

IMPLEMENTACIÓN DE COMPARACIONES ENTRE GRUPOS UTILIZANDO LA ECUACIÓN DE GOMPERTZ Y VARIABLES DUMMY.

Pablo Vázquez-Borsetti¹, Juan Pablo Barreyro¹, Calero Alejandra Daniela¹.

1 Consejo Nacional de Investigaciones Científicas y Técnicas

Palabras clave: Gompertz, variables dummy , desarrollo, vocabulario.

Dada la importancia de la función de Gompertz en las ciencias biomédicas, es crucial contar con un método robusto para comparar grupos. En una búsqueda bibliográfica en PubMed, solo se encontró un trabajo que utiliza variables dummy en conjunto con la función de Gompertz (Djanira Rodrigues Negrão, 2021). Lamentablemente, dicho trabajo no proporciona el código necesario para que un público más amplio pueda realizar este tipo de análisis.

En este trabajo, se presenta una implementación en R para hacer comparaciones entre grupos utilizando la ecuación de Gompertz y variables dummy. La función final desarrollada en este estudio es ligeramente diferente de la utilizada en el trabajo de Negrão.

Función de Gompertz:

$$F(t) = A \cdot e^{(-e^{(-K(t - T))})}$$

Función de Gompertz con variables dummy:

$$(A + cat \times coef_A) \times \exp(-\exp(-(K + cat \times coef_K) \times (t - (T + cat \times coef_T))))$$

cat = variable dummy

Las variables dummy son binarias, tomando el valor de 0 o 1, y se utilizan en modelos estadísticos para representar categorías o atributos cualitativos. Estas variables permiten incluir información categórica en modelos, como la regresión, tratándolas como si fueran numéricas. Los coeficientes asociados (en este caso *coef_A*, *coef_K* y *coef_T*) a las variables dummy en un modelo reflejan el cambio en la variable dependiente al pasar de la categoría de referencia (valor 0) a la categoría representada por la dummy (valor 1). Esto permite interpretar el impacto de pertenecer a una categoría específica en el resultado modelado. Así, un ajuste de la ecuación no lineal donde los coeficientes sean significativamente distintos de 0 indica un efecto significativo del grupo sobre el parámetro en cuestión.

La implementación de este análisis se realizó en el marco de un trabajo sobre el desarrollo de vocabulario, actualmente pendiente de publicación. La base de datos consiste en 838 casos a los que se les evaluaron dos categorías de vocabulario (vocabulario emocional y general) . En el trabajo final, se incluirá la variable "sujeto" como variable aleatoria en un modelo mixto, pero para su utilización como ejemplo en este trabajo se consideraron los datos como si fueran variables independientes.

Para esta comparación, se realizó un paso preliminar de normalización, que consistió en dividir la variable de respuesta de cada grupo por la asymptota máxima (parámetro A) obtenida en el ajuste de la ecuación de Gompertz. Los datos se ajustaron mediante mínimos cuadrados no lineales. El ajuste fue excelente ($A = 0.502 \pm 0.006$, $\text{valor-t} = 83.746$, $p < 2e-16 ***$; $K = 0.149 \pm 0.0148$, $\text{valor-t} = 10.066$, $p < 2e-16 ***$; $T = 5.63 \pm 0.770$, $\text{valor-t} = 7.309$, $p = 4.16e-13 ***$)

indicando que la ecuación de Gompertz modela adecuadamente del desarrollo del vocabulario. El coeficiente del parámetro A no mostró diferencias significativas ($\text{coef_A} = 0.033 \pm 0.049$, valor-t = 0.678, p = 0.498), como era de esperarse debido a la normalización. Sin embargo, los coeficientes de los parámetros K y T mostraron diferencias significativas entre grupos ($\text{coef_K} = -0.108 \pm 0.023$, valor-t = -4.609, p = 4.36e-06 ***; $\text{coef_T} = 7.206 \pm 0.925$, valor-t = 7.783, p = 1.23e-14 ***), lo que indica que el vocabulario emocional se desarrolla antes que el vocabulario general. Se destaca la importancia, en el ámbito de las ciencias biomédicas, de disponer de una herramienta para realizar comparaciones entre grupos en el contexto de la evolución de variables que pueden ser modeladas mediante la función de Gompertz.

Código disponible bajo licencia Creative Commons Zero v1.0 Universal, la base de datos quedará disponible después de la publicación del trabajo.