

1.2 Outline of generalized linear models:

1. Model assumptions

Let Y be a random variable.

Linear models:

$$\mu = E(Y) = x\beta,$$

where

$$x = [x_1 \quad x_2 \quad \cdots \quad x_p].$$

Generalized linear models:

$$g(\mu) = g[E(Y)] = \eta = x\beta,$$

where g is called the **link function**.

In addition, the response Y has a distribution in the exponential family, taking the form

$$f(y, \theta, \phi) = \exp \left\{ \frac{[y\theta - b(\theta)]}{a(\phi)} + c(y, \phi) \right\}.$$

Intuitively, generalized linear model is the “extension” of the linear model. As the distribution is normal and the link function is identity function, the generalized linear model reduces to the linear model.

Example 1 (normal distribution):

$Y \sim N(\mu, \sigma^2)$. Then

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

Therefore,

$$\theta = \mu, \phi = \sigma^2,$$

$$a(\phi) = \sigma^2 = \phi, b(\theta) = \frac{\mu^2}{2} = \frac{\theta^2}{2},$$

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

Example 2 (Poisson distribution):

$Y \sim P(\mu)$. Then,

$$f(y, \theta, \phi) = \frac{\exp(-\mu)\mu^y}{y!} = \exp[y \cdot \log(\mu) - \mu - \log(y!)].$$

Therefore,

$$\begin{aligned}\theta &= \log(\mu), \phi = 1, \\ a(\phi) &= 1 = \phi, b(\theta) = \mu = \exp(\theta), c(y, \phi) = -\log(y!).\end{aligned}$$

Example 3 (binomial distribution):

$Y \sim \frac{B(m, p)}{m}, 0 \leq Y \leq 1 \Rightarrow$ *Binomial distribution in frequency.*

Then,

$$\begin{aligned}f(y, \theta, \phi) &= \binom{m}{my} p^{my} (1-p)^{m-my} \\ &= \exp \left\{ my \cdot \log(p) + (m - my) \cdot \log(1-p) + \log \left[\binom{m}{my} \right] \right\} \\ &= \exp \left\{ \frac{\left[y \cdot \log \left(\frac{p}{1-p} \right) - \log \left(\frac{1}{1-p} \right) \right]}{(1/m)} + \log \left[\binom{m}{my} \right] \right\}\end{aligned}$$

Therefore,

$$\begin{aligned}\theta &= \log \left(\frac{p}{1-p} \right), \phi = \frac{1}{m}, \\ a(\phi) &= \frac{1}{m} = \phi, b(\theta) = \log \left(\frac{1}{1-p} \right) = \log[1 + \exp(\theta)], c(y, \phi) = \log \left[\binom{m}{my} \right].\end{aligned}$$

Example 4 (gamma distribution):

$Y \sim G(\mu, v)$. Then,

$$\begin{aligned}f(y, \theta, \phi) &= \frac{1}{\Gamma(v)} \left(\frac{vy}{\mu} \right)^v \exp \left(-\frac{vy}{\mu} \right) \frac{1}{y} \\ &= \exp \left\{ y \left(\frac{-v}{\mu} \right) + v \cdot \log \left(\frac{vy}{\mu} \right) - \log[\Gamma(v)] - \log(y) \right\} \\ &= \exp \left\{ \frac{\left[y \left(\frac{-1}{\mu} \right) + \log \left(\frac{1}{\mu} \right) \right]}{(1/v)} + v \cdot \log(vy) - \log[\Gamma(v)] - \log(y) \right\}.\end{aligned}$$

Therefore,

$$\theta = \frac{-1}{\mu}, \phi = \frac{1}{v},$$

$$\begin{aligned}
a(\phi) &= \frac{1}{v} = \phi, b(\theta) = -\log\left(\frac{1}{\mu}\right) = -\log(-\theta), \\
c(y, \phi) &= v \cdot \log(vy) - \log[\Gamma(v)] - \log(y) \\
&= \frac{1}{\phi} \cdot \log\left(\frac{y}{\phi}\right) - \log\left[\Gamma\left(\frac{1}{\phi}\right)\right] - \log(y).
\end{aligned}$$

Example 5 (inverse Gaussian distribution):

