

# Modelos Lineales Generalizados: aplicaciones en R

## Modelos Lineales Generalizados: un enfoque aplicado

Ana M. Bianco Jemina García

(anambianco@gmail.com) (jeminagarcia@gmail.com )

### 4. Modelo Lineal Generalizado: Aspectos Generales

# Modelo Lineal Generalizado

Para simplificar, primero pensemos que tenemos una sola observación:  $(Y, x_1, \dots, x_p)$  y asumamos que sigue un GLM.

**Componentes del modelo:** Un GLM posee tres componentes

1. Componente Aleatoria
2. Componente Sistemática
3. Función de Enlace o Link

# Componente Aleatoria

Sea  $Y$  la variable de respuesta asociada a las covariables  $x_1, x_2, \dots, x_p$

- $Y$  tiene f.d. o f.p.p dada por

$$f(y, \theta) = \exp \left\{ \frac{y\theta - B(\theta)}{A(\phi)} + C(y, \phi) \right\},$$

donde  $\theta$  es el parámetro canónico,  $\phi$  es un parámetro auxiliar o nuisance y las funciones  $A()$ ,  $B()$ , y  $C()$  son conocidas.

# Componente Aleatoria: Ejemplos

**1. Normal:**  $Y \sim N(\mu, \sigma^2)$ .

$$\begin{aligned} f(y) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(y-\mu)^2}{\sigma^2}\right) \\ &= \exp\left(\frac{y\mu - \mu^2/2}{\sigma^2} - \frac{1}{2} \left[\frac{y^2}{\sigma^2} + \log(2\pi\sigma^2)\right]\right), \end{aligned}$$

por lo tanto  $\theta = \mu$ ,  $B(\theta) = \frac{\mu^2}{2}$ ,  $\phi = \sigma^2$ ,  $A(\phi) = \phi y$

$$C(y, \phi) = -\frac{1}{2} \left[ \frac{y^2}{\phi} + \log(2\pi\phi) \right].$$

$$E(Y) = \mu$$

# Componente Aleatoria: Ejemplos

**2. Caso Poisson:**  $Y \sim P(\lambda)$ .

$$\begin{aligned}P(Y = y) &= e^{-\lambda} \frac{\lambda^y}{y!} \\ &= \exp(y \log \lambda - \lambda - \log y!)\end{aligned}$$

por lo tanto  $\theta = \log \lambda$ ,  $B(\theta) = e^\theta$ ,  $\phi = 1$ ,  $A(\phi) = 1$  y  $C(y, \phi) = -\log y!$

$$E(Y) = \lambda = e^\theta$$

# Componente Aleatoria: Ejemplos

## 3. Caso Binomial: $Y \sim Bi(n, p)$

Consideremos  $\frac{Y}{n} =$  proporción de éxitos.

$$\begin{aligned} P\left(\frac{Y}{n} = y\right) &= P(Y = ny) = \binom{n}{ny} p^{ny} (1-p)^{n-ny} \\ &= \exp\left(\frac{y \log\left(\frac{p}{1-p}\right) + \log(1-p)}{1/n} + \log\left(\binom{n}{ny}\right)\right) \end{aligned}$$

por lo tanto  $\theta = \log \frac{p}{1-p}$ ,  $B(\theta) = \log(1 + e^\theta)$ ,  $\phi = n$ ,  $A(\phi) = \frac{1}{\phi}$  y

$$C(y, \phi) = \binom{\phi}{\phi y}.$$

$$E\left(\frac{Y}{n}\right) = p = \frac{e^\theta}{1 + e^\theta} = \frac{1}{1 + e^{-\theta}}$$

## Componente Sistemática:

Suponiendo que hay intercept, el vector de covariables es  $\mathbf{x}^t = (1, x_1, x_2, \dots, x_p)$  y da origen al predictor lineal.

$$\eta = \mathbf{x}^t \boldsymbol{\beta},$$

siendo  $\boldsymbol{\beta}$  el vector de parámetros a estimar

## Función de Enlace o Link:

Relaciona las dos componentes  $\mu$  y  $\eta = \mathbf{x}^t \boldsymbol{\beta}$

$$g(\mu) = \mathbf{x}^t \boldsymbol{\beta}$$

Es decir, relaciona al predictor lineal  $\eta = \mathbf{x}^t \boldsymbol{\beta}$  con la esperanza  $\mu$  de la respuesta  $Y$ .

