## *2.4 Measuring the goodness of fit:*

### Goodness of fit:

**Let the log-likelihood function**

$$\boldsymbol{l(\beta) = \sum_{i=1}^{n}\left\{log\left[\binom{m_i}{y_i}\right] + y_i \cdot log(\pi_i) + (m_i - y_i)\cdot log(1-\pi_i)\right\}}$$

$$\boldsymbol{= \sum_{i=1}^{n}\left\{log\left[\binom{m_i}{y_i}\right] + y_i \cdot log\left(\frac{\mu_i}{m_i}\right) + (m_i - y_i)\cdot log\left(\frac{m_i-\mu_i}{m_i}\right)\right\}}$$

**Then, the deviance function is**

$$\boldsymbol{D(y_1, y_2, \cdots, y_n|\tilde{\mu}_1, \tilde{\mu}_2, \cdots, \tilde{\mu}_n) = 2\sum_{i=1}^{n}\left[y_i log\left(\frac{y_i}{\tilde{\mu}_i}\right) + (m_i - y_i) log\left(\frac{m_i - y_i}{m_i - \tilde{\mu}_i}\right)\right]}$$

**where** $\boldsymbol{\tilde{\mu}_1 = [m_i\pi_i]_{\beta=\hat{\beta}}}$.

**Important result:**

**Let the null model involves *p* parameters,** $\boldsymbol{g(\mu_i) = x_i\beta}$ **and** $\boldsymbol{\tilde{\mu}_i = g^{-1}(x_i\hat{\beta}), i = 1, 2, \cdots, n}$. **Then,**

$$\boldsymbol{D(y_1, y_2, \cdots, y_n|\tilde{\mu}_1, \tilde{\mu}_2, \cdots, \tilde{\mu}_n) \xrightarrow{m_i\to\infty} \chi^2_{n-p}}$$

**under the following assumptions:**

1. **The observations are distributed independently, according to the binomial distribution. The case of over-dispersion is not considered.**
2. **The sample size** $\boldsymbol{n}$ **is fixed,** $\boldsymbol{m_i \to \infty}$ **and in fact** $\boldsymbol{m_i\pi_i(1-\pi_i) \to \infty}$ **for** $\boldsymbol{i = 1, 2, \cdots, n}$.