

## LA ESTADÍSTICA EN LA MEJORA DE LA CALIDAD DE LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR

**Jaime Dionísio Cuevas Domínguez**

*jaicueva@uqroo.mx*

División de Ciencias e Ingeniería, Universidad de Quintana Roo  
Boulevard Bahía s/n, esq. Ignacio Comonfort, Col. del Bosque,  
Chetumal, Quintana Roo, C.P. 77019

### RESUMEN

Se presenta una aplicación en la Universidad de Quintana Roo del análisis de *clusters*, con el propósito de motivar el uso del análisis estadístico en los sistemas de mediciones organizacionales.

Palabras clave: Calidad en educación superior—análisis multivariado—análisis *cluster*.

### INTRODUCCIÓN

Como resultado de la implementación de las políticas nacionales de educación superior, las universidades mexicanas realizan esfuerzos para mejorar la calidad de sus instituciones y de sus servicios basados en métodos sistémicos como la planeación estratégica, las evaluaciones institucionales con fines de diagnóstico o acreditación y la implementación de sistemas de gestión de calidad bajo las normas de la ISO 9000.

El éxito de estos métodos en las diferentes etapas de planeación, implementación, seguimiento y mejora, dependen de múltiples factores; pero sin

duda, en gran medida de la implementación de un sistema de indicadores o medidas organizacionales que reflejen en todos los niveles el quehacer académico, como fundamento para la toma de decisiones oportunas por las autoridades y los órganos colegiados.

La Subsecretaría de Educación Superior (SESIC), plantea a las Instituciones de Educación Superior (IES) un grupo de indicadores académicos, que pueden y deben servir de referencia o entrada para satisfacer los requerimientos de información de la SESIC. Sin embargo, es necesario diseñar, implementar, operar y mejorar un sistema de mediciones propio de cada IES, de forma que las necesida-

des de información de las diferentes partes interesadas se incorporen en una estructura coherente.

Es deseable que este sistema de mediciones organizacionales tenga diversas características, pero quizás las más importantes son la factibilidad de ser implementado y mantenido con el nivel de precisión apropiado, y sobre todo que sea útil a los usuarios de forma que propicien una acumulación de experiencias y generación de conocimiento de los complejos procesos educativos. Para lograr estas características en el sistema de mediciones es necesario establecer diferentes métodos o procedimientos; desde cómo identificar y validar las variables o indicadores de los conceptos que deseamos medir, cómo coleccionar los datos sin pérdida de precisión y oportunidad, cómo almacenarlos, cómo analizarlos y hasta cómo presentarlos a los usuarios de forma atractiva y clara, que favorezca la apropiación del sistema.

Los métodos estadísticos pueden ser factores clave de éxito en estas tareas, su amplia variedad permite elegir el método adecuado para revelar o confirmar relaciones y estructuras entre variables o explorar patrones de datos que permitan evaluar el desempeño organizacional y prever escenarios futuros.

Entre los métodos estadísticos, los multivariados tienen una creciente aceptación por la posibilidad de explorar muchas variables al mismo tiempo y la facilidad de disponer de paquetes estadísticos comerciales eficientes.

Este artículo pretende ilustrar una aplicación en la exploración de variables de los procesos educativos en la

educación superior mediante una técnica conocida como análisis de *clusters* (análisis de agrupamientos), con la finalidad de mostrar el potencial de los métodos estadísticos multivariados en los sistemas de mediciones organizacionales.

## **PLANTEAMIENTO DEL PROBLEMA**

Los enfoques modernos de calidad universitaria están basados en la satisfacción de los intereses de los distintos actores, como los inversionistas (representados en México por la SESIC), los estudiantes, los profesores, otros empleados y la sociedad.

Los inversionistas esperan una rentabilidad apropiada, expresada en indicadores de eficiencia y cobertura, sin embargo los intereses de los actores restantes como el caso de los estudiantes, suelen variar y evolucionar de forma diferente entre distintos grupos. El identificar oportunamente estas necesidades o requerimientos y agrupar a los individuos con intereses homogéneos puede ser la mejor alternativa para trazar estrategias o acciones tácticas.

