



Técnicas estadísticas comparativas con mayor potencial de aplicación en ingeniería

Carlos Alberto Contreras Verteramo^{1*}, Jesús Gómez Castellanos¹, Manuel Antonio Arenas Méndez¹

¹Instituto Tecnológico Superior de Pánuco

*contrerc@itspanuco.edu.mx

En el presente trabajo se analizan las características de algunas técnicas estadísticas de carácter comparativo con mayor potencial de aplicación en el área de ingeniería.

Determinación del tamaño de muestra. Al planear un trabajo de investigación, irremediablemente se presentará la pregunta ¿Cuántas son las muestras que debo coleccionar? La condición es buscar un tamaño de muestra lo más pequeño posible para que permita optimizar tiempo y recursos, pero a la vez lo suficientemente grande para que refleje a la población en estudio, es decir, la muestra debe ser representativa de la población.

Comúnmente los muestreos se realizan con la finalidad de estimar un promedio o media poblacional, para ello es necesario establecer la exactitud con que se desea estimar la media, este es un paso esencial para determinar el tamaño de la muestra (Eckblad 1991).

De forma general existen dos grupos de técnicas de muestreo, las aleatorias o probabilísticas y las no aleatorias, la diferencia entre los dos tipos radica en el uso o no del azar para la elección de las observaciones que formarán la muestra. En las muestras aleatorias todos los integrantes de la población tienen las mismas probabilidades de ser seleccionados de manera independiente, lo que significa que la selección de cualquier miembro de la población no debe influir en la selección de cualquier otro miembro, normalmente este tipo de muestras son las que mejor garantizan la representatividad de la muestra (Argibay 2009; Zar 2014).

Al realizar un muestreo aleatorio simple, se utiliza un procedimiento de generación de números al azar para elegir a los sujetos de la población que van a formar parte de la muestra. En un muestreo basado en vías de comunicación, las observaciones aleatorias se registran únicamente en áreas cercanas a carreteras y caminos. En el caso de un muestreo estratificado se divide la población en subpoblaciones o estratos y se selecciona de manera aleatoria la cantidad de observaciones a tomar de cada estrato de acuerdo con la proporción de la población representada por el estrato.

Estadística descriptiva. Los métodos de estadística descriptiva son utilizados para sintetizar datos y presentarlos de manera ordenada, de forma que puedan ser manejados con facilidad. Se emplean para este fin las medidas de tendencia central y de variabilidad o dispersión, organizadas en cuadros y figuras, estas técnicas constituyen la base para la realización de análisis estadísticos cuantitativos (Sokal y Rohlf 2009; Zar 2014).

Al emplear cada uno de los estadísticos descriptivos, la información de una gran cantidad de observaciones se resume en un sólo número que las representa a todas, aunque existe el riesgo de que cada vez que se trate de describir un gran número de datos mediante un sólo indicador, se puede distorsionar la información original o perder detalles importantes. Aún con estas limitaciones, los métodos estadísticos descriptivos son excelentes herramientas para resumir información y pueden permitir comparaciones entre grupos de datos (Trochim 2000).

En la práctica, es necesario asignar filtros a los datos, obtener el número de clases; de acuerdo con el número de clases obtenidas se conforma una tabla de frecuencias que posteriormente se representa gráficamente (Sturges 1926).

Las medidas de tendencia central más comunes son la media con sus variantes aritmética, geométrica y armónica, la mediana y la moda. Es posible obtener también la clase modal y la media ponderada. El cálculo de las medidas de dispersión incluye el rango, la desviación media, la varianza muestral, la desviación estándar muestral y los cuatro cuartiles.

Cuando se analizan dos grupos de datos es recomendable obtener también los valores para covarianza y la correlación.

Mediante este proceso exploratorio el investigador puede visualizar la estructura de los datos, detectar los que pueden presentar imprecisiones u outliers; también le permite resumirlos en

unos cuantos números que caracterizan la muestra de manera eficiente, con mínima pérdida de información.

