^2_{n-1}. Finalmente como $Z_1, Z_2, \ldots Z_n$ son independientes V y Z_1 también lo son. Y el teorema queda demostrado.

TEOREMA 3. Sean $Y_1, \ldots, Y_n$ n variables aleatorias independientes con distribución de probabilidad $N\ \mu_i, \sigma_i^2$, para i = 1,2,...,n y $a_1, a_2, \ldots, a_n$ constantes cualesquiera. Entonces $\sum_{i=1}^{n} a_i Y_i$ se distribuye según una distribución Normal[4].

DEMOSTRACIÓN.

Sea

$$U_i = \frac{Y_i - \bar{Y}}{\sigma_i} \qquad (3.1)$$

la forma estándar de Y_i , de modo que por ser las Y_i 's independientes, también lo serán las $U_i \sim N(0,1)$.

Por otro lado, aplicando las propiedades del valor esperado y de varianza se obtiene que

$$E\left[\sum_{i=1}^{n} a_i Y_i\right] = \sum_{i=1}^{n} a_i \mu_i \quad \text{y} \quad Var\left[\sum_{i=1}^{n} a_i Y_i\right] = \sum_{i=1}^{n} a_i^2 \sigma_i^2 = \sigma^2.$$

Por lo tanto,

$$\frac{1}{\sigma}\left[\sum_{i=1}^{n} a_i Y_i - \sum_{i=1}^{n} a_i \mu_i\right] = \frac{1}{\sigma}\sum_{i=1}^{n} a_i \;\; Y_i - \mu_i \qquad (3.2)$$

y por (3.1)

$$\sigma_i U_i = \;\; Y_i - \mu_i \;\; . \qquad (3.3).$$

4 Es posible que este resultado ya el estudiante lo conozca. En este caso la demostración no tiene por qué incluirse.

Sustituyendo (3.3) en (3.2)

$$\frac{1}{\sigma}\left[\sum_{i=1}^{n} a_i Y_i - \sum_{i=1}^{n} a_i \mu_i\right] = \frac{1}{\sigma}\sum_{i=1}^{n} a_i \left(Y_i - \mu_i\right) = \sum_{i=1}^{n} a_i' U_i$$

donde $a_i' = \frac{a_i \sigma_i}{\sigma}$ y $\sum_{i=1}^{n}\left(a_i'\right)^2 = 1$. Pero por teorema 2, $\sum_{i=1}^{n} a_i' U_i \sim$ N(0, 1). Nótese que la expresión (3.2) de la cual ha surgido la variable $\sum_{i=1}^{n} a_i' U_i \sim$ N(0, 1) es la estandarización de la variable $\sum_{i=1}^{n} a_i Y_i$, por tanto esta última posee una distribución de probabilidad Normal, con media $\sum_{i=1}^{n} a_i \mu_i$ y varianza $\sum_{i=1}^{n} a_i^2 \sigma_i^2 = \sigma^2$.

El teorema que a continuación se enuncia completa la propuesta de este trabajo.

TEOREMA 4. Sean $Y_1, Y_2, \ldots, Y_n$ n variables aleatorias IID ~ $N\left(\mu, \sigma^2\right)$, entonces

$$\bar{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \ , \ \sum_{i=1}^{n} \frac{\left(Y_i - \bar{Y}\right)^2}{\sigma^2} \sim \chi^2_{n-1}$$

y son independientes.

DEMOSTRACIÓN.

