**La naturaleza estadística de la**

**evaluación visual**

Juan Ramón Navarro Flores1

**Introducción:**

Siempre que se va a realizar un ensayo donde para alguna evaluación se va a utilizar una escala, surge la pregunta de si esa evaluación resultante es una variable cualitativa o es una variable numérica continua. Para responder a esta pregunta es necesario hurgar en la historia un poco ya que esto es un debate que existe desde el siglo XIX cuando por primera vez Ernest Heinrich Weber (1795-1878), (fundador de la psicofísica, profesor de la universidad de Halle-Wittemberg) formuló la ley que lleva su nombre y que posteriormente Gustav Theodor Fechner (1801-1887) planteó, en términos matemáticos, la relación entre el estímulo físico y la sensación. La ley de Weber-Fechner dice que la agudeza de la visión es directamente proporcional al logaritmo del estímulo (Wikipedia), y que por ser una función matemática entonces la respuesta necesariamente tiene que ser numérica y así fue juzgada luego por un estudiante de fitopatología.

Cuando James G. Horsfall estudiaba fitopatología (1945) en la estación experimental agrícola de Connecticut, quería “**medir”** la enfermedad, no solo describirla como leve, media o severa por lo que recurrió  a un colega estudiante graduado de otra universidad que se dedicaba a la fisiología de las enfermedades.  Estudiando juntos el problema Horsfall y Barratt tropezaron con el principio de Weber-Fechner lo que rápidamente los llevó a plantear una escala numérica para **“medir”** la enfermedad.  En el artículo que siguió a este trabajo en Phytopathology, los estudiantes exponen que el ojo es una “fotocelda” que lee en logaritmos. En ese mismo sentido, el Dr. Grönevald, emérito de la Universidad Estatal de Iowa, afirmaba en su curso de diseño experimental que “el ojo humano es un instrumento de medida como cualquier otro, quizás no tan preciso pero un instrumento de medida al fin”.

Posteriormente, en 1978 John C. Baird y Elliot Noma publican un libro titulado “Fundamentals of scaling and psychophysics” en que amplían el concepto del Horsfall y Barratt; estos autores postulan que probablemente existen cuatro tipos de curvas estímulo-respuesta: la primera es una función lineal en la que la respuesta aumenta linealmente conforme el estímulo también aumenta linealmente, dentro de este grupo se citan principalmente las estimaciones de proporciones; el segundo tipo es una función logarítmica en la que la respuesta aumenta linealmente mientras que el estímulo aumenta de manera logarítmica; el tercer tipo es una función exponencial en la que la respuesta aumenta logarítmicamente

1 Centro de Investigaciones en Protección de Cultivos (CIPROC), [juan.navarro@ucr.ac.cr](mailto:juan.navarro@ucr.ac.cr)

mientras que el estímulo lo hace linealmente; por último, el cuarto tipo es una función potencial en la que la respuesta y el estímulo aumentan logarítmicamente, conocida como la ley de Steven, y que ha sido verificada para gran cantidad de estímulos tales como estimaciones de área, dulzura, el olor del café, ruido, brillo, etc.

Gran cantidad de autores se pueden citar que están en contra de este concepto; sin embargo, ninguno lo llega a negar o a descartar del todo.  La máxima crítica lanzada a estas mediciones con una escala es que (Herbert, 1982) no existe evidencia experimental de que usar una escala produce un mejor resultado que la evaluación directa de la severidad de una enfermedad.  En el 99, el Dr. Arauz y yo (Navarro y Arauz, 1999) hicimos algo en este sentido, comparamos dos métodos de evaluación de enfermedades del fruto de papaya  con un método de referencia; el método A era la evaluación visual directa y el método B era lo mismo pero la evaluación era asistida por una escala pictórica, basada en una escala numérica que responde al principio de Horsfall y Barratt.  Obtuvimos que con ambos métodos se obtenían resultados similares pero que el método de evaluación visual directa era mucho menos repetible y bastante más variable, esto es: más impreciso. Además, Arauz (comunicación personal) apunta un problema adicional del uso de escalas; cuando por ejemplo se evalúan raíces, no es lo mismo un % de daño en la raíz principal que el mismo % en una raíz secundaria. En este caso sugiere emplear un índice que sería una especie de ponderación entre el grado asignado por la escala (1 equivale a una severidad en un ámbito que va de más de 0 hasta 3%,  2 = entre 3 y 6% , 3 = de 6 a 12%, etc.) y la cantidad de individuos que califiquen dentro de cada clase, todo dividido por el producto del número total de unidades evaluadas por el grado más alto de la escala.

Dicen los críticos (Herbert, 1982) que se debe determinar si las escalas se utilizan para estimar la proporción real de  enfermedad en un fruto o en una planta, o si simplemente las escalas son un instrumento que ayuda a ubicar el material vegetal en categorías  de intensidad de enfermedad.  Hasta el momento no he visto una respuesta satisfactoria a esta crítica de Herbert. Mi opinión es que el número es una apreciación de cantidad desarrollado por nuestra misma mente, y si el número lo consideramos una medida, bien podemos también considerar una medida la apreciación de cantidad que hacemos de una enfermedad, un sabor, un color, un área, etc., en última instancia es la misma mente humana la que está creando ambos: número y apreciación de cantidad, con la salvedad de que se debe utilizar un mayor número de repeticiones para incrementar un poco la precisión del ojo.

Hay sin embargo, situaciones donde se pretende cuantificar criterios descriptivos como por ejemplo, calificar el grado de pedregosidad como: 1, poco pedregoso; 2, pedregoso; 3, muy pedregoso; 4, extremadamente pedregoso, y luego promediar el 1 con el 4 para describir el promedio de la pedregosidad de dos suelos. O fenómenos muy complejos como por ejemplo: el dolor, el apego y el compromiso en los que intervienen otros aspectos adicionales a los sentidos tales como sentimientos, preparación, experiencia, etc; en estos casos si es preferible recurrir a metodologías no paramétricas.

**Literatura citada:**

# Baird, J.C. y Noma, E.J... 1978. Fundamentals of scaling and psychophysics (Wiley series in behavior). John Wiley & Sons. 287 pp.

**Herbert, T.T**. 1982. The rationale for the Horsfall-Barratt plant disease assessment scale. Phytopatholgy 105 72(10):1269.

**Horsfall, J.G., y Barratt, R.W.** 1945**.** An improved grading system for measuring plant diseases. (Abstract.). *Phytopathology* 35:655.

**Navarro, J.R., Arauz, L.F**. 1999. Exactitud y repetitividad de dos métodos para la evaluación de la severidad de enfermedades fungosas en el fruto de la papaya (*Carica papaya*). Agronomía Costarricense 23(1): 89-96. 1999

<http://es.wikipedia.org/wiki/Ley_de_Weber-Fechner>