

## *Validación del modelo Vitamínico de Warr mediante Regresión No paramétrica localmente ponderada*

*Autor: Univ. Deyvis Nina Canaviri*

### *Introducción*

Los modelos lineales son los modelos estadísticos más ampliamente utilizados en el campo de las ciencias del comportamiento debido a su facilidad de uso y versatilidad. Sin embargo, si las dependencias existentes en los datos son de carácter no lineal, estos modelos adolecen de buenos ajustes proporcionando poca y tal vez incorrecta información al investigador. Un ejemplo de ello lo encontramos en la validación del Modelo Vitamínico de Peter Warr (1987). Dicho modelo plantea la posible existencia de relaciones curvilíneas (en forma de U invertida) entre determinadas características ambientales y el bienestar psicológico en el trabajo (BPT). El estudio de la posible existencia de estos tipos de relaciones se ha llevado a cabo tradicionalmente por medio de regresiones múltiples jerárquicas (Warr, 1990b) o mediante análisis de ecuaciones estructurales (De Jonge y Schaufeli, 1998).

Un método estadístico, novedoso en este contexto, que puede permitir este tipo de análisis no lineal son los **modelos de regresiones no paramétricas**, en particular **regresiones localmente ponderadas**. Estos modelos ajustan una ecuación de regresión localmente a partir de una vecindad asociada a cada observación determinada.

Los modelos resultantes suelen ser no lineales y de una complejidad más elevada que los modelos tradicionales de regresión. Sin embargo, como hacemos resaltar en este trabajo, proporcionan mejores ajustes adaptándose al tipo de relaciones definidas. El objetivo del presente trabajo es pues el presentar la metodología de los **modelos de regresiones localmente ponderadas** y aplicar dichos modelos para analizar las relaciones teóricamente curvilíneas postuladas por el modelo de Warr (1987).

### *Modelización estadística*

En el problema de la regresión múltiple, disponemos de  $n$  observaciones de una variable respuesta  $Y$ , y las denotamos por  $y = (x_1, x_2, \dots, x_n)^T$ , medidas sobre  $n$  vectores de diseño  $x^i = (x_{i1}, \dots, x_{ip})$ . Los puntos  $x_i$  suelen ser medidas de variables aleatorias  $X_j$  para  $j = 1, \dots, p$ . El objetivo es modelizar la dependencia de  $Y$  frente a  $X_1, \dots, X_p$  para (a) describir el tipo de dependencia de la variable respuesta con las predictoras, (b) inferir las contribuciones de cada predictor a la hora de explicar  $Y$  y (c) predecir  $Y$  en ciertos valores de  $X_1, \dots, X_p$ . El modelo lineal (LM) viene dado por

$$Y = \alpha + \sum_{i=1}^p \beta_i X_i + \varepsilon \quad (1)$$

donde  $E(\varepsilon) = 0$  y  $\text{var}(\varepsilon) = \sigma^2$ . Este modelo se basa en la hipótesis de que la dependencia de  $E(Y)$  sobre  $X_1, \dots, X_p$  es lineal en cada una de las variables predictoras y la variable respuesta debe seguir una distribución normal. Una forma de generalizar los modelos LM es mediante los modelos de suavizamiento no paramétricos. En este caso, si reemplazamos la ecuación (1) por

$$Y = \alpha + \sum_{i=1}^p f_i(X_i) + \varepsilon \quad (2)$$

donde las  $f_i$  son funciones univariantes arbitrarias, una para cada predictor, obtenemos los modelos aditivos (AM). Como en (1), los errores son independientes de las variables  $X_i$  con esperanza cero y varianza constante. Además,  $E\{f_i(X_i)\} = 0$ . Las correspondientes funciones  $f_i$  en (2) pueden ser calculadas por diferentes métodos: splines cúbicas (ns), B-splines (bs), polinomios (poly), suavizamientos loess (lo) o suavizamientos splines (s). Los métodos ns, bs y poly llevan a ajustes

paramétricos, mientras que lo y s introducen ajustes no paramétricos en el modelo.

