

## PATRONES DE DISTRIBUCION DE CALIFICACIONES EN EL PROFESORADO

### Un estudio sobre profesores de Bachillerato

FLORENTINO BLÁZQUEZ ENTONADO

#### CALIFICACIONES Y PROFESORADO

El problema de la cuantificación de resultados de la evaluación del rendimiento de los alumnos por los profesores es una cuestión que, aunque insistentemente planteada, no encuentra una solución eficaz.

Desde la aguda reflexiones de DOTRENS (1946, p. 125) o los trabajos sistemáticos de WRINKLE (1947), los sistemas de calificaciones, cualquiera que fuera el número de categorías que hubieran incorporado, siempre han sido blanco de severas críticas, generalmente motivadas pro la inexistencia de un patrón común de aplicación o interpretación.

Las escalas de evaluación, sean numéricas o verbales —diferentes según países— exigirían para su validez una intercomparación con el patrón de la curva de la distribución normal. Ocurre, además, que no todos los maestros y profesores entienden de igual modo qué es un sistema de calificaciones ni adecúan las notas a lo que subjetivamente juzgan como más correcto.

Es de todos sabido, además, lo difícil que resulta objetivar una calificación debido a la multiplicidad de papeles que le es obligado interpretar al profesor. Es fácil imaginar, con CARDINET (1984), casos en los que la calificación del profesor se encuentra afectada circunstancialmente:

- por elevar la puntuación a un alumno débil, con el fin de estimularlo;
- concediendo más tiempo a unos alumnos que a otros, al objeto de conseguir su suficiencia;
- tomando en cuenta el trabajo realizado en casa por ciertos alumno y no por todos, para no desfavorecer a los que no se benefician del apoyo familiar suficiente;
- no concediendo malas calificaciones al final de curso, para no incrementar el riesgo de repetición de ciertos alumnos;
- redondeando, en fin, la calificación media hacia arriba o hacia abajo para tomar en consideración aspectos no mensurables de su rendimiento...

Es decir, que puede haber tantas formas de «poner notas» como docentes existen.

La carencia de bases más o menos precisas que actúen como normas estables para los educadores ha determinado una notable pérdida de fiabilidad de las notas escolares, como se reconoce por todos.

*Pautas de calificación*

LAFOURCADE (1972, p. 230s) revisa brevemente antiguos trabajos sobre la «variabilidad» en las pautas de adjudicación de calificaciones:

De ellas conviene que nos fijemos en la de ODELL (1950) que encontró que en centros donde se aplicaban las clásicas cinco letras (A, B, C, D, F), el porcentaje de frecuencia de las más altas variaba del 0 al 40%, las intermedias oscilaban entre el 10 y el 50% y las más bajas del 0 al 25%.

EBEL (1965) destaca cómo algunos estudios han demostrado que diversos cursos de una misma escuela utilizan porcentajes de A y B, que varían de un 63 a un 26%.

Este mismo autor ha verificado que los criterios para usar idénticos sistemas han variado con el tiempo en favor de una mayor liberalidad en la actualidad.

CARTER (1952) descubrió que a igualdad de rendimiento y capacidad, la niñas reciben mejores notas que los varones.

HADDLEY (1954) informaba que los alumnos más aceptados por sus maestros, a igualdad de rendimiento y capacidad, obtienen calificaciones más altas que los menos aceptados.

GARCÍA HOZ (1963), analizando problemas similares en la Universidad de Madrid, comprobó que dos profesores de la misma asignatura con un mismo curso y en un mismo año emitían un juicio absolutamente dispar...

Otras investigaciones señalan que las profesoras tienden a asignar mejores calificaciones a las niñas que a los varones y a ser menos generosas que los maestros. Así ELSBREE [1965] concluye que una nota alta o baja puede estar relacionada con el sexo de quien califica...

Diversos autores reconocen por otro lado la dificultad de superar «ciertas actitudes de sospechada generosidad o mezquindad», producto en algunos casos de rasgos de personalidad con necesidad de reajuste.

Todo ello se debe, a nuestro juicio, a que las diferentes funciones de la evaluación son percibidas de modo diferente por cada profesor, y las actitudes respecto de las mismas son diferentes de acuerdo a las distintas concepciones que poseen sobre el papel de la evaluación en la escuela.

*Hipótesis del trabajo*

CARDINET (1984) ha demostrado que los profesores no desean, conscientemente al menos, impartir sus calificaciones sobre el patrón de la curva normal y, según el resultado de las pruebas, reconocen que pueden distribuir una mayoría de buenas o malas notas. No obstante, el mismo autor reconoce:

«A fin de curso se inquietarían, sin embargo, si obtuviesen resultados excepcionales o inusuales con relación a otros colegas... Temerían estar equivocados en su estimación del nivel de exigencia requerido».

No nos constan referencias, de todos modos, acerca de si existe homogeneidad interna en las distribuciones de calificaciones que los profesores emiten sucesivamente. Es decir, si cada profesor posee un determinado tipo de distribución «personal» que reitere convencionalmente en grupos de alumnos distintos y durante años sucesivos.

Citando a Wallson, REUCHELIN (1961) trata de «*La typologie des correcteurs d'exams*»; mientras que PIERON (1963) habla del «*efecto halo*», producto del esquema o pre-juicio que posee cada calificador, estableciendo que a causa del «inevitable» margen de arbitrariedad en la fijación del punto de discriminación, es el azar quien decide la suerte acerca del apto/no apto o aprobado/suspenso.

Ambos aspectos —efecto halo y amplitud de calificación— es lo que provoca en PIERON tal designación.

A propósito de una investigación de la «Carnegie Corporation» sobre «Conceptos, métodos, técnicas y alcance pedagógico y social de los exámenes y concursos» que cita FERNÁNDEZ PÉREZ (1986), la Comisión francesa informaba, allá por los años cincuenta:

«Las curvas de distribución muestran, por lo general, un aspecto poco regular y bastante diferente de un corrector a otro. Tan pronto se presenta achatadas y casi rectangulares con cimas poco agudas..., como, por el contrario, fuertemente ceñidas a una calificación central..., o, finalmente, manifestando una tendencia a acumularse en torno a dos cimas distintas».

Según nuestra hipótesis *es probable que cada examinador posea una específica amplitud o espectro de distribución de las calificaciones que emite*. Y, efectivamente, la hipótesis que trataremos de contrastar es la de que cada uno de los profesores posee un determinado tipo de distribución de calificaciones, o lo que es lo mismo, que cada profesor califica según una curva determinada.

Este supuesto, de verificarse a través del ajuste de las curvas respectivas, plantearía problemas serios respecto de la propia esencia de la labor calificadora del profesor. En todo caso detectará:

- En qué medida cada profesor posee, previa e independientemente de otras variables, una distribución peculiar, la cual «reparte» en cada grupo, curso o nivel, de modo «significativamente» semejante/no semejante.
- Qué grado o ponderación habrá que conceder a variable profesor en cualquier estudio sobre calificaciones de alumnos.
- Por otro lado, podemos también examinar la evolución del comportamiento del profesor en sus exigencias de evaluación conforme va transcurriendo su vida profesional.

