

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE CIENCIAS

PROCESO DE IDENTIFICACION TAXONOMICA
POR SELECCION DE CARACTERISTICAS CON
MAXIMO CONTENIDO DE INFORMACION

ESTE LIBRO NO DEBE
SALIR DE LA BIBLIOTECA

T E S I S
QUE PARA OPTAR EL TITULO DE
B I O L O G O
PRESENTA

IRENE VASCONCELOS DUEÑAS

MEXICO, D. F.

1970



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A MIS PADRES
Con todo Cariño
y Agradecimiento

Quiero hacer patente mi agradecimiento a la Dra. Guillermina Yankelevich, al Biol. Hermilo Quero Rico, al Biol. Francisco González Medrano, a la Biol. Neli Diego y a Jesús Lara, en quienes encontré valiosa ayuda y consejos para la elaboración de este trabajo.

CONTENIDO

	Pags.
I.- Introducción	1
Antecedentes sobre la Teoría de Información	4
II.- Métodos:	7
botánico y probabilístico	
III.- Resultados	9
IV. Discusión	11
V.- Conclusiones	14
VI.- Apéndice:	
Tabla 1.- Matriz columna y rectangular de las familias de Fanerógamas, y probabilidad de las características dada la familia.	15
Tabla 2.- Cuadro sinóptico de la vegetación localizada en la parte baja del SENECTIONETUM PRAECOSIS del Pedregal de San Angel.	16
Tabla 3.- Lista de características empleadas en la construcción de las matrices, y cantidad de información aportada cada una de ellas.	18
Tabla 4.- Programa de computadora para el proceso de identificación automatizada.	26
Tabla 5.- Listado de algunos diagnósticos obtenidos con la computadora.	29
Tabla 6.- Diagnósticos obtenidos con la computadora para todos los géneros empleados. Se señala exclusivamente las tres familias para las cuales la probabilidad fue más elevada.	30
Fotografías que ilustran el tipo de Flora del área estudiada en el Pedregal de San Angel.	35
VII Referencias.	40

PROCESO DE IDENTIFICACION TAXONOMICA POR SELECCION DE
CARACTERISTICAS CON MAXIMO CONTENIDO DE INFORMACION

Introducción.-

Muchos son, en el momento actual, los estudios que sobre taxonomía cuantitativa han aparecido en la literatura. Algunos de ellos -- fueron elaborados en el siglo pasado, como el de Adanson, contemporáneo de Lineo, quien propuso un conjunto de propiedades básicas que - los individuos deberían reunir para ser agrupados en unidades taxonomicas. En la actualidad todavía algunos procedimientos de clasificacion taxonomica cuantitativa, se basan en los principios Adansonianos (taxonomía numérica) (17).

A pesar de que la taxonomía incluye, además del proceso de clasificacion, el procedimiento de nomenclatura, el de identificacion y el que corresponde al estudio de los procedimientos y bases teóricas que rigen la clasificacion, es sorprendente observar que la mayor - parte de los estudios de taxonomía cuantitativa se han dirigido fundamentalmente al proceso de clasificacion taxonomica propiamente dicha. Así, la aplicacion de técnicas de Cromatografía, Serología, Biología molecular, como también los de taxonomía numérica y algunos otros procedimientos basados en principios estadísticos, se ocupan exlusivamente de crear métodos cuantitativos para evaluar el grado de semejanza, o la magnitud de las diferencias entre individuos o unidades taxonomicas, con objeto de agruparlos en forma objetiva.

Probablemente, de entre los procedimientos cuantitativos mencionados, solamente la taxonomía numérica se ha ocupado, aunque de manera somera, en analizar el problema de la nomenclatura(16) y también en forma importante del establecimiento de las bases teóricas del -- proceso de clasificacion taxonomica (2) (15).

Sin embargo, sorprende encontrarse con la circunstancia de que ninguno de los métodos cuantitativos se ha ocupado de abordar el --- problema de la identificacion de los organismos con respecto a los - grupos taxonomicos ya establecidos o de la identificacion de un grupo taxonomico dentro de otras jerarquías más elevadas.

Afirman algunos autores, que la identificacion es un proceso logico derivado del de la clasificacion, de tal forma que una vez lorgrado el establecimiento de clasificaciones naturales, en base a prol

cedimientos objetivos y reproducibles, la identificación sería un proceso automático basado en los mismos conocimientos (17).

Las consideraciones resumidas en el párrafo anterior nos explican por qué, en la actualidad, una gran parte de los taxónomos continúan trabajando con procedimientos de identificación que no han sido modificados en varias décadas; esto significa, que los investigadores en la rama mencionada, no han aprovechado la tecnología moderna que ha elevado considerablemente el nivel de la eficiencia con la cual se procesan los datos en otros campos de trabajo.

Estas técnicas, fundamentalmente las de computación, ya han sido utilizadas en los aspectos de almacenamiento y recuperación de datos, lo cual en sí, representa una gran ayuda para el taxónomo; los trabajos de clasificación e identificación en el momento actual, requieren sin lugar a duda del uso de una "memoria" complementaria de velocidad de almacenamiento y recuperación de información de capacidad casi ilimitada, como lo es una máquina computadora.

En lo que deseamos hacer hincapié, es en que el taxónomo ha utilizado las ventajas de la computación, en los renglones señalados, pero empleando en el proceso de identificación, las mismas claves ya establecidas que, en algunos casos, son en sí mismas, ineficientes -- (6).

Siendo el problema de identificación taxonómica una etapa de rutina de la cual el investigador en taxonomía no puede substraerse, sería deseable que a ella se dedique el menor tiempo posible, mediante el empleo de las técnicas de computación ya mencionadas, pero también utilizando procedimientos de identificación más eficientes en cuanto a la selección del mínimo número de características necesarias para tal fin y, sobre todo, aquéllas que son las más apropiadas.

El simple hecho de postular la selección de un conjunto de características como el más adecuado, implica el conceder distinta importancia a cada una de ellas.

Este procedimiento ha sido muy ampliamente discutido por diversos autores (2) (17), y lo han juzgado inadecuado en el proceso de clasificación taxonómica. Es evidente que el seleccionar las características para formar grupos taxonómicos implica un prejuicio por parte del investigador, que puede influir en el tipo de agrupación generada. Sin embargo, en el diagnóstico taxonómico, como el mismo Sokal señala (17), es deseable conceder "pesos" diferentes a

las características tomando en consideración aquella o aquellas que son excluyentes con respecto a la identificación de un grupo taxonómico.

La utilización de métodos cuantitativos en la identificación taxonómica ya aparece en la literatura, aun cuando los reportes son extraordinariamente escasos, así Rescigno y Maccacaro en (1960) (12), y Möller en (1961) (8), reportan ensayos de procedimientos cuantitativos basados en la teoría de Información, el primero, y en aspectos estadísticos probabilísticos, el segundo, para la identificación taxonómica.

En el trabajo que a continuación se presenta, se propone un método cuantitativo basado en la teoría de Información, para optimizar el procedimiento de identificación taxonómica a través de la selección de aquellas características cuyo elevado contenido de información sobre un grupo taxonómico dado, permite la selección del conjunto mínimo de ellas que permita hacer un diagnóstico adecuado.

Este procedimiento fue propuesto inicialmente por Yankelevich y col. (20), en un trabajo preliminar, quienes utilizaron algunos estadísticos subjetivos en el proceso de identificación taxonómica de subórdenes de Orthoptera.

El presente estudio, fue realizado en las familias de las Fanérogamas del Pedregal de San Angel. En él, todos los estadísticos necesarios para aplicar el nuevo método propuesto, fueron calculados directamente sobre el material colectado en el área mencionada. La identificación inicial del material recolectado, se hizo aplicando los criterios usuales de la identificación botánica, y más adelante, se comparan estos resultados con los obtenidos por el procedimiento automatizado propuesto. Con base en esta comparación se discute la confiabilidad del método y el aumento en la eficiencia que éste permite.

En virtud de que el trabajo que a continuación se presenta se basa, como ya ha sido indicado, en varios conceptos de la teoría de Información, consideramos pertinente hacer una revisión muy somera sobre varios aspectos dentro de este tema, que serán de utilidad para la comprensión y evaluación de los resultados obtenidos.

ANTECEDENTES SOBRE LA TEORIA DE INFORMACION.

