

# La inferencia estadística en la investigación turística

## *The statistical inference in the tourist investigation*

Alfredo Ascanio<sup>1</sup>

**RESUMEN:** El objetivo de este artículo es demostrar que los hallazgos de estadígrafos de la estadística descriptiva no se pueden trasladar intuitivamente a los parámetros poblacionales o del universo, sin que antes se realicen pruebas o tests de significación estadística.

**PALABRAS-CLAVES:** Estadística descriptiva. Estadística evaluativa. Test de hipótesis.

<sup>1</sup> Profesor titular jubilado del Dpto. de Ciencias Económicas y Administrativas de la Universidad Simón Bolívar, Venezuela. Economista con maestría y doctorado en Ciencia Política. Por más de 12 años fue funcionario internacional en cargos importantes en el Banco Interamericano de Desarrollo, en la Organización de Estados Americanos y en el Organismo Mundial de Turismo. Dirección postal: CCS: 90372, PO BOX 025323, MIAMI, FL, USA, 33102-5323. Correo electrónico: alfredo.ascanio@gmail.com

**ABSTRACT:** The objective of this article is to demonstrate that the findings of parameters of the descriptive statistic cannot be transferred very intuitively to the population findings or universe, without before tests of statistical meaning are made.

**KEY WORDS:** Descriptive statistic. Statistical evaluative. Test of hypothesis.

## Introducción

El investigador en el área del turismo debe ir más allá de la simple descripción de sus hallazgos; es deseable hacer enunciados formulando una hipótesis nula ( $H_0$ ) que debe luego ser comprobada o no y contestar a la pregunta clave: ¿es digna de confianza la aparente diferencia que se ha encontrado en el trabajo de investigación? Para contestar a esta pregunta sólo se puede lograr aplicando los criterios básicos de la estadística evaluativa o inferencial.

Es imposible trabajar con los datos de una población o universo en su totalidad y además los parámetros poblacionales son raramente conocidos, entonces el investigador se ve en la necesidad de trabajar con una muestra representativa del universo; pero al extraer muestras aleatorias de una población, los estadígrafos que se obtienen (promedios, varianza, correlaciones) no necesariamente representan los parámetros de la población.

La única manera de lograr que la muestra sea representativa del universo es utilizando el muestreo aleatorio simple de manera que cada muestra de un tamaño dado tenga exactamente la misma probabilidad de ser elegida y luego calcular la probabilidad de que el valor de cualquiera de esos estadígrafos no rebase los límites establecidos que se obtiene en tablas numéricas ad-hoc y que se relacionan con la curva normal.

Como los investigadores están interesados en demostrar que existen determinadas relaciones entre variables, ello se debe hacer respecto a los parámetros de una población o universo. Como la población es muy grande, es necesario trabajar con las muestras, pero los estadígrafos que se obtengan de allí sólo se pueden trasladar a la población aplicando la estadística evaluativa para conocer la significación estadística o prueba de hipótesis. Es decir, sobre la base de algunas observaciones, debemos reconstruir el fenómeno en su totalidad y ello nos lleva entonces a los tests estadísticos, para saber si la reconstrucción es significativa o no (si es exacta o no lo es).

## La significación estadística

El término significación tiene una gran importancia en la estadística evaluativa y señala que la diferencia entre estadígrafos no puede deberse a un hecho fortuito; por ejemplo, si admitimos que una probabilidad superior al 10% se debe al azar entonces no hay significación y tenemos que aceptar la hipótesis nula ( $H_0$ ) y si la probabilidad se ubica entre 10% y 5% es posible que exista significación estadística, aunque todavía podemos admitir que la hipótesis nula ( $H_0$ ) es dudosa, ya que para que exista significación y la hipótesis nula ( $H_0$ ) se puede rechazar sólo se lograría si la probabilidad aparece entre el 5% y el 1%, claro si la probabilidad es menor del 1% no hay duda que existe una muy alta significación y se admitiría rigurosamente que existen diferencias entre los parámetros analizados, lo cual permite rechazar la hipótesis nula ( $H_0$ ) y aceptar la hipótesis alternativa ( $H_1$ ) o de trabajo.