Es necesario mencionar que los métodos de estadística descriptiva son sólo eso, descriptivos, en ningún caso son válidos para establecer generalizaciones que vayan más allá de los datos incluidos en la o las muestras.

La función de la estadística descriptiva es presentar descripciones cuantitativas de grupos de datos, que pueden ser muy grandes, en forma comprensible y manejable ya que cada estadístico descriptivo resume los datos a una forma más simple. La descripción obtenida a partir del análisis estadístico descriptivo permite comparar los resultados con los de otros trabajos de investigación realizados en condiciones o con métodos similares.

Prueba t. Un procedimiento estadístico inferencial utilizado para determinar si existen diferencias significativas entre las medias de dos muestras es la prueba t de Student. Como en otros métodos estadísticos paramétricos, se asume que la variable dependiente tiene una distribución normal, los datos deben estar en escala de intervalo o de razón y es necesario especificar el nivel de significancia (alfa) (Zar 2014).

El estadístico de prueba para la t de Student es el valor t. Teóricamente este valor representa el número de unidades de desviación estándar que separan las medias de los grupos. Al aplicar una prueba t, el investigador puede establecer, con cierto nivel de confianza que la diferencia observada entre los promedios de los grupos es demasiado grande para deberse al azar. Si el resultado de la prueba produce un valor t con probabilidad de 0.01, es posible afirmar que la probabilidad de encontrar diferencia por causas aleatorias entre los grupos analizados es de 1 cada 100 veces (Sokal y Rohlf 2009; Zar 2014).

Existen cinco factores que podrían indicar si la diferencia entre las medias de dos grupos puede ser considerada significativa estadísticamente. En primer lugar, es posible mencionar que, a mayor diferencia entre medias, mayor probabilidad de encontrar diferencia significativa; el segundo factor es la cantidad de traslape, que se expresa como una función de la variación al interior de los grupos, a menor variación interna del grupo mayor será la probabilidad de encontrar diferencia significativa; el tercer factor es el tamaño de la muestra, al aumentar el número de muestra las medias de los grupos tenderán a ser más estables y representativas; el cuarto factor es el nivel de alfa, mayores valores de alfa requerirán menor diferencia entre

las medias de los grupos para que se pueda reportar como significativa y; como quinto factor, se debe plantear una hipótesis sin dirección, es decir, de dos colas (Colquhoun 1971).

La aplicación de la prueba t requiere que se cumplan los siguientes supuestos: las muestras deben haber sido seleccionadas de forma aleatoria a partir de sus respectivas poblaciones; las poblaciones deben tener distribución normal, ser unimodales, simétricas y asintóticas; las muestras deben presentar igualdad de varianzas.

El procedimiento para realizar una prueba t de Student es calcular la diferencia entre las medias de los dos grupos y dividirla por el error estándar de la diferencia, es decir la desviación de estándar de la distribución de las diferencias.

El intervalo de confianza para una prueba t de dos colas se calcula al multiplicar los valores críticos por el error de estándar, posteriormente se suma y resta de la diferencia de las medias.

Análisis de varianza. La prueba de ANOVA o análisis de varianza es un método comúnmente utilizado para determinar si existe diferencia estadística entre dos o más grupos de datos. Mediante esta prueba se puede establecer si las diferencias se deben efectivamente a diferencias entre las medias de las poblaciones o son producto de variaciones en el muestreo. El análisis de varianza considera tres supuestos básicos que deben cumplirse para aplicar la prueba: 1) las muestras deben ser aleatorias e independientes; 2) deben provenir de poblaciones que se ajustan a la distribución normal y 3) es necesario que las varianzas de las poblaciones sean iguales. Afortunadamente el ANOVA es una prueba robusta que opera de manera correcta aún con una considerable heterogeneidad de varianza, mientras los números de datos por grupo sean iguales o muy similares (Colquhoun 1971).