Existen varias definiciones sobre el concepto de "cantidad de información". Estas han variado de acuerdo con los diversos campos de trabajo en los que este concepto ha sido manejado.

En virtud de que, en el presente trabajo, únicamente se aborda -- el concepto de cantidad de información derivado a partir de la medida de "cantidad de variedad" en un conjunto, la revisión se llevará a cabo exclusivamente bajo este aspecto.

La cantidad de información contenida en un conjunto (3), es una medida de la dificultad con la que se puede identificar un elemento de dicho conjunto, y por consiguiente, también es función del tamaño del mismo.

En el caso de que el conjunto estuviera formado por un sólo elemento, no existiría dificultad para identificarlo y la información -- sería cero. Considérese ahora el caso de un conjunto formado por dos elementos; se requeriría de una pregunta para determinar de cual de los dos elementos se trata. Si los elementos fueran cuatro en el conjunto, el número mínimo de preguntas necesarias para identificar -- alguno de ellos serían 2.

Puede observarse que el número de preguntas, que pudieran considerarse como función directa del "grado de dificultad en la identificación de un elemento", aumenta en uno, a medida que el conjunto duplica sus elementos; es decir, existe una relación logarítmica entre ambos, que puede expresarse de la manera siguiente:

$$n = 2^I$$

$$I = \log_2 n$$

Siendo (n) el número de elementos del conjunto e (I), el número de preguntas necesarias para identificar un elemento o "contenido de información del conjunto".

Estas consideraciones son válidas únicamente si se emplea una -- estrategia óptima en la selección.

La información contenida en un conjunto, puede expresarse como -- contenido total, o como información promedio por elemento. A esta última, se le ha denominado Incertidumbre del conjunto y se denota con la letra H (3).

Cuando los elementos del conjunto, no poseen la misma abundancia

dentro de él, el cálculo de la incertidumbre debe tomar en cuenta el factor de probabilidad ya que el número de preguntas para identificar un elemento no es función exclusiva del número de elementos diferentes, sino también de su abundancia relativa en el conjunto.

En este caso la incertidumbre es un promedio pesado por el factor probabilidad, como se expresa en la siguiente ecuación:

$$H = - \sum_{i=1}^n p_i \log_2 p_i$$

En caso de que los n elementos fueran equiprobables: $p = 1/n$, y por lo tanto:

$$\begin{aligned} H &= - \sum_{i=1}^n 1/n \log_2 1/n \\ &= - n (1/n \log_2 1/n) \\ &= - \log_2 1/n \\ &= \log_2 n \end{aligned}$$

quedando demostrado que para eventos equiprobables los valores de I y H en un conjunto son iguales.

Además de las dos variables ya mencionadas, la incertidumbre es también función de la dependencia probabilística de los elementos; es to es, si éstos no son independientes entre sí, además de su probabilidad "a priori" de presentarse en el conjunto, hay que considerar en los cálculos de H , la probabilidad condicionada de un elemento con respecto a los otros. En este caso, la ecuación para el cálculo de incertidumbre se transforma de la manera siguiente:

$$H(y/x) = - \sum_{x=1}^n \sum_{y=1}^m p(y/x) \log_2 p(y/x) p(x)$$

Siendo $p(x)$ la probabilidad "a priori" de que se presenten los elementos en el conjunto y $p(y/x)$, que se calcula como $p(y) p(x) / p(x)$, es la probabilidad condicionada de que se presente (y) dada la presencia de (x).

Existen varias unidades de información (4), y ellas también son dependientes del campo de trabajo en donde el concepto se ha usado. -

En el presente caso empleamos como unidad el bit o binit (palabra nemotécnica derivada del Inglés Binary unit), que es la más comúnmente usada y es derivada de la definición de cantidad de información que involucra el concepto de variedad.

El bit o binit, como medida de cantidad de información, tiene la ventaja de que, cuando en un conjunto se tiene la mínima variedad -- (dos elementos diferentes) la información es igual a 1 bit ($\log_2 2=1$), o sea también, la mínima cantidad posible. Por otra parte, tratándose de un sólo elemento en el conjunto o sea , siendo cero la variedad, - la información contenida también es cero ($\log_2 1=0$).

Métodos.-

Las plantas utilizadas en el presente trabajo, fueron colectadas en la zona del Pedregal de San Angel, dentro de los límites de la Ciudad Universitaria, y específicamente, del área comprendida en el costado izquierdo del Jardín Botánico exterior.

Debemos hacer notar que, de toda la vegetación presente, nos ocupamos únicamente del estudio de las Fanerógamas. Estas pertenecen exclusivamente a la parte baja de la asociación Senecionetum praecosis, que es una de las más extendidas en el Pedregal (13).

Para estudiar la vegetación del lugar, se recurrió al método denominado "método del cuadro" (18) (11), el cual parece ser uno de los más ampliamente usados para estos estudios.

En dicha zona se escogió un lugar al azar y ahí se procedió al muestreo correspondiente; para ello se delimitó un "cuadro", el cual, también colocado al azar, tuvo una orientación NNE y SSE. Con este método se obtuvo el área mínima, cuya curva se presenta en la Gráfica 1, en el apéndice se muestran algunas vistas de la flora ahí presente. En base a este dato se continuó el muestreo, el que se concluyó con la delimitación de un segundo "cuadro", del cual se colectaron y registraron todas las Fanerógamas ahí presentes; el material fue trabajado con las técnicas habituales de la botánica y pasó a formar parte del Herbario del Instituto de Biología en Ciudad Universitaria.

La colecta se realizó durante la primavera y verano (del 23 de abril al 4 de septiembre de 1969) y la identificación taxonómica del material se llevó a cabo inicialmente con los métodos usuales de la botánica.

Una vez conocida la naturaleza del material, se procedió a la determinación de 2 estadísticos: a) la abundancia relativa a las familias localizadas en el área, y b) la abundancia de los géneros en cada una de las familias.

En base a estos estadísticos, se elaboraron 2 matrices las que aparecen en la Tabla 1 (ver el apéndice).

Una de ellas, la matriz columna señala las probabilidades de encontrar cada una de las familias dentro del área, y la otra matriz rectangular, muestra las probabilidades de las características (anotadas en el encabezado de la matriz rectangular), dadas las familias en la matriz columna. En la Tabla 3 (ver el apéndice), se presenta una

GRAFICA I

CURVA DE AREA MINIMA PARA EL PEDREGAL DE SN. ANGEL

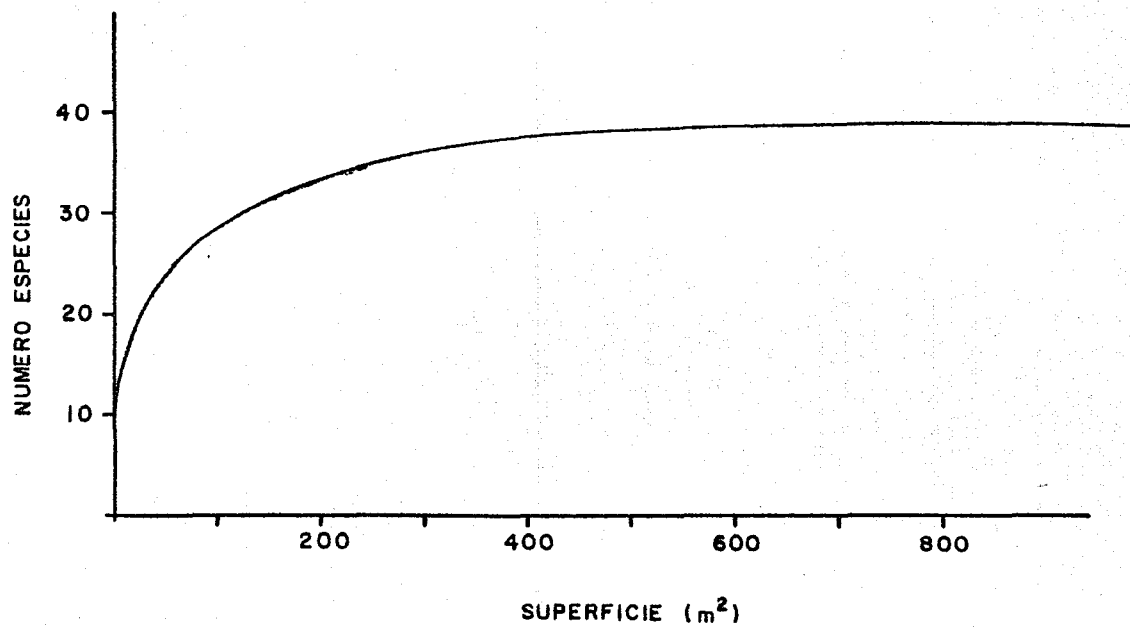


tabla con las características utilizadas, numeradas de acuerdo con el número clave que se les adjudica en la Tabla 1..

La determinación de las probabilidades condicionadas en cuestión, se realizó tomando como base el número de géneros presentes en cada una de las familias que presentaban una característica dada, en relación al total de géneros de la familia.

El valor de la probabilidad condicionada se obtuvo a través de la ecuación de Bayes que a continuación se transcribe:

$$P(Y_i/X) = P(Y_i) \prod P(X/Y_i) / \sum (P(Y_i) \prod P(X/Y_i))$$

en donde $P(Y_i/X)$ es la probabilidad condicionada de una familia (Y_i) dado un conjunto de características (X); $P(Y_i)$ representa la probabilidad a priori de una familia (Y_i); \prod , el producto sobre todos los valores que debajo de la letra se señalan, y \sum , la suma sobre todos los valores que debajo de la letra se indican.

Utilizando los valores de probabilidad a priori y condicionadas obtenidas con los cálculos anteriores, se determinó la incertidumbre a priori de las familias empleando la siguiente ecuación:

$$H(Y_i) = - \sum_{i=1}^n p_i \log_2 p_i$$

donde (p_i) es la probabilidad a priori de una familia (Y_i).

Con los valores de probabilidad condicionada, obtenidos con la ecuación de Bayes, se procedió a calcular la incertidumbre condicionada de las familias, dadas las características que las describen. Este cálculo se llevó a cabo con la siguiente ecuación.

$$H(Y_i/X) = - \sum_{x=1}^n \sum_{y=1}^m P(Y/X) \log_2 P(Y/X) P(X)$$

La diferencia entre los valores de incertidumbre a priori menos la incertidumbre condicionada, se consideró como un índice de la información que aporta una característica en cuestión. Con los valores de información obtenidos de la manera descrita, para cada una de las características, se construyó un histograma de cantidad de información como el que aparece en la figura 1.

Para el proceso automatizado de identificación taxonómica, se construyó un programa para la máquina computadora B 5500, empleando como

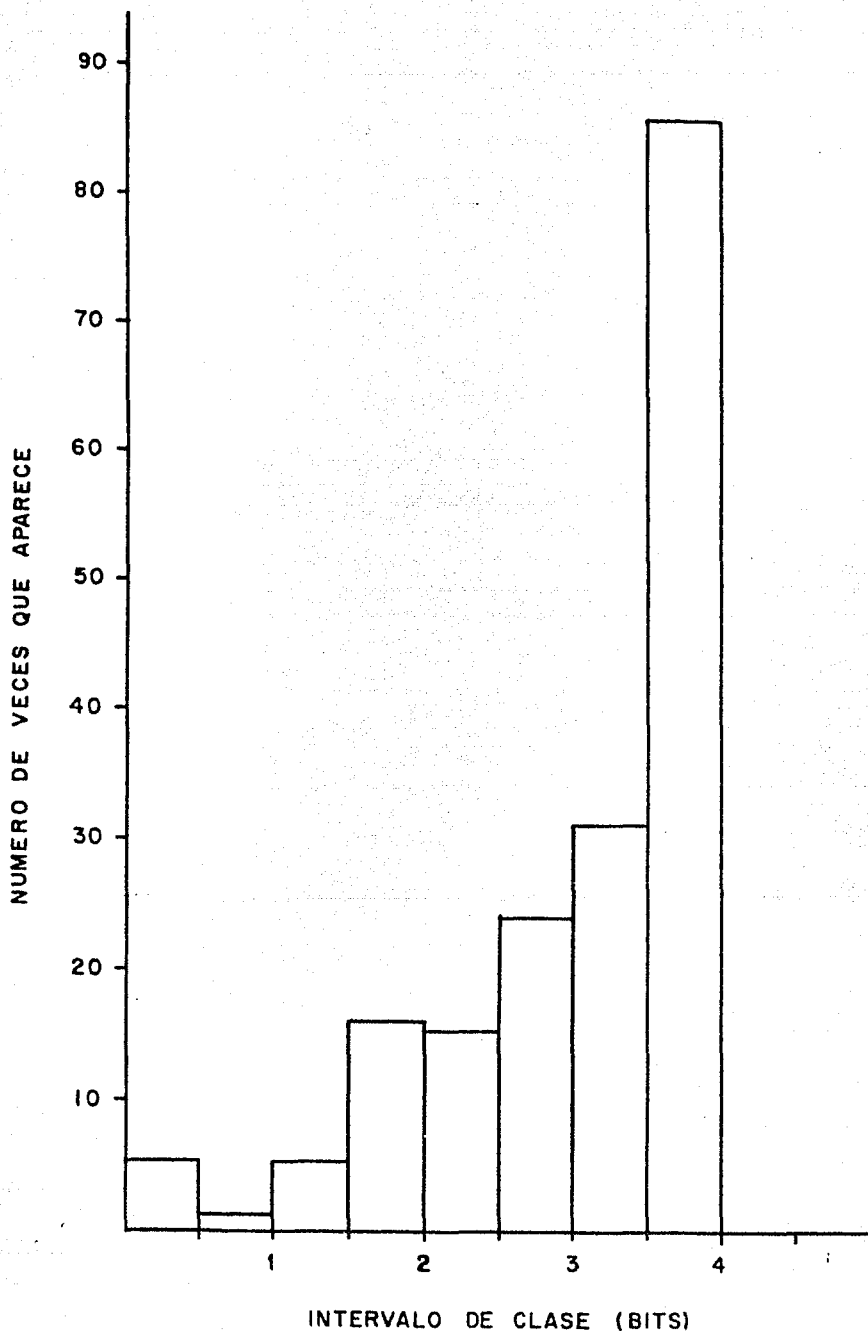


Figura 1 .-

Histograma que relaciona el número de veces que se observa un valor de contenido de información en el grupo de características empleadas.

lenguaje fuente el Algol

Resultados.-

En la Tabla 2 (ver el apéndice) se presenta un cuadro sinóptico con las familias de las Fanerógamas encontradas en la zona estudiada en el Pedregal, así como también los géneros correspondientes a cada una de ellas. Puede observarse, que este cuadro no incluye todas las familias ni todos los géneros que otros autores han reportado para dicha zona (5) (14), sin embargo, como en este trabajo es fundamental el conocimiento de los estadísticos ya nombrados en la parte correspondiente a métodos, nosotros tomamos en consideración, para nuestro estudio, exclusivamente la vegetación que fue colectada. Puede observarse, además, que en la matriz columna de la Tabla 1 no aparecen las familias Portulacaceae, Anacardiaceae y Loganiaceae, en vista de que su probabilidad de aparición en la colecta fue extraordinariamente baja, y por lo tanto, se valoró como cero.

Las características listadas en la Tabla 3 comprenden tanto las observadas en los ejemplares colectados, como las reportadas por algunos autores (14) (5), como pertenecientes a las familias con las que se trabajó y que, por la época del año en que se llevó a cabo dicha colecta, no pudieron observarse directamente en ellos (como lo son las flores y frutos de algunos ejemplares).

Es interesante notar que la matriz rectangular está formada fundamentalmente por valores de cero y de uno; incluso hay características que pudiéramos llamar "claves" que presentan una probabilidad de 1 para cierta familia y cero para todas las demás; así por ejemplo, de este tipo son las características números 6, 10, 13, 27.....

Se presentan también otros caracteres como por ejemplo el 1,9,2, 22, y otros, cuya probabilidad de aparición en las familias, es de valor uno para casi todas ellas.

En la columna del lado derecho de la Tabla 3, se anotan los valores de información calculados para cada una de las características. Estos valores de información se presentan en forma de histograma en la fig 1. Puede notarse en el histograma, la distribución tan amplia de valores de información con que contribuyen las características utilizadas. Estos valores oscilan desde casi cero bits hasta 4 bits de información. Se observa también, que el intervalo de clase con columna más alta, coincide con aquellas características que aportan el máximo contenido de información; esas características corresponden a las de-

nominadas "características claves".

