

## ANÁLISIS ESTADÍSTICO DE RESPUESTAS LIBRES DE UNA ENCUESTA A INGRESANTES EN 1997. PRIMEROS RESULTADOS<sup>1</sup>

**Pagura, José Alberto.**

**Quaglino, Marta Beatriz.**

**Antoni, Elsa Josefina.**

*Instituto de Investigaciones Teóricas y Aplicadas, Escuela de Estadística, Asesoría Pedagógica.*

### 1. INTRODUCCION

La deserción en la FCEyE, como en general en la Universidad, constituye un problema de consideración. En trabajos anteriores se ha cuantificado esta problemática para distintas cohortes, habiéndose podido determinar además, que el mayor volumen de abandonos se produce en las primeras etapas de los estudios. En este sentido, es importante, acercarse de algún modo a la determinación de las causas de deserción, de modo de poder tomar acciones tendientes a disminuirla, al menos en lo que depende de la institución.

En esta línea, se han realizado también varios trabajos. Por ejemplo en 1995 se diseñó una muestra de alumnos en situación de abandono mediante la cual se indagó acerca de causas que podían haber provocado su situación, pero se obtuvo una baja tasa de respuesta dada la dificultad de ubicar a quienes abandonaron, luego de que han dejado la facultad. Esto condujo a pensar en la realización de un estudio prospectivo, teniendo un primer contacto con ingresantes mientras están cursando el primer año de sus estudios, evaluando algunos aspectos que pudieran tener relación con su éxito o fracaso posterior. Luego, a través de un seguimiento identificar a quienes abandonan y observar la existencia de alguna particularidad en lo manifestado en sus inicios, que permita aproximarse a la detección de las dificultades que impidieron la continuación de los estudios en los casos en que éste hubiera ocurrido.

En 1997 se dispuso realizar una encuesta entre ingresantes con la finalidad de obtener información que permitiera acercarse a este objetivo. Dicha encuesta fue aplicada a una muestra de 825 alumnos de tres de las carreras. La selección de la misma se hizo entre quienes estaban cursando materias de primer año, esto eliminó la posibilidad de seleccionar de entre un listado de inscriptos, a quienes solo manifestaron su "deseo" de estudiar en la universidad realizando el trámite del ingreso. El formulario de encuesta fue diseñado con preguntas cerradas y abiertas. Las preguntas abiertas tenían por finalidad en algunos casos, justificar una respuesta cerrada y en otros indagar sobre cuestiones para las que no se podía cerrar a priori la pregunta. También existía la intención de dejarlos expresar libremente sus ideas en algunos aspectos, por si de allí pudiera surgir alguna cuestión vinculada con la deserción, no prevista inicialmente.

Las alternativas tradicionales de tratamiento para las preguntas abiertas o respuestas libres, intentan enumerar la gama de respuestas obtenidas, estableciendo una codificación de modo de poder realizar el análisis estadístico de forma similar al correspondiente a variables cualitativas. Al optar por este camino, si bien se puede encasillar el tratamiento en métodos ampliamente conocidos, se produce una pérdida de la riqueza original de las respuestas. Los métodos estadísticos de análisis de datos textuales permiten estudiar las respuestas tal como fueron dadas, encontrando expresiones más frecuentes, respuestas características de diferentes subgrupos de la muestra, relación entre las diferentes expresiones y otras variables numéricas o cualitativas de la encuesta, etc.

Los primeros artículos publicados sobre estos métodos datan de 1981<sup>2</sup> y tomaron un impulso importante con la expansión del uso y las capacidades de las computadoras.

En este trabajo se presenta el análisis de una de las respuestas de la encuesta como ejemplo

<sup>1</sup> Este trabajo fue parcialmente financiado por el Ministerio de Educación y Justicia, Secretaría de Ciencia y Tecnología, mediante proyecto aprobado por el programa PID 202.

<sup>2</sup> Lebart, Ludovic (1981) Une Procédure d'Analyse Lexicale Ecrite en Langage Fortran. Les Cahiers d'Analyse des Données, vol VI n° 2 pp. 229-243, Dunod, Paris.

de aplicación de esta metodología, a fin de mostrar algunas de sus posibilidades, tratando además de destacar los aspectos que se deben tener en cuenta al realizar estos estudios.

## 2. LA ENCUESTA Y SUS PREGUNTAS ABIERTAS

Es importante destacar en primer lugar que para llegar a los alumnos, se visitaron diferentes comisiones de SIC I en los tres turnos y la comisión de primer año de la Licenciatura en Estadística. En algún momento de la clase convenido previamente con el profesor, se apartaba un espacio para que todos los presentes llenaran el formulario. El tiempo promedio para responderlo fue entre 20 y 30 minutos y la acogida de parte de los alumnos fue positiva en todo momento, al igual que la respuesta de los profesores, que generosamente cedieron una parte de sus clases. Un listado completo de las preguntas de la encuesta se puede encontrar en el trabajo "Una propuesta para la aplicación de análisis de correspondencias múltiples", que se presenta en estas mismas Jornadas. Además de ellas, se incluyeron seis preguntas abiertas. Las dos primeras, referidas a actitudes referentes al cursado del cursillo inicial y las cuatros últimas, más generales, trataban de percibir su valoración respecto de la obtención de un título universitario, sus razones de elección de la carrera y expectativas sobre lo que esperaban que les ofreciera la facultad para obtener un buen rendimiento.

Concretamente, las preguntas abiertas fueron:

- 1) ¿Por qué realizó (o no realizó) el curso preparatorio ofrecido por la Facultad?
- 2) ¿Por qué le parece que el curso le ayudó (o no le ayudó)?
- 3) ¿Cuál es la importancia que para usted tiene la obtención de un título universitario?
- 4) ¿Por qué eligió esta carrera?
- 5) ¿Qué opina de las clases a las que está asistiendo?
- 6) ¿Usted piensa que el rendimiento dependerá sólo de usted, o cree que la facultad puede hacer cosas para que sus posibilidades de éxito aumenten? ¿Qué cosas?

El análisis de estas respuestas, como datos textuales puede encararse por pregunta, o considerando el conjunto de todas ellas como un discurso más completo. A fin de simplificar la exposición, en este trabajo, se presentan solamente los resultados obtenidos al analizar la quinta pregunta dado que en el análisis surgen interesantes particularidades.

## 3. EL ANÁLISIS DE LAS RESPUESTAS LIBRES

Antes de mencionar los resultados obtenidos, es conveniente resumir brevemente los fundamentos metodológicos. El análisis de respuestas libres puede ubicarse en el análisis de datos, metodología basada en las posibilidades de cálculo de la computadora, que parte de los datos y trata con un gran número de dimensiones en forma simultánea. Estas técnicas pueden dividirse en análisis factoriales y de clasificación.