En efecto poseemos datos relativos a profesores que iniciaron o tenían una vida profesional mínima en el momento en que comienzan nuestros datos en 1975. La trayectoria profesional de estos profesores la hemos seguido en muchos casos a lo largo de un decenio, período de tiempo suficiente para poder también apuntar respuestas a preguntas como:

- Si el profesor va disminuyendo o aumentando el nivel de exigencias en el desarrollo de su vida profesional.
- Hasta que punto pueden existir influencias entre los efectos de aquella pretendida curva de distribución de notas por cada profesor y el transcurso de los años de ejercicio.

## MUESTRA Y TRATAMIENTO DE LOS DATOS

Para realizar el estudio nos hemos servido de las actas de DIEZ cursos académicos —1975-1985— de 15 Institutos de Bachillerato de Extremadura, seleccionados de acuerdo a parámetros académicos y sociogeográficos, en los que hemos podido realizar el seguimiento de los profesores firmantes de las mismas.

Dicho seguimiento nos ha permitido trazar las sucesivas curvas de distribución de las calificaciones de cada uno de ellos en los diferentes grupos, cursos o años y comprobar en qué medida resultan homogéneas dichas distribuciones.

Para seleccionar un profesor hemos fijado en un mínimo de veinte los grupos localizados a los que califica y al menos durante 6/7 años de trayectoria profesional.

En cierta medida también se ha procurado elegir especialistas de diversas materias para contrastar si la variable «asignatura impartida» hacía configurar o no algún modelo de curva.

Con todas estas condiciones hemos obtenido datos de TREINTA profesores, a cada uno de los cuales asignamos un número de identificación del 1 al 30, identificando, asimismo, la asignatura o asignaturas que ha impartido.

Debe advertirse que las calificaciones que se han seleccionado corresponden todas a la convocatoria de junio. Sería hartamente imposible reunir datos suficientes en las convocatorias de septiembre ni realizar un tratamiento estadístico riguroso por la escasez de alumnos y calificaciones propias de esa segunda convocatoria.

En el ejemplo del Cuadro 1 se presentan los datos resumidos de un profesor, acompañadas de los valores obtenidos después de aplicar el test de Ji-cuadrado, de acuerdo a agrupaciones bianuales de sus calificaciones en datos directos y porcentajes.

Su respectiva representación gráfica se ha confeccionado a partir de resultados bianualmente agrupados; y un bienio aparece solamente y siempre que se ha podido localizar su firma en las actas de un mínimo de cinco/seis grupos de alumnos en esos dos años.

Asimismo en la representación gráfica aparece la curva promedio y en el margen superior derecho se consigna la nota media de todos los grupos localizados, su desviación típica y los coeficientes de asimetría y curtosis.

*La prueba Ji cuadrado* nos ha parecido el procedimiento adecuado para contrastar la hipótesis y efectuar el análisis cuantitativo. Este estadístico, efectivamente, parece un método válido para comparar resultados obtenidos especialmente con aquellos que son de esperar teóricamente en virtud de alguna hipótesis. Y cuanto más se aproximan los resultados observados a los esperados, tanto más pequeño es el valor Ji-Cuadrado y tanto mayor la concordancia entre los datos observados y la hipótesis que trata de comprobarse.

De todos modos puede parecer pretencioso intentar validar nuestra hipótesis con dicha prueba, ya que tal parámetro vendrá a confirmar, en este caso, si las calificaciones emitidas por cada profesor a lo largo de los períodos indicados son muestras sacadas aleatoriamente de una misma distribución —la promedio o patrón—; es decir, que en los casos en que el Ji-Cuadrado indique homogeneidad podremos afirmar que las fluctuaciones entre las curvas representativas de cada profesor podrían deberse sólo a errores de muestreo, pretensión exagerada para nuestra hipótesis, que se establece en torno a algún tipo de homogeneidad.

Conviene notar, además, que los valores de Ji-Cuadrado se han extraído agrupando las diferentes calificaciones en sólo tres columnas: 1) SUSPENSOS (que comprende las de «MD» e «IN»); 2) APROBADOS (sumas de «SF» y «BI»); y 3) DESTACADOS

que comprende «NT» y «SO»). (Cuadro 1). Tal agrupación se ha realizado porque la prueba Ji-Cuadrado no resulta estable cuando se calcula sobre una tabla de datos en la que cualquiera de las frecuencias experimentales ( $F_o$ ) es menor que 5, caso que ocurre en alguna ocasión en nuestros datos sin agrupar.

Las gráficas, no obstante, se extraen de los % de calificaciones originales o reales.

Para tener una nueva perspectiva, dado que es posible que la especialidad a la que pertenece cada profesor o asignatura que imparte pudiera producir algún tipo de incidencia en la configuración de la distribución de calificaciones, en nuestro trabajo original (BLAZQUEZ, 1987), los hemos agrupado por especialidad.

### *Análisis cualitativo*

Se ha realizado una entrevista con cada uno de los profesores «para conocer su opinión acerca de sus “personales” distribuciones y gráficas». Las cuestiones reiteradas a todos han girado en torno a conocer:

- Si ven como propias las distribuciones que aparecen en su gráfico.
- La opinión que le merecen.
- Si encuentran justificación acerca de alguna curva anómala que pudiera haber en su gráfica.
- Por qué utilizan —o no— determinada calificación y qué sentido tiene para ellos.
- La función que pueden cumplir las calificaciones en el marco de la evaluación..., etc.

En este breve informe no se consignan detalles de los análisis cuantitativos y cualitativos correspondientes a cada profesor. Solamente se presentan los resultados más significativos y de forma global.

### *Resultados*

— El primer y más importante resultado de este trabajo lo constituye la sorprendente analogía que poseen las distribuciones de calificaciones de los profesores sometidos a estudio, confirmando de manera plena la hipótesis inicial. (Véanse gráficas de los profesores n.º 6, 9, 17, 21 y 29, por ejemplo).

— De los treinta profesores objeto de seguimiento, más del cincuenta por ciento de los mismos, —dieciséis—, configuran unas distribuciones acerca de las cuales puede afirmarse que estadísticamente corresponden a la misma muestra; o lo que es lo mismo, que las diferencias que existen entre ellas pueden atribuirse a simples errores de muestreo. Para ser más explícitos: a pesar de provenir de diferentes grupos, cursos y años podría decirse que son iguales.

— En los casos en los que estadísticamente no puede hablarse de absoluta homogeneidad, la representación gráfica que sobre cada profesor se ha realizado, (véanse gráficas de profesores n.º 14, 19 ó 23), muestran unas similitudes tales que, así mismo, permiten concluir que en la mente del profesor existe una determinada forma de distribuir las calificaciones.