En este trabajo nos centramos en modelos AM con suavizamientos tipo loess, es decir, modelos de regresión no paramétrica localmente ponderada. En estos modelos se construye una función de suavizamiento  $s(x)$  de la siguiente forma:

a) Sea un punto  $x_0$ . Buscamos su vecindad  $N(x_0)$ , es decir, el conjunto de puntos más próximos a  $x_0$ . El número de vecinos,  $k$ , se especifica como un porcentaje del número total de puntos. Dicho porcentaje se denomina span.

b) Calculamos la mayor distancia entre  $x_0$  y cualquier punto en su vecindad:

$$\Delta(x_0 = \max_{N(x_0)} |X_0 - X_i|) \quad (3)$$

c) Asignamos pesos a cada punto de  $N(X_0)$  mediante la función

$$w\left(\frac{|X_0 - X_i|}{\Delta(x_0)}\right) \quad (4)$$

donde

$$W(u) \begin{cases} (1 - u^3)^3 & 0 \leq u \leq 1 \\ 0 & \text{en otro caso} \end{cases} \quad (5)$$

d) Calculamos el ajuste de mínimos cuadrados ponderados de  $y$  en la vecindad  $N(X_0)$ .

e) Repetimos este proceso para cada punto  $X_0$ . de las variables predictivas.

La amplitud o span utilizada por defecto es de  $2/3$ , aunque la mayor parte de las veces, a nivel práctico, se especifica una amplitud menor que oscila generalmente entre 0.3 y 0.5. (S-Plus, Statistical Sciences, 1993).

El modelo aditivo AM puede ser estimado mediante el algoritmo iterativo *backfitting*, el cual descansa en el hecho de que si el modelo (2) es correcto, entonces se tiene que para todo  $k$ ,

$$E(Y - \alpha - \sum_{j \neq k} f_j(X_j) | X_k) = f_k(X_k)$$

$$E(Y - \alpha - \sum_{j \neq k} f_j(X_j) | x_k) = f_k(x_k) \quad (6)$$

La convergencia de este algoritmo está garantizada si se tiene en cuenta que se trata de un algoritmo de Gauss-Seidel para la resolución de ecuaciones estimativas. Para justificar este procedimiento, se necesita introducir las funciones de suavizamiento sugeridas por los scatterplots. Es decir, si necesitamos minimizar

$$\sum_{i=1}^n \left\{ Y_i - \sum_{j=1}^p f_j(X_{ij}) \right\}^2 \quad (7)$$

Una forma puede ser añadir un término a (7) que penalice la falta de suavizamiento y actuar por medio de *mínimos cuadrados penalizados* consistente en minimizar

$$\sum_{i=1}^n \{Y_i - \sum_{j=1}^p f_j(X_{ij})\}^2 + \sum_{j=1}^p \lambda_j \int \{f_j''(X_{ij})\}^2 \quad (8)$$

La solución de (8) viene dada por

$$\hat{f}_k = (I + \lambda_k K_k)^{-1} (y - \sum_{j \neq k} \hat{f}_j) \quad (9)$$

donde  $K_k$  son matrices de penalización para cada predictor.

Aplicación: Análisis del modelo vitamínico de Warr Siguiendo el modelo formulado por Warr (1987, 1994) las características ambientales tendrían una influencia sobre el bienestar psicológico del trabajador de la misma manera que las vitaminas influyen sobre el cuerpo humano. En este contexto, dos son los resultados que se pueden observar: (1) un decremento en la salud (Decremento Adicional), que es lo que ocurre en el caso por ejemplo de las vitaminas A y D; (2) una estabilidad (no mejora) en la salud (Efecto Constante) que es el caso de las vitaminas C y E.