En análisis de datos, se trabaja con matrices que suelen mencionarse como "matrices de individuos x variables" dado que las filas de la misma representan a los individuos, personas u objetos bajo estudio (I) y las columnas a las variables que se estudian sobre cada uno (J). En el cuerpo de la tabla aparecen los valores numéricos de esas variables o los códigos de las modalidades si se trata de variables nominales.

Los datos de variables nominales también pueden presentarse en una matriz donde cada columna corresponde a una modalidad y el cuerpo de la tabla contiene 1 o 0 según esté o no presente la modalidad en el individuo. En este caso se dice que la matriz está en forma disyuntiva completa. La dimensión de esta matriz es mayor, es  $I \times M$ , donde M es el número total de modalidades de las J variables. Estos datos referidos a variables nominales, pueden ser resumidos también en una "matriz de Burt", matriz de dimensión  $M \times M$  que contiene las frecuencias con que se presenta cada combinación de modalidades. Estas tablas pueden ser tratadas mediante un análisis de correspondencias para observar las asociaciones existentes y encontrar factores subyacentes que las expliquen.

En el análisis de datos textuales, las palabras o formas gráficas<sup>3</sup> cumplen el papel de las modalidades de una variable nominal, la totalidad de las palabras contenidas en las respuestas aparecen como columnas y cada fila corresponde a una persona. En el cuerpo de la tabla aparecen las frecuencias con que cada individuo utilizó cada palabra en su respuesta libre. A partir de un análisis de correspondencias de esta **tabla léxica**, se pueden apreciar asociaciones entre palabras y aparecerán

<sup>3</sup> Se llama forma gráfica a una sucesión de caracteres no delimitadores.

las grandes ideas presentes en el texto. También se pueden utilizar otras variables nominales que se hayan medido sobre los mismos individuos, como variables suplementarias y aparecerán próximas en los planos factoriales, a las formas gráficas utilizadas por los individuos de esas modalidades.

Otra opción es dividir el conjunto de las respuestas (individuos) en subgrupos previamente establecidos o sugeridos por un análisis de correspondencias, por ejemplo carrera que cursan los alumnos, nivel de instrucción de los padres, etc., o combinaciones de algunas de ellas. A cada uno de esos subgrupos se los llama textos y se analiza la **tabla léxica agregada** formada por las palabras en sus filas y los textos en sus columnas. La comparación de los perfiles léxicos correspondientes a cada modalidad permite encontrar respuestas características de cada uno de ellos.

### 3.1- Tablas a analizar:

Tabla léxica:  $Z = \{z_{ij}\}$   $i = 1 \dots I$ ,  $j = 1 \dots J$ ; donde  $z_{ij}$  es la frecuencia con que la palabra  $j$  ha sido utilizada en la respuesta de la persona  $i$

Tabla léxica agregada:  $T = \{t_{jk}\}$ ; donde  $t_{jk}$  es la frecuencia con que la forma  $j$  aparece en el texto (modalidad de una variable nominal seleccionada)  $k$ .

Si  $X$  es la matriz disyuntiva completa que tiene en sus filas a las respuestas y en sus columnas las  $K$  modalidades que definen los textos, entonces:  $T = Z' X$

Es de mucha utilidad realizar un análisis similar utilizando, en lugar de palabras, segmentos<sup>4</sup>.

### 3.2 - Palabras y respuestas características

Para encontrar palabras características de cada texto o modalidad de una variable nominal elegida, se realiza una comparación entre el perfil léxico del texto, considerando a éste como una muestra del perfil léxico general. De estas comparaciones se obtienen para cada palabra valores test que orientan acerca de cuáles palabras son características de cada texto, ya sea porque se mencionan con una frecuencia relativa mucho mayor que en el corpus (conjunto total de palabras empleadas por todos los respondientes) o mucho menor.

Para elegir respuestas características de cada texto se toman respuestas reales y no artificiales y se pueden considerar dos criterios de elección:

a) Criterio del  $\chi^2$ : cada respuesta se considera como un vector fila cuyas componentes son las frecuencias de cada una de las formas utilizadas en esa respuesta, este vector es el perfil léxico de la respuesta. Un texto es un conjunto de vectores filas y el perfil medio del texto puede calcularse como la media de los perfiles de las respuestas del texto. Se calcula entonces la distancia del perfil léxico de cada respuesta al perfil medio del texto utilizando la distancia  $\chi^2$ .

Las respuestas más cercanas al perfil medio del texto serán las características de ese texto. Este criterio puede mejorarse considerando la distancia de la respuesta a los perfiles medios de los otros textos.

b) Criterio del valor medio: se mencionó que para cada forma se calcula un valor test; mediante este criterio se asigna a cada respuesta el promedio de los valores test de las formas incluídas en ella. La respuesta más característica será aquella que tenga el valor test más alto.

El software SPAD.T desarrollado por CISIA, Francia y que fue utilizado para procesar los datos en este trabajo, permite realizar las depuraciones que sean necesarias, como la eliminación de formas sin contenido semántico (formas herramienta), formas poco frecuentes, definición de formas equivalentes, etc. En sus diferentes instancias es posible obtener:

- listas de palabras y segmentos con sus frecuencias
- análisis de correspondencias de la matriz de respuestas x formas gráficas
- análisis de correspondencias de la matriz de formas gráficas x textos
- palabras, respuestas y segmentos característicos

## 4. RESULTADOS

Mediante la pregunta ¿Que opina de las clases a las cuales está asistiendo? se intentaba evaluar la percepción del alumno acerca de la principal actividad de la institución tratando de descubrir si esa percepción está vinculada con otros aspectos, como cuestiones relacionadas a su capacitación, a su contexto socio-cultural, carrera, u otros. En ninguna manera puede considerarse su respuesta como un aspecto de evaluación de la institución dado que el enfoque en la encuesta era detectar

<sup>4</sup> sucesión de formas separadas por delimitadores débiles, por ejemplo ‘,’ ‘ ‘ etc.

problemas **del alumno**. Además, antes de ser entregados los formularios de encuesta se realizaba una charla introductoría donde se explicaba el propósito de la misma: indagar sobre aspectos que estuvieran relacionados con la deserción, a partir de los altos índices de abandono que eran habituales año tras año.