Existe una cómoda convención que admite trabajar con un nivel de significación del 5%, pues es válido para la mayoría de los hallazgos en la investigación turística. Ahora bien, como toda probabilidad viene expresada en función del número de grados de libertad el cual lleva en cuenta la importancia de las observaciones o datos y funciona siempre como un coeficiente de ponderación, entonces ese grado de libertad o sea el número mínimo de datos que es necesario conocer para reconstruir los parámetros y en la práctica para un dato determinado el número de grados de libertad es igual al número de datos menos uno.

Los estadígrafos más utilizados para hacer test de hipótesis o tests de significación son: el promedio, desviación típica, varianza y coeficiente de correlación simple.

Muchas veces el investigador después de correr un cuestionario a una muestra de turistas para saber, por ejemplo, el dato del promedio de la estadía para un segmento de mercado y el promedio de la estadía para otro segmento, obtiene esos resultados pero lo que interesa es saber si esos valores promedios obtenidos de la muestra es un dato real o ficticio y si aparece también o no en los parámetros de la población de donde se obtuvo la muestra; en resumen: el investigador quiere saber si esos promedios y sus diferencias se pueden utilizar realmente para fijar una determinada política turística promocional.

Si el dato obtenido es digamos 10 noches, ese número 10 puede ser el valor promedio de:  $9 + 10 + 11$ , pero puede ser también un valor promedio de la serie:  $2 + 8$

+ 20; entonces: ¿se puede considerar que ese valor promedio de 10 es representativo de las dos series anotadas anteriormente? Lo anterior sólo lo podemos saber al aplicar un test de significación.

Veamos el siguiente ejemplo para ilustrar el asunto: Supongamos que un investigador entrevista a dos mercados de turistas para evaluar el gasto promedio diario probable sólo de alimentación en hoteles de 5 estrellas de dos muestras segmentadas (turistas de USA entrevistados en el hotel y turistas de USA entrevistados en el Aeropuerto), y que estaría interesado, como condición exigida, que el gasto promedio día sea superior de 27 dólares para poder tomar alguna decisión promocional para ese mercado. Para comprobarlo en un número pequeño de turistas, obtiene 6 datos en cada lugar de la entrevista en A (hotel) y en B (aeropuerto), como aparece en la tabla 1 sobre los datos básicos para este análisis.

Tabla 1 – Gasto promedio diario en alimentación según entrevista en el hotel y en el aeropuerto

Número (Datos)	A (USA)	B (USA)	A menos B = x
1	31	29	+2
2	27	27	0
3	29	25	+4
4	30	29	+1
5	26	28	-2
6	28	24	+4
Totales	171	162	+9
Totales	28,5	27	+1,5

Se observa que el gasto promedio día obtenido para el turista de USA (en A) es de 28,5, mientras que el turista de USA (en B) arroja un valor igual a 27. Resulta, que desde el punto de vista de la estadística descriptiva el dato de USA entrevistado en el hotel sería el mejor y el dato del turista entrevistado en el Aeropuerto no cumple la

condición exigida por el investigador. Pero la pregunta importante es: ¿Se debe admitir que los resultados obtenidos de A y de B son diferentes, como parece que indican los hallazgos, o son idénticos, tomando en consideración los parámetros del universo?

La única manera de dar una respuesta a esta importante pregunta es someter esos hallazgos a un test de significación adecuado. La hipótesis nula ( $H_0$ ) señalaría que: el gasto diario en A y en B son idénticos, o sea que la diferencia entre las entrevistas en A y en B no difiere de cero más que por razones accidentales debido un hecho fortuito.

