GLM

LM:

$$Y_i = \underset{\sim}{x}_i^T \underset{\sim}{\beta} + \epsilon_i \qquad i = 1, 2, \cdots n.$$

$$E(\epsilon_i) = 0 \qquad V(\epsilon_i) = \sigma^2 \qquad \epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$E(Y_i) = \underset{\sim}{x}_i^T \underset{\sim}{\beta} = E(y_i | x_i) \longrightarrow \text{Regression.}$$

Linear model.

$$E(Y_i) = I(\underset{\sim}{x}_i^T \underset{\sim}{\beta})$$ where I is the identity function. $f(x) = x$.

can we go for $E(Y_i) = g(\underset{\sim}{x}_i^T \underset{\sim}{\beta})$ 'g' must be a well representative of the model.

GLM = generalized linear model.

Example: 1. Record of customers in bank ($\underset{\sim}{X}$) whether lone will be given or NOT?

2. Based on your previous search data. a certain movie will be recomanded for you in Netflix, Aamazon / OTT.

3. Iris flower category. devision.

$Y_i \mid x_i$	Prob.
0	$P(Y_i = 0) = 1 - \pi_i$
1	$P(Y_i = 1) = \pi_i$

~~Bernoul~~

$Y_i \mid X_i \sim \text{Bernoulli}(\pi_i)$.

$0 < \pi_i < 1$

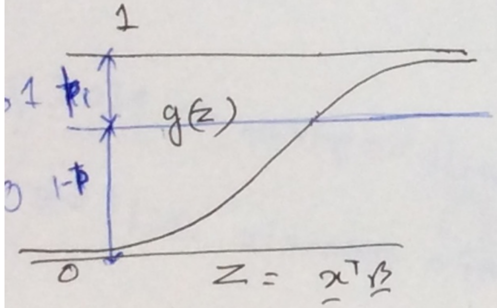
→ poisson.
→ Binomial.
→ multinomial.

$E(Y_i \mid x_i) = \pi_i \in (0,1) \neq \mathbb{R}$.

$V(Y_i \mid x_i) = \pi_i(1-\pi_i) \in (0, 1/4) \neq \mathbb{R}^+$

$\underset{\sim}{x}_i^T \underset{\sim}{\beta} \in \mathbb{R}$ we need $g : \mathbb{R} \to (0,1)$

any cdf (preferably continuous) will work.



$$g(z) = \begin{cases} \dfrac{e^z}{1+e^z} = \dfrac{1}{1+e^{-z}} & \text{(logit model.)} \\ \Phi(z) \quad \text{std. normal cdf.} & \text{(probit model)} \end{cases}$$

$\Phi\left(\frac{z}{\sigma}\right)$ σ small

$\Phi\left(\frac{z}{\sigma}\right)$ σ >>

$$\pi_i = g(\underset{\sim}{x}_i^T \underset{\sim}{\beta}) = \frac{e^{x_i^T \beta}}{1+e^{x_i^T \beta}} = E(Y_i \mid x_i).$$

$$\Leftrightarrow \quad \log_e \frac{\pi_i}{1-\pi_i} = (\underset{\sim}{x}_i^T \underset{\sim}{\beta}) \longrightarrow \text{log-link.}$$

For model fitting or parameter estimation. we can use MLE. ③

~~likelihood~~ / ~~probability~~. Joint density of $y_1, y_2 \cdots y_n$.

$$\prod_{i=1}^{n} f_i(y_i | x_i) = \prod_{i=1}^{n} \left\{ \pi_i^{y_i} (1-\pi_i)^{1-y_i} \right\}.$$

Likelihood function of $\underset{\sim}{\beta}$.

$$L(\underset{\sim}{\beta} | y_1, y_2 \cdots y_n) \equiv \prod_{i=1}^{n} \left\{ \pi_i^{y_i} (1-\pi_i)^{1-y_i} \right\}$$

$$\Leftrightarrow \log_e \left[L(\underset{\sim}{\beta} | \underset{\sim}{y}) \right] = \sum_{i=1}^{n} \left[y_i \log_e \left(\frac{\pi_i}{1-\pi_i} \right) \right] + \sum_{i=1}^{n} \log_e (1-\pi_i).$$

$$\equiv L(\underset{\sim}{\beta}) = \sum_{i=1}^{n} y_i \left[\exp(x_i^T \underset{\sim}{\beta}) \right] + \sum_{i=1}^{n} \log \left[1 + \exp(x_i^T \beta) \right]$$

$$\frac{\partial L(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}} = \underset{\sim}{0} \Rightarrow$$ solve with Newton-Raphson method.