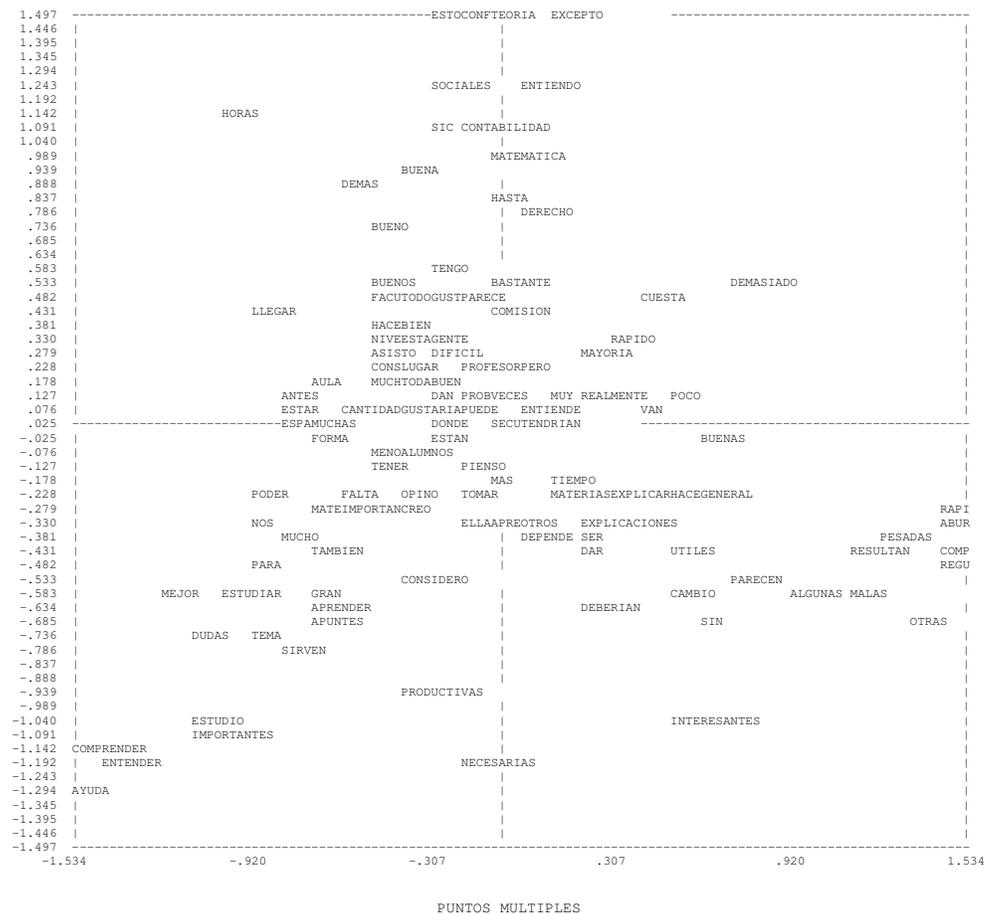
#### 4.1 -Depuración del corpus

Como primer paso, se procedió a depurar el corpus eliminando las formas herramienta<sup>5</sup> y las poco frecuentes, reteniendo para el análisis aquellas que aparecieron en más de 5 oportunidades. Como resultado, quedaron 141 formas gráficas.

#### 4.2 -Los temas generales

El análisis de correspondencias de la matriz de respuestas x formas gráficas, proporciona un primer plano factorial con características que se puede observar en el gráfico siguiente.

**Gráfico 1. Proyecciones de las formas gráficas en el plano de los primeros dos factores**



POINT * VU *	ABSCISSE * APPROCHEE	ORDONNEE * APPROCHEE	POINT * CACHE *	ABSCISSE * APPROCHEE	ORDONNEE * APPROCHEE
*ESTOY *	-.20	1.45	*GUSTA *	-.20	1.45
*CONFORME *	-.10	1.45	*PRACTICA *	-.10	1.45
*COMISION *	.00	.38	*EXPLICAN *	.00	.38
*HACE *	-.41	.33	*MUCHOS *	-.41	.33
*NIVEL *	-.41	.28	*PARTE *	-.41	.28
*ESTA *	-.31	.28	*PORQUE *	-.31	.28
*MAYORIA *	.31	.23	*TIENEN *	.31	.23
*BUEN *	-.20	.13	*DICTAN *	-.20	.13
*BUEN *	-.20	.13	*EJEMPLO *	-.20	.13
*ENTIENDE *	.10	.03	*HAY *	.10	.03
*ENTIENDE *	.10	.03	*TODOS *	.10	.03
*SECUNDAR *	.00	-.03	*TANTO *	.00	-.03
*SECUNDAR *	.00	-.03	*+--- *	.00	-.03
*EXPLICAC *	.31	-.38	*NADA *	.31	-.38

NOMBRE DE POINTS DOBLES = 14

<sup>5</sup> Son formas vacías de contenido semántico



El programa brinda el siguiente listado que ayuda a interpretar los ejes:

AIDE A L INTERPRETATION DES AXES	PROJECTION DES POINTS SUR L-AXE	
	1	
-----		-----
-----		COMPRENDER
-----		AYUDA
-----		ENTENDER
-----		MEJOR
-----		IMPORTANTES
-----		DUDAS
-----		ESTUDIO
-----		HORAS
-----		ESTUDIAR
-----		LLEGAR
-----		-----
-----		DEMASIADO
-----		ALGUNAS
-----		RESULTAN
-----		MALAS
-----		PESADAS
-----		OTRAS
-----		ABURRIDAS
-----		RAPIDAS
-----		COMPLETAS
-----		REGULARES
	2	
-----		-----
-----		AYUDA
-----		COMPRENDER
-----		NECESARIAS
-----		ENTENDER
-----		IMPORTANTES
-----		INTERESANTES
-----		ESTUDIO
-----		PRODUCTIVAS
-----		REGULARES
-----		SIRVEN
-----		-----
-----		SIC
-----		HORAS
-----		ENTIENDO
-----		SOCIALES
-----		ESTOY
-----		GUSTA
-----		TEORIA
-----		EXCEPTO
-----		CONFORME
-----		PRACTICA
	3	
-----		-----
-----		CANTIDAD
-----		GENTE
-----		MALAS
-----		ESPACIO
-----		MUCHA
-----		ALUMNOS
-----		CONSEGUIR
-----		PUEDE
-----		FALTA
-----		LUGAR
-----		-----
-----		ENTIENDO
-----		TEORIA
-----		ESTOY
-----		SIRVEN
-----		UTILES
-----		ENTENDER
-----		PRACTICA
-----		COMPRENDER
-----		AYUDA
-----		CONFORME

En un extremo del primer factor se ubican las palabras comprender-ayuda-entender-mejor-importantes-dudas-estudio y otras del mismo tenor, que indican porque es bueno ir a clase, mientras que en el otro extremo están regulares-completas-rápidas-aburridas-pesadas-otras-malas-resultan-algunas, las que resultan opiniones adversas. Quienes utilizaron palabras críticas para algunas clases, indicaron que las clases “eran buenas pero”, “o algunas buenas y otras no”; por esta razón aparece la palabra algunas en esa ubicación.

