

Uso metodológico de las tablas de contingencia en la Ciencia Política

Fecha de recepción: 2 de noviembre de 2004. Fecha de aprobación: 2 de mayo de 2005.

*Miguel Ángel Sánchez Ramos**

RESUMEN

Este trabajo presenta la importancia que tienen las tablas de contingencia como herramienta metodológica en la construcción de la Ciencia Política. Se expone que a partir de una tabla de contingencia se puede explorar relaciones que apoyen la explicación de una relación entre variables de estudio o de investigación. Se ofrecen algunos ejemplos que ilustran el uso de la tabla para arrancar de ahí una explicación que se apoye en la teoría. Asimismo se hace un repaso rápido a los coeficientes de asociación que se basan en la tabla de contingencia.

INTRODUCCIÓN

El proceso de construcción del conocimiento de lo social y político exige el carácter científico; es decir, que la investigación se promueva y se realice con características científicas. La necesidad de observar un método científico es ahora una prioridad en la Ciencia Política, deseamos que esta se constituya como tal y encuentre el reconocimiento de los investigadores sociales y no sociales, a la vez que consolide la universalidad de su espacio.

La Ciencia Política como tal data de finales del siglo XIX con el famoso tratado de Gaetano Mosca (Cisneros; 2000: 41), aunque reconoce el mismo Isidro Cisneros que es después de la segunda guerra mundial que en Estados Unidos tiene auge el desarrollo de la Ciencia Política. También Norberto Bobbio (1987: 256) centra el nacimiento de la Cien-

* Coordinador de la Licenciatura en Ciencias Políticas y Administración Pública en la Unidad Académica Profesional Amecameca de la Universidad Autónoma del Estado de México. Maestro en Gobierno y Asuntos Públicos por la UNAM.

cia Política en la segunda mitad del siglo XIX, enfatizando que es en el siglo XX donde se separa la producción de los estudios políticos del tradicional enfoque del derecho. De esta forma la Ciencia Política busca su identidad y autonomía.

La Ciencia Política se debe distinguir de la Filosofía Política porque son espacios del saber peculiares, aun cuando traten el mismo objeto de estudio, a la primera le interesa entenderlo y explicarlo, mientras que a la Filosofía le interesa prescribir lo que debe ser.

Hoy, la Ciencia Política se ha transformado de manera radical, trasladándose en el estudio empírico de ‘lo que ocurre’ en la política y preocupándose sólo de modo indirecto de aquello que ‘debe ser’ en la política. Por lo tanto, la primera distinción analítica que es posible formular es entre Ciencia Política y Filosofía Política (Cisneros, 2000: 41).

Así la Ciencia Política ha ido construyendo su propio espacio y objeto de estudio a través de recurrir a la metodología que le permita comprender y explicar el hecho. Para lograr esto los estudiosos de la Política, preocupados por colaborar en la formación de la ciencia y por hacerla, se apoyan en el método propio de las ciencias empíricas. Por lo tanto, la Ciencia Política es conocida como una ciencia empírica.

...la Ciencia Política en sentido estricto y técnico designa a la ‘ciencia empírica de la política’, o a las ‘ciencias de la política’ conducida según la metodología de la ciencia empírica más desarrollada, como es el caso de la física, de la biología, etc. Cuando hoy se habla de desarrollo de la Ciencia Política se hace referencia a las tentativas orienta-

das, con mayor o menor éxito pero que intentan obtener una gradual acumulación de resultados, a promover el estudio de la política hasta alcanzar el nivel de ciencia empírica (rigurosamente entendida) (Bobbio, 1987: 255).

La concepción de Ciencia Política como ciencia empírica es reciente, en la actualidad se está introduciendo el uso de esta metodología para fincar la Ciencia Política, existen algunas excepciones que se localizan en el desarrollo de las ideas políticas, como es el caso del estudio y obra de Aristóteles cuando hace una revisión de las constituciones de más de 100 estados, Maquiavelo quien es considerado un precursor de la Ciencia Política por haber eliminado la carga valorativa y moral en el estudio del hecho político, Montesquieu, y Tocqueville que realiza su investigación en la misma arena donde se desarrollan las acciones políticas, sus relaciones entre las instituciones y demás fenómenos.

Ahora bien para construir esa Ciencia Política, se requiere del uso de técnicas que permitan obtener la información más fiel posible del desenvolvimiento del hecho político. Dentro de esas técnicas destaca la observación directa, ya sea intensiva o extensiva. Esto produce una gran cantidad de datos que debe procesar y sistematizar el investigador, el cual puede hacer uso de los métodos cuantitativos.

Se entiende por método cuantitativo todo aquel procedimiento racional que sistematiza para el análisis los datos cuantitativos obtenidos de la observación principalmente; su identificación con la esta-

dística es elevada aunque no solamente ésta integra a aquél. Por ejemplo, a parte de la estadística está la simulación, la programación lineal, la ruta crítica, la asignación de recursos, control de inventarios, investigación de operaciones, teoría de juegos y otros más, que producen herramientas para el análisis de datos cuantitativos que a su vez pueden facilitar la explicación de un fenómeno o hecho determinado.

Aquí, para el presente trabajo importa destacar el método estadístico, como una herramienta técnica que auxilia al investigador de la política. Nos proponemos, por lo tanto, describir los métodos estadísticos más recurrentes para el análisis de los datos de una y de dos variables; revisar la importancia que tienen las tablas de contingencia para explicar la relación que se da entre dos variables, una independiente y otra dependiente; y, ejemplificar algunos coeficientes de asociación entre las variables descritas.

Se pretende demostrar que la utilidad metodológica de las tablas de contingencia facilitan la compresión de la relación que guardan dos variables que inciden en la reproducción de un fenómeno político y que a partir de aquí, con el auxilio de la teoría, el politólogo puede explicar el fenómeno.

Para dar cumplimiento a estos propósitos y al supuesto a demostrar, se hizo una revisión bibliográfica, lo más reciente posible así como del medio, hoy por hoy, de comunicación e información más efectivo del mundo: Internet. También se procedió a trabajar con un paquete estadístico (SPSS:

Paquete estadístico para las Ciencias Sociales, por sus siglas en inglés.) para la demostración de los coeficientes y la realización de los cálculos respectivos para los casos hipotéticos que se manejan en el desarrollo de la siguiente explicación.

La exposición, consiste en un primer apartado que presenta a la estadística como disciplina y sus principales métodos. Posteriormente, se trabaja la tabla de contingencia como técnica estadística para el manejo de dos variables. Por último, se presenta una conclusión y reflexión sobre la bondad y limitantes que presentan las tablas de contingencia.

BREVE PRESENTACIÓN DE LA ESTADÍSTICA

Las teorías tratan de dar sentido a los hechos de la realidad y explicarlos. La investigación es la parte creativa de la ciencia. Procura establecer la relación entre variables. La estadística trata de poner a prueba las hipótesis y usa el concepto de error como ruido. Hay una relación entre la medición, los errores de medida y la estadística (Elorza, 2000: 3)

La estadística como la disciplina matemática encargada de la recopilación, organización, presentación, análisis e interpretación de datos, se convierte en un auxilio indispensable para un investigador que busca la información en la realidad, en los hechos, en lo que está pasando.

En la mayoría de las veces, la investigación se refiere a datos de personas, electores, partidos políticos, presupuesto, gasto por persona, ingreso *per cápita*, edad, sexo,

nivel educativo, votos, y otras mediciones que hacen necesario el uso de la estadística como la herramienta que permitirá organizar, presentar, analizar e interpretar esa información con rigurosidad científica.

La estadística, como se ha mencionado, contempla que en toda investigación se pueden cometer errores, sobre todo precisa dos tipos de error: el sistemático y el aleatorio. El sistemático es el que se comete cuando el instrumento de medición no es preciso, por omisión o por exceso en su construcción. Sin embargo, este error se puede corregir si se compara o se aplica una regla que exista para configurar la precisión del instrumento. Por ejemplo, un instrumento de medición en la Ciencia Política es un cuestionario, entonces este debe ser hecho para medir lo que debe medir, y las preguntas deben estar en un contexto tanto teórico como empírico. Debe desprenderse de un marco conceptual que operativizado permita acercarnos al objeto de estudio.

En cuanto al error aleatorio, es el que se comete por accidente, por una falta ocasional por la influencia del medio en el que se aplica la medición, como puede ser presión del sujeto a quien se cuestiona, el estado de ánimo del encuestador o del encuestado, medio ambiente, entre otros.

La estadística conoce de estos errores y la forma de tratarlos. Para el error sistemático contempla una serie de técnicas de investigación y de inferencias que contrastan diferencias entre los valores y un estándar. El error aleatorio se elimina si se usa la estadística, la prueba de significancia¹ tiene

como propósito asegurar que nuestros resultados no se deben al azar, y en sentido contrario, debemos tomar la decisión de rechazar o deshacerse de esos datos que contiene error.

Recordemos que la construcción del conocimiento científico tiene como principal referente la utilización del método científico que permite conocer la realidad observable, y consiste en la formulación de una pregunta como centro de investigación de la realidad, con base en la teoría ya existente, buscando dar soluciones a los problemas planteados.

En el método científico se recopila información, se ordena y se prepara para el análisis. Parte de la observación de una situación determinada en una población específica, de donde es necesario pensar en la posibilidad de recurrir a la muestra para facilitar el estudio, optimizar costos y esfuerzos. Después de esa observación se plantea un problema a investigar, que por lo general se identifica con una pregunta. Una vez planteado el problema de investigación, surge la hipótesis² que es una respuesta tentativa a la pregunta de investigación. La hipótesis intenta explicar la relación causa-efecto entre los hechos. La hipótesis debe ser contrastada y permitir que cualquier otro investigador pueda hacer la contrastación.

Las hipótesis indican lo que estamos buscando o tratando de probar y pueden definirse como explicaciones tentativas del fenómeno investigado formuladas a manera de proposiciones... Las hipótesis pueden ser más o menos generales o precisas, e involucrar dos o más variables, pero en cual-

quier caso son sólo proposiciones sujetas a comprobación empírica, a verificación en la realidad. (Hernández, 1998: 74)

Para Hempel, la hipótesis:

Es cualquier enunciado que esté sometido a contrastación, con independencia de si se propone describir algún hecho o evento concreto o expresar una ley general o alguna otra proposición más compleja (Hempel, 1982: 33)

Como puede observarse a partir de estas dos concepciones de hipótesis, ambas coinciden en que son proposiciones que están sujetas a la comprobación o la contrastación con la realidad, de donde se sabrá si son verdaderas o falsas.

La contrastación de la hipótesis implica el uso de un método. Para facilitar la contrastación, debe haber una operacionalización de términos; es decir, la hipótesis debe ser operacionalizable, que pueda tener su referente en la realidad, que haga accesible a lo empírico.

Debemos resaltar que Bartolini hace énfasis en la construcción de la hipótesis cuando dice que las hipótesis son:

Juicios sobre la relación causal existente entre un número reducido de variables independientes y la o las variables dependientes (Bartolini, 1996: 62)

Tenemos que si una hipótesis debe ser contrastada y además señala una relación causal, debe buscarse una forma idónea que permita hacer esa comprobación de causalidad para cumplir con la contrastación.

Para facilitar la contrastación de la hipótesis, debe recurirse a un método de control

operativo. Bartolini en su artículo sobre la "Metodología de la investigación política" (1996), sostiene que los métodos de control operativo más usados en la Ciencia Política son el experimental, estadístico, comparado y el histórico. Entonces, un método de control operativo para la comprobación de hipótesis en la investigación científica es el método estadístico.

El método de control operativo es el ejercicio científico de transformación de variables independientes. Stefano Bartolini, explica la sustentación de dicho proceso:

El proceso de comprobación de las hipótesis puede y debe someterse a algunas reglas metodológicas. Respecto a cualquier hipótesis de relación causal entre varias variables se presentan, de hecho constantemente dos problemas: 1) cómo aislar una condición causal de la otra de tal manera que se conozca el papel causal independiente de cada variable; y 2) cómo controlar si otros tipos de variables no tenidas en cuenta directamente en la hipótesis influyen en la relación que se estudia y cómo lo hacen. Este proceso, que permite ganar control empírico sobre posibles fuentes de variación de las variables independientes, se basa, en cada tipo de investigación científica, en la transformación en parámetros –o sea, en constantes que no influyen en la relación- de algunas de las variables operativas que se sospecha influyen en las variables dependientes. Mediante la transformación de variables operativas en constantes se actúa de tal modo que la mayor parte de las posibles condiciones causales no varíen, de manera que la influencia de una o algunas condiciones –a las que en cambio se deja variar- pueda aislarse y analizarse. Repitiendo esta operación de transformación de variables en

constantes de modo sistemático para todas las variables que se sospecha puedan influir en la variable dependiente, la explicación se hace más afinada y generalizada (Bartolini, 1996: 62).

Como hemos dicho, es el mismo Bartolini quien ubica tres métodos de control operativo, donde uno de ellos es el que deseamos presentar en este apartado: el estadístico. El método estadístico permite controlar la covariación³ entre las variables que se evalúan. Esto nos da pauta para saber si existe relación entre esas variables.

El funcionamiento de este método se basa en la variación de casos o valores de una variable.

Si se dispone de un gran número de casos, el método estadístico permite la aplicación de técnicas matemáticas al universo o a una muestra de él para obtener los mismos efectos de transformación de variables operativas en parámetros que ya hemos citado como esencial esfuerzo cognitivo (Bartolini, 1996: 65).

Las técnicas estadísticas facilitan el establecimiento de la tipología de la covariación entre las variables. Antes de explicar una de estas técnicas: tabla de contingencia, me referiré a mencionar los grandes grupos de análisis estadístico, atendiendo al número de variables que se deseen manejar.

ANÁLISIS ESTADÍSTICO

El análisis estadístico de los datos, debe atender a tres factores:

a) El nivel de medición de las variables.⁴

b) La manera como se hayan formulado las hipótesis.

c) El interés del investigador. (Hernández, 1998: 342)

Atendiendo al nivel de medición, el análisis estadístico se divide en tres grandes tipos: univariado, bivariado y multivariado.⁵

La característica del análisis univariado es la descripción de una variable por vez, en donde se utilizan las frecuencias o distribución de frecuencias. Se identifica este análisis con lo denominado estadística descriptiva.

La descripción de una variable inicia desde el ordenamiento de la misma si se refiere al tipo numérico. Por supuesto que me refiero a los datos que se obtuvieron en la recopilación, los cuales se pueden presentar a través de tablas o de gráficas, enumerar su promedio aritmético (media), mediana, moda, desviación estándar, valores mínimos y máximos, su rango, su coeficiente de variación. Todas estas medidas sólo aportan atributos de una variable, de ahí que se le llame univariado, es decir de una variable. Por lo tanto, este análisis sólo describe las particularidades que guarda una variable.

La tabla de distribución de frecuencia y la gráfica, son las formas por excelencia para presentar datos. El histograma es un gráfico que se utiliza para variables numéricas continuas, representa la distribución de frecuencia que tiene cada dato o grupo de datos.

Un ejemplo de una tabla de distribución es la tabla No. 1, que representa la distribu-

ción de fondos públicos para el año 2000, como monto al que tuvieron derecho los partidos políticos conforme a ley.

TABLA NO. 1
DISTRIBUCIÓN DE FONDOS PÚBLICOS PARA EL AÑO 2000

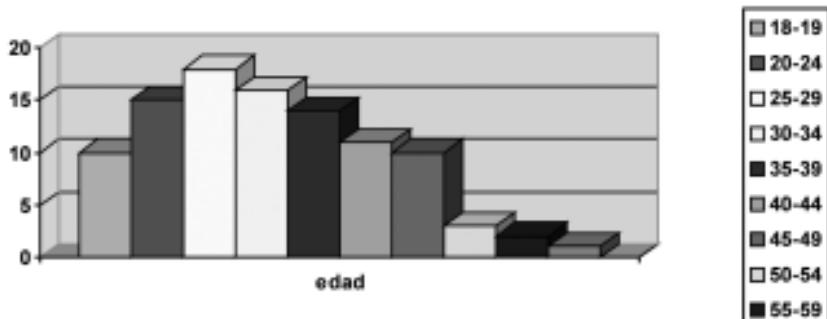
<i>Partido</i>	<i>Monto en pesos</i>
PRI	891 439 746
PAN	657 664 912
PRD	639 913 858
PVEM	229 062 582
PT	205 961 412
PCD	52 480 850
PCD	52 480 850
PSN	52 480 850
PARM	52 480 850
PAS	52 480 850
PDS	52 480 850
TOTAL	2 938 927 614

FUENTE: Crespo (2000:458)

La tabla N° 1, es un clásico de la presentación de datos en donde el único propósito es demostrar o enumerar la frecuencia, cantidad o característica de la variable. En el ejemplo, la variable es el partido político del cual se da a conocer la cantidad de dinero a la que tiene derecho como entidad de interés público para su campaña electoral y su mantenimiento normal. No tiene otra finalidad y, por lo tanto, su propósito no es relacionar variables.

Otro ejemplo, es el gráfico N° 1, que también representa la distribución de una variable, en este caso es la edad por grupo de edad, de un supuesto donde el eje de las abscisas contiene a los grupos de edad y el de las ordenadas a los porcentajes de frecuencia de cada grupo. Sólo se describe una característica del dato. Se identifica la clase con mayor frecuencia o la menor y el tipo de distribución que tiende a tomar la variable.

GRÁFICO NO. 1
DISTRIBUCIÓN DE LA EDAD POR GRUPO



FUENTE: Elaboración propia

En el caso de variables categóricas que son las más comunes en la investigación politológica, se utilizan medidas como la moda o la mediana según sea nominal u ordinal la variable. En este tipo de variables no procede todo tipo de medida estadística. Por ejemplo, cuando en una encuesta de preferencia partidista se le cuestiona al ciudadano cuál es la intención de su voto, la respuesta es considerada como una información categórica, y su valor representa una información nominal a la que estadísticamente sólo se le puede obtener una proporción, que es el resultado de dividir el número total de preferencias a favor de un partido entre el número total de casos.

El análisis bivariado trabaja con las variables independientes y la dependiente. Esto se hace apareando una independiente contra la dependiente y posteriormente otra independiente con la dependiente, así sucesivamente. Entonces, el análisis bivariado, de dos variables, una independiente y otra dependiente, busca la relación que pueda existir entre esas variables. En consecuencia, la información es cruzada, una parte proviene de la variable independiente y la otra pertenece a la dependiente.

En un análisis bivariado, por lo tanto, pueden estarse enfrentando dos variables categóricas, dos numéricas o una y una. Atendiendo al tipo de variable (numérica o no numérica) debe elegirse el tipo de prueba que se aplicará para sustentar una confrontación de hipótesis estadística. En términos generales se recomienda que si las variables son numéricas y

tienen una distribución normal debe aplicarse una prueba paramétrica, pero si las variables son categóricas o no tienen distribución normal, entonces la prueba es de tipo no paramétrica.⁶

En el caso de que las variables sean categóricas, se deben usar las tablas de contingencia en las cuales se distribuyen las variables en forma de columna y en forma de filas.

El análisis multivariado se caracteriza por el tratamiento que hace de trabajar simultáneamente con dos o más variables independientes contra una o más variables dependientes.

Los métodos estadísticos multivariados más recurridos son: la regresión lineal múltiple, la regresión logística, análisis discriminante, análisis multivariado de varianza (manova), análisis de factores, entre otros.

Como la pretensión de este trabajo es destacar la importancia de la utilidad que tienen las tablas de contingencia, estaremos enfatizando sobre el análisis bivariado.

TABLAS DE CONTINGENCIA

Para facilitar el análisis bivariado, se recomienda el uso de las tablas de doble entrada, conocidas en estadística como las tablas de contingencia.

Las tablas de contingencia se integran por filas o renglones y columnas, las cuales forman celdas o casillas en donde se registra la frecuencia absoluta o relativa de cada una de las categorías analizadas. En la tabla de contingencia, entonces, encontramos en una celda la información de la frecuencia que existe en los casos en que se comparte la unidad de análisis. En la tabla N° 2, se ejemplifica una tabla de contingencia en donde se especifica el tipo de voto que se emite, según una encuesta hipotética de los Estados Unidos de Norteamérica.

TABLA N° 2
VOTO PRESIDENCIAL POR TIPO DE RAZA Y PARTIDO POLÍTICO

Voto Partido político	Tipo de raza		Total
	blancos	negros	
	demócratas	135	
republicanos	890	16	906
Total	1349	151	1500

FUENTE: Elaboración propia a partir de datos de Manheim (1999: 346)

Comprendamos lo que la tabla encierra en sí, está compuesta por renglones y columnas, en las cuales existe un encabezado o concepto, para el caso de los renglones se identifican los partidos políticos: demócrata o republicano; mientras que en las columnas se representa al tipo de raza: blanca o negra. Por lo tanto, el renglón contiene información de una variable: voto de partido político; en la columna está incluida otra variable que es tipo de raza. En el interior de la tabla de contingencia se forman cuatro celdas, las cuales contemplan información del número de ciudadanos que tienen la intención de votar a favor de uno de los dos partidos y esto depende del tipo de raza. Así tenemos que 459 ciudadanos blancos tienen la intención de votar por los demócratas y 890 de los ciudadanos blancos lo prefieren por los republicanos, mientras que los negros votan en 135 por los demócratas y 16 por los republicanos.

Como se desprende de esta tabla, no sólo se refiere a información de una variable como lo hizo la tabla de distribución de frecuencias, sino que aquí se ofrece la frecuencia que comparten dos variables al mismo tiempo. Esta es una gran ventaja de la tabla de contingencia que ofrece información sistematizada y lista para poder facilitar al investigador la búsqueda de relación entre esas dos variables.

Es tal vez el modelo de tabla más común utilizado en la investigación contemporánea de la Ciencia Política y proporciona la base para un cierto número de cálculos estadísticos... Las tablas de contingencia se basan más directamente en las hipótesis y están estructuradas de tal manera que facili-

tan el examen de las relaciones entre las variables. (Manheim, 1999: 346)

La tabla de contingencia debe ser utilizada de forma tal que su lectura acerque a la explicación de la hipótesis, por lo tanto, es recomendable que se aparezcan siempre, o en un reporte, bajo un mismo formato o modelo de la tabla; es decir, se puede dar lectura a las columnas como variable independiente y a los renglones como variable dependiente. Este es el caso de la tabla N° 2.