Así el promedio del conjunto es igual a + 1,5, pero necesitamos dos datos adicionales para hacer un test como son: la desviación estándar y el número de grados de libertad, en este caso:  $6 - 1 = 5$

La desviación estándar o típica de todas las observaciones se calcula con la fórmula:

$$s = \sqrt{\sum (x - \bar{x})^2 / N - 1}$$

También la desviación estándar se puede estimar por el método simplificado; es decir, con la raíz cuadrada de la sumatoria al cuadrado de todas las diferencias encontradas en la última columna del cuadro anterior, multiplicado por la sumatoria de todas las diferencias al cuadrado entre 6 datos, y todo ello entre los grados de libertad igual a  $6 - 1 = 5$ , como sigue:

$$s = \frac{[\sqrt{2^2 + 0^2 + 4^2 + 1^2 + (-2)^2 + 4^2}] - (2 + 0 + 4 + 1 - 2 + 4)^2 / 6}{6 - 1}$$

$$s = \frac{\sqrt{41 - (81 / 6)}}{5}$$

$$s = \frac{\sqrt{27,5}}{5}$$

$$s = \sqrt{5,5}$$

$$s = 2,345$$

Ahora tenemos que hacer un test de significación con el test “t” de Student y con su fórmula, o sea:

$$t = \frac{1,5}{2,345 / \sqrt{6}}$$

$$t = \frac{1,5}{2,345 / 2,445}$$

$$t = 1,5 / 0,959 = 1,564$$

Recordemos que el número de grados de libertad es 6 menos 1 = 5. Ahora buscaremos en la tabla “t” el valor crítico con 5 grados de libertad y obtenemos el porcentaje de probabilidad para alcanzar o pasar la prueba con el valor encontrado igual a:  $t = 1,564$ , el cual arroja la probabilidad de 10% a 20%, pero no la probabilidad del 5% que era nuestro criterio de aceptabilidad o para admitir que existe un adecuado nivel de significación.

Así pues el investigador no puede señalar que los resultados entre los dos segmentos de mercado entrevistados en los dos lugares sean diferentes, o sea que no se puede admitir que el resultado de la entrevista en A (en el hotel) es superior al resultado de la entrevista en B (en el aeropuerto). El resultado real para la población en su conjunto consiste en señalar que, en este caso, es indiferente para una política promocional relativa al gasto turístico considerar el sitio donde se obtenga la información, pues las diferencias que aparecieron en las muestras pudieron ser accidentales. En resumen se acepta la hipótesis nula ( $H_0$ ).

Pero qué hubiese pasado si al realizar la encuesta con otras dos muestras obtene-  
mos los hallazgos, como sigue en la tabla 2.

Tabla 2 – Gasto promedio diario en alimentación según entrevista en el hotel y en el aeropuerto

Número (Datos)	A (USA)	B (USA)	A menos B = X
1	21	17	+4
2	20	18	+2
3	20	18	+2
4	22	16	+6
5	16	14	+2
6	21	13	+8
Totales	120	96	+24
Promedios	20	16	+4

Al hallar la desviación típica o estándar de los aumentos del gasto diario de los dos segmentos de turistas, obtenemos:

$$s = \sqrt{\frac{32}{5}}$$

$$s = \sqrt{6.4}$$

$$s = 2,530$$

$$t = 4 / 2,53 / \sqrt{6}$$

$$t = 4/2,53/2,45$$

$$t = 4/ 1,033 = 3,872$$

El valor calculado de “t” es entonces 3,872, y este dato con el grado de libertad igual a  $6 - 1 = 5$ , en la tabla “t” nos muestra que existe un 2% de probabilidad de que se logre la diferencia entre los gastos de los dos segmentos de turistas según el lugar de la entrevista.

El investigador puede entonces concluir que el gasto promedio diario del turista que procede de USA y que es entrevistado en el hotel, es superior al que es entrevistado en el aeropuerto, y entonces la política promocional se debe basar en los hallazgos encontrados en los hoteles como criterio básico y prioritario.