El segundo factor ubica en un extremo palabras que se refieren a utilidad y opiniones de las clases, en otro extremo se observan nombres de asignaturas y algunas opiniones; esto significa que hay un grupo de respuestas que hizo referencia a cada materia en particular.

En el tercer factor aparece un nuevo tema que es el de la “incomodidad” que perciben algunos alumnos, los que contestan con palabras como cantidad-gente-espacio-mucha-alumnos-

conseguir-falta-lugar-malas, es decir que vinculan los inconvenientes de espacio con lo que perciben como una clase infructífera.

En el otro extremo de este factor aparecen opiniones favorables a las clases y menciones a su utilidad.

#### **4.3 - Los temas generales y su relación con variables de la encuesta**

Para enriquecer la interpretación anterior, que constituye un primer paso en el análisis, se consideran algunas de las variables de la encuesta para observar su asociación con las respuestas obtenidas.

Estas se proyectan como variables suplementarias sobre los espacios generados por el análisis de correspondencias de la tabla léxica.

Las variables consideradas fueron:

- carrera que cursa
- número de horas que trabaja
- en cuánto tiempo espera terminar la carrera
- está en condiciones académicas en contabilidad, según las pruebas de ingreso?
- está en condiciones académicas en matemática, según las pruebas de ingreso?
- sexo
- grupo de pertenencia.

Esta variable corresponde al grupo en que fueran ubicados los alumnos luego de un procedimiento de clasificación automática realizado a partir del análisis de correspondencias de las preguntas cerradas de la encuesta.

Estos grupos responden a características que pueden resumirse a partir de si el alumno logró aprobar los exámenes de ingreso (en condiciones o no), de su declaración sobre el entrenamiento para el estudio (con problemas o no), y de su rendimiento en el ciclo anterior dado por sus notas en quinto año (bajo rendimiento anterior o no)

Se muestran en el Gráfico 2 de la página siguiente, las proyecciones de las modalidades de respuesta de estas variables en el plano de los dos primeros factores.

A través de la ubicación de las mismas y la comparación con lo observado en el Gráfico 1, se observan las asociaciones que existen con cierto tipo de respuestas.

#### **4.4 - Respuestas características por grupo**

Se obtuvieron respuestas características para dos de las variables nominales: grupo al que pertenece y carrera que cursa.

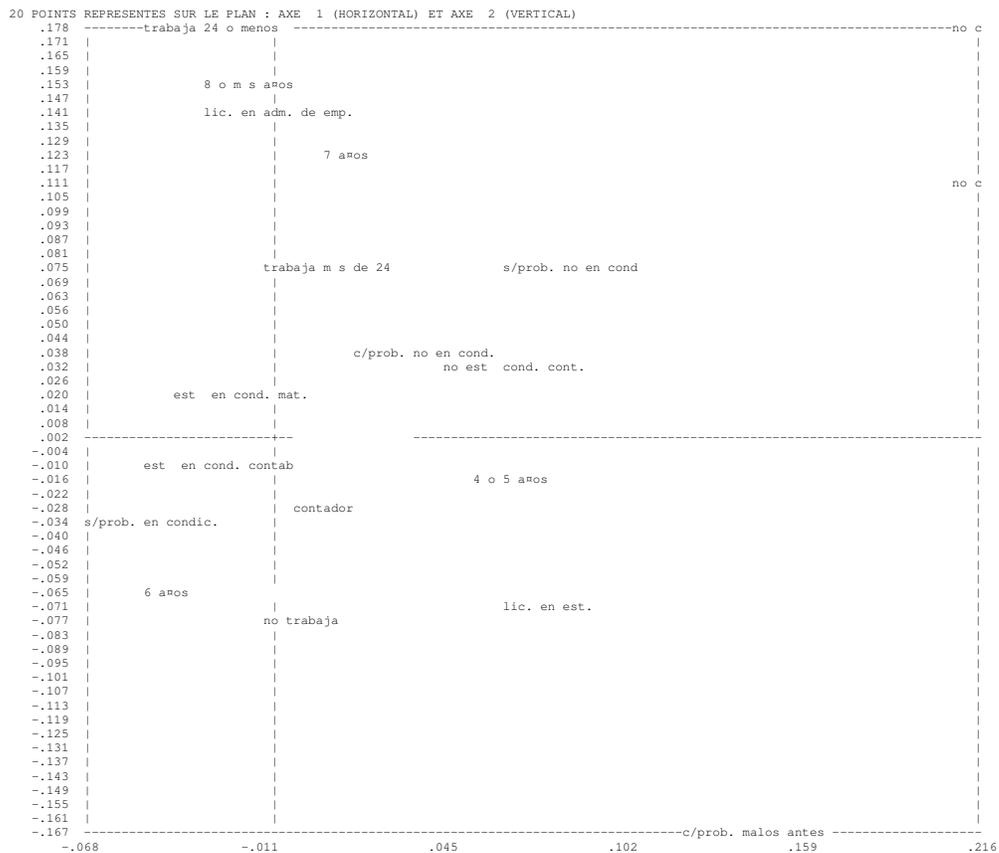
La primera de ellas, se eligió porque la constitución de los grupos sugiere diferente riesgo de abandono en cada uno y por lo tanto interesa conocer las particularidades de cada uno de ellos.

La segunda se eligió en virtud de las asociaciones que se pudieron visualizar en el análisis de correspondencias, lo que hizo suponer que las respuestas entre las diferentes categorías estarían muy bien diferenciadas.

La Tabla 1, permite apreciar la longitud de las respuestas y la riqueza de las mismas.

Las respuestas más largas corresponden a los que están en condiciones, pero las más ricas en cuanto a variedad de palabras, son las del tercer grupo.