El problema que se aborda en este trabajo es el de cómo determinar los intereses y expectativas de calidad de los estudiantes de una universidad, teniendo en cuenta que posiblemente esos intereses no sean compartidos.

Para su solución, se explica en qué consiste el análisis de *clusters*, apoyado en un ejemplo hipotético, y posteriormente se presenta su aplicación

en la Universidad de Quintana Roo, con el propósito de conocer en primer término cuáles son los atributos o variables de calidad que los estudiantes perciben como de mayor importancia, y posteriormente realizar un agrupamiento de los individuos, con las variables más representativas.

**ANÁLISIS DE CLUSTERS**

El análisis de *clusters* es la técnica adecuada para realizar agrupamientos de individuos o de variables, mediante una comparación de la similitud de la medición de las variables que representen los intereses de los individuos.

El análisis de *clusters* tiene decenas de versiones que se pueden usar para determinar los agrupamientos; las diferencias entre ellas consisten: en la forma de medir la similitud de los valores de las variables asignadas por los individuos, y los algoritmos utilizados para realizar los agrupamientos.

Para ilustrar de una manera sencilla cómo funciona el análisis de *clusters*, suponga que se mide la percepción de cuatro estudiantes (*A*, *B*, *C* y *D*) sobre la disponibilidad de los libros en la biblioteca (variable *X1*) y la calidad de los servicios de computo (*X2*), la percepción u opinión de los estudiantes puede tomar cinco valores codificados en una escala ordinal donde 1= Pésimo, 2= Malo, 3= Neutro, 4= Bueno, 5= Excelente. Las respuestas hipotéticas de los cuatro estudiantes se resumen en la Tabla 1.

Para calcular la medida de similitud se pueden utilizar varios métodos

Tabla 1. Datos

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>X1</i>	4	1	5	2
<i>X2</i>	1	4	2	4

como medidas de correlación, medidas de distancia y otras medidas de asociación. En este ejemplo se utiliza la distancia más corta, es decir la distancia euclídea, en donde las distancias más pequeñas indican mayor similitud. Estas distancias pueden medirse gráficamente en el diagrama de dispersión siguiente o calcularse analíticamente.

El análisis de *clusters* agrupa a los individuos (o variables) con medidas de similitud más homogéneas, en este caso con las distancias más cortas. En este ejemplo (Figura 1) es evidente que se forman dos grupos: el primero por los estudiantes *B* y *D*, que están satisfechos con los servicios de cómputo, pero no con la disponibilidad de libros en la biblioteca, y un segundo grupo formado por los estudiantes *A* y *C*, que tienen opiniones contrarias. Estos dos grupos tienen una distancia corta entre los individuos, y una distancia mayor entre los grupos, como puede observarse en la Figura 1 o en el *dendograma* de la Figura 2.

Es evidente que el número de agrupaciones convenientes es de dos conglomerados, de los cuales, el primero requiere mayor atención en la disponibilidad de libros en la biblioteca y el segundo requiere una mayor calidad de los servicios de computo. Si en un paso posterior se identifica el perfil de cada

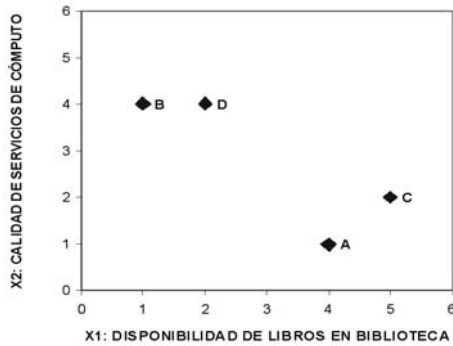


Figura 1. Gráfico de dispersión.