Definamos $U_i = \left(Y_i - \mu\right)/\sigma$, de lo cual se obtiene que $Y_i = \sigma U_i + \mu$. Esta última expresión permite expresar la media de Y como

$$\bar{Y} = \sigma \bar{U} + \mu \qquad (3.4)$$

y la suma de cuadrados de las diferencias de cada Y_i respecto a su media adopta la forma

$$\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2 = \sigma^2 \sum_{i=1}^{n}\left(U_i - \bar{U}\right)^2 . \qquad (3.5)$$

Por ser las U_i las variables Y_i estandarizadas, las primeras son también independientes al serlo las Y_i . Sean $a_1 = a_2 = \cdots = a_n = \frac{1}{\sqrt{n}}$, de lo cual $\sum_{i=1}^{n} a_i^2 = 1$. Por teorema 2 se tiene que

$$\mathbf{Z_1 = \sum_{i=1}^{n} a_i U_i = \sum_{i=1}^{n} \frac{1}{\sqrt{n}} U_i = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} U_i = \frac{1}{\sqrt{n}} n\bar{U} = \sqrt{n}\bar{U}} \qquad \text{y}$$

$$\mathbf{V = \sum_{i=1}^{n} U_i^2 - Z_1^2 = \sum_{i=1}^{n} U_i^2 - n\bar{U}^2 = \sum_{i=1}^{n}\left(U_i - \bar{U}\right)^2} .$$

Por el propio teorema 2 se obtiene inmediatamente que $Z_1 \sim$ N(0,1), V ~ χ^2_{n-1} y que además estos son independientes.
Sustituyendo $\bar{U}$ en (3.4) se obtiene que

$$\bar{Y} = \sigma \bar{U} + \mu = \sigma \frac{Z_1}{\sqrt{n}} + \mu \ ,$$

quedando demostrado que la media muestral depende de $Z_1 = \frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$, siendo ésta la expresión estandarizada de $\bar{Y}$ y por tanto la distribución de la media muestral es $N\left(\mu, \frac{\sigma^2}{n}\right)$. Por otra parte $V = \sum_{i=1}^{n} (U_i - \bar{U})^2$ y por (3.5)

$$V = \sum_{i=1}^{n} (U_i - \bar{U})^2 = \frac{1}{\sigma^2} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 .$$

De donde se obtiene que por ser Z_1 independiente de V, también lo son $\bar{Y}$ y $\frac{1}{\sigma^2}\sum_{i=1}^{n} (Y_i - \bar{Y})^2$, distribuyéndose este último con una distribución Chi-Cuadrado con n-1 grados de libertad. Finalmente el teorema está demostrado.

CONCLUSIONES.

Se ha logrado exponer la demostración de la independencia de la media y la varianza muestrales bajo el supuesto de una muestra aleatoria de la distribución Normal y las correspondientes distribuciones de la media y la varianza muestrales. La herramienta matemática fundamental que se consideró ha sido la de una transformación ortogonal. En este propio trabajo se muestra un resumen de los resultados fundamentales asociados con las matrices ortogonales. Para facilitar la demostración de la independencia de los estadísticos antes señalados, así como sus respectivas distribuciones, se utilizaron tres teoremas.

BIBLIOGRAFÍA.

1. Bru, R., Clement, J. J, Mas, J. y Urbano, A. (2004): Álgebra Lineal. 1ª reimpresión 2ª edición. Alfaomega Grupo Editor, S. A. de C. V. México.
2. Casella, G. and Berger, R. (2002): Statistical Inference. 2nd edition. Duxbury. USA. Pág. 222.
3. Hogg, R. V. and Tanis, E. A. (2006): probability and statistical Inference. 7th ed. Pearson prentice Hall. USA.
4. Kalbfleisch, J. G. (1985): Probability and Statistical Inference: Probability, Vol. 1. Springer. USA.
5. Kurosh, A. G. (1994): Curso de Álgebra Superior. 5ª reimpresión. Editorial Limusa S.A. de C. V. Grupo Noriega Editores. México.
6. Lang, S. (2004): Linear Álgebra. Springer. 3rd edition. USA.
7. Rao, C. R. (2002): Linear Statistical Inference and Its Applications. 2nd edition. John Wiley. USA.
8. “Student” (W. S. Gosset) (1908): The probable error of a mean. Biometrika 6, pág. 1-25.