**Gráfico 2. Proyecciones de variables suplementarias en el plano de los primeros dos factores**



POINTS MULTIPLES ( 100 AU MAXIMUM )

POINT	VU	ABSCISSE APPROCHEE	ORDONNEE APPROCHEE	POINT	CACHE	ABSCISSE APPROCHEE	ORDONNEE APPROCHEE
4 o 5 años		.06	-.02	no est cond. mat.		.06	-.02

NOMBRE DE POINTS DOUBLES = 1

**Tabla 1: Longitudes medias de respuestas, número de palabras diferentes, porcentaje de palabras diferentes con respecto al tamaño del corpus y número de encuestados.**

Grupo	Longitud media de respuesta	Número de palabras distintas	Porcentaje de palabras distintas con respecto al total de palabras del texto	Número de respuestas
sin problemas-no están en condiciones	14,2	131	5,14	179
con problemas-están en condiciones	17,5	136	2,65	293
con problemas-bajo rendimiento anterior	11,7	87	14,52	51
sin problemas-están en condiciones	18,0	136	2,5	302

Las respuestas características de cada grupo, por cada uno de los criterios, son proporcionadas por el SPAD.T en la siguiente forma:



**Tabla 2. Respuestas características según criterio de frecuencia de palabras.**

SELECTION DES INDIVIDUS OU REPONSES CARACTERISTIQUES (CRITERE DE FREQUENCE DES MOTS)

-----  
 TEXTE NUMERO 1 g1 = s/prob. no en cond  
 -----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
1.157 --	1 SON MUY UTILES .
1.057 --	2 ME GUSTAN TODAS .
.931 --	3 ME PARECE MUY UTILES . APRENDES MUY RAPIDO.
.919 --	4 MUY BUENAS
.919 --	5 MUY BUENAS .
.919 --	6 MUY BUENAS .
.919 --	7 MUY BUENAS .
.919 --	8 MUY BUENAS.

-----  
 TEXTE NUMERO 2 g2 = c/prob. no en cond.  
 -----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
.798 --	1 ALGUNAS BUENAS OTRAS MALAS .
.661 --	2 ESTA BIEN , SON BUENAS .
.638 --	3 ALGUNAS ME GUSTAN , OTRAS NO .
.583 --	4 CREO QUE ES COMO TODO ALGUNAS ME GUSTAN, OTRAS NO, ALGUNAS ENTIENDO, OTRAS NO.
.532 --	5 ALGUNAS SON BUENAS Y OTRAS NO
.532 --	6 ALGUNAS BUENAS OTRAS NO LO SON
.532 --	7 ALGUNAS SON BUENAS Y OTRAS REGULARES
.505 --	8 ME GUSTAN , ALGUNAS ME RESULTAN MAS FACILES QUE OTRAS , PERO SON IMPORTANTES .

-----  
 TEXTE NUMERO 3 g3 = c/prob. bajo rendimiento anterior  
 -----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
.714 --	1 SIRVEN PARA APRENDER
.686 --	2 ALGUNAS BUENAS .
.686 --	3 MUY BUENAS .
.500 --	4 OPINO QUE EN GENERAL SON BUENAS. HAY ALGUNAS MAS UTILES Y MEJORES QUE OTRAS
.497 --	5 PIENSO QUE HAY ALGUNAS CLASES MAS LLEVADERAS QUE OTRAS Y HAY OTRAS QUE SON MUY DIFICILES.
.489 --	6 QUE HAY MUCHA GENTE , NO ENTRAMOS.
.457 --	7 SON MUY BUENAS
.368 --	8 DEPENDE DE QUE MATERIA ; PERO POR LO GENERAL SON BASTANTE BUENAS

-----  
 TEXTE NUMERO 4 g4 = s/prob. en condic.  
 -----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
1.422 --	1 ESTOY CONFORME .
1.029 --	2 SON MUY PRODUCTIVAS
.906 --	3 AYUDAN A ENTENDER VARIOS TEMAS
.804 --	4 AYUDAN MUCHO A COMPRENDER LOS TEMAS . SON PARTICIPATIVAS . SON IMPORTANTES .
.772 --	5 ME SON MUY PRODUCTIVAS
.762 --	6 QUE ES MUY BUENA .
.753 --	7 MUY BUENAS .
.718 --	8 SE DICTAN LOS TEMAS "MUY" VELOZMENTE .



**Tabla 3. Respuestas características según criterio de  $\chi^2$ .**

SELECTION DES INDIVIDUS OU REPONSES CARACTERISTIQUES (CRITERE DU CHI-2)

-----

TEXTE NUMERO 1 g1 = s/prob. no en cond  
-----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
.921 -- 1	QUE SON MUY BUENAS, QUE LOS PROFESORES EXPLICAN BIEN Y QUE SIRVEN MUCHISIMO PARA LUEGO RENDIR LAS MATERIAS
.923 -- 2	LO QUE ME PARECE QUE ESTA MAL ES TENER QUE ESTAR VINIENDO UNA HORA ANTES COMO MINIMO PARA PODER ASISTIR A ALGUNAS CLASES , EJEMPLO : CONTABILIDAD. LAS OTRAS ME PARECEN MUY BUENAS . CONTABILIDAD TAMBIEN , PERO ME PARECERIA MEJOR SI NO HUBIESE TANTOS ALUMNOS , O SINO QUE HUBIESE MAS BANCOS .
.931 -- 3	SON MUY BUENAS Y MUY EXPLICITAS .
.931 -- 4	MATEMATICA MUY ACELERADA DEBIDO A QUE EL PROGRAMA ES DE CORTO TIEMPO PERO ASI MISMO TODAS SON MUY INTERESANTES
.932 -- 5	ME PARECEN BUENAS , MAS ME GUSTAN CUANDO ENTIENDO RAPIDO A LAS CLASES QUE DAN .
.932 -- 6	QUE ESTAN MAL ORGANIZADAS,SE DAN LOS TEMAS POR ALTO,QUE NO SE PUEDE PRESTAR ATENCION DEBIDO A LA GRAN CANTIDAD DE ALUMNOS QUE ASISTEN Y LA FALTA DE ESPACIO FISICO.DEBERIA HABER UN TRATO MAS PERSONALIZADO YA QUE ESTE NO EXISTE
.933 -- 7	SON MUY ABULTADAS EN NUMERO DE PERSONAS , Y LOS TEMAS SE DAN MUY RAPIDO , LAS PRACTICAS SON MUY POCAS DEBERIA HABER MAS EJERCICIOS .
.938 -- 8	ESTAN MUY BIEN PREPARADAS Y MUY BIEN DADAS POR LOS PROFESORES .