A diferencia del modelo lineal clásico, aquí introducimos una función uno-a-uno continua y diferenciable,  $g(\mu)$ , tal que

$$\mathbf{x}^t \boldsymbol{\beta} = g(\mu).$$

## Función de Enlace o Link:

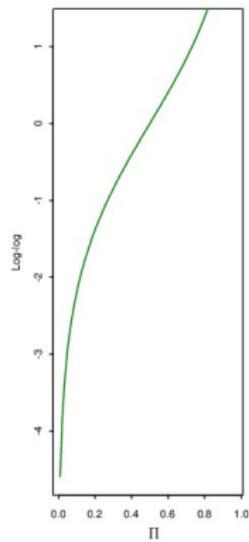
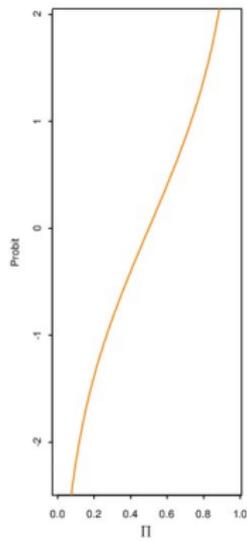
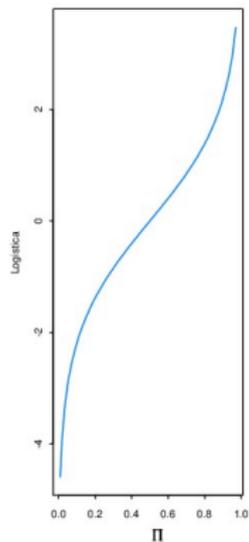
Como la función  $g$  es biyectiva podremos invertirla, obteniendo:

$$\mu = g^{-1}(\mathbf{x}^t \boldsymbol{\beta}) = g^{-1}(\mathbf{x}^t \boldsymbol{\beta}) = h(\mathbf{x}^t \boldsymbol{\beta}).$$

En el caso Binomial, por ejemplo, tenemos que  $\mu \in (0, 1)$  y el link tiene que mapear sobre la recta real. Suelen usarse 3 links:

1. Logit:  $\mathbf{x}^t \boldsymbol{\beta} = \log \frac{\mu}{1-\mu}$
2. Probit:  $\mathbf{x}^t \boldsymbol{\beta} = \Phi^{-1}(\mu)$
3. Complemento log-log:  $\mathbf{x}^t \boldsymbol{\beta} = \log(-\log(1 - \mu))$

# Funciones de Enlace o Link:



# Familia exponencial

La variable de respuesta  $Y$  tiene f.d. o f.p.p dada por

$$f(y, \theta) = \exp \left\{ \frac{y\theta - B(\theta)}{A(\phi)} + C(y, \phi) \right\},$$

donde para un GLM la función  $B$  es de especial importancia ya que describe la relación entre la media y la varianza de la distribución.

En efecto, se puede verificar que

$$\mu = E(Y) = B'(\theta) \quad \text{y} \quad \text{Var}(Y) = A(\phi)B''(\theta).$$

Así, por ejemplo en el caso Poisson:

$$B(\theta) = \exp(\theta) \Rightarrow B'(\theta) = B''(\theta) = \exp(\theta)$$

## log-Likelihood: Pensemos en una sola observación....

$$L(\theta, y) = \exp \left\{ \frac{y\theta - B(\theta)}{A(\phi)} + C(y, \phi) \right\}$$

luego, abusando de la notación

$$\ell = \log(L(\theta, y)) = \frac{y\theta - B(\theta)}{A(\phi)} + C(y, \phi)$$

En un GLM,  $\theta$  es función del vector de parámetros  $\beta$ . Luego, para hallar el EMV vemos a  $\ell$  como función de

$$\mathbf{b} = (b_0, b_1, \dots, b_p)$$

Calculemos la derivada de  $\ell$  respecto de cada  $b_j$  e igualemos a 0:

$$\frac{\partial \ell}{\partial b_j} = 0 \quad j = 0, 1, \dots, p$$

Obtenemos

$$\frac{\partial \ell}{\partial b_j} = \frac{Y - \mu}{\text{Var}(Y)} \quad \frac{\partial \mu}{\partial \eta} \quad x_j = 0$$