$\underset{\sim}{\hat{\beta}}$ will be a <u>consistent estimator of $\underset{\sim}{\beta}$</u>.

$\underset{\sim}{\hat{\beta}}$ dont have any closed form and hence cannot be shown as unbiased estimator.

$Y = \underset{\sim}{x}^T \underset{\sim}{\beta} + \epsilon$. $\quad \epsilon \sim N(0, \underline{\underline{\sigma^2}})$

if σ^2 is no more a fixed number.

Delta method:

Let x be a random variable $\quad E(x) = \mu \quad V(x) = \sigma 2$

σ^2 may be a function of μ.

$Y = g(\underset{\sim}{x})$ where 'g' is a smooth function.

$Y \approx g(\mu) + g'(\mu)(x - \mu) + \cdots$ ~~Linear approximation~~

$E(Y) \approx g(\mu)$

$V(\underset{\sim}{Y}) \approx (g'(\mu))^2 \sigma 2$

$\approx \underline{(g'(\mu))^2 h(\mu)} =$ Constant (we want).

$\boxed{\text{Let } \sigma^2 = h(\mu)}$

$v(x) = \sigma 2$

$Y = a + bx$

$v(Y) = b^2 \sigma 2$

$$c^2 = (g'(\mu))^2 h(\mu) > 0$$ as it is variance.

$$\Rightarrow \quad g'(\mu) = \frac{c}{\sqrt{h(\mu)}}$$

$$\Rightarrow \quad \frac{d\,g(\mu)}{d\mu} = \frac{c}{\sqrt{h(\mu)}} \quad \Rightarrow \quad g(\mu) = \int \frac{c}{\sqrt{h(\mu)}}\, d\mu + c_1$$

M.S. Bartlett, 1947.

proportion

$$\sigma^2 = p(1-p) = h(p) \quad \Rightarrow \quad g(p) = \int \frac{c}{\sqrt{p(1-p)}}\, dp + c_1$$

$$= C \sin^{-1}\sqrt{p} + c.$$

$\sigma^2 \propto E(y)(1-E(y))$ $\qquad y^* = \sin^{-1}\sqrt{y}$

$\sigma^2 \propto E(y)$ $\qquad y^* = \sqrt{y}$.

$\sigma^2 \propto (E(y))^2$ $\qquad y^* = \log_e y$

$\sigma^2 \propto (E(y))^3$ $\qquad y^* = y^{-1/2}$

$\sigma^2 \propto (E(y))^4$ $\qquad y^* = 1/y$.

Variance stabelizing transformation.
C.R. Rao
Linear Statistical Infere...

GLM

LM: $Y_i = \underset{\sim}{x}_i^T \underset{\sim}{\beta} + \epsilon_i \quad i = 1, 2, \cdots n.$

$E(\epsilon_i) = 0 \quad V(\epsilon_i) = \sigma^2 \quad \epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

$E(Y_i) = \underset{\sim}{x}_i^T \underset{\sim}{\beta} = E(y_i | x_i) \longrightarrow$ Regression.

Linear model.

$\downarrow$

$E(Y_i) = I(\underset{\sim}{x}_i^T \underset{\sim}{\beta})$ where I is the identity function. $f(x) = x.$

canwe go for $E(Y_i) = g(\underset{\sim}{x}_i^T \underset{\sim}{\beta})$ 'g' must be a well representative of the model.

GLM = generalized linear model.

Example: ① Record of customers in bank $(\underset{\sim}{X})$ whether lone will be given or NOT ?

② Based on your previous search data. a certain movie will be recomanded for you in Netflix, Aamazon / OTT.

③ Iris flower category. devision.