Las columnas del histograma, con escaso contenido de información (las que van entre 0 y 1,5 bits) corresponden a aquellas que aparecen con la misma probabilidad en casi todas las familias.

El programa que se utilizó para el proceso de identificación taxonómica automatizada, se presente en la Tabla 4 (ver el apéndice). El diagnóstico que se realiza a través de este programa se basa en el cálculo, con la ecuación de Bayes de la probabilidad de que un género pertenezca a cada una de las familias consideradas, es decir, no es un diagnóstico de tipo determinístico.

En la Tabla 5 se muestran algunos ejemplos de los diagnósticos obtenidos con la máquina para los géneros trabajados; y en la Tabla 6, se muestra una lista, en donde aparecen todos los ejemplares y las probabilidades más altas que tiene cada uno de los géneros de pertenecer a las familias por diagnosticar.

Para la identificación automática se emplearon únicamente la quinta y sexta columnas de caracteres, representadas en el histograma. El diagnóstico que se tomó como el acertado, fué el de la familia que presentaba la máxima probabilidad. Puede observarse que, de los géneros sometidos a la identificación, 97% de ellos fueron correctamente identificados.

DISCUSION. -

La decisión de estudiar las plantas Fanerógamas del Pedregal para llevar a cabo el presente trabajo, se hizo en función del fácil acceso para el muestreo del material y, además, se contaba con el asesoramiento de expertos en la taxonomía de esta área *. Ya hemos referido que los métodos empleados para muestreo son muy variados y que dicha variación depende del autor y del tipo de estudio que se realice así como de la finalidad que se persiga. En nuestro caso, el método empleado, "método del cuadro", es adecuado para las finalidades que aquí se persiguen pues nos permitió obtener resultados satisfactorios; además, el método postulado en el presente trabajo para diagnóstico automatizado, muestra ductilidad suficiente para poder ser aplicado a cualquier otro conjunto de vegetales o animales de cualquier zona.

El empleo del teorema de Bayes para diagnóstico diferencial, fue postulado inicialmente por Takahashi (19) en diagnóstico médico para características no exclusivas (Síntomas); en nuestro estudio también se hace la misma consideración: las características que describen una familia pueden, en parte presentarse en individuos de otras familias,

Posteriormente, este mismo método ha sido aplicado por otros autores, también en diagnóstico clínico automatizado (9) (10), los cuales obtuvieron un nivel de confiabilidad muy adecuado en sus resultados.

El uso de diagnóstico probabilístico como el que aquí proponemos, pudiera ser utilizado por el taxónomo como un estimador que proporcione información preliminar sobre el grupo taxonómico al que un individuo pertenece. El hecho de poder obtener una escala de probabilidad para todos los grupos manejados, permite al investigador hacer una evaluación preliminar, aún en el caso en el que las probabilidades obtenidas fueran muy semejantes para varios grupos. Este último resultado, además, puede orientar al investigador sobre la insuficiencia en el número de características utilizadas o sobre la inadecuada selección de ellas como ya ha sido reportado en otros trabajos (19) (20).

Nos parece importante hacer hincapié en que, el método propuesto para identificación, emplea también para el diagnóstico la probabilidad a priori de encontrar un individuo de una cierta familia. Es evidente que el conocimiento de este estadístico es importante

* Diego, N. y Quero, H. comunicación personal.

para el investigador, y muy probablemente, lo utiliza subjetivamente en el proceso habitual de identificación.

Como puede observarse en este trabajo, las características empleadas por el taxónomo para propósitos de identificación no son el mínimo que pudiera ser utilizado. En nuestro caso, el uso de 52 características, únicamente, de las 193 encontradas, son suficientes para lograr una identificación correcta de 63 géneros de los 65 utilizados a pesar de que las características empleadas, no fueron las "claves" en ningún caso.

La adición de un par de características "claves" correspondientes a las familias incorrectamente diagnosticadas originará un grupo de características suficiente, aunque no mínimo, para propósitos de diagnóstico.

Las características de tipo "clave" fueron excluidas del estudio ya que, como ha sido discutido anteriormente por Yankelevich y col. (20), la presencia de estas características permite realizar simplemente un diagnóstico de tipo lógico, que no necesariamente es el adecuado por el método empleado por nosotros, que se justifica únicamente en casos de grupos taxonómicos que no posean características "claves" como en algunas de las familias de este estudio o, en circunstancias en las cuales las características "claves" sean difíciles de observar, ya sea porque la época del año no ayude para tal fin, o por el arduo trabajo que en sí su observación implique, o porque la confiabilidad de la observación sea muy baja.

El conocimiento de la información que aporta cada carácter a las familias por diagnosticar, a nuestro modo de ver, representa para el investigador las siguientes ventajas: a) permite seleccionar aproximadamente el número mínimo de características necesarias para una identificación adecuada. Esto hace que el proceso de identificación taxonómica - labor rutinaria para el taxónomo - se haga más eficiente; b) el conocimiento del contenido de información de las características permite seleccionar aquéllas cuyo contenido de información acerca de las familias sea máximo y, por lo tanto, son las más adecuadas para tal propósito. Esto concuerda con lo que se ha observado en el histograma de la Fig. 1, en donde las características con máximo contenido de información son las "claves", que como ya se indicó, son también las más adecuadas en caso de existir, éstas y estar accesibles para el investigador.

El contenido de información que aportan las características es un parámetro que pudiera ser utilizado en la selección de ellas a través del establecimiento de un umbral de información por parte del especialista, esto es, el umbral representaría la mínima cantidad de información que una característica cualquiera debiera contener para ser tomada en consideración; así, por ejemplo, en nuestro caso, un umbral adecuado sería 2.5 bits, y todas aquellas características cuyo contenido de información es menor a este valor, pudiera ser automáticamente desecheda.

A nuestro criterio, una de las contribuciones importantes de este trabajo, es también el programa ya experimentado y optimizado para identificación taxonómica. Este programa está a disposición inmediata de cualquier investigador, dentro del tema, que desee ponerlo en práctica; con pequeñas modificaciones, pudiera también ser utilizado para cualquier otro grupo ya sea zoológico o botánico.

Hay que tomar en cuenta que, a pesar de que el programa, como ya se señaló, es adecuado para el diagnóstico de cualquier otro grupo taxonómico, el investigador debe conocer los estadísticos reales mencionados anteriormente, o cuando menos los estadísticos subjetivos que, aunque no son del todo confiables, permiten diagnósticos aceptables como lo demuestran los trabajos de Negrete y col. (9) (10) en diagnóstico métrico, y los de Yankelevich y col. (20), y Yankelevich (19), en identificación taxonómica.

Por último, creemos necesario insistir en que el valor de información calculado por este procedimiento, es exclusivamente un estimador de información, el cual para su cálculo real requeriría de las probabilidades condicionadas correspondientes a todas las posibles combinaciones de características que generan grupos exclusivos, lo cual representa un trabajo casi imposible de realizar, aún contando con máquinas de cálculo rápido.

Conclusiones:

1) el conocimiento de la información que aportan las características de los individuos, al grupo taxonómico al que pertenecen, permite la selección de un grupo suficiente de ellas para una identificación taxonómica adecuada.

2) el conocimiento de dicho parámetro, permite además seleccionar las características más apropiadas para los propósitos de identificación taxonómica ya que se sugiere que las seleccionadas sean aquellas con máximo contenido de información.

3) la ecuación de Bayes permite un diagnóstico taxonómico correcto, que puede representar una estimación preliminar o una ayuda para el taxónomo en sus trabajos rutinarios de identificación. Este procedimiento tiene además la ventaja de evitar la apreciación subjetiva del que identifica.

4) el programa para el proceso de identificación taxonómico automatizado que en este trabajo se propone, permite un aumento en la eficiencia del investigador en taxonomía, con base en economía del tiempo que dedica a dicha labor, además de la probabilidad del manejo simultáneo de una gran cantidad de datos, que de otra manera no sería posible.

5) se insiste en la necesidad ya señalada en otro trabajo previo (9) de evaluar el grado de dificultad en la obtención de las características, Este parámetro, junto con el que nosotros manejamos (información aportada por las características) permitiría una selección más acertada del conjunto adecuado de características suficientes para un diagnóstico óptimo.