# log-Likelihood

$$\ell(\theta, y) = \frac{y\theta - B(\theta)}{A(\phi)} + C(y, \phi)$$

Recordemos que  $\theta$  es función de  $\mu = E(Y)$ ,  $\mu$  es función de  $\eta$ , y  $\eta$  es función de  $\mathbf{b}$ .

Usando la regla de la cadena, resulta

$$\frac{\partial \ell}{\partial \mathbf{b}_j} = \frac{\partial \ell}{\partial \theta} \frac{\partial \theta}{\partial \mu} \frac{\partial \mu}{\partial \eta} \frac{\partial \eta}{\partial \mathbf{b}_j}$$

## ¿Cuánto vale cada derivada?

Como

$$\begin{aligned}\mu &= E(Y) = B'(\theta) \quad y \\ \text{Var}(Y) &= A(\phi)B''(\theta) \\ \eta &= \mathbf{x}'\boldsymbol{\beta} \\ g(\mu) &= \eta\end{aligned}$$

obtenemos

$$\begin{aligned}\frac{\partial \ell}{\partial \theta} &= \frac{y - B'(\theta)}{A(\phi)} = \frac{y - \mu}{A(\phi)} \\ \frac{\partial \theta}{\partial \mu} &= \frac{1}{B''(\theta)} = \frac{1}{\text{Var}(Y)} \\ \frac{\partial \mu}{\partial \eta} &= \text{depende de la función de enlace} \\ \frac{\partial \eta}{\partial b_j} &= x_j,\end{aligned}$$

Luego, resulta

$$\frac{\partial \ell}{\partial b_j} = \frac{y - \mu}{\text{Var}(Y)} \frac{\partial \mu}{\partial \eta} x_j$$

## log-Likelihood:

Supongamos ahora que tenemos una muestra de vectores independientes  $(Y_1, x_{11}, \dots, x_{1p}), \dots, (Y_n, x_{n1}, \dots, x_{np})$  que siguen un GLM.

El log likelihood será una suma con componentes como la que calculamos y para hallar el EMV resolveremos el sistema

$$\frac{\partial \ell}{\partial b_j} = \sum_{i=1}^n \frac{y_i - \mu_i}{V_i} \frac{\partial \mu_i}{\partial \eta_i} x_{ij} = 0 \quad j = 0, 1, \dots, p \quad (1)$$

Para resolver (1) usaremos el método iterativo de Newton-Raphson tal como ya vimos:

$$\beta^{(k+1)} = \beta^{(k)} - [\ell''(\beta^{(k)})]^{-1} \ell'(\beta^{(k)})$$

# Fisher-scoring

En el método de **Fisher-scoring** se propone utilizar  $E\left(\frac{\partial^2 \ell}{\partial b_k \partial b_j}\right)$  en lugar de  $\frac{\partial^2 \ell}{\partial b_k \partial b_j}$  con el fin de obtener resultados más estables.

Por lo tanto, en cada paso iterativo usaríamos

$$- \sum_{i=1}^n V_i^{-1} \left( \frac{\partial \mu_i}{\partial \eta_i} \right)^2 x_{ij} x_{ik}$$

que en forma matricial podemos escribir como:

$$- \mathbf{X}' \mathbf{W} \mathbf{X}$$

siendo  $\mathbf{W} = \text{diag} \left( V_i^{-1} \left( \frac{\partial \mu_i}{\partial \eta_i} \right)^2 \right)$ .