-----

TEXTE NUMERO 2 g2 = c/prob. no en cond.  
-----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
.926 -- 1	SE ENSEÑA BIEN , PERO A VECES ALGUNOS NO SE ANIMAN PREGUNTAR Y SE QUEDAN CON LA DUDA .
.927 -- 2	LAS CLASES SON MUY BIEN DESARROLLADAS, PERO DEBERIA HABER SALONES APROPIADOS PARA GRUPOS NUMEROSOS.
.930 -- 3	SON BASTANTE EXTENSAS Y ACLARAN MUCHA NEBULOSA QUE CREARON CONCEPTOS DE LA SECUNDARIA PERO A VECES ME PIERDO Y NO ENTIENDO. HAY CLASES DE CONSULTA PERO MI TRABAJO NO ME PERMITE ASISTIR . EN SIC , LAS CLASES SON BUENAS; TAL VEZ VAYAN MUY RAPIDO . PERO TODOS LOS PROFESORES QUE ME TOCARON ME GUSTAN COMO SE DESEMPEÑAN.
.932 -- 4	EN ALGUNAS VOY MUY BIEN,PERO EN OTRAS ME CUESTA UN POCO
.934 -- 5	ALGUNAS CLASES ESTAN BIEN DICTADAS,PERO ALGUNOS PROFESORES DICTAN MUY RAPIDO Y NO SE LLEGA A CAPTAR BIEN
.935 -- 6	A VECES ME DESESPERAN , PORQUE NO ENTIENDO ALGUNOS TEMAS , EN OTRAS , ME PONEN CONTENTO DE TERMINAR LA CLASE.
.937 -- 7	ALGUNAS SON LLEVADERAS,POR EJEMPLO DERECHO PRIVADO,SIC1 Y CS. SOCIALES.PERO LO QUE ES MATEMATICA TEORICA NO SE SI NO EXPLICA BIEN O SOY YO EL QUE NO ENTIENDE MUCHO NO ME GUSTA,NO SE SI SON LOS TEMAS U OTRA COSA
.941 -- 8	LAS CLASES SON BUENAS,EXCEPTO CUANDO HAY COMPAÑEROS QUE HABLAN TODA LA CLASE Y NO DEJAN QUE UNO SE CONCENTRE A ESCUCHAR AL PROFESOR,PORQUE MOLESTAN HABLANDO,PERO LO DEMAS ESTA TODO BIEN

-----

TEXTE NUMERO 3 g3 = c/prob. bajo rendimiento anterior  
-----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
.886 -- 1	OPINO QUE EN GENERAL SON BUENAS. HAY ALGUNAS MAS UTILES Y MEJORES QUE OTRAS
.891 -- 2	CREO MEJOR DICHO, MUCHOS ALUMNOS POR AULA, YO TENGO DIFICULTAD CON EL HORARIO, SI LLEGO TARDE NO ENCUENTRO ESPACIO Y NO HAY ASIENTO. LAS CLASES, HAY PROFESORES QUE ENSEÑAN MUY BIEN, OTROS NO
.912 -- 3	PIENSO QUE HAY ALGUNAS CLASES MAS LLEVADERAS QUE OTRAS Y HAY OTRAS QUE SON MUY DIFICILES.
.934 -- 4	EN LAS CLASES A LAS CUALES ASISTO , EL PROFESOR NO TRABAJA COMODAMENTE , ESO ES PERCIBIDO POR EL ALUMNO , EL ESPACIO FISICO ES EL MOTIVO PRINCIPAL DE LAS MIGRACIONES , NO HAY QUE BUSCARLE LA VUELTA.
.942 -- 5	DEPENDE DE QUE MATERIA ; PERO POR LO GENERAL SON BASTANTE BUENAS
.948 -- 8	SON BUENAS, PERO ENTRE TANTA GENTE NO SE PUEDE APRENDER MUCHO
.949 -- 9	OPINO QUE SON IMPORTANTES PARA APRENDER
.952 -- 10	QUE HAY MUCHA GENTE , NO ENTRAMOS.

-----

TEXTE NUMERO 4 g4 = s/prob. en condic.  
-----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU	CARACTERISTIQUE
.906 --	1	MI OPINION AL RESPECTO DE LAS CLASES QUE ASISTO : ES QUE ME SIRVEN MUCHO PORQUE HAY TEMAS TANTO DE CONTABILIDAD COMO DE MATEMATICA QUE NO LOS DI EN EL SECUNDARIO , O QUE NO LOS RECUERDO Y LAS EXPLICACIONES DEL PROFESOR SON MUY UTILES Y NECESARIAS PARA ENTENDER LOS TEMAS DEL PARCIAL.
.909 --	2	LAS CLASES SON MUY BUENAS, ME GUSTAN MUCHO. LOS PROFESORES EXPLICAN MUY BIEN Y ESTOY MUY CONFORME CON LAS CLASES , YA QUE ENTIENDO TODO.
.926 --	3	HASTA AHORA LAS CLASES DONDE ESTOY ASISTIENDO ME PARECEN MUY BUENAS YA QUE APRENDO MUCHISIMO , PERO EXISTE UNA EN LA QUE NO ESTOY MUY CONFORME PORQUE LA PROFESORA NO ES TAN DADA COMO LAS DEMAS QUE TENGO Y QUE SON MUY BUENAS Y DE LAS CUALES SE APREDE MUY BIEN .
.927 --	4	HASTA AHORA SON MUY BUENAS Y LOS PROFESORES QUE TENGO EXPLICAN CLARAMENTE Y ENTENDIBLE . ESTOY MUY CONFORME CON ESTOS PROFESORES.
.928 --	5	AYUDAN MUCHO A ENTENDER LOS TEMAS AUNQUE A VECES HA QUE COMPLETAR ESE TEMA CON BIBLIOGRAFIA BUSCADA POR NOSOTROS
.929 --	6	ME PARECEN MUY BUENAS , EL PROFESOR SE EXPRESA EN UNA FORMA MUY CLARA , ESTOS SON IMPORTANTES PARA ENTENDER LOS TEMAS DE LAS PUBLICACIONES Y ASI PODER PREPARARSE PARA LOS PARCIALES .
.931 --	7	ALGUNAS ESTAN MUY BIEN DICTADAS. PERO OTRAS SON MUY LENTAS Y SOLO SE LEEN LOS CUADERNILLOS EN LUGAR DE PROFUNDIZAR MAS EN EL TEMA
.935 --	8	EN LAS COMISIONES QUE ME ASIGNARON LOS PROFESORES SON MUY BUENOS, ES DECIR QUE SE ENTIENDE MUY BIEN . PERO EL NUMERO DE ALUMNOS ME HACE PERDER MUCHO TIEMPO YA QUE TENGO QUE VENIR UNA HORA ANTES PARA CONSEGUIR UN LUGAR DONDE SENTARME.

