Maximum Likelihood estimation for binomial categorical data:

In practice the parmeter values for binomial distribution are unknown. Using sample data, we estimate the parameters.
The maximum likelihood estimate of a parameter is the parameter value for which the probability of the observed data takes its greatest value. It is the parameter value at which the likelihood function takes its maximum.

$$f(x) = {}^nc_x\, p^x q^{n-x}$$
$$= {}^nc_x\, p^x (1-p)^{n-x}$$

likelihood function $\Rightarrow L(\cdot) = \prod_{i=1}^{n} ({}^nc_x)\, p^{\Sigma x_i} (1-p)^{n-\Sigma x_i}$

log likelihood $\Rightarrow \ln L(\cdot) = \ln\left(\prod_{i=1}^{n} ({}^nc_x)\right) + \Sigma x_i \ln(p) + (n-\Sigma x_i)\ln(1-p)$

taking derivative w.r.t p & equate it to zero

$$\frac{d \ln L(\cdot)}{d(p)} = 0 + \Sigma x_i\left(\frac{1}{p}\right) + (n-\Sigma x_i)\frac{1}{1-p}(-1) = 0$$

$$\frac{\Sigma x_i}{p} - \frac{(n-\Sigma x_i)}{1-p} = 0$$

$$\frac{(1-p)\Sigma x_i - pn + p\Sigma x_i}{p(1-p)} = 0$$

$$\frac{\Sigma x_i - \cancel{p\Sigma x_i} - pn + \cancel{p\Sigma x_i}}{p(1-p)} = 0$$

$$\frac{\Sigma x_i - np}{p(1-p)} = 0$$

$$\Sigma x_i - np = 0$$

$$\Sigma x_i = np$$

$$p = \frac{\Sigma x_i}{n}$$

So

$$\boxed{\hat{p} = \frac{\Sigma x_i}{n}}$$ in our notation $$\boxed{\hat{\pi} = \frac{\Sigma x_i}{n}}$$

Scanned with CamScanner

For most GLM, calculation of ML parameter estimates is computationally complex. So software uses an algorithm known as Newton-Raphson Algorithm. For the GLM's ML estimation, it is sometimes called iteratively reweighted least squares. The algorithm starts at an initial guess for the parameter values that maximize the likelihood function. The ML estimators are approximately normal for large samples.

⇒ **Inference for logistic regression:**

We have studied how logistic regression helps describe the effects of a predictor on a binary response variable. Now we will discussed statistical inference for the model parameters.

Confidence Interval for the slope (β_1):

The large-sample Wald confidence interval for the parameter β_1 in the logistic regression model, $\log(\pi(x)) = \beta_0 + \beta_1 x$ is.

$$\hat{\beta}_1 \pm Z_{\alpha/2}(SE)$$

For the odd ratio (e^{β_1})

$$e^{\hat{\beta}_1 \pm Z_{\alpha/2}(SE)}$$

where $Z_{\alpha/2}$ is the value for the standard normal density curve.

Significance Testing:

For the logistic regression model, $H_0 : \beta_1 = 0$ states that the probability of success is independent of x. **Wald test** statistics are simple. For large samples:

$$Z = \frac{\hat{\beta}_1 - \beta_1}{SE} = \frac{\hat{\beta}_1}{SE} \quad \because \overset{H_0:}{\beta_1 = 0}$$

has a standard normal distribution when $\beta_1 = 0$. Refer Z to the standard normal table to get a one-sided or two-sided P-value. Equivalently, for the two sided $H_1: \hat{\beta}_1 \neq 0$,

$$Z^2 = \left(\frac{\hat{\beta}_1}{SE}\right)^2$$

has a large sample chi-squared null distribution with $df = 1$.

Although the wald test is adequate for large samples, the **likelihood-ratio test** is more powerful and more reliable for sample sizes often used in practice. The likelihood ratio test statistic compares the maximum (L_0) of the log-likelihood function when $\beta_1 = 0$ to the maximum (L_1) of the log-likelihood function for unrestricted β. The test statistic

$$LR = -2(L_0 - L_1)$$

follows chisquare distribution with $df = 1$.

Scanned with CamScanner