APENDICE

Tabla 2

EMBRYOPHYTA SIPHONOGAMA

ANGIOSPERMAE

Monocotyledoneae		
Ordenes	Familias	Géneros
Glumiflorae	Gramineae	<u>Stipa sp.</u> , <u>Leptochloa sp.</u> , <u>Setaria sp.</u> , <u>Tripsacum sp.</u> , <u>Panicum sp.</u> , <u>Paspalum sp.</u>
	Cyperaceae	<u>Cyperus sp.</u>
Farinosae	Commelinaceae	<u>Commelina sp.</u> , <u>Tinantia sp.</u>
	Liliaceae	<u>Calochortus sp.</u> , <u>Allium sp.</u> , <u>Echcandia sp.</u>
Liliiflorae	Amaryllidaceae	<u>Agave sp.</u> , <u>Manfreda sp.</u>
	Dioscoreaceae	<u>Dioscorea sp.</u>
Microspermae	Orchidaceae	<u>Habenaria sp.</u> , <u>Spiranthes sp.</u>
Dicotyledoneae		
Piperales	Piperaceae	<u>Peperomia sp.</u>
Centrospermae	Amaranthaceae	<u>Iresine sp.</u>
	Portulacaceae	<u>Portulaca sp.</u>
	Caryophyllaceae	<u>Drymaria sp.</u>
Rhoedales	Cruciferae	<u>Lepidium sp.</u>
	Crassulaceae	<u>Echeverria sp.</u> , <u>Tillaea sp.</u>
Rosales	Leguminosae	<u>Eysenhardtia sp.</u> , <u>Phaseolus sp.</u> , <u>Crotolaria sp.</u>

Ordenes	Familias	Géneros
Geraniales	Oxalidaceae	<u>Oxalis</u> sp.
	Burseraceae	<u>Bursea</u> sp.
	Euphorbiaceae	<u>Euphorbia</u> sp.
Sapindales	Anacardiaceae	<u>Schinus</u> sp.
	Sapindaceae	<u>Cardiospermum</u> sp.
Rhamnales	Vitaceae	<u>Cissus</u> sp.
Malvales	Malvaceae	<u>Anoda</u> sp.
Parietales	Begoniaceae	<u>Begonia</u> sp.
Opuntiales	Cactaceae	<u>Opuntia</u> sp., <u>Mammillaria</u> sp.
Myrtiflorae	Lythraceae	<u>Cuphea</u> sp.
	Onagraceae	<u>Oenothera</u> sp.
Umbeliferae	Umbeliferae	<u>Arracacia</u> sp.
Plumbaginales	Plumbaginaceae	<u>Plumbago</u> sp.
Contortae	Loganiaceae	<u>Buddleja</u> sp.
	Asclepiaceae	<u>Gonolubus</u> sp.
	Convolvulaceae	<u>Ipomea</u> sp., <u>Evolvulus</u> sp.
Tubiflorae	Hydrophylaceae	<u>Wigandia</u> sp.
	Labiatae	<u>Salvia</u> sp.
Rubiales	Rubiaceae	<u>Bouvardia</u> sp.
		<u>Parthernium</u> sp., <u>Dahlia</u> sp.
		<u>Baccharis</u> sp., <u>Conyza</u> sp.
Campanulatae	Compositae	<u>Gnaphalium</u> sp., <u>Taetes</u> sp.
		<u>Verbesina</u> sp., <u>Senecio</u> sp.
		<u>Bidens</u> sp., <u>Ageratum</u> sp.
		<u>Fubatorium</u> sp.



Tabla 3

# Clave de las Carac.	CARACTERISTICAS	INFORMACION APORTADA POR LAS CARACTERISTICAS.
22	Hierbas	0.0110
46	hierbas anuales	0.9532
80	hierbas perennes	2.7189
72	hierbas volubles	2.3685
75	plantas leñosas	1.8885
78	árboles balsamíferos, corteza con canales balsamíferos.	3.8779
104	plantas laticíferas	3.8779
124	plantas suculentas	2.4857
125	plantas sin hojas	3.8779
126	grupos de agujones presentes	3.8779
152	arbustos trepadores con los troncos flexibles	3.4037
1	flores actinomorfas	0.2768
51	flores cigomorfas	1.8353
106	flores anodinas	3.8779
108	flores trímeras	2.2009
139	flores hexámeras	3.8779
162	flores tetrámeras	3.1073
163	flores pentámeras	1.3915
29	flores polígamo-dioicas	3.1145
2	flores hermafroditas	0.0608
15	flores unisexuales	1.8986
43	flores dioicas	2.9771
140	flores monoicas	3.4339

115	flores femeninas con perigonio	2.4903
109	flores masculinas con estambres numerosos y dispuestos en cabe <u>z</u> uela	3.8779
159	flores masculinas con perianto infundibuliforme	3.5667
3	perigonio corolino	1.8761
10	perigonio tubular (corto)	3.8779
27	perigonio de pétalos separados hasta la base	3.8779
28	perigonio infundibuliforme	2.5894
53	sépalos y pétalos del mismo color	3.1115
83	quilla presente	3.8779
96	hipsófilas presentes	3.8779
97	Perigonio tubular infundibuliforme	2.8902
116	vilano presente	3.8779
141	sepálos de prefloración valvada	3.8779
147	cáliz de 4-5 divisiones, adheren <u>t</u> es al ovario	3.8779
171	cáliz de sépalos soldados entre sí	2.2562
54	pétalos de igual longitud que los sepálos	3.8779
55	labelo presente	3.8779
64	corola gamopétala tubular	3.0457
68	corola gamopétala en forma de campa <u>n</u> na	3.1895
93	canales resiníferos presentes	3.8779
117	Corola bi o unilabiada tubular	3.8779
120	corolas periféricas alargadas en - lígulas	3.8779

123	pétalos y sépalos unidos formando un eje largo	3.8779
130	corola acampanada e infundibuliforme	2.5723
142	corola de prefloración contorneada	3.8779
150	corola de prefloración valvada	3.8779
159	glumas presentes	3.5667
172	corola de pétalos unidos por su base a una columna estaminal	3.8779
160	2-5 estilos	3.2048
4	ovario ínfero	2.1092
5	ovario trilocular, tricarpelar	1.8912
6	ovario con muchos o pocos óvulos en cada celda	3.8779
11	3 estambres presentes	2.3711
12	estilo tripartido	1.8196
20	6 estambres presentes	1.8536
21	ovario súpero	0.3598
30	estaminodios presentes	2.5529
35	ovario tricarpelar, unilocular, con un sólo óvulo	2.0179
36	2 estigmas presentes	2.9984
44	2-3 estigmas	3.8779
56	ginostegio corto sin pie	3.8779
59	ginostegio presente	3.5272
60	polinia presente	3.5272
61	ovario tricarpelar, unilocular, con varios óvulos presentes.	3.5084

62	3 placentas parientales	3.8779
65	4-5 estambres	3.1055
66	ovario bicarpelar, bilocular	1.9701
42	ovario bicarpelar, unilocular	3.8107
70	estambres unidos a la base de la corola	2.9606
76	Ovario con óvulos péndulos en cada división	3.8779
81	disco presente	2.0149
82	ovario excéntrico	3.1768
84	estambres opuestos a las divisiones del perigonio	3.1575
85	ovario con funículo manifiesto	3.4723
88	placentación central	3.8779
94	ovario semi-infero	3.8779
99	anteras biloculares	2.9606
101	estigma peltado	3.8779
110	Dorso del carpelo estriado	3.8779
111	placenta bipartida	3.8779
112	3 estilos cortos	2.5170
118	estilo bipartido	3.1218
132	4 estambres didínamos	3.8779
135	10-codo estambres presentes	3.1411
143	estambres en doble número que los sépalos	3.8779
92	los estambres más o menos unidos en la base.	3.8779
145	5 estambres presentes, y opuestos a las divisiones de la corola	3.1145