Con esta introducción lo que queremos dejar bien explícito es la importancia de comprobar si los datos de la estadística descriptiva (los estadígrafos) obtenidos de muestras son significativos (se pueden generalizar a los parámetros de la población) para poder tomar decisiones sin correr riesgos innecesarios.

Pero también muchas veces tenemos que comparar datos de segmentos de turistas que difieren de su lugar de origen para saber qué segmento es superior uno del otro en base al criterio del gasto diario promedio. Tomemos el ejemplo que nos permita comparar las propiedades de ese gasto promedio día para dos tipos de turistas: el que proviene de USA (A) y el que proviene de Europa (B) (Tabla 3).

Tabla 3 – Gasto promedio al día del turista norteamericano (A) y del turista europeo (B)

Para A = x	Para B = z	x al cuadrado	z al cuadrado
17	17	289	289
19	18	361	324
20	18	400	324
24	16	576	256
18	14	324	196
22	13	484	169
120	96	2434	1558
20	16	-	-

Hemos visto que los totales son: 120, 96, 2434 y 1558 y los valores promedios del gasto día son: 20 y 16.

Para determinar la estimación de la desviación típica de las dos muestras de datos, lo haremos de esta manera: 1) estimamos la varianza con el método simplificado, que nos arroja el valor de 5,6 y luego la desviación típica, es decir: la raíz de 5,6, como aparece enseguida:

$$s^2 = \frac{1}{N + N^l - 2} \left[ \sum (x^2) - \frac{(\sum x)^2}{N} + \sum (x)^2 - \frac{(\sum x)}{N} \right]$$

$$s^2 = 1 / 6 + 6 - 2x \left[ 2434 - \frac{120^2}{6} + 1558 - \frac{96^2}{6} \right]$$

$$s^2 = 1 / 10 (56) = 5,6$$

$$s = \sqrt{5,6}$$

$$s = 2,366$$

Ahora bien, el valor calculado de “t” es igual a:

$$t = 20 - 16 / 2.366 * \sqrt{6 * 6 / 6 + 6}$$

$$t = 4 / 2.366 * \sqrt{36 / 12}$$

$$t = 1.691 * \sqrt{3}$$

$$t = 1.691 * 1.732$$

$$t = 2,929$$

En este caso el número de grado de libertad es igual a  $= 6 + 6 - 2 = 10$

Entonces para 10 grados de libertad y un nivel de significación del 5% en la tabla “t” obtenemos el valor crítico o teórico de 2,228. Como en nuestro cálculo hemos obtenido el valor de “t” igual a 2,929, o sea mayor que el dato de la tabla, entonces se debe concluir que la diferencia entre los dos valores promedios es significativo, pues sólo existe el 2% de probabilidad de que las dos muestras pertenezcan a la misma población y sus parámetros; entonces el segmento de turistas A (USA) es superior al segmento del turista B (Europa) en cuanto a las propiedades del gasto promedio día.

## Otras pruebas de hipótesis

### La comparación de varianzas

En la investigación turística a veces se necesita comparar los estadígrafos de dispersión como por ejemplo la varianza, o sea la desviación típica elevada al cuadrado. Dos series de muestras pueden presentar dispersiones y entonces se necesita conocer si son idénticas o diferentes. La prueba adecuada para realizar esto es el test de Fischer y además con el grado de libertad para cada muestra (número de datos menos la unidad). La tabla de Fischer establece los valores críticos para diferentes niveles de significación y para diversos grados de libertad.

Supongamos que existen dos regiones turísticas con diferentes ventas de un mismo producto turístico de aventura y que según los datos recopilados para el primer semestre del año y sus variaciones son como se señala de inmediato. Lo que se desea saber es si la variabilidad de estas ventas es la misma en las dos regiones o si es más mayor en la Región A que en la Región B. Para simplificar el cálculo pongamos números reducidos para calcular con mayor facilidad las varianzas de cada región como aparece en la tabla 4.