El primer grupo da opiniones favorables y, al examinar las respuestas características por el criterio de  $\chi^2$  aparecen menciones a la falta de espacio o a inconvenientes de infraestructura. Este grupo no está en condiciones académicas, si bien no manifiesta tener problemas. El segundo grupo, que reconoce problemas aunque ha aprobado los cursillos, expresa que algunas clases son buenas y otras no, y las respuestas características seleccionadas por el criterio  $\chi^2$ , reflejan la presencia de problemas para “entender”, y los generados por la falta de espacio. El tercer grupo tiene también opiniones favorables a algunas clases, y desfavorables a otras; también expresan los problemas de infraestructura, pero no se observan comentarios que permitan evidenciar la presencia de problemas en el estudio, que si se notaron en las preguntas cerradas. En el cuarto grupo, se observan respuestas favorables a las clases, y una clara manifestación de conformidad, acorde con la ausencia de problemas en el estudio que manifestaron en las preguntas cerradas.

#### 4.5 - Respuestas características por carrera

Las respuestas de los alumnos de la Lic. en Estadística son más largas y con un porcentaje alto de palabras diferentes (Tabla 4). El bajo porcentaje de palabras distintas que se observa en las respuestas de los alumnos de la carrera de Contador, es coherente con la homogeneidad de las mismas, notada en las respuestas características.

**Tabla 4: Longitudes medias de respuestas, número de palabras diferentes, porcentaje de palabras diferentes con respecto al tamaño del corpus y número de encuestados.**

Carrera	Longitud media de respuesta	Número de palabras distintas	Porcentaje de palabras distintas con respecto al total de palabras del texto	Número de respuestas
Contador	16,1	136	1,43	589
Lic. en Adm. de Empresas	17,6	132	3,95	190
Lic. en Estadística	19,2	95	10,76	46

Las respuestas características obtenidas por ambos criterios, se muestran en las tablas 5 y 6.



**Tabla 5. Respuestas características por carrera, según criterio de frecuencia de palabras.**

SELECTION DES INDIVIDUS OU REPONSES CARACTERISTIQUES (CRITERE DE FREQUENCE DES MOTS)

-----

TEXTE NUMERO 1 cont = contador

-----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
1.341 -- 1	MUY BUENAS .
1.341 -- 2	MUY BUENAS
1.341 -- 3	MUY BUENAS .
1.341 -- 4	MUY BUENAS .
1.341 -- 5	MUY BUENAS.
1.341 -- 6	MUY BUENAS .
1.341 -- 7	MUY BUENAS
1.341 -- 8	MUY BUENAS .
.998 -- 9	SON MUY PRODUCTIVAS
.894 -- 10	SON MUY BUENAS

-----

TEXTE NUMERO 2 lae = lic. en adm. de emp.

-----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
.565 -- 1	ESTAN BIEN
.565 -- 2	ESTAN BIEN
.438 -- 3	BUENAS, ALGUNAS MUY RAPIDAS (MATEMATICA TEORIA)
.414 -- 4	LAS CLASES DE CONTABILIDAD Y MATEMATICA SON FUNDAMENTALES Y MUY UTILES. PERO LAS CLASES CUATRIMESTRALES (SOCIALES Y DERECHO PRIVADO) SON MUY NUMEROSAS, LO QUE LAS HACE UN POCO DENSAS Y ABURRIDAS.
.396 -- 5	EN GENERAL SON MUY BUENAS , EN ESPECIAL SIC , DERECHO PRIVADO, MATEMATICA , NIVEL DEMASIADO ACELERADO
.352 -- 6	ME GUSTAN MUCHO.
.348 -- 7	MATEMATICA: MUY DENSA. CONTABILIDAD:ME GUSTAN LAS CLASES Y COMO LAS DA LA PROFESORA.
.335 -- 8	ALGUNAS CLASES COMO INT. A LAS CIENCIAS SOCIALES RESULTAN BASTANTE INTERESANTES. DERECHO ES UN POCO PESADO. SIC ME RESULTA SENCILLO. MATEMATICA HAY QUE SEGUIRLA

-----

TEXTE NUMERO 3 lest = lic. en est.

-----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
1.496 -- 1	SON MUCHAS HORAS .
.841 -- 2	SON IMPORTANTES PORQUE ASI ADQUIRIMOS CONOCIMIENTOS .
.449 -- 3	ME CUESTAN UN POCO , CONSIDERO QUE SON IMPORTANTES , ( DE GRAN AYUDA ) .
.361 -- 4	OPINO QUE TENGO MUCHAS HORAS DE CLASES POR DIA A COMPARACION DE CASI TODAS LAS CARRERAS .
.342 -- 5	DENTRO DE TODO SON BUENAS . PORQUE AUNQUE SON MUCHAS HORAS , HAY TEORIA Y PRACTICA .
.330 -- 6	ME PARECE QUE SON CLASES IMPORTANTES , PORQUE EN ELLAS ADEMAS DE DAR PARTES TEORICAS , TAMBIEN TENEMOS CLASES PRACTICAS , QUE NOS PERMITEN ENTRE OTRAS COSAS SACARNOS DUDAS .
.311 -- 7	ME PARECE QUE LA CANTIDAD DE HORAS QUE SE DAN EN ESTADISTICA SON DEMASIADAS EN PROPORCION A LAS HORAS DE LAS DEMAS CARRERAS DE ESTA FACULTAD.
.309 -- 8	SON BUENAS Y ME INTERESAN PORQUE TIENEN PRACTICA Y TEORIA .

**Tabla 6. Respuestas características por carrera, según criterio de  $\chi^2$**

SELECTION DES INDIVIDUS OU REPONSES CARACTERISTIQUES (CRITERE DU CHI-2)

-----

TEXTE NUMERO 1 cont = contador

-----

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
.857 -- 1	SON MUY BUENAS,PERO SOLO ALGUNAS
.887 -- 2	QUE SE APRENDE , PERO NO HAY UNA RELACION MUY DIRECTA ENTRE EL ALUMNO Y EL PROFESOR .
.891 -- 3	QUE SON MUY BUENAS Y APROVECHABLES .
.891 -- 4	MUY BUENAS .
.891 -- 5	SON MUY BUENAS
.891 -- 6	MUY BUENAS .
.891 -- 7	MUY BUENAS
.891 -- 8	SON MUY BUENAS



TEXTE NUMERO 2 lae = lic. en adm. de emp.