Warr postula que un efecto similar a éste puede ocurrir en relación con las características del trabajo (que actuarían como vitaminas en este caso) y el bienestar psicológico de los empleados. Así, algunas características pertenecerían al grupo de las vitaminas C y E (salario, seguridad y posición social valorada), mientras que el resto lo haría al de las vitaminas A y D. Estas relaciones se muestran en la figura 1.

Concretamente, y centrándonos en las relaciones curvilíneas planteadas, el autor señala que dicho tipo de relación con el BPT (placer, ansioso-contento y deprimido-entusiasmado) se producirá con 7 de las variables ambientales (Warr, 1998; Cifre, 1999): Oportunidad de uso de habilidades, oportunidad de control, retroalimentación de la tarea, oportunidad de contactos interpersonales, supervisión de apoyo, metas generadas extrínsecamente y variedad de las tareas.

## **Diseño de la investigación**

### **Muestra**

La muestra está compuesta por 297 trabajadores de 12 empresas del Sector Cerámico de Castellón. La edad media de la muestra es de 32.94 años (D.T. = 8.49), y los hombres conforman el 67% de la muestra. El nivel de estudios queda dividido del siguiente modo: el 50.12% tiene estudios primarios, el 33.8% tiene estudios secundarios, mientras que sólo un 16.08% cuenta con algún título universitario.

### **Procedimiento**

Pase de una batería de cuestionarios por parte de un responsable del departamento de Recursos Humanos de las empresas participantes del estudio. La recogida de datos se realizó entre Mayo de 1996 y Diciembre de 1997.

### **Instrumentos**

#### **Bienestar psicológico**

*Satisfacción laboral (eje 1).* Cuestionario S10/12 validado en la muestra del estudio (Cifre, 1999). *Escalas de ansioso-contento (eje 2) y de deprimido-entusiasmado (eje 3).* Se ha empleado la validación (Cifre, 1999) de la adaptación al castellano (Lloret y Tomás, 1994) de las dos escalas desarrolladas por Warr (1990a).

## **Características del trabajo**

De todas las características del trabajo señaladas por Warr, se han seleccionado en el presente trabajo aquellas escalas de medición de estas características que han demostrado ser válidas y fiables en estudios anteriores (Cifre, 1999). Las escalas definitivas utilizadas en el presente trabajo han quedado configuradas del siguiente modo: 1) *Oportunidad de control:* Se ha empleado la escala de Job Authority de Van de Ven y Ferry (1980). 2) *Oportunidad de uso de habilidades:* Escala monoítem propuesta por Warr (1987). 3) *Metas generadas externamente:* Se ha evaluado a partir de medidas del nivel de *sobrecarga* percibida en el trabajo, que hace referencia al grupo de demandas intrínsecas del trabajo de Warr (1987), a través de la adaptación a una muestra española de la escala de sobrecarga de rol realizada por el equipo WOSY (Work Socialization of Youth; ver Bravo, Salanova, Ripoll, Hernández y Zornoza, 1994). 4) *Variedad:* Ha sido evaluada por la escala Van de Ven y Ferry (1980) de variedad. 5) *Claridad ambiental:* La validación de dos escalas de medida (Feedback del Job Diagnostic Survey (JDS) de Hackman y Oldham (1975) y la adaptación española (Bravo y otros, 1994) de la escala de Ambigüedad de rol de Rizzo, House y Lirtzman) dio lugar a dos nuevas subescalas: a) Claridad extrínseca del trabajo, o grado en que los supervisores y Compañeros de la cual proporcionan retroalimentación al trabajador acerca de su desempeño de su trabajo y b) Claridad intrínseca del trabajo, o retroalimentación acerca de cómo deben realizarse las tareas proporcionada por el propio trabajo. 6) *Supervisión de apoyo:* Se ha empleado una versión reducida del cuestionario de Liderazgo de Fleishman (1957) compuesto 9 ítems que configuran dos sub escalas: a) Apoyo del supervisor centrado en la tarea (4 ítems) y b) Apoyo del supervisor centrado en la persona (5 ítems). 7) *Oportunidad de contactos interpersonales:* Se ha evaluado a partir de dos ítems de la escala de «Trato con otros» del JDS (Hackman y Oldham, 1975) y la escala de «oportunidad de hacer amistades» de Hackman y Lawler (1971).