Tabla 4 – Ventas de un producto turístico de aventura para dos regiones específicas

Mes	Región A	Región B	Valor al cuadrado para A	Valor al cuadrado para B
Enero	30	42	900	1764
Febrero	1	8	1	64
Marzo	34	40	1156	1600
Abril	17	36	289	1296
Mayo	45	51	2025	2601
Junio	22	43	484	1849
Totales	149	220	4855	9174

$s$  al cuadrado para A =  $4855 - 149$  al cuadrado / 6 y entre 5 =  $1155 / 5 = 231$

$s$  al cuadrado para B =  $9174 - 220$  al cuadrado / 6 y entre 5 =  $1107 / 5 = 221$

La relación de F se establece comparando los hallazgos del resultado mayor de “s” entre el menor, es decir:  $F = 231 / 221 = 1,045$  y recordemos que el grado de libertad es  $6-1=5$  para las dos muestras.

Si ahora vamos a la tabla “F” de Fischer para 5 grados de libertad, allí encontramos estos valores según sus porcentajes de probabilidad de 0,20; 0,10; 0,05; 0,01; 0,001 secuencialmente así: 2,2; 3,5; 5,1; 11,0; y 29,8.

El valor calculado de F fue de 1,045, el cual corresponde según la Tabla de Fischer a un nivel de significación superior al 20% porque es menor que 2,2; entonces se infiere que no hay diferencias significativas entre las dos varianzas y que las variaciones de ventas entre las dos regiones se pueden considerar como las mismas en el universo total.

## La comparación global de frecuencias

Aquí el investigador lo que desea conocer es si la frecuencia observada de un fenómeno es significativamente igual a la frecuencia teórica prevista, o si por lo contrario estas dos frecuencias presentan una diferencia significativa para un nivel de significación dado.

El test para estos cálculos se denomina Chi cuadrado y es muy utilizado en estudios de mercado en el campo del turismo y en especial al determinar preferencias de los turistas por algún lugar determinado. Los datos se presentan en tablas de  $1 \times 2$  o en tablas de  $2 \times 2$  o incluso en tablas de  $2 \times N$  veces. Veamos un ejemplo simple para ilustrar este test.

Una estadística de problemas acontecidos a los turistas en dos alojamientos hoteleros (A y B) muestran que de 102 problemas, 59 problemas han tenido lugar en el Hotel A y 43 en el Hotel B. La hipótesis nula ( $H_0$ ) del investigador es que no existe relación entre el número de problemas por el hecho de que ocurran en el hotel A o en hotel B.

Lo que sigue de inmediato es saber si esa hipótesis nula carece de fundamento y se puede rechazar. Este test se hace con la prueba Chi cuadrado. Lo primero que parece lógico es que si no existe relación entre el número de problemas y los hoteles, deberían repartirse por igual los problemas entre los dos hoteles, digamos unas frecuencias esperadas o teórica igual a:  $51$  y  $51 = 102$  problemas. Veamos las frecuencias observadas (a y b) y las frecuencias esperadas (a prima y b prima):

$$a = 59$$

$$b = 43$$

$$a \text{ prima} = 51$$

$$b \text{ prima} = 51$$

Con estos datos podemos estimar la prueba de Chi cuadrado:

$$X^2 = [(a - a \text{ prima}) - 0,5]^2 / a \text{ prima} + [(b - b \text{ prima}) - 0,5]^2 / b \text{ prima}$$

$$X^2 = [(59 - 51) - 0,5]^2 / 51 + [(43 - 51) - 0,5]^2 / 51$$

$$X^2 = 7,5^2 / 51 + 7,5^2 / 51 = 1,103 + 1,103 = 2,206$$

Los grados de libertad es igual a 2 lugares menos 1 = 1

La tabla teórica Chi cuadrado, para diferentes niveles de significación y diversos grados de libertad, nos permite interceptar 1 grado de libertad y para un nivel de significación del 5% y entonces encontrar el valor crítico de 3,841. Dado que en nuestro estimado el valor encontrado fue de 2,206, o sea menor que el dato crítico, podemos admitir que la hipótesis nula ( $H_0$ ) es correcta, o sea: que no existe razón para suponer que se produzcan más problemas en el hotel A que en el hotel B, si se conocieran los parámetros de la población.