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
.871 --	1 LAS CLASES DE CONTABILIDAD Y MATEMATICA SON FUNDAMENTALES Y MUY UTILES. PERO LAS CLASES CUATRIMESTRALES (SOCIALES Y DERECHO PRIVADO) SON MUY NUMEROSAS, LO QUE LAS HACE UN POCO DENSAS Y ABURRIDAS.
.890 --	2 CONTABILIDAD ME GUSTA MUCHO Y LA CLASE SE HACE LLEVADERA. MATEMATICA ME RESULTA UN POCO PESADA, SOBRE TODO LA TEORIA: HAY MUCHOS ALUMNOS Y LA CLASE NO SE LLEGA A COMPRENDER BIEN. LAS CUATRIMESTRALES NO SON MUY DIDACTICAS Y SE HACEN LARGAS
.906 --	3 ALGUNAS COMO LA DE CONTABILIDAD ME PARECEN MUY BUENAS PERO EN OTRAS DONDE ASISTE MUCHA GENTE, ME PARECEN MUY POCO PRODUCTIVAS
.907 --	4 ALGUNAS CLASES COMO INT. A LAS CIENCIAS SOCIALES RESULTAN BASTANTE INTERESANTES. DERECHO ES UN POCO PESADO. SIC ME RESULTA SENCILLO. MATEMATICA HAY QUE SEGUIRLA
.909 --	5 SON BUENAS, MUY BIEN EXPLICADAS Y SOBRE TODO SIRVEN PARA CUANDO UNO TENGA QUE ESTUDIAR PARA RENDIR, YA QUE PARA MI NO ES PROBLEMA ASISTIR PORQUE ME GUSTAN MUCHO. CREO QUE ES MENTIRA QUE NO ES NECESARIO VENIR, AL CONTRARIO ES MUY IMPORTANTE.
.910 --	6 PARA MI SON MUY UTILES YA QUE NUNCA HABIA VISTO CON PROFUNDIDAD NI CIENCIAS SOCIALES NI DERECHO PRIVADO Y POR ESO CREO QUE ES MUY IMPORTANTE ASISTIR A LAS CLASES
.911 --	7 POR SUERTE ASISTO A CLASES EN DONDE MIS PROFESORES EXPLICAN EXCELENTEMENTE LAS MATERIAS (EN CONTABILIDAD Y MATEMATICA) Y ME SIRVEN DE MUCHO Y SON ESENCIALES PARA LA COMPRESION Y EL EMPUJE PARA ESTUDIAR LA MATERIA.
.916 --	8 LAS CLASES SON MUY BIEN DESARROLLADAS, PERO DEBERIA HABER SALONES APROPIADOS PARA GRUPOS NUMEROSOS.

TEXTE NUMERO 3 lest = lic. en est.

CRITERE DE CLASSIFICATION	REPONSE OU INDIVIDU CARACTERISTIQUE
.877 --	1 LAS CLASES SON MUY IMPORTANTES SOLO QUE ALGUNAS SON MAS ENTENDIBLES QUE OTRAS, DE ACUERDO A LOS PROFESORES.
.908 --	2 DENTRO DE TODO SON BUENAS. PORQUE AUNQUE SON MUCHAS HORAS, HAY TEORIA Y PRACTICA.
.910 --	3 ME PARECE QUE SON CLASES IMPORTANTES, PORQUE EN ELLAS ADEMÁS DE DAR PARTES TEORICAS, TAMBIEN TENEMOS CLASES PRACTICAS, QUE NOS PERMITEN ENTRE OTRAS COSAS SACARNOS DUDAS.
.930 --	4 PODRIAN SER MEJORES. LOS PROFESORES PODRIAN SER MAS CLAROS Y PRACTICOS Y MAS LENTOS EN SUS ENSEÑANZAS. HAY MATERIAS QUE SE HACEN MUY PESADAS EN CLASE YA SEA POR LA CARGA HORARIA EN ESA MATERIA O EN EL DIA.
.932 --	5 SON IMPORTANTES PORQUE ASI ADQUIRIMOS CONOCIMIENTOS.
.933 --	6 SON IMPORTANTES Y A LAS CUALES SE DEBE ASISTIR PORQUE SON IMPRESCINDIBLES PARA LAS DISTINTAS MATERIAS. TODAVIA MUY BIEN NO ME ADECUO A ELLAS, YA QUE SON LA ANTITESIS A LAS CLASES DADAS EN LA SECUNDARIA.
.937 --	7 CREO QUE LOS PROFESORES DEBERIAN SER MAS PRACTICOS EN CUANTO A SUS EXPLICACIONES, Y DEBERIAN SER MAS CLAROS Y PAUSADOS.
.938 --	8 OPINO QUE LAS CLASES SON MUY IMPORTANTES, YA QUE SI NO ASISTO NO SABRIA NADA SOBRE LA CARRERA.

Las diferencias de respuestas características entre carreras son muy notables; en la carrera de Contador solo se encuentran, con los dos criterios de selección, respuestas del tipo "son muy buenas". En el caso de los alumnos de la licenciatura en administración, los conceptos acerca de las clases son en general buenos y mencionan algunas de las materias, pero aparecen los comentarios sobre la "falta de lugar". En la Licenciatura en Estadística las opiniones de las clases son buenas y aparece como distintivo de las demás carreras las cuestiones relativas a la cantidad de horas de clase.

## 5. COMENTARIOS FINALES

Las técnicas de análisis estadístico de datos textuales en encuestas, resultan apropiadas en la investigación de temas sobre los que no se pueden prever las respuestas, como explicación de alguna pregunta cerrada o donde no se quiere perder la espontaneidad de las expresiones. Las posibilidades de relacionar las respuestas libres con las variables de la encuesta abre aún más el panorama de explotación de los datos permitiendo ordenar conceptos utilizados por los entrevistados que no surgen de una simple descripción de los mismos. Las facilidades proporcionadas por el software específico pone al alcance de los investigadores la posibilidad de aplicación de estas técnicas.



Los resultados que se han presentado, aunque parciales, permiten vislumbrar la utilidad de estos métodos en el análisis concreto de la información recogida para esta investigación. Se continúa actualmente con el análisis exhaustivo de la encuesta, combinando su aplicación con la de otras técnicas de análisis de datos.

#### **BIBLIOGRAFÍA.**

- BECUE BERTAUT, Mónica (1991) *Análisis de Datos Textuales. Métodos Estadísticos y Algoritmos*. CISIA, París
- CRIVISQUI, Eduardo M. (1993) . *Análisis factorial de Correspondencias, un instrumento de investigación en las Ciencias Sociales*. Laboratoire de Methodologie du Traitement de Donnés. Université Libre de Bruxelles.
- GREENACRE, Michael y JÖRG, Blasius. (1994) *Correspondence Analysis in the Social Sciences. Recents developments and applications*. Academic Press.
- LEBART, Ludovic, MORINEAU, Alain y PIRON, Mari. (1995) *Statistique exploratoire multidimensionnelle*. Editorial DUNOD. París.
- LEBART, Ludovic, SALEM, André. (1994) *Statistique Textuelle*. DUNOD. París.