Sequential Probability Ratio Testing

LRT and Neyman-Pearson testing of hypothesis are based on fixed constant sample size. In these types of test number of random variables has been supposed to be chosen before experimentation starts. Also we fix α and try to minimize β.

In fixed sample set up, when we change the model in one or more directions, the tradition process ~~is~~ fail.

One alternative to traditional method is Wald's sequential probability ratio testing (SPRT). In SPRT, sample size (n) is a random variable. The principal feature of such a procedure is a sampling scheme which lays down a rule under which one decides at each stage of the sampling whether to stop or to continue sampling.

Suppose we want to test the hypothesis, $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ for the distribution with p.d.f. $f(x;\theta)$. For any positive integer m, the likelihood function of a sample $x_1, x_2, \ldots, x_m$ from the population with p.d.f. $f(x;\theta)$ is given by

$$L_{1m} = \prod_{i=1}^{m} f(x_i, \theta_1) \text{ when } H_1 \text{ is true}$$

and $$L_{0m} = \prod_{i=1}^{m} f(x_i, \theta_0) \text{ when } H_0 \text{ is true.}$$

and the likelihood ratio λ_m is given by

$$\lambda_m = \frac{L_{1m}}{L_{0m}} = \prod_{i=1}^{m} \frac{f(x_i, \theta_1)}{f(x_i, \theta_0)} \quad ; \quad m = 1, 2, \ldots$$

The SPRT of testing H_0 against H_1 is defined as follows,

At each stage of the experiment (at the m th trial for any integral value m), the likelihood ratio λ_m is computed.

(i) If $\lambda_m \geq A$, we terminate the process with the rejection of H_0.

(ii) If $\lambda_m \leq B$, we terminate the process with the acceptance of H_0 and

(iii) If $B < \lambda_m < A$ we continue sampling by taking an additional observation.

Here A and B are the constants which are determined by the relation

$$A = \frac{1-\beta}{\alpha}, \quad B = \frac{\beta}{1-\alpha}$$

where α and β are the probabilities of type I error and type II error respectively.

For computational ease, one can use

$$\log \lambda_m = \sum_{i=1}^{m} \log \frac{f(x_i, \theta_1)}{f(x_i, \theta_0)}$$

$$= \sum_{i=1}^{m} z_i$$

Where $z_i = \log \frac{f(x_i, \theta_1)}{f(x_i, \theta_0)}$.

In terms of z_i's, SPRT is defined as follows:

(i) If $\sum z_i \geq \log A$, reject H_0

(ii) If $\sum z_i \leq \log B$, reject H_1

(iii) If $\log B < \sum z_i < \log A$ continue sampling by taking an additional observation.

Example Let X be a discrete random variable having the p.m.f.

$$f_\theta(x) = \begin{cases} \theta^x (1-\theta)^{1-x} & \text{if } x = 0,1 \\ 0 & \text{o.w.} \end{cases}$$

In this case, if $X_1, X_2, \ldots$ are independent and identically distributed random variables.

Define $S_m = \sum_{i=1}^{m} X_i$

$$L_{1,m} = \theta_1^{\sum x_i} (1-\theta_1)^{m - \sum x_i}$$

$$L_{0,m} = \theta_0^{\sum x_i} (1-\theta_0)^{m - \sum x_i}$$

$$\frac{L_{1m}}{L_{0m}} = \left(\frac{\theta_1}{\theta_0}\right)^{\sum x_i} \left(\frac{1-\theta_1}{1-\theta_0}\right)^{m - \sum x_i}$$

or $\log \frac{L_{1m}}{L_{0m}} = \sum x_i \log \frac{\theta_1}{\theta_0} + (m - \sum x_i) \log \frac{1-\theta_1}{1-\theta_0}$

Now SPRT procedure will be as follows:

$\boxed{1}$ Continue taking observations as long as

$$\log \frac{\beta}{1-\alpha} < \log \frac{L_{1m}}{L_{0m}} < \log \frac{1-\beta}{\alpha}$$

$\boxed{2}$ Accept H_0 if $\log \frac{L_{1m}}{L_{0m}} \le \log \frac{\beta}{1-\alpha}$

$\boxed{3}$ Reject H_1 if $\log \frac{L_{1m}}{L_{0m}} \ge \log \frac{1-\beta}{\alpha}$

$\boxed{1'}$ $\Rightarrow$ $$\log \frac{\beta}{1-\alpha} < \Sigma x_i \log \frac{\theta_1}{\theta_0} + (m - \Sigma x_i) \log \frac{1-\theta_1}{1-\theta_0} < \log \frac{1-\beta}{\alpha}$$

$\Rightarrow$ $$a_m < \Sigma x_i < r_m$$

where $$a_m = \frac{\log \frac{\beta}{1-\alpha}}{\log \frac{\theta_1}{\theta_0} - \log \frac{1-\theta_1}{1-\theta_0}} + m \frac{\log \frac{1-\theta_0}{1-\theta_1}}{\log \frac{\theta_1}{\theta_0} - \log \frac{1-\theta_1}{1-\theta_0}}$$

and $$r_m = \frac{\log \frac{1-\beta}{\alpha}}{\log \frac{\theta_1}{\theta_0} - \log \frac{1-\theta_1}{1-\theta_0}} + m \frac{\log \frac{1-\theta_0}{1-\theta_1}}{\log \frac{\theta_1}{\theta_0} - \log \frac{1-\theta_1}{1-\theta_0}}$$

we reject H_0 if $\Sigma x_i > r_m$

we accept H_0 if $\Sigma x_i < a_m$.

$\boxed{\text{Note}}$ The sampling process may be carried out graphically. As soon as the point (m, s_m) lies on or below the line a_m or above the line for r_m, sampling will be stopped.