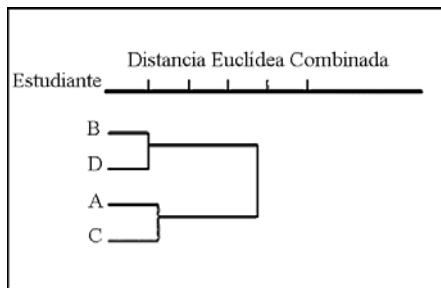


Figura 2. Dendrograma.

grupo es posible obtener mayor información y tomar mejores decisiones.

El lector puede concluir en este punto que no es necesario una técnica multivariada para realizar el análisis, ya que con un análisis univariado o bivariado podría llegarse a las mismas conclusiones; sin embargo, con mayor número de variables y de individuos no serían tan obvios los agrupamientos.

Debe ser claro que las técnicas univariadas no se contraponen a las multivariadas, por el contrario, las complementan y auxilian en la exploración inicial para la selección o diseño de un método adecuado o en la interpretación de los resultados o establecimiento de nuevas hipótesis.

### Aplicación

El primer paso a realizar es conocer cuáles atributos o variables de calidad del proceso educativo requieren mayor atención desde la perspectiva de los estudiantes de la Universidad de Quintana Roo y un segundo paso consiste en determinar si existen necesidades específicas para grupos de estudiantes determinados.

Para la identificación de los atributos o variables a medir esto puede lograrse mediante técnicas de consulta a los interesados y expertos, o partiendo de referencias preestablecidas. En este trabajo se utilizan las recomendaciones del modelo European Foundation Quality Management (EFQM) de los elementos a evaluar para la satisfacción de los estudiantes (Martín, 2001) y de un estudio sobre las condicionantes de la calidad desde el punto de vista estudiantil realizado en la Universidad de Salamanca, España (González, 2003).

Los atributos o variables a evaluar son:

- Formación para acceder a un buen empleo.
- Disponibilidad de información de los programas de estudio.
- Claridad de los objetivos de la carrera.
- Mercado de trabajo de su carrera.
- Ambiente.
- Competencia de los profesores.
- Métodos innovadores de enseñanza.
- Formación práctica.
- Requisitos y formas de titula-

- ción.
- Disponibilidad de libros.
  - Proceso de reinscripción.
  - Servicios recibidos en la biblioteca.
  - Disponibilidad de servicios de cómputo.
  - Prácticas de laboratorio y campo.
  - Consideración de opinión de los estudiantes en procesos de evaluación.
  - Participación en decisiones de la Universidad.

En otoño del 2004 se aplicó una encuesta a una muestra representativa de 150 estudiantes de la Universidad de Quintana Roo, utilizando un muestreo por conglomerados de los grupos de licenciatura de la unidad Chetumal. Se

utilizó una escala ordinal de cinco valoraciones de Likert desde el uno como totalmente en desacuerdo al cinco como totalmente de acuerdo. Al realizar un análisis de confiabilidad de la encuesta, se obtiene un alfa de Cronbach<sup>1</sup> de 0.83, lo que indica que el instrumento de medición es confiable.

En la Tabla 2 se encuentran ordenadas las valoraciones de los estudiantes de mayor a menor, en donde se observa que desean una mayor participación en las decisiones; un enfoque más hacia la práctica y la mejora de los servicios de laboratorio, reinscripciones y cómputo. Sin embargo se puede extraerse mayor información con análisis adicionales, por lo que se utilizará el análisis multivariado para tal fin. Es interesante explorar variables o di-

Tabla 2. Promedios de valoraciones

	Número de casos	Media aritmética	Desviación estándar
Profesores	150	3.75	.98
Objetivos	150	3.73	1.02
Ambiente	150	3.71	.85
Programas	150	3.51	.87
Biblioteca	150	3.41	1.13
Innovación	150	3.39	.93
Titulación	150	3.23	.99
Mercado	150	3.19	1.07
Libros	150	3.19	1.13
Evaluación	150	3.17	1.15
Cómputo	150	3.01	1.15
Reinscripción	150	3.01	1.07
Laboratorio	150	2.99	.97
Práctica	150	2.96	1.01
Decisiones	150	2.80	1.13

<sup>1</sup> Este coeficiente es uno de los más usados para determinar la confiabilidad de una encuesta o cuestionario, y está basado en la consistencia interna del instrumento; esto es se obtiene del promedio de los coeficientes de Pearson o de las covariancias, dependiendo si las puntuaciones están o no estandarizadas. Un coeficiente de Cronbach cercano a uno indica una consistencia adecuada, contrariamente cercana a cero indica una confiabilidad pobre.

menciones representativas, mediante un análisis de *clusters*. En la Figura 3 se representa un dendograma, donde se observa cómo las variables se agrupan. Los grupos de variables están señalados por líneas horizontales que apuntan a la variable y que está a su vez unida por una línea vertical que apunta hacia la distancia entre *clusters*. Se puede observar cómo se relacionan las variables que representan las opiniones sobre el *ambiente* y el *profesor*, uniéndose seguidamente con *innovación*, para incorporar sucesivamente a *programas*, *objetivos* y *mercado*, los cuales forman un grupo (A) que está asociado con la médula del proceso de formación desde sus principales condiciones de aprendizaje. Por otra parte, se observa cómo las variables relacionadas con los servicios de soporte como: laboratorio, biblioteca, cómputo, *reinscripción* también se van agrupando y se unen con otras como la participación de los alumnos en las *decisiones* hasta formar un grupo (B). La última en agruparse de una manera aislada es la

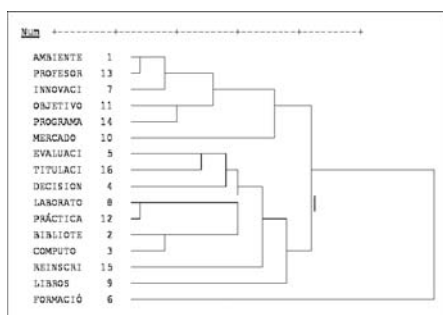


Figura 3. Dendograma del análisis de *clusters*. Opinión de estudiantes de la UQROO. Distancia *cluster* combinada.

variable denominada “*formación* para acceder a un buen empleo”, lo que indica que no pertenece a este instrumento, pero por otra parte (no se muestra en la gráfica) es la que más alta calificación recibió, lo que indica que es una variable común en todos los estudiantes.

### Distancia *cluster* combinada

En la Figura 4 se presenta una proyección en dos componentes principales y se confirma que las variables se agrupan en dos dimensiones: la dimensión relacionada con la enseñanza-aprendizaje y la relacionada con los servicios de soporte.

En la Figura 5 se muestra el perfil promedio de los grupos de variables

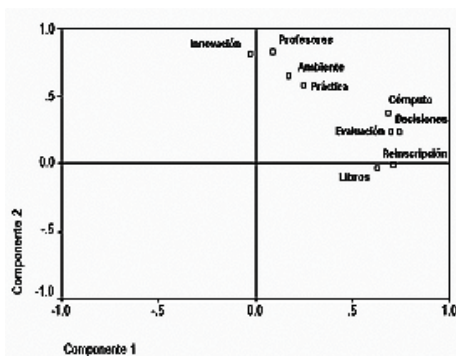


Figura 4. Opinión de estudiantes proyectadas en componentes.

así formadas, observándose que en ambos grupos las evaluaciones más frecuentes fueron 3 y 4, pero en el grupo 1 predominaron las evaluaciones de 4, mientras en el grupo 2 predominaron las evaluaciones de 3; las evaluaciones de 5 fueron más frecuentes en el grupo 1 que en el grupo 2,

ocurriendo lo inverso con las evaluaciones de 1 y 2.