**Resultados**

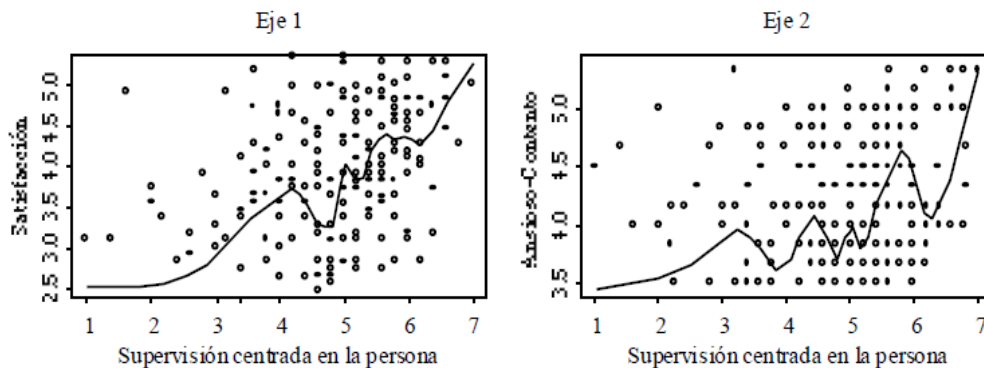
En la tabla 1 se muestran los resultados de cada predictor (característica del trabajo) para cada variable predicha (dimensión del BPT). En dicha tabla se informa sobre el número de observaciones

que se utilizan para construir el modelo, la complejidad del mismo (nº de parámetros), el span y el grado de explicación (coeficiente de determinación).

*Tabla 1*  
Resumen de los valores obtenidos con el modelo de regresión local ponderado

	Eje 1: Satisfacción	Eje 2: Ansioso-contento	Eje 3: Deprimido-entusiasmado
1. Oportunidad de control	N = 241 N de parámetros: 15 Span: .15 R <sup>2</sup> = 14%	N = 279 N de parámetros: 14 Span: .15 R <sup>2</sup> = 10%	N = 279 N de parámetros: 14 Span: .15 R <sup>2</sup> = 9%
2. Oportunidad para el uso de habilidades	N = 245 N de parámetros: 4.6 Span: .60 R <sup>2</sup> = 16%	N = 286 N de parámetros: 4.6 Span: .60 R <sup>2</sup> = 2%	N = 286 N de parámetros: 4.6 Span: .60 R <sup>2</sup> = 3%
3. Sobrecarga de rol	N = 246 N de parámetros: 5.7 Span: .50 R <sup>2</sup> = 10%	N = 284 N de parámetros: 5.7 Span: .50 R <sup>2</sup> = 12%	N = 284 N de parámetros: 5.7 Span: .50 R <sup>2</sup> = 4%
4. Variedad	N = 240 N de parámetros: 8.7 Span: .30 R <sup>2</sup> = 5%	N = 275 N de parámetros: 8.7 Span: .30 R <sup>2</sup> = 4%	N = 275 N de parámetros: 8.7 Span: .30 R <sup>2</sup> = 5%
5. Claridad ambiental intrínseca	N = 224 N de parámetros: 11.6 Span: .30 R <sup>2</sup> = 17%	N = 257 N de parámetros: 11.5 Span: .30 R <sup>2</sup> = 6%	N = 257 N de parámetros: 11.5 Span: .30 R <sup>2</sup> = 9%
6. Claridad ambiental extrínseca	N = 238 N de parámetros: 8.5 Span: .50 R <sup>2</sup> = 14%	N = 275 N de parámetros: 8.4 Span: .50 R <sup>2</sup> = 10%	N = 275 N de parámetros: 8.4 Span: .50 R <sup>2</sup> = 10%
7. Apoyo del supervisor centrado en la persona	N = 232 N de parámetros: 12.8 Span: .30 R <sup>2</sup> = 32%	N = 264 N de parámetros: 12.6 Span: .30 R <sup>2</sup> = 19%	N = 264 N de parámetros: 12.6 Span: .30 R <sup>2</sup> = 13%
8. Apoyo del supervisor centrado en la tarea	N = 220 N de parámetros: 12.9 Span: .30 R <sup>2</sup> = 13%	N = 252 N de parámetros: 11.7 Span: .30 R <sup>2</sup> = 6%	N = 252 N de parámetros: 11.7 Span: .30 R <sup>2</sup> = 6%
9. Oportunidad de contactos interpersonales	N = 245 N de parámetros: 12.2 Span: .30 R <sup>2</sup> = 12%	N = 283 N de parámetros: 12.1 Span: .30 R <sup>2</sup> = 4%	N = 283 N de parámetros: 12.1 Span: .30 R <sup>2</sup> = 5%