— El estudio por asignaturas nos permite realizar algunas precisiones:

a) Que las similitudes se producen por igual, e independientemente de la especialidad a que pertenezcan.

b) Que se producen agrupaciones particularmente semejantes entre profesores de la misma especialidad.

c) En Matemáticas, sin embargo, podrían verse en la muestra ciertas tendencias análogas (no más que el esperado signo de calificaciones «severas», como lo muestran las gráficas de los profesores n.º 21, 26, y especialmente la 16), al contrario que en las asignaturas «marías», que presentan una tendencia «suave» (gráficas de prof. n.º 4, 13 y 30).

d) La referencia a la edad del profesor o los años de ejercicio no nos permite concluir que se den mayores o menores analogías en uno u otro caso.

e) Tampoco se podría concluir mayor severidad en los profesores de menor edad o mayor tolerancia en los veteranos, si bien habrían de extraerse muestras más significativas por sectores para poder concluir con mayor precisión en torno a la creencia general que habla de «Sancho el Fuerte, Sancho el Justo y Sancho Panza...».

f) En las apreciaciones personales solicitadas a los profesores, éstos no son conscientes de expedir tan similares distribuciones, teniendo más bien la creencia de todo lo contrario.

g) Se constata asimismo una opinión generalizada de que los niveles de conocimiento de los alumnos están descendiendo.

h) Algunas de las calificaciones de la escala se utilizan muy escasamente. Así ocurre con «SO», que para algunos es usada como excepción, y con «M.D» a la que se le ha conferido de modo muy generalizado más un valor de sanción moral que de calificación académica. Ambas referencias pueden constatarse en las gráficas presentadas.

i) Por último, de nuestras entrevistas con los profesores se deduce una notable discrepancia de criterios y hasta de función de las propias calificaciones.

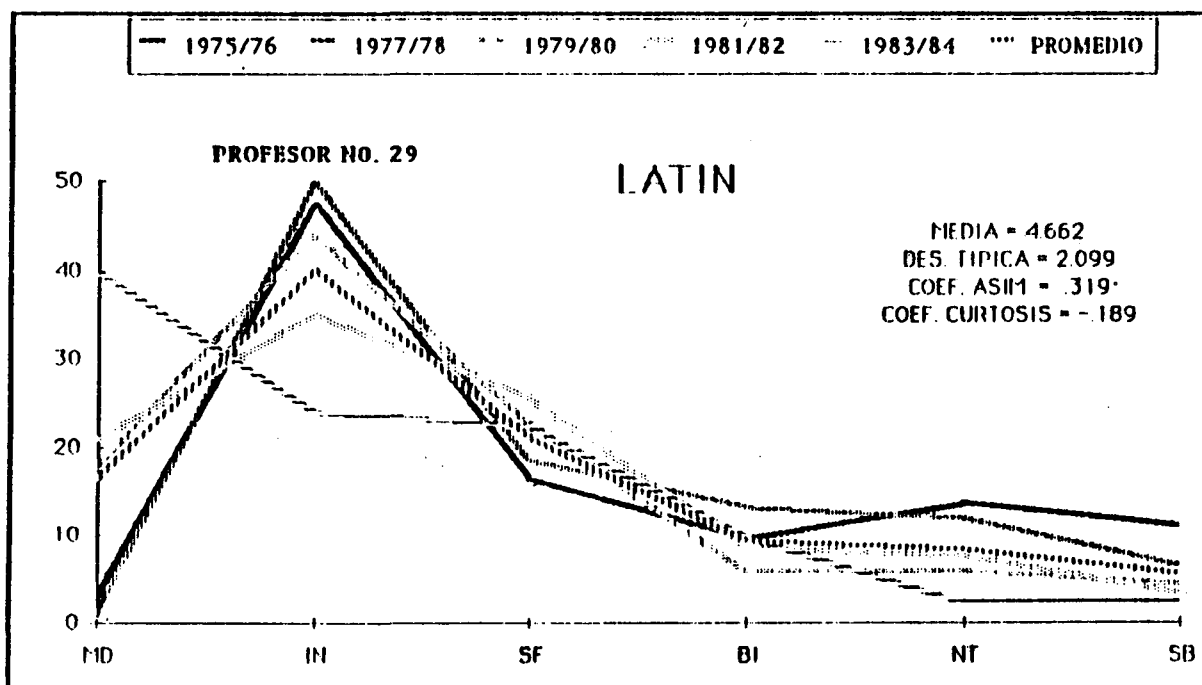
FLORENTINO BLÁZQUEZ ENTONADO  
Universidad de Extremadura

## BIBLIOGRAFIA

- ALVIRA, F. (1983) «Perspectiva cualitativa/perspectiva cuantitativa en la metodología sociológica». *R.I.S.*, 22, 53-75.
- ALVIRA, F. (1985) «La investigación evaluativa: una perspectiva experimentalista». *R.I.S.*, 29, 129-141.
- ANGUERA, M. T. (1985) «Posibilidades de la Metodología Cualitativa versus Cuantitativa». *Revista de investigación educativa*, v. 3, 6, 339-345.
- BARTOLOMÉ, M. (1985) «Incidencia en la práctica educativa de la investigación sobre el rendimiento». *Revista de investigación educativa*, v. 3, 6, 145-154.
- BARTOLOMÉ, M. (Coord.) y otros (1982): *Modelos de investigación educativa* I.C.E. Universidad de Barcelona.
- BONBOIR, A. (1974) *La docimología: problemática de la evaluación*. Madrid. Morata.
- CARDINET, J. y ALLAL, L. (1978) «La mesure des resultats de l'enseignement». *Techniques d'instruction*, 33, 7-11.
- CARDINET, J. (1980) «L'elargissement de l'évaluation». *Education et Recherche*, 1, 15-34.
- CARDINET, J. (1984) «Reflexions d'enseignants sur l'évaluation des élèves». *I.R.D.P.* Neuchatel.
- COOK, T. D. y REICHARDT, Ch. S. (1986) *Métodos cualitativos y cuantitativos en investigación evaluativa*. Madrid. Morata.
- CRONBACH, L. J. (1980) *Toward Reform of Program Evaluation*. S. Francisco. Sage.
- DIXON, V. J. y BROWN, M. B. (1979) *Biomedical Computer Programs*. Los Angeles. Univ. California Press.
- DOTTRENS, R. (1946) *Hay que cambiar la educación* B. Aires. Guadalupe.
- EBEL, R. L. (1954) «Procedures for the analysis of classroom tests». *Educ. and Psychol. Measurement*, vol. XIV.
- EBEL, R. (1977) *Fundamentos de la evaluación educacional*. B. Aires. Guadalupe.
- EISNER, E. (1981) *The Methodology of Qualitative Evaluation*. Stanford. Unpublish.
- ESCUADERO, J. M. (1982) «El fracaso escolar: hacia un modelo de análisis», en BARTOLOMÉ, M. y otros, *Modelos de investigación educativa*, o. cit.
- FERMÍN, M. (1971) *La evaluación, los exámenes y las calificaciones*. B. Aires. Kapelusz.
- FERNÁNDEZ, M. (1986) *Evaluación y cambio educativo: el fracaso escolar*. Madrid. Morata.
- GARCÍA HOZ, V. (1963) *Normas elementales de pedagogía empírica*. Madrid. Escuela Española.
- GARANTO, J., MATEO, J. y RODRÍGUEZ, S. (1985) «Modelos y técnicas de análisis del rendimiento académico». *Revista de Educación*, 277, 171-199.
- GIMENO, J. y PÉREZ, A. (1983) *La enseñanza: su teoría y su práctica*. Madrid. Akal.
- HUBERMAN, M. y MILES, M. (1983) «L'analyse des données qualitatives: quelques techniques de reduction et de rep.». *I.R.P.D.* v. 83, 3.
- HUBERMAN, M. (1983) *S'évaluer pour s'illusionner? Promesses et écueilles de l'évaluation*. *I.R.P.D.* Neuchatel.
- KETELE, J. M. (1981) «Representations qu'ont de l'école les parents, élèves et professeurs de secondaire». *Les Sciences de l'éducation*, 2-3, 59-96.
- LANDSHEERE, G. de (1973) *Evaluación continúa y exámenes: manual de docimología*. B. Aires. Ateneo.
- NEVO, D. (1983) «The Conceptualization of Educational Evaluation». *Review of Educational Research* 1, 117-128.
- O.C.D.E. (1975) *Los indicadores de resultados en los sistemas de enseñanza*. París. O.C.D.E.
- ORDEN, A. de (1985) «Hacia una conceptualización del producto educativo». *Revista de Investigación educativa*, 3, 271-283.