# Fisher-scoring

Finalmente, si:

$$\begin{aligned}\mathbf{W}^{(t)} &= \text{diag} \left( V_i^{-1} \left( \frac{\partial \mu_i}{\partial \eta_i} \right)^2 \right) \\ (\mathbf{V}^{(t)})^{-1} &= \text{diag}(V_i^{-1})\end{aligned}$$

resulta un **cuadrados mínimos repesados iterados (IRLS)**

$$\beta^{(t+1)} = \left( \mathbf{X}' \mathbf{W}^{(t)} \mathbf{X} \right)^{-1} \mathbf{X}' \mathbf{W}^{(t)} \mathbf{z}^{(t)}$$

donde  $\mu = \mu^{(t)}$  y  $\eta = \eta^{(t)}$  y

$$\mathbf{z}^{(t)} = \eta + \frac{\partial \eta}{\partial \mu} (Y - \mu)$$

# Intervalos de Confianza y Tests de Hipótesis

Los resultados de Fahrmeir y Kaufmann (1985) que enunciamos para regresión logística son más generales y abarcan el comportamiento de los EMV para un GLM bajo condiciones de regularidad.

Para  $n$  suficientemente grande, una aproximación razonable es

$$(\hat{\beta}_n - \beta) \stackrel{(a)}{\approx} N(\mathbf{0}, \mathbf{V}(\hat{\beta}_n)),$$

siendo

$$\mathbf{V}(\hat{\beta}_n) = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}.$$

y la estimaremos por

$$\hat{\mathbf{V}}(\hat{\beta}_n) = (\mathbf{X}'\mathbf{W}(\hat{\beta}_n)\mathbf{X})^{-1}.$$

# Inferencia

Como antes, tenemos que si nos interesa  $\mathbf{a}^t\boldsymbol{\beta}$ , una aproximación razonable para  $n$  suficientemente grande es

$$(\mathbf{a}^t\widehat{\boldsymbol{\beta}} - \mathbf{a}^t\boldsymbol{\beta}) \stackrel{(a)}{\sim} N(\mathbf{0}, \mathbf{a}^t\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}})\mathbf{a})$$

Por ejemplo, si nos interesa testear

$$H_0 : \beta_j = 0 \quad \text{vs.} \quad H_1 : \beta_j \neq 0$$

la regla de decisión

$$\text{Rechazamos } H_0 \text{ si } \left| \frac{\widehat{\beta}_j}{\sqrt{\widehat{\mathbf{V}}(\widehat{\boldsymbol{\beta}}_j)}} \right| > z_{\alpha/2}$$

resulta en un test de nivel aproximado  $\alpha$ .

Esto es parte del summary que nos devuelve R.

# Bondad de Ajuste

Consideremos la comparación de dos modelos anidados: la diferencia entre los dos modelos es que la componente lineal de un modelo se obtiene anulando alguno de los parámetros del otro.

En otras palabras, el modelo más simple en  $H_0$  corresponde a un caso particular de  $H_1$ , modelo más general.

## Bondad de Ajuste

Consideremos la comparación de dos modelos anidados: la diferencia entre los dos modelos es que la componente lineal de un modelo se obtiene anulando alguno de los parámetros del otro.

En otras palabras, el modelo más simple en  $H_0$  corresponde a un caso particular de  $H_1$ , modelo más general.

Si el modelo más simple ajusta a los datos tan bien como el más general, entonces, en virtud del principio de parsimonia no rechazaremos  $H_0$ .

Si el modelo más general ajusta significativamente mejor, rechazaremos  $H_0$  en favor de  $H_1$ , que corresponde al modelo más complejo. Para realizar estas comparaciones deberemos usar medidas de *bondad de ajuste*.

## Bondad de Ajuste

Consideremos la comparación de dos modelos anidados: la diferencia entre los dos modelos es que la componente lineal de un modelo se obtiene anulando alguno de los parámetros del otro.

En otras palabras, el modelo más simple en  $H_0$  corresponde a un caso particular de  $H_1$ , modelo más general.

Si el modelo más simple ajusta a los datos tan bien como el más general, entonces, en virtud del principio de parsimonia no rechazaremos  $H_0$ .

Si el modelo más general ajusta significativamente mejor, rechazaremos  $H_0$  en favor de  $H_1$ , que corresponde al modelo más complejo. Para realizar estas comparaciones deberemos usar medidas de *bondad de ajuste*.