La variable independiente, en el supuesto que contempla la tabla N° 2, es el tipo de raza, es decir, se supone que según el tipo de raza, es el tipo de voto que se ofrece a los partidos políticos en los Estados Unidos.

La tabla, dijimos está integrada por celdas o casillas, las cuales tienen una información que hemos denominado como frecuencia y que más exactamente debe ser denominada como frecuencia observada. Esta es la cantidad de casos que pertenecen a una categoría de una variable independiente y a su vez a otra que es la dependiente, por ejemplo el valor 459 son los casos observados que son blancos y contestaron a la pregunta "Si hoy fueran las elecciones para tal presidente de los Estados Unidos, ¿por qué partido votaría?", a favor del partido demócrata.

Además de la frecuencia observada, en cada casilla se puede obtener la frecuencia esperada y las frecuencias marginales. Estas últimas son porcentajes con relación al total marginal de la columna o del renglón, según la necesidad de lo que se deseé explicar. Las frecuencias esperadas se obtienen dividiendo el total del renglón donde

se encuentra la casilla entre el gran total y al resultado se le multiplica por el total de la columna respectiva de la casilla. Por ejemplo, la frecuencia esperada de la celda demócratas y blancos, es el resultado de $(594/1500) * 1349 = 534.2$. Estas frecuencias esperadas nos indican la distribución que se espera o teórica suponiendo que no hay asociación entre las dos variables que se analizan, es decir, se supone que son independientes las dos variables. El resultado de los cálculos respectivos se puede observar en la tabla N° 3.

La interpretación de esta tabla auxilia bastante para comprender y explicar al fenómeno político o social que tenemos enfrente. En este caso hipotético, se puede decir que del total de blancos, el 34% prefiere al partido demócrata mientras que el resto (66%), lo hacen por el republicano, en otras palabras, este partido es mas fuerte en la raza blanca ya que los negros votan por este partido (republicano) en un 10.6%, contra un 89.4% que los negros emiten a favor de los demócratas.

Así también se puede hacer la explicación en el sentido de renglón, donde se desprende que los demócratas tienen una recepción de preferencias en 77.3% provenientes de blancos y un 22.7% de negros, contra un 98.2% que los republicanos reciben de los blancos y un 1.8% que reciben de los negros. Así se afirma que existe una sospecha del grado de asociación o dependencia entre estas dos variables. En ambos sentidos, renglón y columna, son los republicanos quienes reciben el voto mayoritario de los blancos, por lo tanto a partir de esta observación se debe recurrir a una prueba de ajuste o de asociación y solicitar también el nivel de asociación si es que existe.

Las tablas de contingencia, en consecuencia nos ofrecen información más precisa para poder explicar un fenómeno, estas tablas superan a las de distribución de frecuencia porque tienen un cruce de información entre las variables. Las tablas de contingencia son, además, imprescindibles para hacer un análisis estadístico para demostrar asociación entre variables.

TABLA NO. 3.
VOTO PRESIDENCIAL POR TIPO DE RAZA Y PARTIDO POLÍTICO
(VALORES OBSERVADOS, ESPERADOS Y MARGINALES)

		Tipo de raza		Total
		blancos	negros	
Voto por partido político	Demócratas	Observado	459	135 594
		Esperado	534.2	59.8 594.0
		% marginal del renglón	77.3%	22.7% 100.0%
	Republicanos	% marginal de la columna	34.0%	89.4% 39.6%
		% marginal del total	30.6%	9.0% 39.6%
		observado	890	16 906
Total	Republicanos	Esperado	814.8	91.2 906.0
		% marginal del renglón	98.2%	1.8% 100.0%
		% marginal de la columna	66.0%	10.6% 60.4%
	Total	% marginal del total	59.3%	1.1% 60.4%
		observado	1349	151 1500
		esperado	1349.0	151.0 1500.0
		% marginal de 1 renglón	89.9%	10.1% 100.0%
		% marginal de la columna	100.0%	100.0% 100.0%
		% marginal del total	89.9%	10.1% 100.0%

FUENTE: Elaboración propia a partir de resultados emitidos por el SPSS

Habíamos comentado que el análisis bivariado se apoya en el uso de las tablas de contingencia para ir enfrentando cada variable independiente contra la variable dependiente. Para exemplificar esto puede verse la tabla N° 4, que se refiere al mismo caso de la preferencia partidista pero ahora contra otra variable que puede considerarse en el supuesto como otra independiente que está asociada a la preferencia del voto. Ahora, en la tabla N° 4, se hace referencia a la situación económica y el tipo de voto que se ofrece a las dos opciones de partidos.

En esta nueva tabla se puede comentar que el voto según la posición económica del elector se distribuye así: el elector con posición económica baja vota más por el partido republicano; el que tiene una situación económica alta, vota también en mayor frecuencia por el republicano pero no tan grande la diferencia con respecto al de la baja. Entonces,

frente a esta situación surge una cuestión ¿Qué tan grande debe ser la diferencia para considerar que existe una asociación entre las variables?. Para responder esta pregunta se debe recurrir al análisis estadístico y en concreto, a las pruebas de significado, para tener mayor certidumbre en la relación que guardan las variables.

TABLA NO. 4
VOTO POR PARTIDO POLÍTICO Y POSICIÓN ECONÓMICA

		<i>Posición económica</i>		<i>Total</i>
		baja	alta	
<i>voto por partido político</i>	demócratas	255	339	594
	republicanos	525	381	906
<i>Total</i>		780	720	1500

Este tipo de tablas son muy usadas por los analistas de las elecciones y sobre todo cuando se aplica una encuesta para conocer la tendencia electoral. Gustavo Ernesto Emmerich (2000), en su artículo "Condicionamientos y motivaciones del voto ciudadano en 1999" hace una exposición basada en el uso de tablas de contingencia para ser más explícito. Un ejemplo concreto al que se alude, es el referido a las intenciones del voto según el género del ciudadano mexiquense. Esto puede observarse en la tabla N° 5.

TABLA NO. 5
INTENCIÓNES DE VOTO EN LA ELECCIÓN DE GOBERNADOR DEL ESTADO DE
MÉXICO EN 1999, SEGÚN EL GÉNERO DEL CIUDADANO

Género	PAN/PVEM	PRD/PT	PRI	NS/NC	Total
	N°	N°	N°	N°	N°
Masculino	172	102	189	51	514
Femenino	185	88	265	59	597
Total	357	190	454	110	1111

Fuente: Emmerich (2000: 29)

La tabla N° 5 que Emmerich utilizó para demostrar una hipótesis, según el usual, que dice que los mayores porcentajes de voto entre las mujeres son para el PRI, y que el PRD obtiene la mayoría de sus votos de los varones y que el PAN es inmune a la variable género. En consecuencia, Emmerich recurre a la tabla de contingencia para analizar las intenciones de voto por género y confirma su hipótesis.

Así podemos seguir dando ejemplos de tablas de contingencia, pero pasemos a comentar las pruebas estadísticas que se aplican con base en las tablas de contingencia para cuando se pretende demostrar una hipótesis alterna⁷ de asociación o de dependencia entre dos variables.

La estadística inferencial se encarga de realizar los estudios respectivos para la proyección de la muestra a la población, la inferencia la hace a través de la estimación y del contraste de hipótesis. Para nuestro interés particular del presente trabajo, debemos revisar el contraste de hipótesis.