$Y \sim IG(\mu, \sigma^2)$. Then

$$\begin{aligned}
f(y, \theta, \phi) &= \sqrt{\frac{1}{2\pi\sigma^2 y^3}} \exp\left[\frac{-(y - \mu)^2}{2\mu^2 \sigma^2 y}\right] \\
&= \exp\left[\frac{-(y^2 - 2y\mu + \mu^2)}{2\mu^2 \sigma^2 y} - \frac{1}{2} \cdot \log(2\pi\sigma^2 y^3)\right] \\
&= \exp\left\{\left[\frac{y\left(\frac{-1}{2\mu^2}\right) + \frac{1}{\mu}}{\sigma^2} - \frac{1}{2\sigma^2 y} - \frac{1}{2} \cdot \log(2\pi\sigma^2 y^3)\right]\right\}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\theta &= \frac{-1}{2\mu^2}, \phi = \sigma^2, \\
a(\phi) &= \sigma^2 = \phi, b(\theta) = \frac{-1}{\mu} = -\sqrt{-2\theta}, \\
c(y, \phi) &= \frac{-1}{2} \left[\frac{1}{\sigma^2 y} + \log(2\pi\sigma^2 y^3) \right] = \frac{-1}{2} \left[\frac{1}{\phi y} + \log(2\pi\phi y^3) \right]
\end{aligned}$$

2. Properties of generalized linear model

Important Result:

Let the response Y has a distribution in the exponential family, taking the form

$$f(y, \theta, \phi) = \exp\left\{\frac{[y\theta - b(\theta)]}{a(\phi)} + c(y, \phi)\right\}.$$

Then,

1. $E(Y) = \mu = b'(\theta)$
2. $Var(Y) = a(\phi)b''(\theta)$

[Derivation:]

Let

$$l(y, \theta, \phi) = \log[f(y, \theta, \phi)]$$

Then,

$$E \left[\frac{\partial l(Y, \theta, \phi)}{\partial \theta} \right] = 0$$

and

$$-E \left[\frac{\partial^2 l(Y, \theta, \phi)}{\partial \theta^2} \right] = E \left[\frac{\partial l(Y, \theta, \phi)}{\partial \theta} \right]^2.$$

Thus,

$$\frac{\partial l(y, \theta, \phi)}{\partial \theta} = \frac{y - b'(\theta)}{a(\phi)}.$$

Then

$$E \left[\frac{\partial l(Y, \theta, \phi)}{\partial \theta} \right] = E \left[\frac{Y - b'(\theta)}{a(\phi)} \right] = \frac{E(Y) - b'(\theta)}{a(\phi)} = 0$$

and this gives

$$E(Y) = b'(\theta)$$

Also,

$$\frac{\partial^2 l(y, \theta, \phi)}{\partial \theta^2} = \frac{-b''(\theta)}{a(\phi)}.$$

Then

$$\begin{aligned} -E \left[\frac{\partial^2 l(Y, \theta, \phi)}{\partial \theta^2} \right] &= \frac{b''(\theta)}{a(\phi)} = E \left[\frac{\partial l(Y, \theta, \phi)}{\partial \theta} \right]^2 = E \left[\left(\frac{Y - b'(\theta)}{a(\phi)} \right)^2 \right] \\ &= \frac{E(Y - \mu)^2}{a^2(\phi)} \\ &= \frac{Var(Y)}{a^2(\phi)}. \end{aligned}$$

and hence

$$\frac{b''(\theta)}{a(\phi)} = \frac{Var(Y)}{a^2(\phi)} \Rightarrow Var(Y) = a(\phi)b''(\theta).$$

Example 1 (normal distribution, continue):

$Y \sim N(\mu, \sigma^2)$. Then

$$\theta = \mu, \phi = \sigma^2, a(\phi) = \sigma^2 = \phi, b(\theta) = \frac{\mu^2}{2} = \frac{\theta^2}{2}.$$

Therefore

$$E(Y) = b'(\theta) = \theta = \mu$$

and

$$Var(Y) = a(\phi)b''(\theta) = \phi = \sigma^2.$$

Example 2 (Poisson distribution, continue):