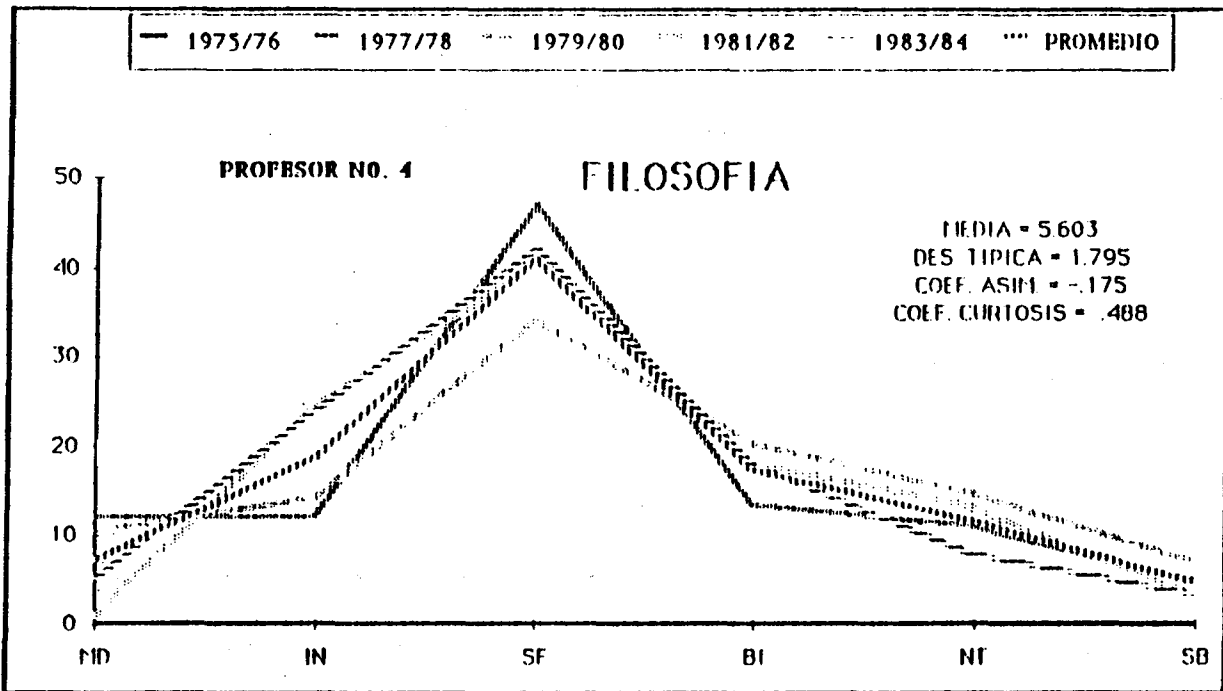
- PAQUAY, L. (1985) «Les axes paradigmatiques des recherches su l'évaluation des innovations scolaires». *Les Sciences de l'Education*, 4, 3-34.
- PFISTER, Ch. (1975) «La validité de la note scolaire». *Publications Universitaires Europeennes*, v. XI. Berna. Lang.
- PIERON, H. (1963) *Examens et docimologie*. París. P.U.F.
- REICHARDT, Ch. y COOK, T. D. (1982) «más allá de los métodos cualitativos versus cuantitativos». *Estudios de Psicología*, 11, 41-53.
- REUCHELIN, M. (1961) «La tipologie des correcteurs des examens». *B.I.N.O.P.*, 17, 2 serie.
- RODRÍGUEZ DIÉGUEZ, J. L. (1979) *Técnicas de evaluación educativa*. Valencia. Nau Llibres.
- RODRÍGUEZ ESPINAR, S. (1985) «Modelos de investigación sobre el rendimiento académico». *Revista de investigación educativa*, 6, 285-303.
- SHAVELSON, R. y STERN, P. (1983) «Investigación sobre el pensamiento pedagógico del profesor», en GIMENO y PÉREZ, *La enseñanza.*, o. c., 372-419.
- SIMPSON, R. H. (1969) *Teacher self-evaluation*. Londres. McMillan.
- STUFFLEBEAM, D. L. y WEBSTER, W. J. (1980) «An Analysis of alternative approaches to evaluation», *Educ. Evaluation and Policy Analysis*. v. 2-3, 5-20.
- TEJEDOR, F. J. (1984) *Análisis de Varianza aplicado a la investigación en Pedagogía y Psicología*. Madrid. Anaya.
- TOURON, J. y ARRIETA, J. (1984) «Validez predictiva de las calificaciones de E. M. y Selectividad para la Universidad» en AGUIRRE, J. (Ed.) *La selectividad a debate*. Madrid, Universidad Autónoma.
- VIDAL, C. (1984) *Críteris d'avaluació a l'ensenyament secundari*. Barcelona. CEAC.

DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO

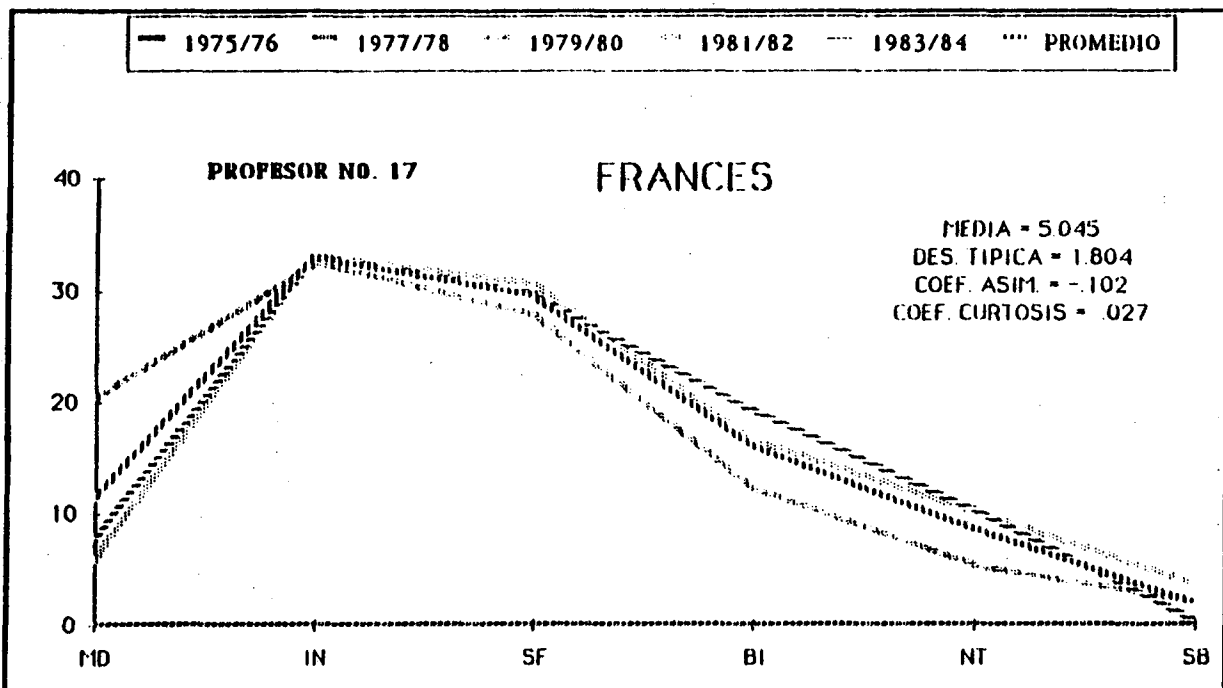




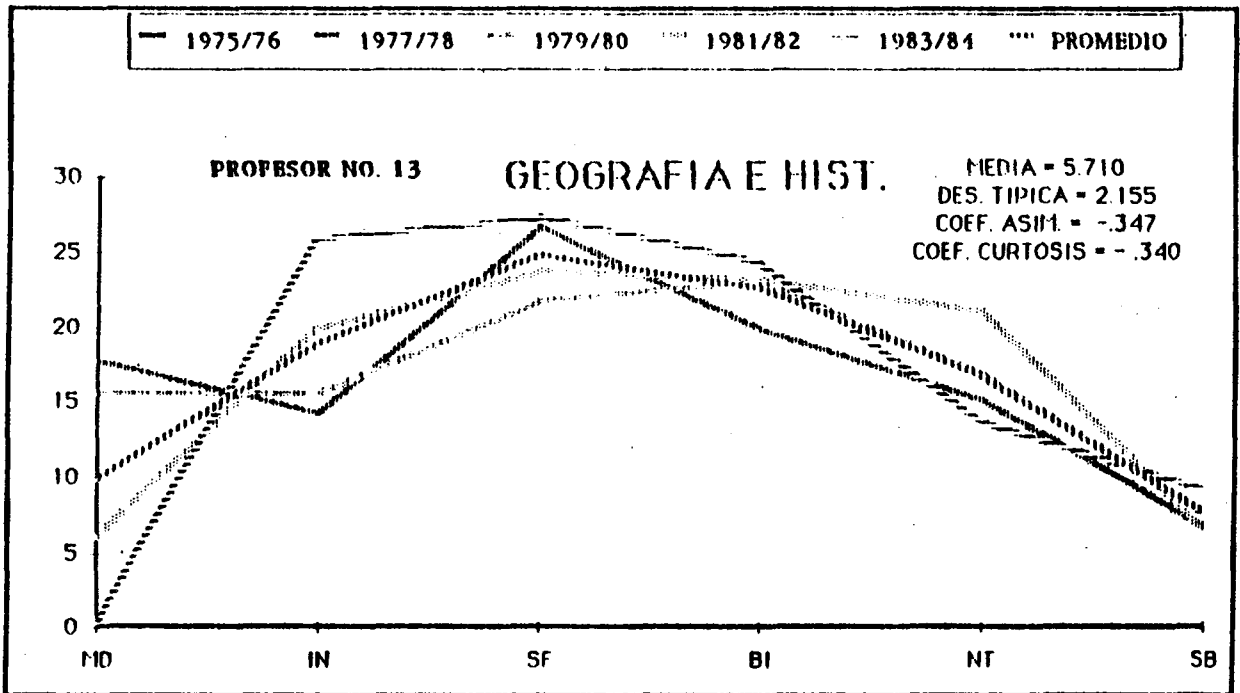
DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



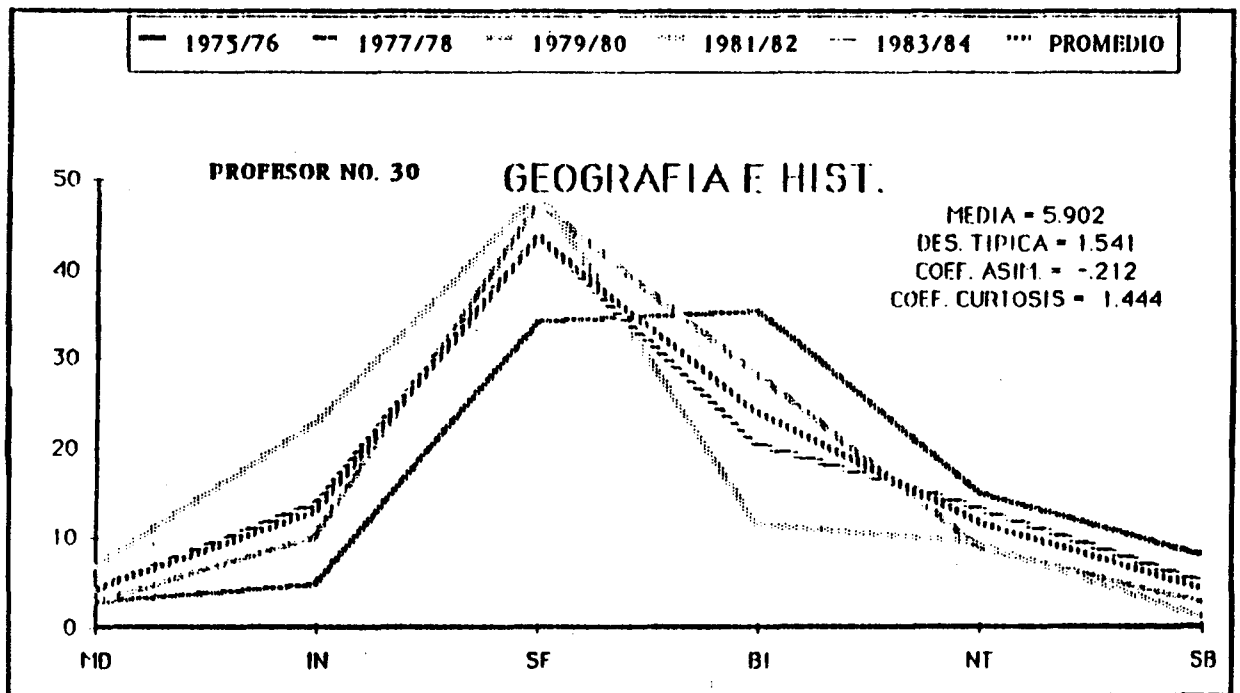
DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