## La comparación entre coeficientes de correlación lineal

La correlación lineal entre dos variables se puede estimar recurriendo a las sugerencias de Student y Fischer cuando afirmaron que: "si el número de pares de datos es pequeño (menos de 20 pares de datos), se puede determinar la significación de la correlación lineal calculando el valor del índice "t", según la siguiente fórmula:  $t = r / \text{Raíz de } 1 \text{ menos } r \text{ al cuadrado} \times \text{Raíz del número de grados de libertad}$ ". En este caso el grado de libertad es igual al número de pares de datos que se comparan menos dos.

Para ilustrar este estimado supongamos lo siguiente: el gerente de marketing de un hotel desea conocer si existe una relación directa entre los gastos anuales de publicidad y las pernoctaciones vendidas anualmente. Estos datos son como siguen en unidades reducidas para facilitar el cálculo:

Pernoctaciones vendidas al año: 32; 54; 95; 15; 164; 180

Gastos de publicidad: 8; 22; 17; 27; 36; 33

Para determinar si el aumento de los gastos de publicidad provoca un aumento proporcional de las ventas, vamos a calcular el coeficiente de correlación lineal “r” como aparece en la tabla 5.

Tabla 5 – Pernoctaciones vendidas anuales y gastos anuales publicitarios

Ventas (x)	Publicidad (y)	x – x promedio	y – y promedio	(x – x promedio) al cuadrado	(y – y promedio) al cuadrado	(x – x promedio) (y menos y promedio)
32	8	-81,7	-15,8	6674,9	9,6	+1290,0
54	22	-59,7	-1,8	3564,1	3,2	+107,5
95	17	-18,7	-6,8	349,7	46,2	-122,9
157	27	+43,3	+3,2	1874,9	10,2	+138,8
164	36	+50,3	+12,2	2530,1	14,8	+613,7
180	33	+66,3	+9,2	4395,7	84,6	+610,0
282	143	-	-	19389,4	542,6	2883,8
113,7	23,8	-	-	-	-	-

Los valores promedios fueron:  $682 / 6 = 113,7$  y  $143 / 6 = 23,8$

Ahora con los valores de la tabla podemos calcular el coeficiente de correlación lineal “r”:

$r = \text{Sumatoria de los valores elevados al cuadrado} / \text{Raíz de la multiplicación de los dos valores elevados}$

$$r = 2.883,8 / \text{Raíz de } 19389,4 \times 542,6$$

$$r = 2883,8 / 3244 = 0,889$$

Según la estadística descriptiva el estadígrafo coeficiente de correlación lineal es elevado (89%), e incluso el coeficiente de determinación es aceptable o sea:  $0,889 \times 0,889 = 0,79$  (79%). Pero no basta con este conocimiento, es necesario determinar la significación estadística de estos datos a nivel poblacional de donde se obtuvo la muestra y según el número de grados de libertad que en este caso es igual a:  $6 - 2 = 4$ .

Si observamos en la tabla el test crítico del coeficiente de correlación para 4 grados de libertad, veremos que el dato calculado 0,889 se ubica entre el 1% y el 2% de probabilidad, es decir mucho más bajo que el nivel de significado aceptable del 5% y por ello podemos inferir que la correlación positiva entre los gastos de publicidad y las ventas es estadísticamente significativa, o sea que las ventas crecen proporcionalmente a los gastos de publicidad en la población como un todo.

## Consideraciones finales

Decía Sierra Bravo (1983, p. 531) que: [...] los resultados de las investigaciones sociales se refieren normalmente a muestras de la población investigada y no a la población misma. Si bien se suele hacer la generalización de que dichos resultados también son válidos para el universo o población de que se trate, siempre se plantea la duda fundamental de si es admisible esta generalización.