El análisis de varianza de una vía es un método para analizar la variación entre muestras y la variación al interior de las muestras. Se le llama de una vía porque analiza una sola variable o factor independiente. Es considerado un método estadístico muy útil para llevar a cabo comparaciones entre dos o más medias poblacionales (Hernández et al. 1996; Zar 2014).

Cuando el estadístico de prueba (F) indica que se debe aceptar la hipótesis nula (H_0), se concluye que, la diferencia entre las medias de las muestras ocurre debido a variaciones en el muestreo. Por el contrario, cuando se rechaza la hipótesis nula, se asume que la diferencia entre las medias no se debe únicamente al azar y por lo tanto las muestras, o al menos algunas de ellas, provienen de distintas poblaciones (Sokal y Rohlf 2009).

En el caso del ANOVA de dos vías, las fórmulas son parecidas a las utilizadas en el de una vía, pero se integra el cálculo por renglones, adicional al de columnas, donde se presentan las posibles combinaciones de dos tratamientos o factores (Sokal y Rohlf 2009; Zar 2014).

Se considera que en general el ANOVA es una prueba útil y confiable, únicamente en caso de observar severas desviaciones de los supuestos sería recomendable utilizar análisis alternativos de tipo no paramétrico como la prueba de Kruskal-Wallis.

El análisis de varianza es una prueba que permite analizar dos o más muestras de manera simultánea, permitiendo mayor estabilidad que pruebas que parcializan el análisis de los datos como la t de Student o la prueba z.

El análisis de varianza permite optimizar el esfuerzo, permitiendo al investigador determinar si existen o no diferencias significativas entre pares de medias, evitándole realizar más comparaciones en caso de no existir, de cualquier manera no sería válido comparar todas las medias de dos en dos, en un mismo planteamiento, ya que se incrementa la probabilidad de que algunas diferencias resulten significativas por azar, presentándose la posibilidad de rechazar la hipótesis nula cuando realmente no hay diferencias.

Entre las limitaciones podemos mencionar que el ANOVA solo establecerá la existencia de una diferencia significativa entre las medias de al menos dos grupos, pero no explicará cual es el par que difiere en sus medias. Si existe un requisito de datos detallados, el despliegue de los procesos estadísticos de seguimiento adicionales ayudará a descubrir qué grupos difieren en el valor medio. Normalmente, ANOVA se utiliza en combinación con otros métodos estadísticos para realizar las comparaciones múltiples.

El ANOVA también presupone que el conjunto de datos se distribuye uniformemente, ya que solo compara medias. Si los datos no se distribuyen a lo largo de una curva normal y existen valores atípicos, esta prueba no será el procedimiento adecuado para interpretar los datos.

De manera similar, asume que las desviaciones estándar son iguales o similares entre los grupos. Si existe una gran diferencia en las desviaciones estándar, la conclusión de la prueba podría ser inexacta.

Pruebas no paramétricas. Cuando se analizan los datos recabados en una investigación, es de vital importancia seleccionar un método de análisis adecuado que permita llegar a

conclusiones correctas. En ocasiones los datos no se apegan a una distribución normal o bien, sólo se cuenta con un tamaño de muestra pequeño, condiciones que no permiten emplear técnicas estadísticas paramétricas. En estos casos se puede recurrir a otros métodos, llamados no paramétricos, fáciles de implementar y que funcionan bien, incluso con reducidos tamaños de muestra (Hollander et al., 2014).

Comparación de dos muestras. La prueba U de Mann-Whitney es un procedimiento estadístico no paramétrico para comparar dos muestras independientes cuyas variables se encuentran en escala ordinal, pero no se apegan a una distribución normal.

La prueba de Wilcoxon emplea rangos, se utiliza para comparar la mediana de dos muestras relacionadas o pareadas y determinar si se presentan diferencias entre ellas. Para aplicarla no es necesario que los datos se apeguen a una distribución, la variable dependiente debe ser ordinal.

Tanto la prueba de U como la de Wilcoxon son buenas opciones cuando no es posible utilizar la prueba t de Student (Altman 1991; Corder y Foreman 2009).