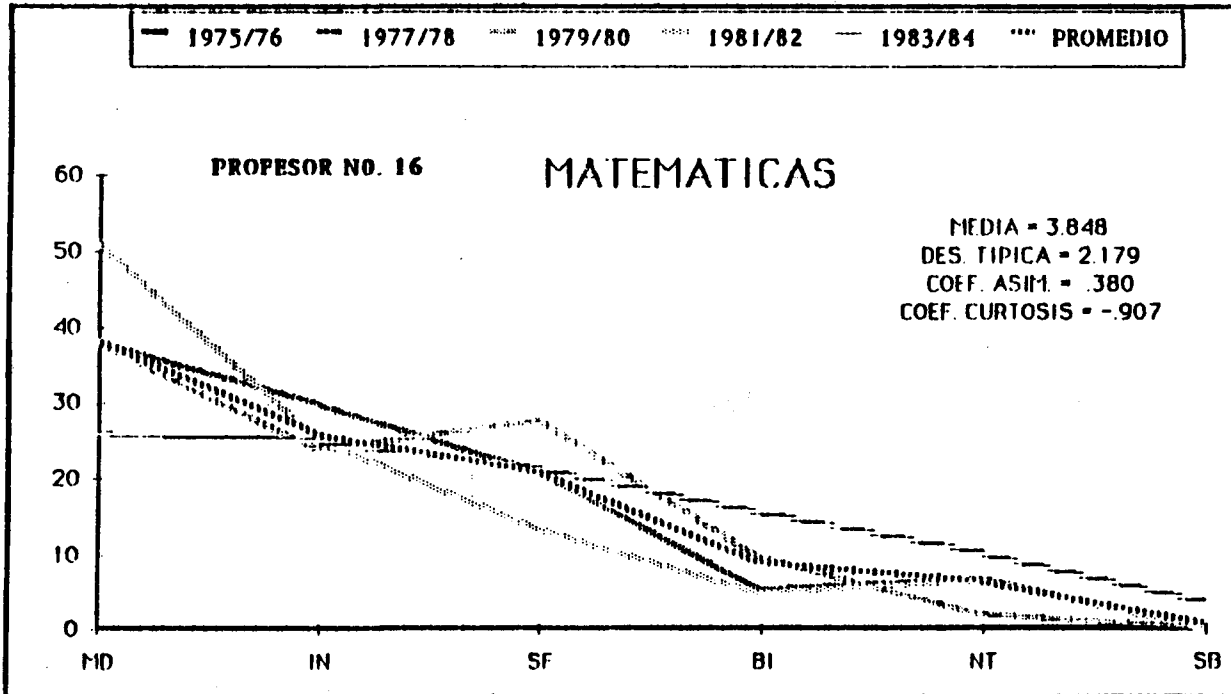
DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



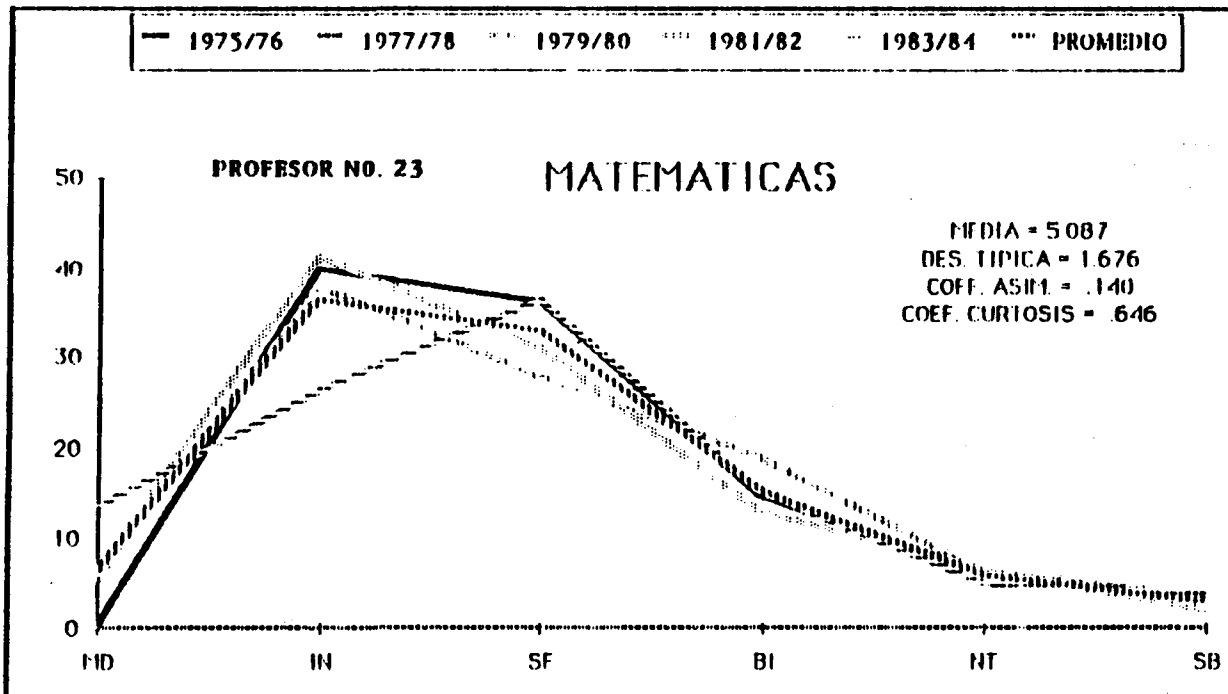
DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