Dijimos que la duda surge porque es posible que los resultados obtenidos en la muestra se puedan deber a un hecho fortuito o accidental y no al hecho de que los estadígrafos encontrados en la muestra se puedan trasladar al universo. Por eso mismo es necesario buscar la validez de los hallazgos, con el fin de saber si los estadígrafos de una muestra se pueden generalizar hacia la población de donde la muestra se tomó. Esta búsqueda de la validez se denomina **significación estadística** y es un cálculo necesario y fundamental para no correr riesgos de hacer aseveraciones infundadas.

Esto también lo corroboró Guillermo Briones (1982, p. 187) al señalar que las pruebas de significación se deben emplear cuando se trabaja con datos que provienen de muestras probabilísticas, siempre que el marco muestral sea perfecto y el universo sea relevante, o sea con las características apropiadas para someterlo a verificación de hipótesis.

Es bueno señalar que los expertos en la metodología cualitativa etnográfica señalan que en este campo se trabaja con mini-paradigmas y con sus propios presupuestos lógicos internos apoyados en valores, tradiciones, roles que se van regularizando para explicar la conducta individual y grupal de una manera adecuada (Martínez, 2000, p. 1).

En este campo los eventos tienen el significado para quienes están en ese medio social o en ese contexto y la relación que consiguen es holística para ser interpretada de acuerdo a criterios vivenciales. Es decir, la pregunta básica es: ¿cuál es la cultura del grupo?, como lo hacen los antropólogos; o bien, ¿cuál es la filosofía o esencia del fenómeno?, como lo hacen los filósofos; o en todo caso, ¿cuál es el significado de la acción humana, según el contexto?, como lo hacen los analistas de contenidos con su hermenéutica; o también, ¿cuál es el procedimiento para superar la situación?, como lo hacen los psicólogos sociales e incluso la perspectiva etnometodológica, cuando el sociólogo trata de conocer de qué manera la gente le da sentido a sus actividades diarias, para comportarse de una manera socialmente aceptable.

En estos estudios cualitativos se acepta una muestra intencional y el investigador tiene que ser muy agudo para poder lograr su evidencia racional o validez empírica, basándose en la coherencia interna y externa, en la comprensión, en la capacidad predictiva, en la precisión conceptual, originalidad, simplicidad, y en aplicación práctica cuando existen contrastes y potencia heurística. La validez aparece al tener una imagen clara y representativa de una realidad y si se pueden aplicar a grupos similares. Incluso, si el estudio se puede repetir con el mismo método sin alterar los resultados, entonces surge la confiabilidad de lo investigado.

Como hemos visto la manera de conocer con propiedad lo que se investiga es muy diferente si se trata de un abordaje cuantitativo o bien cualitativo. Cada investigador, según el planteamiento del problema y el marco teórico tendrá que tomar una decisión del método para realizar su trabajo.

## Referencias

BRIONES, Guillermo. *Métodos y técnicas de investigación para las ciencias sociales*. México: Editorial Trillas, 1982.

MARTÍNEZ, Miguel. *Metodología cualitativa*, Caracas: INESCO-USB, 2000.

MENTHA, Gerald. *Los tests estadísticos aplicados a la empresa*. Bilbao: Deusto, 1964.

ASCANIO, Alfredo. La inferencia estadística en la investigación turística. *Revista Hospitalidade*, São Paulo, ano V, n. 1, p. 131-146, jun. 2008.

RUNYON, Richard ; HABER Autrey. *Estadística para las ciencias sociales*. USA: Addison-Wesley Iberoamericana, 1992.

SIERRA BRAVO, R. *Técnicas de investigación social: teoría y ejercicios*. Madrid: Paraninfo, 1983.

Artigo recibido en 04 de abril de 2008

Aprobado para publicación en 29 de abril de 2008