**GENERALIZED LINEAR AND LOG-LINEAR MODELS**

El GLM se define por tres componentes:

1. Variantes aleatorias independientes de respuesta ; se asume que estas variantes siguen una distribución de probabilidad de la familia exponencial, con valor esperado ​, que, en los modelos logarítmicos lineales, es la media del logaritmo de las frecuencias esperadas en las celdas; la variable de respuesta, , es la primera parte del componente aleatorio del GLM; la segunda parte es el residual.
2. Un predictor lineal basado en las variables predictoras ​ y los correspondientes parámetros, este es el componente sistemático de un GLM.
3. Una función de enlace monótona, , que relaciona el predictor lineal con la respuesta esperada,

Ejemplos de GLM incluyen el modelo logit binomial para datos binarios, el modelo Poisson, y el modelo lineal general para variables de resultado continuas. La función de enlace natural para la distribución de Poisson es la función logarítmica. Resulta en una componente línea

donde ​ es el vector -ésimo de la matriz de diseño .

Los modelos logarítmicos lineales usualmente se expresan por los parámetros involucrados en un modelo. Por ejemplo, un modelo logarítmico lineal para la clasificación cruzada de las variables A y B que considera solo los efectos principales de ambas variables, o en otras palabras, el modelo de independencia de A y B, se expresa como:

donde es la constante del modelo, es el parámetro -ésimo de los parámetros de la variable A (con ​ indicando el número de categorías de A), y ​ es el parámetro -ésimo de los parámetros de la variable B (con indicando el número de categorías de B).

Un modelo a menudo utilizado como base es el modelo nulo, . La comparación de modelos puede realizarse utilizando la diferencia entre los coeficientes de razón de verosimilitud  entre el modelo objetivo y el modelo base. El coeficiente es:

donde ​ es la frecuencia observada para la celda es la frecuencia esperada para esta celda, e recorre todas las celdas de la clasificación cruzada. Cuando se evalúa un solo modelo, el de Pearson puede ser la mejor opción, que es:

Tanto como están distribuidos aproximadamente como con: