With EM Algorithm to Do Tests of Hypotheses

Xiangzhong Fang
School of Mathematical Science, Peking University
Beijing, China
xzfang@math.pku.edu.cn

1. Introduction and Method

In statistics, a fundamental problem is tests of hypotheses. For small sample case, it is needed to find a suitable function of both of random variables and parameters. But the suitable function is not easy to find for many situations. This paper aims to present a general method for tests of hypotheses and using EM-type algorithms to perform the calculation.

Suppose the random variables, Z, follow a joint probability distribution, P_θ, upon which inference will be based; here θ takes its values in a set, Θ; Let $g(\theta)$ be a real-valued function on Θ. Consider hypotheses

$$H_0 : g(\theta) \geq c, \quad H_1 : g(\theta) < c,$$

where c is a constant.

Denote the MLE of θ as $\hat{\theta}$. For any real number r, define

$$\overline{P}(r) = \sup\{P_\theta\left(g(\hat{\theta}) < r\right) : \theta \in \Theta \text{ and } g(\theta) \geq c\}. \tag{1}$$

Given significance level $\alpha \in (0,1)$, a real r_0 is to be find, such that for $r \leq r_0$, $\overline{P}(r) \leq \alpha$, whereas $\overline{P}(r) > \alpha$ if $r > r_0$. Then we reject the null hypotheses if $g(\hat{\theta}) < r_0$, accept it if $g(\hat{\theta}) \geq r_0$.

EM algorithm is just used to find r_0. Any EM-type algorithms could be used, just because the probability $P_\theta\left(g(\hat{\theta}) < r\right)$ can be regarded as the likelihood function of random variable, $I(g(\hat{\theta}) < r)$, where $I(\cdot)$ is the indicator function.

Let $f(y|\theta)$ be a density for the complete-data y. Starting with an initial value $\theta^{(0)} \in \Theta$, the algorithm finds $\overline{P}(r)$, by iterating between the following two steps ($t = 0, 1, \ldots$):

E step: Impute the unknown complete-data loglikelihood $L(\theta|Y) = \log f(Y|\theta)$ by its conditional expectation given the current estimate $\theta^{(t)}$:

$$Q\left(\theta|\theta^{(t)}\right) = E_{\theta^{(t)}}\left(L(\theta|Y)\,|g(\hat{\theta}) < r\right),$$

M step: Determine $\theta^{(t+1)}$ by maximizing the imputed loglikelihood $Q\left(\theta|\theta^{(t)}\right)$:

$$Q\left(\theta^{(t+1)}|\theta^{(t)}\right) \geq Q\left(\theta|\theta^{(t)}\right), g(\theta) \geq c, \theta \in \Theta.$$

So by the general theory of EM algorithm, $P_{\theta^{(t)}}\left(g(\hat{\theta}) < r\right)$ is increasing in $t(= 1, 2, \ldots)$, and converge to $\overline{P}(r)$. Since $\overline{P}(r)$ is increasing in r, it is easy to find r_0.

2. Example:Tests about the Ratio Between Two Proportions

In medicine and industry, in order to assess the efficacy of a new treatment, comparison of proportions in two groups is often used. Let S_1 and S_2 be two statistically independent binomial random variables with parameters n_1, p_1 and n_2, p_2, respectively. Denote $\hat{g} = (S_1/n_1)/(S_2/n_2)$. Consider the hypotheses:

$$H_0 : p_1/p_2 \geq c \quad \longleftrightarrow \quad H_1 : p_1/p_2 < c.$$

By the proposed method, after simple ratiocination, we get

$$p_1^{(t+1)} = \frac{E_{p^{(t)}}(S_1|\hat{g} < r) + \lambda}{n_1 + \lambda}, \quad p_2^{(t+1)} = \frac{E_{p^{(t)}}(S_2|\hat{g} < r) - \lambda}{n_2 - \lambda},$$

where λ is the solution of the following equation

$$\frac{E_{p^{(t)}}(S_1|\hat{g} < r) + \lambda}{n_1 + \lambda} \cdot \frac{n_2 - \lambda}{E_{p^{(t)}}(S_2|\hat{g} < r) - \lambda} = c, \quad -E_{p^{(t)}}(S_1|\hat{g} < r) < \lambda < E_{p^{(t)}}(S_2|\hat{g} < r).$$

REFERENCES

Meng, X.L., Dyk, V.D., (1997) The EM algorithm-an old folk-song sung to a fast new tune. *J.R. Statist. Soc. B*, **59**, 511-567.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) Maximum likelihood from incomplete data via EM algorithm (with discussion). *J.R.Statist.Soc. B*, **39**,1-38.

Liu, C. and Rubin, D.B. (1994) The ECME algorithm:a simple extension of EM and ECM with fast monotone convergence. *Biometrika*, **81**, 633-648.

Fessler, J.A. and Hero III, A.O. (1994) Space alternating generalized expection maximization algorithm. *IEEE Trans. Siginal Process.*, **42**, 2664-2677.

RESUME Xiangzhong Fang is an associate professor in School of Mathematical Science of Peking University. He holds a BSc degree (1983) in Mathematics from Heilongjiang University, and a MSc degree(1991) and PhD degree(1996) in Statistics from Peking University.