El modelo más complejo de todos es el que se llama **saturado**, en el que cada observación se predice por sí misma,  $\hat{y}_i = y_i$ , mientras que el más simple de todos es el **nulo** con un solo parámetro, la intercept, y en el que todas las observaciones se predicen con el promedio,  $\hat{y}_i = \bar{y}_i$ .

# Deviance

La medida de bondad de ajuste más común en GLM es la **deviance**.

Sean  $\ell(\hat{\boldsymbol{\mu}}, \boldsymbol{\phi}, \mathbf{y})$  el log-likelihood del modelo de interés y  $\ell(\mathbf{y}, \boldsymbol{\phi}, \mathbf{y})$  el log-likelihood del modelo saturado.

La deviance o devianza se define como

$$D = 2(\ell(\mathbf{y}, \boldsymbol{\phi}, \mathbf{y}) - \ell(\hat{\boldsymbol{\mu}}, \boldsymbol{\phi}, \mathbf{y})) = -2(\ell(\hat{\boldsymbol{\mu}}, \boldsymbol{\phi}, \mathbf{y}) - \ell(\mathbf{y}, \boldsymbol{\phi}, \mathbf{y})).$$

Notemos que hallar el el EMV equivale a minimizar la **deviance**.

En las distribuciones, como la normal por ejemplo, que tienen un parámetro de escala  $\phi$ , se define la **deviance o devianza escalada** como  $D^* = D/\phi$ , que es la medida que se usa para inferencia.

En las distribuciones como la binomial o la Poisson,  $\phi = 1$ , entonces las dos medidas deviance coinciden.

# Deviance

## Caso Normal

Recordemos que la densidad normal es

$$f(y) = \exp\left(\frac{y\mu - \mu^2/2 - y^2/2}{\sigma^2} - \frac{1}{2}\log(2\pi\sigma^2)\right)$$

Entonces,

$$D = \frac{2}{\sigma^2} \sum_{i=1}^n (y_i^2 - y_i^2/2 - y_i^2/2) - (y_i \hat{\mu}_i - \frac{1}{2} \hat{\mu}_i^2 - \frac{1}{2} y_i^2) = \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \hat{\mu}_i)^2.$$

Aquí se ve que en el caso normal el EMV coincide con el de mínimos cuadrados.

# Bondad de Ajuste

**Caso Binomial**  $\frac{Y}{n} =$  proporción de éxitos.

$$P\left(\frac{Y}{n} = y\right) = P(Y = ny) = \exp\left(\frac{y \log\left(\frac{p}{1-p}\right) + \log(1-p)}{1/n}\right) + \log\left(\binom{n}{ny}\right)$$

entonces

$$\begin{aligned} D &= 2 \sum_{i=1}^n n_i \left[ \frac{y_i}{n_i} \left( \log \frac{y_i/n_i}{1 - y_i/n_i} - \log \frac{\hat{\pi}_i}{1 - \hat{\pi}_i} \right) + \right. \\ &\quad \left. \log\left(1 - \frac{y_i}{n_i}\right) - \log(1 - \hat{\pi}_i) \right] \\ &= 2 \sum_{i=1}^n \left[ y_i \log \frac{y_i}{\hat{\mu}_i} + (n_i - y_i) \log \frac{n_i - y_i}{n_i - \hat{\mu}_i} \right] \end{aligned}$$

# Aplicaciones a Test de Hipótesis

Las hipótesis relativas al parámetro  $\beta = (\beta_0, \dots, \beta_p)^t$  de longitud  $p$  pueden testarse usando el estadístico de Wald y su distribución asintótica.

Si queremos testear las hipótesis

$$H_0 : \beta = \beta^0 \quad \text{vs.} \quad H_1 : \beta \neq \beta^0$$

podemos usar que para  $n$  suficientemente grande

$$(\hat{\beta} - \beta^0)^t (\text{Var}(\hat{\beta}))^{-1} (\hat{\beta} - \beta^0) \stackrel{(a)}{\approx} \chi_p^2.$$

## Modelos anidados

Para decidir entre dos modelos anidados, un enfoque alternativo es comparar la bondad del ajuste de los modelos involucrados.