VALIDACIÓN DEL MODELO VITAMÍNICO DE WARR MEDIANTE REGRESIÓN NO PARAMÉTRICA LOCALMENTE PONDERADA



En general, el porcentaje de varianza explicada es relativamente bajo pero siempre mayor que el obtenido con los modelos lineales. Destaca sin embargo el alto porcentaje explicación en el caso de la relación entre la supervisión centrada en la persona y el eje de satisfacción (32%). En las figuras 2 y 3 se representan los Modelos no lineales ajustados para dos de las variables predictoras. En general, los resultados gráficos parecen mostrar que sólo las variables de variedad (frente a los tres ejes de medida del BPT) y de contactos interpersonales respecto al eje de satisfacción parecen ir en la línea de los resultados hipotetizados por el modelo de Warr.

## Discusión y conclusiones

Como ya se ha señalado anteriormente, el modelo teórico de Peter Warr hipotetiza una relación curvilínea (relación DA) entre las siete variables tenidas en cuenta en el presente estudio y los tres ejes de medida del bienestar psicológico en el trabajo. Esta hipótesis justifica el uso de regresiones no lineales.

Los resultados muestran sin embargo un pobre ajuste de los datos a los modelos encontrados (aunque mejorando el obtenido con modelos lineales según Cifre (1999)). Sin embargo, según Warr (1990a, 1994) estos resultados no deberían sorprendernos, puesto que estas relaciones no lineales postuladas tendrán una mayor posibilidad de aparecer cuando en la muestra recogida se tengan

en cuenta trabajos que recojan todos los extremos de todas las características del trabajo estudiadas. Por lo tanto, la homogeneidad (con bajos valores de desviaciones típicas) encontrada en nuestra muestra puede ser la principal causa de no haber encontrado los resultados esperados.

Desde el punto de vista metodológico, el presente tipo de regresiones presenta la ventaja, con respecto a los modelos lineales, de obtener información local ajustada a los datos empíricos. El modelo lineal pondera globalmente los datos perdiendo ajustes finos que son considerados por estos modelos no paramétricos.

## Referencias

Fleishman, E.A. (1957). The leadership opinion questionnaire. En R.M. Stogdill y A.E. Coons (eds.): *Leader Behavior: Its description and measurement*. Bureau of Business Research, Ohio State University, Columbus, Ohio.

Warr, P. (1998) Well-being at the workplace. En D. KAHENEMAN, E. DIENER y N. SCHWARZ (Eds): *Understanding Quality of Life: Perspectives on Enjoyment and Suffering*. Nueva York: Russell Sage.

Dedicado: A mi Familia y Amigos que siempre me apoyan

*“No mido el éxito de un hombre por lo alto que llega,  
sino por lo alto que rebota cuando toca fondo”*

*General Patton*