La prueba de Kolmogorov-Smirnov se utiliza para comprobar si la muestra se ajusta a una distribución determinada o para comparar la distribución observada con la esperada, emplea las frecuencias relativas acumuladas de las dos distribuciones. Es necesario que los datos se presenten en escala ordinal (Corder y Foreman 2009).

La prueba del signo es utilizada para probar hipótesis sobre de la mediana de una población. Se basa en la distribución Binomial con probabilidad de éxito $p=1/2$, debido a que la probabilidad de que una observación sea mayor o menor que la mediana es 0.5. Su cálculo se realiza determinando las diferencias de las observaciones respecto a la mediana, contándose los signos positivos y negativos. La principal limitación de esta prueba es que no toma en cuenta la magnitud de las observaciones (Rodríguez 2008).

Comparación de más de dos muestras. La prueba de Kruskal-Wallis es un método no paramétrico que sirve para determinar si varios grupos de datos provienen de la misma población. Es muy similar al ANOVA, pero en este caso la normalidad no es un supuesto y los datos son reemplazados por categorías; puede considerarse análoga a una prueba de Mann-Whitney, para 3 o más grupos de datos (Corder y Foreman 2009; Zar 1984).

Una prueba no paramétrica similar al ANOVA de medidas repetidas es la prueba de Friedman, se basa en el ordenamiento de los datos y asignación de un rango. Para esta prueba los datos no necesitan una distribución específica y la variable dependiente debe ser ordinal. Se utiliza para comparar más de dos rangos (medianas) y determinar si las diferencias se deben o no al azar (Altman 1991; Corder y Foreman 2009).

Las pruebas no paramétricas no requieren la distribución de la población caracterizada en función de los parámetros μ y σ , por tanto, son muy útiles cuando los datos se desvían de la normalidad y también cuando el número de datos es pequeño. Sin embargo, este tipo de pruebas no está libre de todo supuesto acerca de los datos, al menos requieren que las observaciones hayan sido colectadas de manera aleatoria e independiente.

Literatura citada.

- Argibay, J. C. 2009. Muestra en investigación cuantitativa. Subjetividad y procesos cognitivos. 13(1):13-29.
- Eckblad, J. W. 1991. How Many Samples Should Be Taken? *BioScience*. 41(5):346-348.
- Millspaugh, J. J. and B. E. Washburn. 2003. Within-sample variation of fecal glucocorticoid measurements. *General and Comparative Endocrinology*. 132:21–26.
- Sturges, H. 1926. The choice of a class-interval. *J. Amer. Statist. Assoc.* 21:65–66.
- Trochim, W. 2000. *The Research Methods Knowledge Base*, 2nd Edition. Atomic Dog Publishing, Cincinnati, Ohio, U.S.A.
- Colquhoun, D. 1971. *Lectures on Biostatistics, An introduction to statistics with applications in Biology and Medicine*. Clarendon press, Oxford, Great Britain. 425 p.
- Hernández, J. A., A. Borges y G. Ramírez. 1996. Análisis de la robustez del ANOVA en el caso de tamaños muestrales distintos y no proporcionales frente a los contrastes de aleatorización. *Psicológica*, (17):127-141.
- Sokal, R. R. and F. J. Rohlf. 2009. *Introduction to biostatistics*. Dover Edition. Dover publications, Inc. Mineola, New York, U.S.A.
- Altman D. G. 1991. *Practical statistics for medical research*. Chapman and Hall/CRC. 623p.
- Corder, G. and D. Foreman. 2009. *Nonparametric statistics for non-statisticians*. Wiley. 247p.
- Hollander, M., D. Wolfe and E. Chicken. 2014. *Nonparametric statistical methods*. Wiley. 819p.
- Rodríguez, F. 2008. *Estudio de métodos no paramétricos*. Informe de Pasantías. U.N.A. Caracas, Venezuela.

Zar, J. 2014. Biostatistical Analysis. 5a Ed. Pearson Education Limi