Además, el agrupamiento obtenido puede servir para otros fines; para el

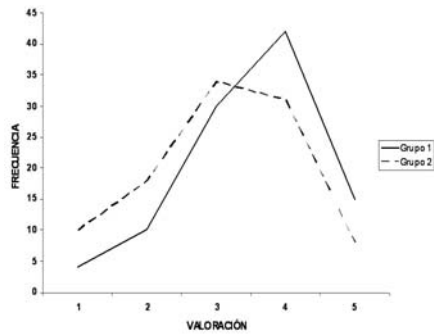


Figura 5. Perfiles de los grupos 1 y 2.

caso que nos ocupa, nos permite seleccionar una o más variables representativas para realizar un análisis de *clusters* y poder identificar grupos de estudiantes con valoraciones homogéneas. Aunque no existe un método exacto para seleccionar las variables, se seleccionan las que puedan representar a sus vecinos más cercanos; en este caso se escogieron las variables de *ambiente*, *cómputo* y *práctica*. Como el análisis de *clusters* es muy sensible a las variables seleccionadas se recomienda realizar varios análisis con la finalidad de encontrar la mayor cantidad de información posible.

Al realizar el análisis con las variables en cuestión se forman tres grupos. En la Tabla 3 se puede observar cada grupo con las valoraciones de cinco atributos (por cuestión de intentar dar claridad al artículo no se incluyeron todos los atributos): el grupo uno consta de 85 estudiantes (de una muestra de 150), que en promedio valoran

positivamente a los maestros y al ambiente creado en clase, en cambio sobre los servicios de cómputo y de reinscripciones están indiferentes; el grupo dos consta de 65 estudiantes claramente insatisfechos de los servicios de cómputo, del proceso de reinscripciones y también del aspecto de la práctica en la enseñanza; el grupo 3 sólo tiene un estudiante.

Es útil identificar posteriormente el perfil de los grupos dos y tres, para tener elementos para trazar las acciones diferenciadas que permitan mejorar los servicios que se reflejarían en la percepción del estudiante.

La Figura 6 muestra los perfiles de cada uno de estos tipos de alumnos, comparándolos con el perfil promedio total. Se aprecia de nuevo la

Tabla 3. Valoraciones promedio de tres grupos de estudiantes.

Grupo		Profes.	Amb.	Cómp.	Práctica	Reinscrip.
1	Media	4,14	4,08	3,74	3,37	3,21
	N	84	84	84	84	84
2	Media	3,23	3,22	2,11	2,42	2,78
	N	65	65	65	65	65
3	Media	5,00	5,00	1,00	4,00	1,00
	N	1	1	1	1	1
Total	Media	3,75	3,71	3,01	2,96	3,01
	N	150	150	150	150	150

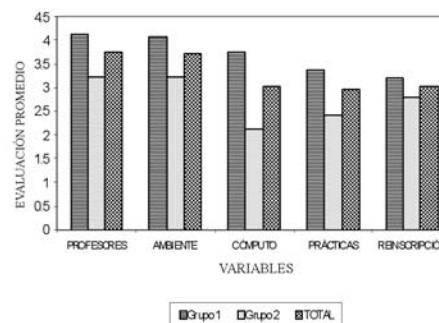


Figura 6. Perfiles de alumnos.



superioridad en calificaciones de las variables del primer grupo (que en este caso son *profesores y ambiente*), pero en general los alumnos del segundo tipo fueron más exigentes.

## CONCLUSIONES

Se ha pretendido mostrar que mediante el uso del análisis estadístico multivariado es posible obtener información que permita tomar decisiones más precisas. Para lograr lo anterior se utiliza y sugiere la siguiente metodología: identificar los atributos o variables de interés, elaborar el instrumento de medición estimando su precisión o confiabilidad, explorar las relaciones, seleccionar las variables representativas, realizar los agrupamientos de individuos y finalmente identificar los perfiles de los grupos.