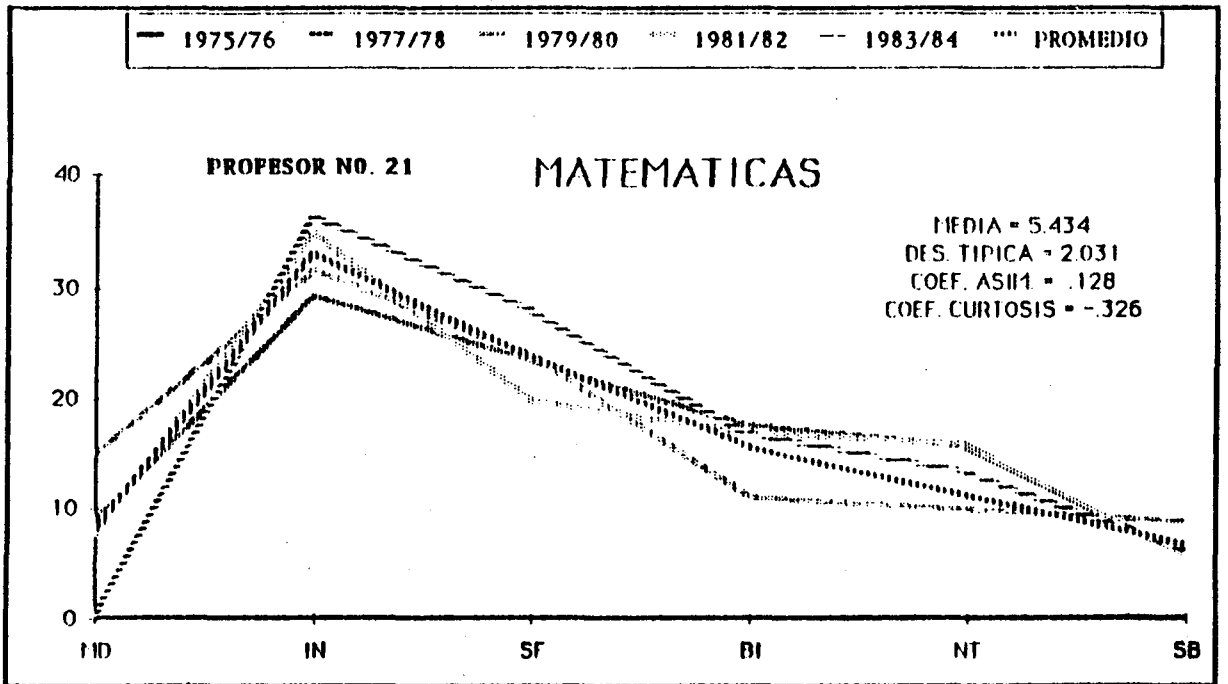
DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



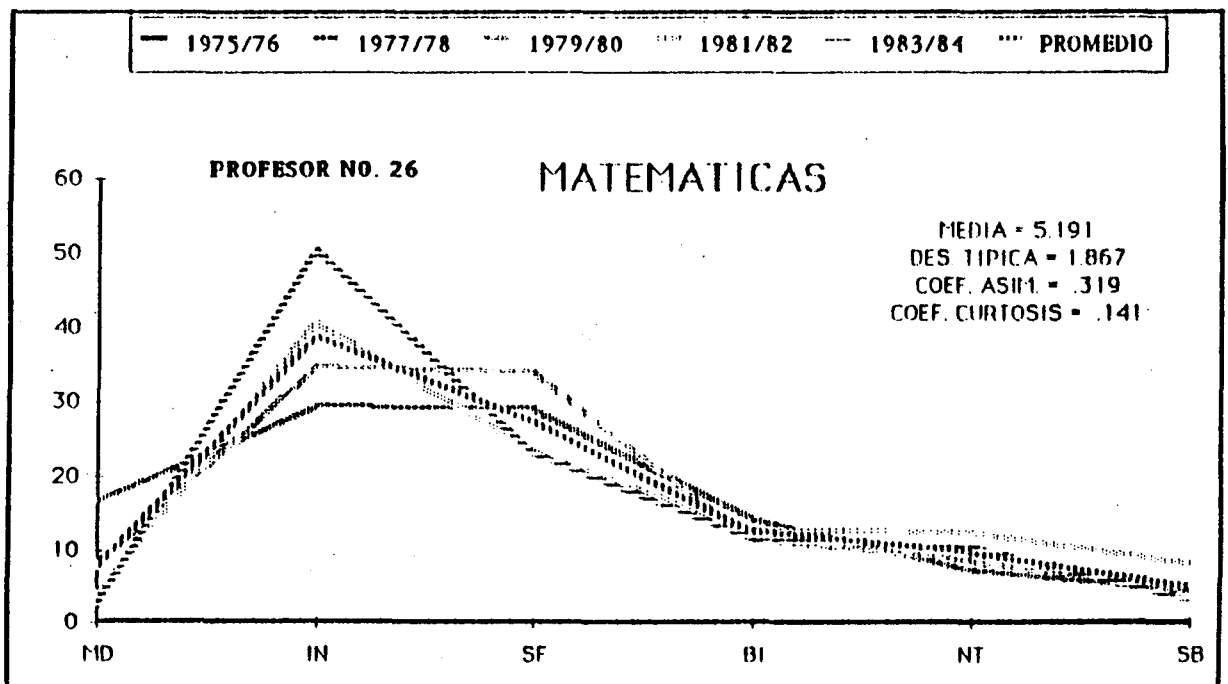
DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



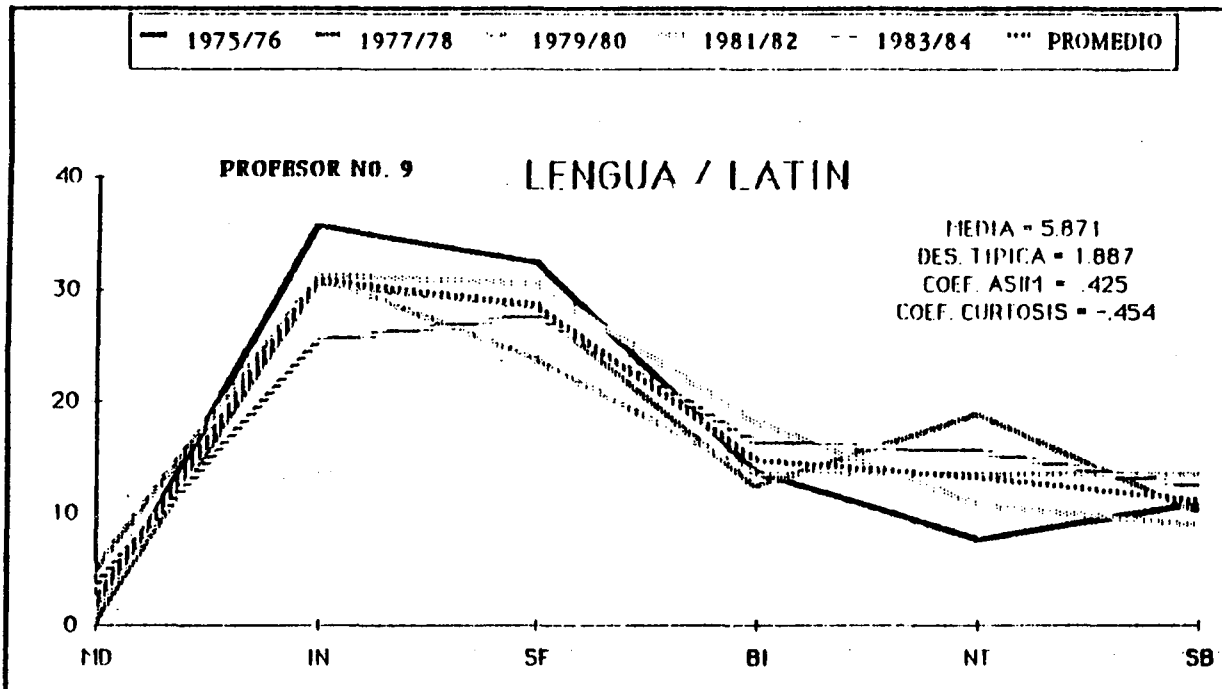
DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