Un contraste de hipótesis, en estadística significa someter a prueba la hipótesis nula, la que señala ausencia de diferencia o de asociación. Lo que se desea es rechazar la hipótesis nula y aceptar la alterna. También se conoce esto como una prueba de significado.

Si la diferencia entre lo que se espera bajo la hipótesis nula y lo que se observa en la muestra es demasiado grande para atribuirse en forma razonable a la oportunidad, se desprecia la hipótesis nula. Si la diferencia entre lo que se espera y lo que se observa es tan chica que bien se puede atribuir a la oportunidad, se dice que los resultados no son

estadísticamente significativos (Freund, 1993: 349).

La prueba de significado o contraste de hipótesis consta de una serie de pasos:

1. Formulación de hipótesis nula y alterna
2. Establecimiento del nivel de significado con el que se hará la prueba, lo más común es el 1%, 5% y 10%. Se le identifica con α .
3. Se establece el criterio de aceptación o de rechazo.
4. Se realiza el cálculo con el estadístico o estadígrafo correspondiente a la prueba que se quiera aplicar.
5. Se toma la decisión
6. Se formula una conclusión.

Aunque ahora con los programas estadísticos computarizados, se puede obviar la parte de cálculos para realizarlos a través del paquete. Sólo que en un paquete estadístico cuenta la probabilidad que tiene el valor en presentarse, se le conoce como p-valor.

Una prueba idónea para aplicar al caso de asociación de variables es la prueba de Ji cuadrado de Pearson.⁸

La prueba de Ji cuadrado se aplica al caso de que se disponga de una tabla de contingencia con r filas y c columnas correspondiente a la observación de muestras de dos variables X e Y, con r y c categorías respectivamente. Se utiliza para contrastar la hipótesis nula:

H_0 : Las variables X e Y son independientes

Si el p-valor asociado al estadístico de contraste es menor que α , se rechazará la hipótesis nula.

tesis nula al nivel de significación α (Ferrán, 1996: 141).

La prueba estadística de Ji-cuadrada parte del establecimiento de la hipótesis nula que los criterios de dos muestras son independientes, contra la hipótesis alterna: los criterios son dependientes. Así tenemos que si $\alpha = 0.05$ y p-valor es menor a este, entonces se rechaza la hipótesis nula y se acepta la alterna.

Para comprender mejor esta prueba, pasemos a un ejemplo con los datos proporcionados en la misma tabla de contingencia que se representa con los datos de la tabla N° 2.

1. H_0 . No hay asociación entre el voto a partido y el tipo de raza

Ha. Hay asociación entre el voto a partido y el tipo de raza.

2. $\alpha. 0.05$

3. Si p-valor < 0.05 , se rechaza H_0 .

4. Según el resultado emitido por el SPSS v. 10. p-valor = 0.000, lo que significa que es mucho más pequeña que α .

5. Decisión: Se rechaza la H_0 .

6. Conclusión: Con un nivel de confianza del 95% se afirma que el tipo de voto está asociado al tipo de raza de quien lo emite.

De esta forma se demuestra que según el ejemplo hipotético con el que se inició esta exposición, el tipo de voto que el elector estadounidense emite depende de su tipo de raza, o está asociado al tipo de raza.

Veamos el mismo problema sólo que ahora con los resultados de la tabla N° 4, donde la variable tipo de partido sigue y cambia la de tipo de raza por situación económica. Recuérdese que se ha dicho que el tipo de voto es la variable dependiente y el tipo de raza o la situación económica son las variables independientes.

1. H_0 . No hay asociación entre el voto a partido y la situación económica

Ha. Hay asociación entre el voto a partido y la situación económica.

2. $\alpha. 0.05$

3. Si p-valor < 0.05 , se rechaza H_0 .

4. Según el resultado emitido por el SPSS v. 10. p-valor = 0.000, lo que significa que es mucho más pequeña que α .

5. Decisión: Se rechaza la H_0 .

6. Conclusión: Con un nivel de confianza del 95% se afirma que el tipo de voto está asociado a la situación económica del elector.

Si el valor de p-valor es mayor que α , entonces se acepta la hipótesis nula y se concluye que no hay asociación entre esas variables, o bien que la información con la que se cuenta no permite confirmar la asociación que se pretende demostrar.

Una vez que se sabe de la existencia de la asociación, surge una pregunta ¿de qué tamaño es la relación o asociación que se establece entre estas dos variables?, para responder esta cuestión es necesario recurrir a los coeficientes de asociación.

Existen medidas de asociación que indican el grado de relación asociativa que guardan las variables.

Una medida de asociación es la *Phi*. Es usada en tablas de 2 x 2 para escala nominal. Es una medida basada en Ji-cuadrada. Toma valores entre 0 y 1, indican mínimo y máximo grado de asociación. Es decir, para 1 se dice que hay una asociación fuerte entre las dos variables, mientras que para 0, la asociación es nula.

Phi presenta el inconveniente de que puede alcanzar valores superiores a 1 si se usa tabla de $r \times c$, por eso solo debe ser usada en tablas de 2 x 2.

Entonces, si *Phi* es una medida del grado de asociación entre dos variables dicotómicas que se basan en Ji-cuadrada, toma valor próximo a 0 para indicar no asociación entre las variables y valores próximos a 1 para indicar una fuerte asociación.

El *riesgo relativo* es también una medida del grado de asociación entre dos variables dicotómicas, que compara el producto de las frecuencias en la diagonal principal de la tabla con el producto de las frecuencias en la diagonal opuesta mediante el cociente entre ambos. En consecuencia, toma valores positivos y, si las variables son independientes, su valor será próximo a 1 (Ferrán, 1996: 147).

Otra medida de asociación es la *V de Cramer*, se utiliza en una tabla de contingencia de escala nominal $r \times c$. Sus valores se encuentran entre el 0 y 1, dando valores próximos a 0 poca asociación y valores próximos a 1 fuerte asociación. Se dice que la *V de cramer* tiende a subestimar el grado de asociación entre las variables.

El *coeficiente de contingencia* se considera una extensión de *Phi*, dicho coeficiente toma valores entre 0 y un valor C (máximo), donde si r y c son el número de categorías de cada una de las variables: $C(\text{máximo}) = \min(r-1, c-1) / 1 + \min(r-1, c-1)$

Los valores cercanos a 0, entonces, indican no asociación mientras que los próximos a C (máximo) será una fuerte asociación.

La medida *Lambda*, es una de asociación cuyos parámetros son igual que las anteriores, el 0 y 1.

Comparten el error cometido en la predicción de los valores de una de las variables a partir de los valores de la otra con el error cometido en la predicción cuando únicamente se considera la información de la propia variable, tomando valores entre 0 y 1 (Ferrán, 1996: 151)

La *Lambda*, es una medida que es simétrica y asimétrica, es decir, mide la asociación sin enterar cual es a variable dependiente, a esto es lo que conoce como simétrica. Pero cuando considera a cada una de las variables como la dependiente y da un valor de asociación, se llama asimétrica. Por lo tanto, esta medida es más oportuna incluso para definir el tamaño de la asociación.

La fórmula para calcular Lambda es:

$$\lambda_{yx} = (\mathfrak{R}_{\text{máx.}} f_{jk} - \text{máx. } f_k) / (n - \text{máx. } f_k)$$

donde

$\mathfrak{R}_{\text{máx.}} f_{jk}$ = Suma de los valores en cada columna (variable independiente);

$\text{máx. } f_k$ = valor máximo en los totales de las filas (variable dependiente) y n es el tamaño de la muestra.

