Tenemos un modelo $Y_i = Bo+B1X+E1.Asumimosqueelerrortienedistribuciónnormal N(0, sigmacua$

Distribución de la variable respuesta (parte aleatoria) = Yi N($u_i, sigmacuadrado)Predictorlineal(parte$ $N_i = Bo+B1*X_i Enlace(funcióndeenlace) = u_i = N_i Elmodelolinealesbienflexible.Puedemodelarunmor$

OJO: imaginemos que quieres modelar la salario con edad. Este sigue una curva no lineal concava. Podemos pensar el modelo como el siguiente: Parte aleatoria Yi N(u_i,σ cuadrado) Predictor lineal = $N_i = B0, B1 * Xi + B2*X_i$ al cuadrado Enlace = $u_i = N_i$

Podrisa incluso partir la función concava en particiones y modelar la función mediante funciones indicadoras.

Para cada valor de Xi tienes una distribución de la Yi (ver gráfico inkscape).

Imaginemos que queremos modelar conteo. Entonces nuestra variable aleatoria tendría distribución Y pois($u_i), nuestropredictorlinealserían_i = B0 + B1X_i)ylafuncióndeenlaceseríal$

Modelo logístico: Y_i $Bernoulli(u_i); N_i = Bo + B1_x i; log(u_i/(1 - u_i)) = Ni.$

Esto da pie a indicar los modelos lineales generalizados. En dónde Y_i $FE_f amilia-exponencial_(u_i$,$\phi$), en donde ϕ es un parámetro de dispersión. La familia exponencial es la siguiente: Normal, Gamma, Normal Inversa, Poisson, Bernoulli. Doblemente diferenciable y monótona.

Modelo aditivo generalizado: Asumo que los efectos de las covariables son aditivos.

Todos los modelos que se han mostrado hasta ahora sirven para modelar la media. Por otro lado, la regresión cuantílica sirve para modelar cuantiles.

Se tiene un modelo Yi = B0 + B1 * Xi1 + B2 * Xi2 + ... + Bk Xik + Ei , i= 1, 2, ... , n. Es más práctico evaluar todo de forma matricial.

Se tiene entonces Y = (Y1, Y2, ... , Yn)T, B = (B0, B1, ... , Bk) , X = (1, X11 .... , X1k . . . . 1, Xn1, ... , Xnk)

y E = (e1, e2 ... ek)T

Entonces todo se puede evaluar como Y = XB + E; E N(0, σ cuadrado I -identidad-) El estimador de mínimos cuadrados ordinarios es: $\sum_{i_n}$ Ei al cuadardo = (E)T E = (Y-XB)T (Y-XB) LS(B) = (Y)TY - 2(Y)T XB + (B)T (X)T XB.

Entonces, derivamos la suma de los cuadrados en relación a β y sale = -2(X)T Y + 2 (X)T X B = 0 Luego tenemos que ((X)T X)) -1 (X)T X B = ((X)T X) -1 (X)T Y se cancela el argumento de la izquierda y se tiene que B = ((X)T X) -1 (X)T Y

Luego tenemos que verificar que el estimador es insesgado y con varianza mínima. A este B hay que hallar la esperanza. Utilizamos el TB2 - prueba 2. y obtenemos que: E(B estimado) = E(((X)T X) -1 (X)T Y) = ((X)T X) -1

ver pantalla de celular. dos imágenes.

El VIF es el factor de inflación de varianza (1/(1-R2 j)).

Vamos a probar que el estimador de B de minimos cuadrados ordinarios es el mejor que hay.

Ver la segunda pantalla de celular. Es un huevo!! XD