62	3 placentas parientales	3.8779
65	4-5 estambres	3.1055
66	ovario bicarpelar, bilocular	1.9701
42	ovario bicarpelar, unilocular	3.8107
70	estambres unidos a la base de la corola	2.9606
76	Ovario con óvulos péndulos en cada división	3.8779
81	disco presente	2.0149
82	ovario excéntrico	3.1768
84	estambres opuestos a las divisiones del perigonio	3.1575
85	ovario con funículo manifiesto	3.4723
88	placentación central	3.8779
94	ovario semi-ífero	3.8779
99	anteras biloculares	2.9606
101	estigma peltado	3.8779
110	Dorso del carpelo estriado	3.8779
111	placenta bipartida	3.8779
112	3 estilos cortos	2.5170
118	estilo bipartido	3.1218
132	4 estambres didínamos	3.8779
135	10-ode estambres presentes	3.1411
143	estambres en doble número que los sépalos	3.8779
92	los estambres más o menos unidos en la base.	3.8779
145	5 estambres presentes, y opuestos a las divisiones de la corola	3.1145

50	estambres en balancín	3.8779
146	óvulo basal	3.8779
148	estambres insertos en el tubo de la corola	2.8676
151	estambres insertos al lado del disco carnoso	3.8779
155	estilopodio presente	3.8779
156	carpóforo presente.	3.8779
164	estambres tetradínamos	3.8779
165	ovario con falso tabique membranoso	3.8779
166	placentación parietal	2.5278
173	los filamentos de los estambres - se unen formándose así un tubo, - pero su parte apical queda libre (anteras)	3.3309
174	estilo dividido en tantas ramas - como carpelos hay, o en el doble	3.8779
7	plantas con bulbo	2.4244
16	plantas con tubérculo subterráneo	3.2698
170	rizoma rastrero estolonífero	3.8779
149	raíz fasciculada y carnosa	3.8779
23	plantas con rizoma	3.0519
38	raíz envuelta en una vaina	3.8779
52	plantas con fibras radicales	3.7371
17	tallos volubles	3.1055
31	tallos nudosos	3.1466
39	tallos huecos	3.3399
47	tallo tricuetro	3.8779
113	tallos carnosos	2.7327
127	plantas acaulecentes	3.8779
133	tallo prismático (cuadrangular)	3.8779

8	hojas en roseta basilar	1.9908
18	hojas alternas	1.7557
24	hojas sentadas	1.9529
25	hojas paralelinervias	1.6776
32	hojas envainadoras	2.8489
40	lígula presente	2.9984
48	hojas cilíndricas	3.8779
73	estípula presente	2.3170
79	hojas compuestas	1.9195
89	hojas opuestas	1.4794
90	hojas acorazonadas	2.3752
114	hojas reticulínervias	3.8779
137	hojas pinnadas	2.7356
153	zarcillos presentes	3.4037
154	zarcillos e inflorescencias se oponen a las hojas	3.8779
157	pecíolo y vaina	3.8779
168	hojas glabras	1.4403
9	inflorescencia presente	0.5246
19	flores en panícula	2.0218
26	flores solitarias	2.9279
33	inflorescencia cimosa	2.8489
34	flores desnudas	3.4469
41	las glumas con las flores deter_ minan la espiguilla	2.9984
49	flores protegidas por brácteas	3.8779
58	inflorescencia es espádice	3.5745
69	prefloración valvada	3.1895

74	flor panículo- cimosa	3.1575
86	inflorescencia en espiga	2.9833
87	flor en cabezuela o corimbo	3.1604
102	flores verdosas	3.5708
121	cabezuela homógama	3.8779
122	cabezuela heterógama	3.8779
129	prefloración plegada y contorneada	2.8992
131	flores agrupadas en cincinos	3.8779
138	inflorescencia racimosa	3.3673
144	flores en las axilas de las hojas	2.8857
158	flor dispuesta en umbela	1.5688
119	debajo de la inflorescencia hay ho- jas involucales	3.8779
103	flores agrupadas en ciatos	3.8779
13	fruto tricuetro con aristas aladas	3.8779
14	fruto una baya	2.9784
37	fruto cariopsis, con embrión de di- latación lateral	3.8779
45	fruto nuez tricuetra	3.8779
57	cápsula alargada	3.5745
63	fruto capsular	1.1087
77	fruto drupáceo	3.8779
100	fruto aquenio	3.3692
128	cápsula subglobosa	3.8779
136	fruto una legumbre	3.8779
167	fruto una silícula	3.8779
91	hojas papiraceas	3.8779
105	semillas con arilo	3.8779

134	epidermis con pubescencia glandulosa	3.8779
71	epidermis con pubescencia urticante	3.8779
161	pubescencia simple	2.8793
169	tegumento formado por pubescencia estrellada	3.8779

TABLA 4

26

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

CENTRO DE CALCULO ELECTRONICO

COMPILADOR ALGOL BURROUGHS B5500 NIVEL 5 LUNES, 1/26/70, 3 26 PM.

REGIN

1 0000
PRINCIPIO DEL SEGMENTO *****

2

COMMENT	APLICACION DEL TEOREMA DE BAYES A LA TAXONOMIA	2	0000
	NUMERICA, IRENE VASCONCELOS, JESUS LARA ;	3	0000
	INTEGER NFAM, NCARAC, NCASOS, CAMAX, I, J, K, KK ;	4	0000
	INTEGER ARRAY CASOS[1:86, 1:20] ;	4 1	0000
	REAL DENOMI, PIA ; BOOLEAN CHICHOLARA ;	5	0002
	REAL ARRAY PROENI[1:34], PRENSI[1:39, 1:193] ;	6	0002
	PI[1:34], DIAG[1:34] ;	7	0006
	FILE IN TAXDEN(2,10) ;	8	0009
	FILE OUT TAXSAL 17(2,17) ;	9	0012
	FORMAT: FOTAX1(26I3), FOTAX2(11F7.5) ;	10	0016

PRINCIPIO DEL SEGMENTO *****

3

	FOTAX3("PROBABILIDAD A PRIORI DE LAS FAMILIAS"),	11	0016
	FOTAX4(13,F7.5, 13,F7.5, 13,F7.5, 13,F7.5, 13,F7.5,	12	0016
	13,F7.5, 13,F7.5, 13,F7.5, 13,F7.5, 13,F7.5,	12 1	0016
	13,F7.5, 13,F7.5), FOTAX5("PROBABILIDADES	12 2	0016
	CONDICIONADAS"), FOTAX6(213,F7.5, 213,F7.5,	13	0016
	213,F7.5, 213,F7.5, 213,F7.5, 213,F7.5, 213,F7.5,	13 1	0016
	213,F7.5, 213,F7.5, 213,F7.5), FOTAX7("DIAGNOSTICO"),	14	0016
	FOTAX8 ("ERROR", 7F15.6) ;	14 1	0016
	FOTAX9 ("DIVIDE ENTRE CERO") ;	14 2	0016

EL SEG 3 TIENE 97 PALABRAS, SIG, SEG, ES EL 2

LIST TEMADELARA (I,PROEN [I]),TEMADEIRENE	15	0016
(I,J, PRENSI[I,J]) ;	16	0023
LABEL GUILLE ;	16 1	0033
READ (TAXDEN,FOTAX1,NFAM, NCHARAC,NCASOS,CAMAX) ;	17	0033
WRITE (TAXSAL, FOTAX1, NFAM, NCHARAC, NCASOS, CAMAX) ;	17 1	0045
READ (TAXDEN,FOTAX2, FOR I+1 STEP 1 UNTIL NFAM DO PROEN [I])	18	0057
WRITE (TAXSAL, FOTAX3) ;	19	0070
FOR I + 1 STEP 1 UNTIL NFAM DO	20	0073
WRITE (TAXSAL, FOTAX4, TEMADELARA) ;	21	0075
FOR I+1 STEP 1 UNTIL NFAM DO	22	0080
READ (TAXDEN,FOTAX2,	22 1	0082
FOR J + 1 STEP 1 UNTIL NCHARAC DO PRENSI [I,J]) ;	23	0084
WRITE (TAXSAL, FOTAX5) ;	23 1	0098
FOR I + 1 STEP 1 UNTIL NFAM DO	23 2	0101
FOR J + 1 STEP 1 UNTIL NCHARAC DO	23 3	0102
IF PRENSI [I,J] = 0 THEN PRENSI [I,J] + 0.00001 ;	23 4	0103
FOR I + 1 STEP 1 UNTIL NFAM DO	24	0113
WRITE (TAXSAL, FOTAX6, FOR J+1 STEP 1 UNTIL NCHARAC DO	25	0116
(I,J,PRENSI[I,J]))	26	0121
FOR I +1 STEP 1 UNTIL NCASOS DO	27	0134
READ(TAXDEN, FOTAX1,	27 1	0136
FOR J+1 STEP 1 UNTIL CAMAX DO CASOS [I,J]) ;	27 2	0138
FOR I+1 STEP 1 UNTIL NCASOS DO BEGIN	28	0152
WRITE(TAXSAL, FOTAX1, FOR J+1 STEP 1 UNTIL CAMAX DO	29	0153
CASOS [I,J]) ;	29 1	0158
DENUMI + 0 ;	30	0166
FOR J+1 STEP 1 UNTIL NFAM DO	31	0167
BEGIN PIA + 1 ;	32	0168
CHUCHOLARA + TRUE ;	33	0168

```

FOR K + 1 STEP 1 UNTIL CASOS(I,K) SO AND (IF K<20 THEN TRUE
ELSE
CHUCHOLARA + FALSE ) DO
PIA + PIAXPRENSI(I,J,CASOS(I,K))
IF NOT CHUCHOLARA THEN GO TO GUILLE
PI(I,J) + PIA * PROCN(I)
DENOMI + DENOMI + PI(I,J) END J
WRITE(TAXSAL,FOTAX)
IF DENOMI = 0 THEN BEGIN WRITE(TAXSAL,FOTAX)
GO TO GUILLE END
DENOMI + 1/DENOMI
FOR J + 1 STEP 1 UNTIL NFAK DO
BEGIN
DIAG(I,J) + PI(I,J) * DENOMI
WRITE(TAXSAL,FOTAX,1,DIAG(I,J)) END J
GUILLE! WRITE(TAXSAL,IPAGE)
END J
END.

```

34	0169
34 2	0176
34 3	0177
35	0178
36	0184
37	0185
38	0188
39	0192
39 1	0195
39 2	0199
39 3	0200
40	0201
41	0203
42	0203
43	0206
43 1	0217
44	0222
45	0224

EL SEG 2 TIENE 227 PALABRAS, SIG. SEG. ES EL 1

INPUT(4) ES EL SEGMENTO NUMERO 0004 DIRECCION EN LA PRT ES 0056
 BLOCK CONTROL ES EL SEGMENTO NUMERO 0005 DIRECCION EN LA PRT ES 0005
 INPUT(5) ES EL SEGMENTO NUMERO 0006 DIRECCION EN LA PRT ES 0054
 ALGOL WRITE ES EL SEGMENTO NUMERO 0007 DIRECCION EN LA PRT ES 0014
 ALGOL READ ES EL SEGMENTO NUMERO 0008 DIRECCION EN LA PRT ES 0015
 ALGOL SELECT ES EL SEGMENTO NUMERO 0009 DIRECCION EN LA PRT ES 0016

EL SEG 1 TIENE 2 PALABRAS, SIG. SEG. ES EL 0

EL SEG 10 TIENE 69 PALABRAS, SIG. SEG. ES EL 0

NUMERO DE ERRORES DETECTADOS = 0, TIEMPO DE COMPILACION = 17 SEGUNDOS

LONGITUD DE LA PRT = 51 PALABRAS; LONGITUD DEL TOTAL DE SEGS. = 395 PALABRAS;

LONGITUD DE DISCO = 21 SEGS; NUMERO DE SEGMENTOS DEL PROGRAMA = 10

REQUERIMIENTO ESTIMADO DE MEMORIA = 4500 PALABRAS

TABLE 5

69-28 0 0 DIAGNOSTICO	61-22 0 0 DIAGNOSTICO	32 33-53 0 DIAGNOSTICO	36 42 65 68 87100118-38 DIAGNOSTICO	160-49 0 DIAGNOSTICO
10.00002	10.00007	10.00001	10.00000	10.00001
20.00001	20.00004	20.00000	20.00000	20.00001
30.00003	30.00008	30.99999	30.00000	30.00001
40.00001	40.00003	40.00000	40.00000	40.00000
50.00001	50.00003	50.00000	50.00000	50.00000
60.00000	60.00001	60.00000	60.00000	60.00000
70.00000	70.07138	70.00000	70.00000	70.00000
80.00000	80.00000	80.00000	80.00000	80.00000
90.00001	90.00003	90.00000	90.00000	90.00000
100.00001	100.00002	100.00000	100.00000	100.00000
110.00000	110.00000	110.00000	110.00000	110.00000
120.00000	120.00000	120.00000	120.00000	120.00000
130.00001	130.00004	130.00000	130.00000	130.00001
140.00002	140.00005	140.00000	140.00000	140.82214
150.00000	150.00000	150.00000	150.00000	150.00000
160.00000	160.92792	160.00000	160.00000	160.00000
170.00000	170.00000	170.00000	170.00000	170.00000
180.00001	180.00002	180.00000	180.00000	180.00000
190.00000	190.00000	190.00000	190.00000	190.00000
200.00000	200.00000	200.00000	200.00000	200.00000
210.00000	210.00001	210.00000	210.00000	210.00000
220.00000	220.00001	220.00000	220.00000	220.00000
230.09300	230.00000	230.00000	230.00000	230.00000
240.00000	240.00001	240.00000	240.00000	240.17776
250.00000	250.00000	250.00000	250.00000	250.00000
260.00000	260.00001	260.00000	260.00000	260.00000
270.00000	270.00001	270.00000	270.00000	270.00000
280.00000	280.00000	280.00000	280.00000	280.00000
290.00000	290.00001	290.00000	290.00000	290.00000
300.90677	300.00003	300.00000	300.00000	300.00000
310.00000	310.00017	310.00000	310.00000	310.00003

Tabla 6

<u>Géneros</u>	<u>Familias</u>	<u>Probabilidades</u>
<u>Stipa sp.</u>	Gramineae	1.00000 0.00000 0.00000
<u>Leptochloa sp.</u>	Gramineae	1.00000 0.00000 0.00000
<u>Setaria sp.</u>	Gramineae	1.00000 0.00000 0.00000
<u>Panicum sp.</u>	Gramineae	1.00000 0.00000 0.00000
<u>Paspalum sp.</u>	Gramineae	1.00000 0.00000 0.00000
<u>Tripsacum sp.</u>	Gramineae	1.00000 0.00000 0.00000
<u>Cyperus sp.</u>	Cyperaceae	0.00000 1.00000 0.00000
<u>Calochortus sp.</u>	Orchidaceae Liliaceae Convolvulaceae	0.00000 0.99998 0.00002
<u>Echcandia. sp.</u>	Orchidaceae Liliaceae Convolvulaceae	0.00000 0.99998 0.00002
<u>Allium sp.</u>	Orchidaceae Liliaceae Convolvulaceae	0.00000 0.99998 0.00002
<u>Manfreda sp.</u>	Gramineae Amaryllidaceae Dioscoreaceae	0.00008 0.99992 0.00000

<u>Agave sp.</u>	Gramineae	0.00008
	Amaryllidaceae	0.99992
	Umbeliferae	0.00000
<u>Dioscorea sp.</u>	Dioscoreaceae	1.00000
		0.00000
		0.00000
<u>Habenaria sp.</u>	Orchidaceae	1.00000
		0.00000
		0.00000
<u>Spiranthes sp.</u>	Orchidaceae	1.00000
		0.00000
		0.00000
<u>Peperomia sp.</u>	Piperaceae	1.00000
		0.00000
		0.00000
<u>Iresine sp.</u>	Amaranthaceae	1.00000
		0.00000
		0.00000
<u>Lepidium sp.</u>	Cruciferae	1.00000
		0.00000
		0.00000
<u>Echeverria sp.</u>	Commelinaceae	0.00087
	Crassulaceae	0.99904
	Plumbaginaceae	0.00004
<u>Tillaea sp.</u>	Comelinaceae	0.00087
	Crassulaceae	0.99904
	Plumbaginaceae	0.00004
<u>Phaseolus sp.</u>	Leguminosae	0.00000
		1.00000
		0.00000
<u>Crotalaria sp.</u>	Leguminosae	1.00000
		0.00000
		0.00000
<u>Eysenhardtia sp.</u>	Leguminosae	1.00000
		0.00000
		0.00000

<u>Euphorbia</u> sp.	Euphorbiaceae	0.92792
	Orchidaceae	0.07138
	Compositae	0.00017
<u>Cardiospermum</u> sp.	Sapindaceae	1.00000
		0.00000
		0.00000
<u>Cissus</u> sp.	Vitaceae	1.00000
		0.00000
		0.00000
<u>Begonia</u> sp.	Begoniaceae	1.00000
		0.00000
		0.00000
<u>Cuphea</u> sp.	Lythraceae	0.99999
	Asclepiadaceae	0.00001
	Convolvulaceae	0.00000
<u>Oenothera</u> sp.	Rubiaceae	0.90677
	Onagraceae	0.09300
	Compositae	0.00005
<u>Arracacia</u> sp.	Umbeliferae	1.00000
		0.00000
		0.00000
<u>Plumbago</u> sp.	Plumbaginaceae	0.99986
	Rubiaceae	0.00010
	Umbeliferae	0.00004
<u>Gonolobus</u> sp.	Asclepiadaceae	1.00000
		0.00000
		0.00000
<u>Ipomea</u> sp.	Convolvulaceae	1.00000
		0.00000
		0.00000
<u>Evolvulus</u> sp.	Convolvulaceae	0.00000
		1.00000
		0.00000
<u>Wigandia</u> sp.	Hydrophyllaceae	0.99953
	Commelinaceae	0.00043
	Plumbaginaceae	0.00002
<u>Bouvardia</u> sp.	Rubiaceae	1.00000
		0.00000
		0.00000

<u>Oxalis</u> sp.	Oxalidaceae	0.82214
	Umbeliferae	0.17776
	Compositae	0.00003
<u>Bursera</u> sp.	Burseraceae	1.00000
		0.00000
		0.00000
<u>Anoda</u> sp.	Malvaceae	0.99970
	Leguminosae	0.00029
	Compositae	0.00001
<u>Mammillaria</u> sp.	Cactaceae	1.00000
		0.00000
		0.00000
<u>Opuntia</u> sp.	Cactaceae	1.00000
		0.00000
		0.00000
<u>Verbesina</u> sp.	Compositae	1.00000
		0.00000
		0.00000
<u>Dahlia</u> sp.	Compositae	1.00000
		0.00000
		0.00000
<u>Parthanium</u> sp.	Compositae	1.00000
		0.00000
		0.00000
<u>Senecio</u> sp.	Compositae	1.00000
		0.00000
		0.00000
<u>Baccharis</u> sp.	Compositae	1.00000
		0.00000
		0.00000
<u>Bidens</u> sp.	Compositae	1.00000
		0.00000
		0.00000
<u>Conyza</u> sp.	Compositae	1.00000
		0.00000
		0.00000

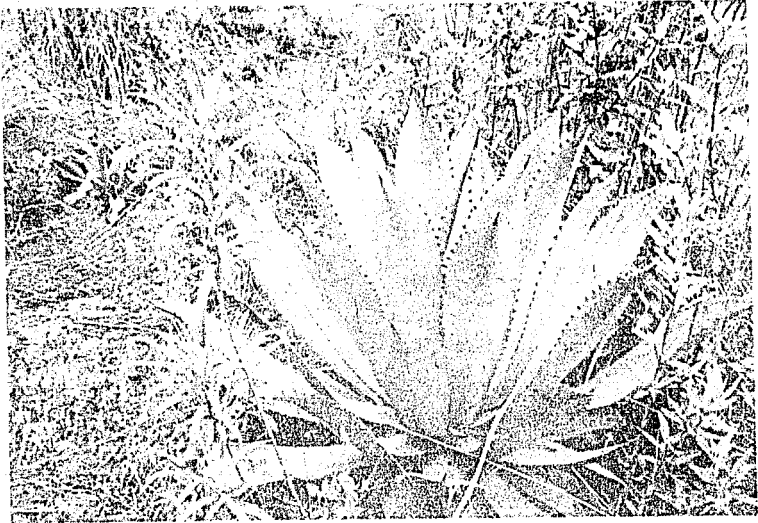
<u>Ageratum sp.</u>	Compositae	1.00000 0.00000 0.00000
<u>Eupatorium sp.</u>	Compositae	1.00000 0.00000 0.00000
<u>Gnaphalium sp.</u>	Compositae	1.00000 0.00000 0.00000
<u>Tagetes sp.</u>	Compositae	1.00000 0.00000 0.00000
<u>Commelina sp.</u>	Commelinaceae: Gramineae Liliaceae	0.99999 0.00001 0.00000
<u>Tinantia sp.</u>	Commelinaceae: Celastraceae Liliaceae	0.99999 0.00001 0.00000
<u>Salvia sp.</u>	Labiataeae Rubiaceae Compositae	0.03191 0.13830 0.82979

VISTAS DE LA FLORA DEL PEDREGAL

- 1) Vista general de la parte baja del Senecionetum praecosis. Puede observarse Senecio sp., (Compositae), Schinus sp., -- (Anacardiaceae) y varias gramíneas.
- 2) Se observan ejemplares de Dahlia sp. (Compositae).
- 3) Varios ejemplares de compuestas, y también Agave sp. (Amaryllidaceae).
- 4) Aparece parte de la delimitación de los "cuadros", se nota la predominancia de gramíneas y compuestas.
- 5) Panorama general del área estudiada.
- 6) Aparecen ejemplares de Commelina sp. (Commelinaceae), Phaseolus sp. (Leguminosae), Oxalis sp. (Oxalidaceae), así como también otros pequeños de Verbesina sp. (Compositae).
- 7) Podemos notar en ésta, la abundancia de las gramíneas; entre ellas sobresalen algunos ejemplares como Verbesina sp. (Compositae).
- 8) Opuntia sp. (Cactaceae).



3



4



5



6



7



8



REFERENCIAS

- 1.- Abramson, N. Teoría de Información y codificación. Ed. Paraninfo, - 1966.
- 2.- Cain, A.J. and Harrison, G.A. An analysis of the Taxonomist's Judgment of affinity. Proc. Zool. Soc. London, 131: 85-98, 1958.
- 3.- Edwards, E. Information transmission. Ed. Chapman S. Hall, 1964.
- 4.- Goldman, S. Information theory. Prentice Hall, 1955.
- 5.- Lawrence, H.M. Taxonomy of Vascular Plants. Mc. Millan Co., New York, 730 pp. 1958.
- 6.- Metcalf, Z.P. The construction of Keys. Systematic Zool., 3: 38-45, 1954.
- 7.- Michener, C.D. and Sokal, R.R. A quantitative approach to a problem in classification. Evolution, 11: 130-162, 1957 (Citado por Sokal.)
- 8.- Müller, citado por Sokal.
- 9.- Negrete, M. J. Olivares L. y Solís, C.P. El uso de estadísticos - subjetivos en la simulación del Diagnóstico Médico por medio de - Computadoras. Bol. Inst. Estud. Med. Biol. 24: 07-108, 1966.
- 10.- Negrete, M.J. Olivares, L. y Solís, C.P. Valoración de la llamada Experiencia acumulada. Noveno Congreso Nacional de Ciencias Fisiológicas. Veracruz, México. p.103, 1966.
- 11.- Oosting, H.J. Ecología vegetal. Ed. Aguilar, Madrid, p.43, 1951.
- 12.- Rescigno, A. and Maccacaro, G.A. The Information content of Biological Classifications. In C. Cherry, (ed.), 1960.
- 13.- Rezdowski, J. Vegetación del Pedregal de San Angel. Tesis profesional. IPN, México. 1954.
- 14.- Sánchez, S.O. La Flora del Valle de México. Ed. Herrero, México - p. 1969.
- 15.- Sneath, P.H. The application of Computers to Taxonomy. J. gen. Microbiol. 17: 201-226, 1957.
- 16.- Sneath, P.H. and Sokal, R.R. Numerical Taxonomy. Nature, 193: 855-886 1962. (citado por Sokal.)
- 17.- Sokal, R.R. and Sneath, P.H. Principles of Numerical Taxonomy. W.H. Freeman Co. H. London, 289pp. 1963.
- 18.- Weaver, J.E. Ecología vegetal. Ed. Acme Agency, Buenos Aires p.11 12, 15, 22, 1950.
- 19.- Yankelevich, G. Aplicación de la Teoría de Información a la Biología, Tesis doctoral, UNAM, México, 1968.
- 20.- Yankelevich, G. y Negrete M.J. El uso del contenido de Información de las características en la Identificación Taxonomica automatizada Bol. Est. Med. Biol. México, 26 (en prensa), 1969.