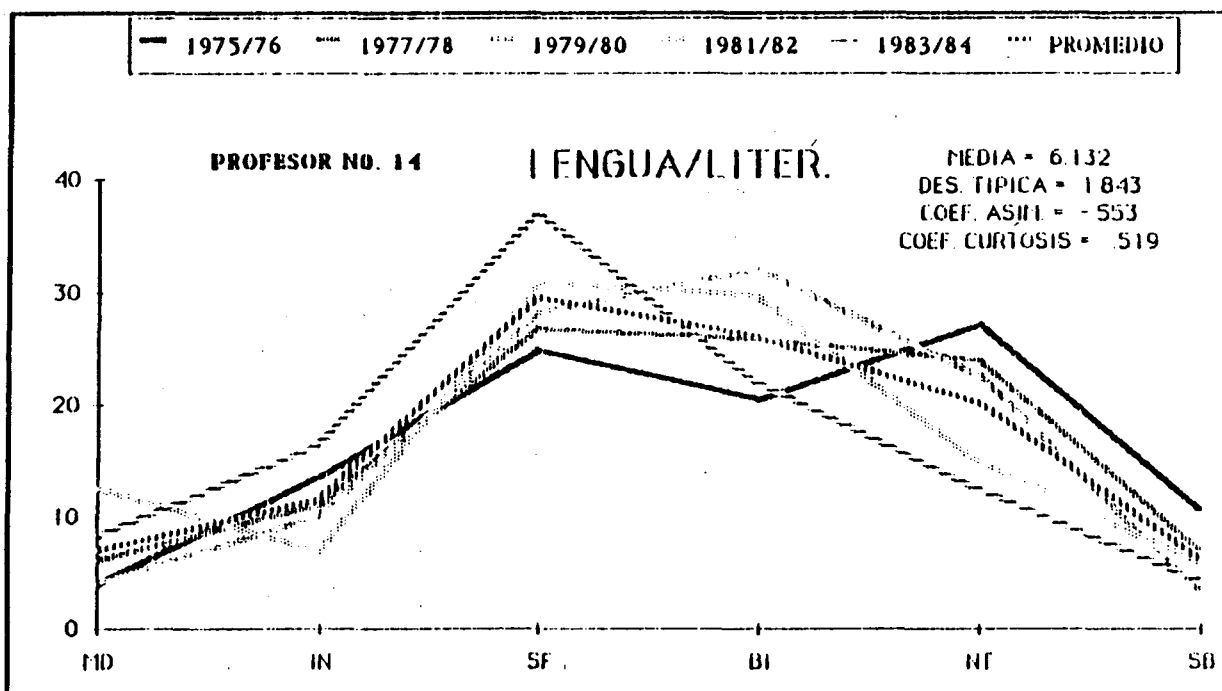
DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



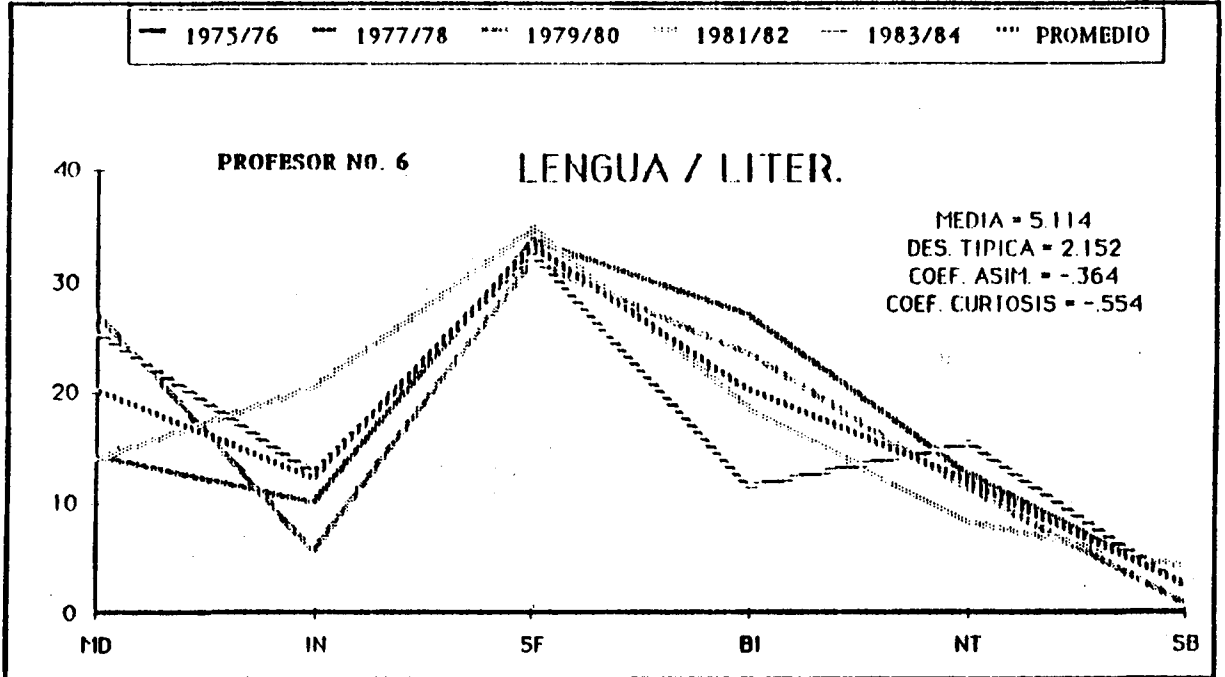
DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO



DISTRIBUCION BIANUAL DEL PORCENTAJE DE CALIFICACIONES EN JUNIO