Veamos un ejemplo que se recupera de una exposición que presenta Guillermo Briones (1998: 246), donde se somete a prueba la hipótesis de que el estado civil (variable dependiente) está asociada al tipo de origen de las personas: urbano o rural (variable independiente). En la tabla N° 6, se resumen los valores de las muestras respectivas.

TABLA NO. 6
ESTADO CIVIL SEGÚN ORIGEN

Estado civil	urbano	rural	Total
Casado	385	190	575
Separado	105	25	130
Viudo	50	185	235
Soltero	200	305	505
Total	740	705	1445

El valor de *Lambda* es $\lambda = (385 + 305 - 575) / (1445 - 575) = 0.132$

Esto significa que con esta información disponible se puede predecir que el 13.20% del estado civil de las personas está asociado a su tipo de origen.

También encontramos el *Coeficiente de incertidumbre*, este es más complejo en su interpretación aun cuando sus valores están también entre 0 y 1 y se usa para tablas de r x c.

Miden la proporción de incertidumbre reducida en la predicción de los valores de una de las variables a partir de los valores de la otra, frente a la predicción cuando únicamente se considera la información de la propia variable, tomando valores entre 0 y 1 (Ferrán, 1996: 152).

Para la escala nominal también encontramos el coeficiente *Kappa* para r x r, es decir, los posibles valores de las dos variables son los mismos. La tabla de contingencia debe ser cuadrada, el mismo número de renglones y de columnas. Sus valores se localizan entre el -1 y 1, que indican, respectivamente, mínimo y máximo grado de acuerdo entre los valores de las dos variables. Así valores de *Kappa* cercanos a -1 son total desacuerdo, los próximos a 1, total acuerdo y los próximos a 0, independencia.

Una vez comentados los coeficientes de asociación para variables nominales, recuperemos el ejemplo donde en ambas tablas: 3 y 4, fue aceptada la asociación. Ahora veamos que nos representan los coeficientes de cada una.

Para el caso donde la variable voto a partido fue puesta frente a tipo de raza y aceptado que hay asociación, tenemos que es una tabla de 2 x 2, que reporta los siguientes coeficientes:

	COEFICIENTE	VALOR
Phi	.341	
V de Cramer	.341	
Coeficiente de Contingencia	.323	
Lambda (simétrica)	.160	
Voto (dependiente)	.200	
Raza (dependiente)	.000	
Coeficiente de incertidumbre (simétrico)	.122	
Voto (dependiente)	.090	
Raza (dependiente)	.186	

Los datos nos demuestran que la asociación es pequeña, es decir, sus valores de los coeficientes están muy poco por arriba de 0, por lo tanto, explican una asociación débil.

Para el caso donde la variable voto a partido fue puesta frente a situación económica y aceptado que hay asociación, tenemos que es una tabla de 2 x 2, que reporta los siguientes coeficientes:

	COEFICIENTE	VALOR
Phi	.147	
V de Cramer	.147	
Coeficiente de Contingencia	.145	
Lambda (simétrica)	.064	
Voto (dependiente)	.000	
Sit. Ecc. (dependiente)	.117	
Coeficiente de incertidumbre (simétrico)	.016	
Voto (dependiente)	.016	
Sit. Ecc. (dependiente)	.016	

Entre estas dos variables, la asociación es aún más débil, sin embargo, cuando los coeficientes son asimétricos, si aceptan que la situación económica sea una variable dependiente, aun cuando su valor esté muy próximo a 0.

Antes de terminar, debemos decir que también existen medidas de asociación para variables que son ordinales, donde sobresalen: *Gamma*, *Tau-b de Kendall*, *Tau-c de Kendall* y *D de Somers*.

"*Gamma* es una medida del grado y tipo de asociación entre dos variables cualitativas en escala ordinal, que toma valores en el intervalo (-1, 1). Valores próximos a 1 indican fuerte asociación positiva: a medida que aumentan los valores de una variable, aumentan los de la otra; valores próximos a -1 indican fuerte asociación negativa: a medida que aumentan los valores de una variable, disminuyen los de la otra, y valores próximos a 0 indican no asociación positiva ni negativa, lo que no significa que no pueda existir otro tipo de asociación. Aunque su valor es sencillo de interpretar, *Gamma* presenta el inconveniente de que puede alcanzar valores 1 o -1 en situaciones en las que la asociación no es total" (Ferrán, 1996: 155)

La *Tau-b de Kendall* se dice que es un coeficiente de extensión de *Gamma*, dado que su interpretación es igual, empero, la *Tau-b de Kendall* sólo alcanza valores de -1 y 1 para situaciones de total asociación. Su inconveniente es que sólo se aplica a tablas cuadradas, es decir, en aquellas variables que tienen el mismo número de categorías.

En el caso de variables con distinto número de categorías, se utiliza la *Tau-c de Kendall*, que es una corrección de la *Tau-b de Kendall*. Sus límites se encuentran entre -1 y 1. Se dice que este coeficiente subestima la asociación.

La *D de Somers* actúa para versión asimétricas y también para la simétrica. Toma valores entre -1 y 1, se interpretan como en los coeficientes anteriores. También tiene el inconveniente de ofrecer valores de 1 y -1 cuando la asociación no es total.

En la tabla N° 4 se detecta que son variables una nominal y la otra, la independiente es ordinal dado que se refiere a la situación económica categorizada como baja y alta, por lo tanto es más conveniente obtener coeficientes propios para variables nominales y ordinales.

Los resultados de estos coeficientes se presentan a continuación:

COEFICIENTE	VALOR
D de Somers	-.147
Tau-b de Kendall	-.147
Tau-c de Kendall	-.144

Nótese como este tipo de coeficientes son más oportunos, pues nos indican hasta el tipo de asociación que guardan las variables, el signo negativo nos señala que mientras cambia la variable de la situación económica de baja a alta, el voto disminuye de demócrata a republicano.

Existe un coeficiente más, el *Eta*, que es una medida recomendable para el análisis de una asociación entre una variable de intervalo y una cualitativa o nominal. Sus valores se encuentran entre 0 y 1. Los valores cercanos a 0 indican que el comportamiento de las variables son independientes, cercano a 1 indicarán alta dependencia.

Si las variables en cuestión son ambas numéricas o de intervalo, entonces lo más apropiado estadísticamente es una correlación o regresión. Empero, en la mayor parte de los problemas que se analizan en la Ciencia Política cuando menos una variable es de tipo nominal.

Por lo que se puede apreciar, las tablas de contingencia son una herramienta metodológica que auxilia al investigador, no sólo al politólogo sino también al administrador y cualquier científico social. Las tablas de contingencia como técnica estadística es un gran recurso para someter a prueba hipótesis de asociación o de dependencia muy recurrentes en la explicación social y política.

La tabla de contingencia se constituye a partir de las frecuencias observadas por cada una de las variables que se quieren analizar, teniendo una intersección de frecuencias que corresponden a las categorías de cada variable. Las frecuencias observadas para cada celda o intersección entre las categorías de las variables, se obtienen con mucha facilidad a través de los paquetes (software) estadísticos como el caso del SPSS, con esto se optimiza esfuerzo, tiempo y dinero. Además, recordemos que el SPSS no sólo hace el cruce de frecuencias sino que también está programado para rendir las pruebas de significado, los coeficientes de asociación y hacer cómputos que se requieran.

Por mucho, las tablas de contingencia, superan a las simples tablas de frecuencia; no por esto estas últimas no se usen, pues estas son útiles para la descripción de un

hecho, fenómeno o situación en particular; empero, las tablas de contingencia sirven para buscar, explorar explicaciones más complejas.