$Y \sim P(\mu)$. Then,

$$\theta = \log(\mu), \phi = 1, a(\phi) = 1 = \phi, b(\theta) = \mu = \exp(\theta).$$

Therefore

$$E(Y) = b'(\theta) = \exp(\theta) = \mu$$

and

$$\text{Var}(Y) = a(\phi)b''(\theta) = \exp(\theta) = \mu.$$

Example 3 (binomial distribution, continue):

$Y \sim \frac{B(m, p)}{m}, 0 \leq Y \leq 1 \Rightarrow$ *Binomial distribution in frequency.*

Then,

$$\theta = \log\left(\frac{p}{1-p}\right), \phi = \frac{1}{m},$$

$$a(\phi) = \frac{1}{m} = \phi, b(\theta) = \log\left(\frac{1}{1-p}\right) = \log[1 + \exp(\theta)].$$

Therefore

$$E(Y) = b'(\theta) = \frac{\exp(\theta)}{1 + \exp(\theta)} = p$$

and

$$\begin{aligned} \text{Var}(Y) &= a(\phi)b''(\theta) = \frac{1}{m} \left[\frac{\exp(\theta)}{1 + \exp(\theta)} - \left(\frac{\exp(\theta)}{1 + \exp(\theta)} \right)^2 \right] = \frac{p - p^2}{m} \\ &= \frac{p(1-p)}{m}. \end{aligned}$$

Example 4 (gamma distribution, continue):

$Y \sim G(\mu, \nu)$. Then,

$$\theta = \frac{-1}{\mu}, \phi = \frac{1}{\nu}, a(\phi) = \frac{1}{\nu} = \phi, b(\theta) = -\log\left(\frac{1}{\mu}\right) = -\log(-\theta),$$

Therefore

$$E(Y) = b'(\theta) = \frac{-1}{\theta} = \mu$$

and

$$\text{Var}(Y) = a(\phi)b''(\theta) = \frac{1}{\nu} \cdot \frac{1}{\theta^2} = \frac{\mu^2}{\nu}.$$

Example 5 (inverse Gaussian distribution, continue):

$Y \sim IG(\mu, \sigma^2)$. Then

Therefore,

$$\theta = \frac{-1}{2\mu^2}, \phi = \sigma^2, a(\phi) = \sigma^2 = \phi, b(\theta) = \frac{-1}{\mu} = -\sqrt{-2\theta},$$

Therefore

$$E(Y) = b'(\theta) = \frac{1}{\sqrt{-2\theta}} = \mu$$

and

$$Var(Y) = a(\phi)b''(\theta) = \frac{-1}{2} \cdot \frac{1}{(-2\theta)^{3/2}} \cdot (-2) \cdot \sigma^2 = \frac{1}{(-2\theta)^{3/2}} \cdot \sigma^2 = \mu^3 \sigma^2.$$

3. Link function

Canonical link function:

Let the response Y has a distribution in the exponential family, taking the form

$$f(y, \theta, \phi) = \exp \left\{ \frac{[y\theta - b(\theta)]}{a(\phi)} + c(y, \phi) \right\}$$

with link function $g(\mu) = g[E(Y)] = \eta = x\beta$. As $\eta = \theta$, the link function is called **canonical link**.

Commonly used link functions:

The canonical links for the following distributions are

1. Normal distribution:

$$\eta = g(\mu) = \mu.$$

2. Poisson distribution:

$$\eta = g(\mu) = \log(\mu).$$

3. Binomial distribution:

$$\eta = g(\mu) = \log\left(\frac{\mu}{1-\mu}\right) = \text{logit}(\mu).$$

4. Gamma distribution:

$$\eta = g(\mu) = \frac{-1}{\mu}.$$

5. Inverse Gaussian distribution:

$$\eta = g(\mu) = \frac{-1}{2\mu^2}.$$

Other useful link functions:

1. Probit link:

$$\eta = g(\mu) = \Phi^{-1}(\mu).$$

2. Complementary log-log link:

$$\eta = g(\mu) = \log[-\log(1 - \mu)].$$

3. Power family of links:

$$\eta = g(\mu) = \begin{cases} \mu^\lambda, \lambda \neq 0 \\ \log(\mu), \lambda = 0. \end{cases}$$