Sea  $M_q$  un modelo con  $q$  componentes y  $M_p$  un modelo con  $p > q$  componentes de manera tal que  $M_q \subset M_p$  o dicho de otra forma:

$M_q$  se obtiene poniendo algunas componentes de  $M_p$  en cero.

Consideremos las siguientes hipótesis:

$$H_0 : M_q \quad \text{vs.} \quad H_1 : M_p$$

Como maximizar el log-likelihood en un espacio más pequeño no puede llevar a un máximo mayor, tenemos que

$$D(\mathbf{y}, \hat{\boldsymbol{\mu}}_p) \leq D(\mathbf{y}, \hat{\boldsymbol{\mu}}_q)$$

por lo tanto tiene sentido ver como se comporta la diferencia

$$\Delta D = D(\mathbf{y}, \hat{\boldsymbol{\mu}}_q) - D(\mathbf{y}, \hat{\boldsymbol{\mu}}_p)$$

## Modelos anidados

Si testeamos  $H_0$  vs.  $H_1$  usando la diferencia de los estadísticos de cociente del logaritmo de la verosimilitud tenemos

$$\begin{aligned}\Delta D^* &= D_q^* - D_p^* = \frac{D_q - D_p}{\phi} = \\ &= 2 [\ell(\mathbf{y}, \mathbf{y}) - \ell(\hat{\mu}_q, \mathbf{y})] - 2 [\ell(\mathbf{y}, \mathbf{y}) - \ell(\hat{\mu}_p, \mathbf{y})] \\ &= 2 [\ell(\hat{\mu}_p, \mathbf{y}) - \ell(\hat{\mu}_q, \mathbf{y})] .\end{aligned}$$

$\Delta D$  se comporta asintóticamente como una  $\phi\chi_{p-q}^2$  ya que bajo  $H_0$  tendríamos que

$$\Delta D \stackrel{(a)}{\sim} \phi\chi_{p-q}^2$$

Por lo tanto, si el valor observado de  $\Delta D^*$  fuera mayor que el percentil  $\chi_{p-q, \alpha}^2$  rechazaríamos a  $H_0$  en favor de  $H_1$ , bajo el supuesto de que  $H_1$  da una mejor descripción de los datos.

## Criterio de Akaike

El criterio de Akaike (AIC) es un método muy usado para seleccionar y comprar modelos (no necesariamente anidados!!): **el modelo con el valor de AIC más pequeño es el modelo elegido.**

Este es un criterio basado en la verosimilitud, pero que penaliza por el número de parámetros.

Esencialmente AIC se computa como

$$-2 * \log - \text{likelihood} + k * npar$$

donde  $k = 2$  para el AIC habitual.

AIC es un criterio que busca un compromiso entre dos componentes:

- el ajuste (medido a través del likelihood)
  - la complejidad del modelo (medida a través del nro. de parámetros).
- Para modelos con igual cantidad de parámetros el criterio decide por el likelihood.

# Criterio de Akaike

Si elegimos  $k = \log(n)$  resulta para la versión bayesiana llamada BIC (Bayesian Information Criterion).

Notar que AIC penaliza un parámetro extra menos que BIC

## Ejemplo

Collett (1991) reporta los datos de un experimento sobre toxicidad en gusanos de la planta de tabaco dosis de *pyrethroid trans-cypermethrin* al que los gusanos empezaron a mostrar resistencia. Grupos de 20 gusanos de cada sexo fueron expuestos a por 3 días al *pyrethroid* y se registró el número de gusanos muertos o knockeados en cada grupo.

Los resultados se muestran en la siguiente tabla.

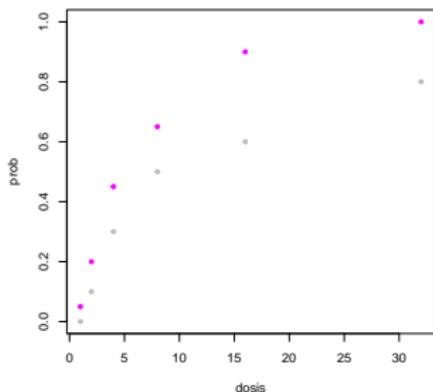
	dosis ( $\mu\text{g}$ )					
sexo	1	2	4	8	16	32
Machos	1	4	9	13	18	20
Hembras	0	2	6	10	12	16

Cuadro: Gusanos del tabaco

Usamos  $\log_2(dosis)$ , dado que las dosis son potencias de 2.

```
> ldose<- rep(0:5,2)
> numdead<- c(1,4,9,13,18,20,0,2,6,10,12,16)
> sex<- factor(rep(c("M","F"),c(6,6)))
> SF<- cbind(numdead,numalive=20-numdead)

> probas=numdead/20
> plot(2^ldose, probas,type="n",xlab="dosis",ylab="prob")
> lines(2^ldose[sex=="M"],type="p", probas[sex=="M"],col=6,pch=20)
> lines(2^ldose[sex=="F"],type="p", probas[sex=="F"],col=8,pch=20)
```



## Ejemplo

De acuerdo con lo que vemos, queremos investigar la posibilidad de que haya diferentes rectas para los dos sexos. Para ello plantearemos y ajustaremos el modelo

$$\text{logit}(\pi) = \beta_0 + \beta_1 \text{sex} + \beta_2 \text{ldose} + \beta_3 \text{sex}:\text{ldose}$$

Por ejemplo, si  $\text{sex} = M$ , para  $\text{ldose} = 3$  tendríamos

$$\text{logit}(\pi_{3,M}) = \beta_0 + \beta_1 + 3(\beta_2 + \beta_3)$$

en cambio si  $\text{sex} = F$ , para  $\text{ldose} = 3$

$$\text{logit}(\pi_{3,F}) = \beta_0 + 3\beta_2$$

```
> salida.gusanos<- glm(SF~sex*ldose, family=binomial)
> summary(salida.gusanos)
```

Call:

```
glm(formula = SF ~ sex * ldose, family = binomial)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.39849	-0.32094	-0.07592	0.38220	1.10375

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	-2.9935	0.5527	-5.416	6.09e-08	***
sexM	0.1750	0.7783	0.225	0.822	
ldose	0.9060	0.1671	5.422	5.89e-08	***
sexM:ldose	0.3529	0.2700	1.307	0.191	

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 124.8756 on 11 degrees of freedom  
Residual deviance: 4.9937 on 8 degrees of freedom  
AIC: 43.104

Number of Fisher Scoring iterations: 4

```
> anova(salida.gusanos, test="Chisq")
```

### Analysis of Deviance Table

Model: binomial, link: logit

Response: SF

Terms added sequentially (first to last)

	Df	Deviance	Resid.	Df	Resid. Dev	Pr(>Chi)
NULL				11	124.876	
sex	1	6.077		10	118.799	0.0137 *
ldose	1	112.042		9	6.757	<2e-16 ***
sex:ldose	1	1.763		8	4.994	0.1842

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## ¿qué es cada cosa?

```
anova(salida.gusanos, test="Chisq")
```

```
Analysis of Deviance Table
```

```
Binomial model
```

```
Response: SF
```

```
Terms added sequentially (first to last)
```

	Df	Deviance	Resid. Df	Resid. Dev	Pr(Chi)	
NULL			11	124.8756		
sex	1	6.0770	10	118.7986	0.0136955	= 1-pchisq(6.0770,1)
ldose	1	112.0415	9	6.7571	0.0000000	= 1-pchisq(112.0415,1)
sex:ldose	1	1.7633	8	4.9937	0.1842088	= 1-pchisq(1.7633,1)

```
> salida.gusanos<- glm(SF~sex+ldose, family=binomial)
> summary(salida.gusanos)
```

Call:

```
glm(formula = SF ~ sex + ldose, family = binomial)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.10540	-0.65343	-0.02225	0.48471	1.42944

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	-3.4732	0.4685	-7.413	1.23e-13	***
sexM	1.1007	0.3558	3.093	0.00198	**
ldose	1.0642	0.1311	8.119	4.70e-16	***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 124.8756 on 11 degrees of freedom  
Residual deviance: 6.7571 on 9 degrees of freedom  
AIC: 42.867

Number of Fisher Scoring iterations: 4