CONCLUSIONES FINALES

La política en su desarrollo como ciencia ha pasado por varias etapas y por diferentes escuelas y enfoques que tienen la intención de formar del estudio de la política una ciencia. Se destaca la teoría política, la corriente normativa, la institucional, la conductista y, recientemente la de la elección racional.

La Ciencia Política en últimos tiempos, pasando primero en Europa donde se destaca Inglaterra, en Estados Unidos y muy después en Latinoamérica, se está iniciando un trabajo de carácter científico para el análisis de lo político. Se han introducido elementos de reflexión para impulsar el carácter científico que exige en esta época la política. Debemos tratar de formar hipótesis que puedan ser contrastadas en la realidad, y que no con discurso se quieran sostener. Las hipótesis constituyen el marco o la guía que conduce la investigación y sobre lo que debemos probar.

Si se logra que en la política se compagine la teoría y la realidad, se estará produciendo un avance significativo en la formación de la Ciencia Política. Para esto se requiere el uso de metodologías adecuadas.

Para contrastar una hipótesis es imprescindible su estructura con base en la operativización de los términos y conceptos, para facilitar la comprobación empíri-

ca y el sustento teórico de lo que se desea comprobar. Por lo tanto, se debe dar entrada con apertura amplia a la nueva conceptualización de la metodología para la Ciencia Política, debe comprenderse el uso de los métodos de control y diferenciarlos de los enfoques teóricos necesarios para la explicación de un fenómeno.

Los métodos de control operativo tienen por finalidad el controlar la forma, recursos y procedimiento para la comprobación de hipótesis. Si no incluimos un método de control, no habrá una guía clara que permita reducir factores de confusión en la investigación y por lo tanto ofrecer conclusiones que no sean verdaderas.

Uno de esos métodos de control operativo es el estadístico, el cual, por su puesto que se basa en la estadística. Esta como una disciplina, tiene una estructuración sólida para analizar las variables y datos según la clasificación de estos en nominales, ordinales o de intervalo y, además según el número de variables que se quieran integrar en el análisis. Si se trata de una variable promueve el uso de técnicas de análisis univariado; si son dos variables se utilizan técnicas de análisis bivariado; y son más de dos entonces existe el análisis multivariado.

En este trabajo nos propusimos hablar de la tabla de contingencia como herramienta elemental de las técnicas de análisis bivariado. En estas tablas de contingencia, como tablas de doble entrada compuestas por renglones y columnas permite integrar información que se lee en dos sentidos, horizontal y vertical, dando un trato, si así

se quiere, a una variable como dependiente y la otra como independiente.

Con esto el científico político puede acceder a una explicación de los fenómenos o hechos que está estudiando basándose en datos reales que obtuvo directamente o que tienen otra fuente de recopilación.

Al principio de este trabajo se estableció como hipótesis a comprobar que la utilidad metodológica de las tablas de contingencia facilitan la compresión de la relación que guardan dos variables que inciden en la reproducción de un fenómeno político y que a partir de aquí, con el auxilio de la teoría, el politólogo puede explicar el fenómeno. Consideramos, por tanto, que con algunos ejemplos muy reducidos, porque no se trataba de ofrecer múltiples casos, sino de explicar como se integra una tabla de contingencia y que tiene una ventaja sobre la de distribución de frecuencia y que por su estructura facilita entender la relación que guardan dos variables. Esto, en consecuencia garantiza al investigador una forma sencilla de buscar explicación del fenómeno relacionado.

Las tablas de contingencia puede usarse en problemas de elecciones para conocer la intención del voto y quién vota, los resultados electorales y las condiciones geográficas de los votantes, la situación económica de los pueblos y su nivel de desarrollo político, el tipo de voto parlamentario y la materia de la votación por partido o en conjunto, las condiciones de una comunidad y el tipo de demanda para elaborar una política pública, el nivel de rendimiento escolar y la situación económica de los estu-

diantes o su posición alimenticia o hábitos de estudio. En fin, la enumeración de casos de aplicación puede ser amplia y variada pero seguro estoy de que las tablas de contingencia serán de gran utilidad para el investigador, el cual sólo requiere de la elaboración de una base de datos. La base de datos no es una limitante, dado que todo investigador cuenta con ella, entonces puede construir su tabla de contingencia de forma rápida e incluso someterla a pruebas de significado para probar la independencia de las variables.

Sólo debemos reconocer que las tablas de contingencia y, en general, el análisis bivariado tiene limitantes, ya que su base de análisis es el enfrentamiento o trato de variables de dos en dos. Una variable independiente contra una dependiente, mientras que el análisis multivariado lo puede generar con varias dependientes y otro tanto de independientes. Esta es una limitante seria pues la mayoría de los fenómenos sociales y políticos no son fenómenos determinísticos, por el contrario, se deben a una serie de factores y bajo determinadas circunstancias, entonces el análisis bivariado no permite contemplar esa dinámica de influencia de factores, pero el multivariado si lo permite porque para eso fue construido. En esta materia, multivariado, se ha y se está produciendo para ser aplicado, todavía no se ha generalizado su uso, sobre todo los pioneros son los estadounidenses en este tipo de análisis estadístico.

Empero, el análisis bivariado y las tablas de contingencia son una buena herramienta para el análisis de la asociación entre

variables, y pueden ser partes de una inicial exploración de los posibles factores que influyen en la manifestación de un fenómeno político que después sea explicado a través del multivariado. Solo es preciso decir que esta técnica no tiene la intención de sustituir a la teoría. El dato es un recurso que no sustituye a la teoría, sólo tiene la función de facilitar la comprensión de problemas o de su relación para poder explicarlo con base en la argumentación teórica que se maneje y se conozca.

Deseo que este trabajo contribuya a una comprensión de una técnica estadística que se constituye en una herramienta metodológica para el politólogo.

NOTAS

¹ La significancia o significación estadística es la posibilidad de que una asociación que se ha advertido entre dos variables, basada en el análisis de una muestra, provenga del azar y, por tanto, no exista en la población en su conjunto; Es decir, que los resultados son producto del azar y no de la realidad. Esto se sabe a través de una prueba estadística conocida como prueba de significado o contraste de hipótesis. La probabilidad más común en esta prueba es del 5% de Fisher, pero esto no significa que no se pueda aplicar cualquier otro valor.

² Para que una hipótesis sea digna de una investigación debe reunir lo siguientes requisitos mínimos: 1) debe referirse a una situación social real; 2) las variables deben ser comprensibles, precisas y lo más concretas posibles; 3) la relación de las variables debe ser clara y verosímil; 4) los términos de las hipótesis y la relación planteada entre ellas,

deben ser observables y medibles, o sea tener referentes en la realidad; 5) deben estar relacionadas con técnicas disponibles para probarlas. (Hernández, 1998: 78)

³ Covariación son las relaciones que indican que dos o más conceptos tienden a cambiar simultáneamente, a medida que una variable aumenta la otra aumenta o disminuye. Esta debe diferenciarse de la causalidad en que esta indica que los cambios producidos en una o más variables llevarán a que se produzcan cambios en una o más variables diferentes o conllevarán dichos cambios. Como esta precisión es compleja, Manheim y Rich (1999: 38) recomiendan que para declarar una causalidad debe haber seguridad del cumplimiento de cuatro condiciones: 1) que la causa y el efecto postulados cambien juntos, o covarien; 2) que la causa preceda al efecto; 3) que se pueda determinar un vínculo causal entre causa y efecto supuestos, y, 4) que la covarianza de los fenómenos de causa y efecto no se deba a su relación simultánea con un tercer factor.