Para la mayoría de las técnicas estadísticas la identificación de las variables de interés es de suma importancia, puesto que requieren representar lo que se supone representan, es decir tener validez de contenido. Esta etapa adquiere mayor relevancia para el análisis *cluster* debido a que los agrupamientos se realizan de acuerdo a las puntuaciones de esas variables. Por lo anterior se sugiere sustentar teóricamente las variables, partiendo de estudios previos o el uso de instrumentos de medición establecidos, y con técnicas como juicio de expertos complementado con otras técnicas estadísticas validar el instrumento. Del mismo modo es recomendable estimar en qué medida las

puntuaciones son replicables si no cambian las circunstancias en las que se aplica la encuesta, es decir calcular la confiabilidad del instrumento, para lo cual es posible utilizar distintas estrategias, siendo la más común el alfa de Cronbach.

Con la ayuda del paquete estadístico SPSS se aplicó la metodología para el caso particular de la Universidad de Quintana Roo y se obtienen las siguientes conclusiones y recomendaciones:

- a) El instrumento utilizado en este trabajo obtuvo un coeficiente de confiabilidad global de 0.83 que sugiere que el instrumento es replicable, y es posible mejorar este índice calculando con SPSS la contribución de cada variable en la consistencia interna global y analizando la posible pérdida de representatividad o validez de contenido al eliminar a la variable que menos contribuya en la consistencia interna.
- b) Al explorar con el análisis *cluster* los agrupamientos entre las variables sugieren que las variables forma dos segmentos, uno relacionado con el proceso directo de aprendizaje en el aula, y otro con los servicios de apoyo, para profundizar en estas relaciones se utiliza el análisis factorial exploratorio, con resultados similares que sugieren dos dimensiones subyacentes, estos resultados permiten seleccionar las variables más representati-



vas de ambas dimensiones o grupos, para facilitar la segmentación de individuos. Para el lector que desee confirmar las dimensiones subyacentes, se comenta que es posible realizar estudios posteriores con la técnica de análisis factorial confirmatoria con el paquete estadístico LISREL.

- c) Al aplicar nuevamente el análisis *cluster* para individuos con variables representativas de las dos dimensiones subyacentes, se encuentra que existen dos grandes grupos de estudiantes.
- d) El perfil de los grupos muestra que uno de ellos está francamente insatisfecho de los servicios educativos recibidos, principalmente en lo relativo a los servicios de apoyo.
- e) El anonimato de las encuestas no permitió extraer mayor información sobre el perfil de los grupos, el potencial aumentaría si se caracteriza mejor esos grupos, si el grupo insatisfecho tiene algunas características homogéneas de estado civil, escuelas de procedencia, nivel económico, estado laboral u otras, permitiría al nivel adecuado trazar mejores acciones y estrategias.
- f) Las desventajas del punto anterior son salvables si se rediseña el instrumento agregándole información no sólo de percepciones, sino de características de los encuestados.
- g) A modo de conclusión general, se debe reconocer que la mayo-

ría de las técnicas multivariadas son de carácter exploratorio, pero aplicadas considerando los supuestos y las restricciones pueden proporcionar información oportuna en la mejora de la calidad de la educación.

## BIBLIOGRAFÍA

- Dransfield B., Fisher N., Vogel N. - 1999, *Using statistics and statistical thinking to improve organizational Performance*. International Statistical Review núm. 67, International Statistical Institute.
- González López Ignacio, 2003, *Determinación de los elementos que condicionan la calidad de la Universidad: Aplicación práctica de un análisis factorial*. Revista electrónica de investigación y evaluación educativa, vol. 9, núm.1, España.
- Hair Joseph. Anderson Ralph, Tatham Ronald, Black William, 1999, *Análisis Multivariante*, 5a edición, Prentice Hall, Iberia, Madrid, España.
- Martín E. 2001, *Gestión de instituciones educativas inteligentes*, Mc Graw Hill interamericana de España.
- Levy Jean-Pierre, 2001, *Análisis multivariable para las ciencias sociales*, Prentice Hall, España.