⁴ Existen cuatro niveles de medición de las variables que tiene que ver con el tipo de las mismas: nivel nominal, nivel ordinal, nivel intervalar y nivel de razón. De esta clasificación depende también el tipo de análisis y tipo de prueba que se aplique. Puede consultarse Ferrán (1996), García (1999).

⁵ Véase al respecto Rodríguez (2000); García (1999); Ferrán (1996); Manheim (1999).

⁶ Al respecto puede consultarse Siegel (1998), García (1999), Ferrán (1996).

⁷ La hipótesis alterna es una alternativa ante la hipótesis nula, la cual en cierto modo, es el reverso de la de investigación. La hipótesis nula es la que niega la diferencia entre variables, o la que niega la asociación o dependencia entre las mismas, es decir, está

por la independencia o no asociación; no así la alterna que es la propone una asociación o dependencia. Un ejemplo de esto es cuando la hipótesis alterna dice, el candidato del partido A obtendrá el 60% de votos de los ciudadanos, la hipótesis nula afirma que el candidato del partido A no obtendrá el 60% de los votos de los ciudadanos. Otro ejemplo, la hipótesis nula dice que no hay asociación entre el voto para los partidos y el tipo de raza, la hipótesis alterna dice que si hay asociación entre el tipo de voto para los partidos y el tipo de raza.

⁸ El estadístico χ^2 -cuadrado de Pearson se construye a partir de las diferencias entre las frecuencias observadas y las esperadas bajo la hipótesis de independencia. Su estadígrafo es $\chi^2 = \sum (f_{\text{observada}} - f_{\text{esperada}})^2 / f_{\text{esperada}}$. Donde los grados de libertad están dados por $gl = (r-1)(c-1)$.

BIBLIOGRAFÍA

Bartolini, Stefano (1996), "Metodología de la investigación política" en S. Bartolini, L. Morlino, A. Panebianco y G. Pasquino, *Manual de ciencia política*, (Tr. Pilar Chavarri, Ma. Luz Morán y A. Ruiz de Azúa) España, Alianza Editorial.

Bobbio, Norberto (1987), "Ciencia política" en Norberto Bobbio y Incola Matteucci *Diccionario de Política*, t. 1, México, 5^a ed., Siglo XXI.

Briones, Guillermo (1998), *Métodos y técnicas de investigación para las ciencias sociales*, México, 3^a ed., Editorial Trillas.

Canavos C., George (1985), *Probabilidad y estadística. Aplicaciones y métodos*, (Tr. Edmundo Urbina) México, Mc Graw Hill.

Cisneros, Isidro H. (2000), "Ciencia Política" en Laura Baca Olamnedi y otros, *Léxico de*

- la Política*, México, Coeditado por FLACSO, SEP-Conacyt, Fundación Heinrich Böll y FCE.
- Cortes, Fernando y Rosa María Rubalcava (1987), *Métodos estadístico aplicados a la investigación en ciencias sociales. Análisis de asociación*, México, El Colegio de México.
- Crespo, José Antonio (2000), "La reforma electoral pendiente" en *Política y gobierno*, núm 2, vol. VII, México, CIDE.
- Elorza Pérez-Tejada, Haroldop (2000), *Estadística para las ciencias sociales y del comportamiento*, México, 2^a ed. Oxford University Press.
- Emmerich, Gustavo (2000), "Condicionamientos y motivaciones del voto ciudadano en 1999" en Miguel Sánchez Ramos y Rafael Cedillo (coords.), *El proceso electoral mexiquense '99*, Toluca, UAEM.
- Ennis, Philip H. (1985), "La importancia del contexto en la decisión electoral" en Francois Chazel, Raymon Boundon y Paul Lazarsfeld, *Metodología de las ciencias sociales. III. Análisis de los procesos sociales*, Barcelona, 2^a ed., Editorial Laia.
- Everitt, Brian S. y Graham Dunn (1998), *Applied multivariate data analysis*, New York, Halsted Press.
- Ferrán Aranaz, Magdalena (1996), *SPSS para Windows. Programación y análisis estadístico*, México, Mc Graw Hill.
- Freund, John E. y Richard Manning Smith (1993), *Estadística*, México, 4a ed., Prentice may Hispanoamericana.
- García Ferrando, Manuel (1999), *Socioestadística. Introducción a la Estadística en Sociología*, España, Alianza Editorial.
- Glaser, William (1985), "Los mecanismos de la participación electoral" en Francois Chazel, Raymon Boundon y Paul Lazarsfeld, *Metodología de las ciencias sociales. III. Análisis de los procesos sociales*, Barcelona, 2^a ed., Editorial Laia.
- Henry, Andrew y James Short Jr. (1985). "Suicidio, estatuto y ciclos económicos" en Francois Chazel, Raymon Boundon y Paul Lazarsfeld *Metodología de las ciencias sociales. III. Análisis de los procesos sociales*, 2^a Edición, Editorial Laia, Barcelona.
- Hernández Sampieri, Roberto y otros (1998), *Metodología de la investigación*, México, McGraw Hill.
- Jonson, Dallas E. (2000), *Métodos multivariados aplicados al análisis de datos*, (Tr. Brooks Cole Publishing Company), México, International Thompson Editores.
- Lazarsfeld, Paul (1985), "Los efectos recíprocos de las variables estadísticas" en Francois Chazel, Raymon Boundon y Paul Lazarsfeld, *Metodología de las ciencias sociales. III. Análisis de los procesos sociales*, Barcelona, 2^a ed., Editorial Laia.
- Manheim, Jarol B. y Richard C. Rich (1999), *Ánalisis político empírico. Métodos de investigación en ciencia política*, (AR. Rosendo Gallego) España, Alianza Editorial.
- Mendelhall, William (1982), *Introducción a la probabilidad y la estadística*, (Tr. Carlos Segami) Estados Unidos de Norte América, Wadsworth International Iberoamericana.
- Miller, W.L. (1997), "Los métodos cuantitativos" en David Marsh y Gerry Stoker (coords), *Teoría y métodos de la ciencia política*, (Tr. Jesús Cuellar Menezo) Madrid, Alainza Editorial.
- Rasmussen, Shelley (1992), *An introduction to statistics with data analysis*, USA, Brooks Cole, California.

Siegel, Sydney y N. John Castellan (1998),
Estadística no paramétrica aplicada a las ciencias de la conducta, México 4^a ed., Editorial Trillas,.

Tufte R., Edward (1977), *Data analysis for politics and policy*, USA, Prentice-Hall.

Zeisel, Hans (1997), *Dígalo con números*, México, 3^a ed., FCE.

INTERNET

Geofysisk Institutt, [en línea], s/f [consultado durante 2004]. Disponible en <http://www.gfi.uib.no/~nilsg/kurs/notes/nodes57.html>

Rodríguez Yáñez, Jorge R. (editor), *StadHelp* [en línea], 2000 [consultado durante 2004]. Disponible en <http://www.medal.org.ar/stadhelp>

Departamento de Estadística e Investigación, Universidad de Valladolid, *Diplomatura en Estadística* [en línea], s/f [consultado durante 2004]. Disponible en <http://www.est.cie.uva.es/docencia/diplomatura/>

Portal de la Universidad de Barcelona, [en línea], 2004, [consultado durante 2004]. Disponible en <http://www.ub.es>

Portal de SPSS, [en línea], 2004, [consultado durante 2004]. Disponible en <http://www.spss.com>