数理解析研究所講究録
第526巻 1984年 149-158

不良品の見落しがある検査について

長岡技術科学大学　藤野和建

製品ロットの検査で，検査方法が不完全なために不良品の見落しがおこるとき，この見落しの確率と，見落されて市場にでる不良品の個数を推定する方法を論ずる。

この2つの量が推定できるためには，いくつかの製品が少なくとも2度検査される必要がある。ここではまず，最初の検査で良品とされたもののみを再検査する場合を論じ，次に全製品を再検査する場合を考える。

1. 良品のみ再検査の場合

N個の製品よりなるロットを全数検査し，そこで良品とみなされたもの全数を再度検査する。

不良品の総数がMで，そのそれぞれの検出される確率が一定値θに等しいとすれば，1回目，2回目の検出される不良品の個数 x_1, x_2 の同時分布は

$$P(x_1, x_2; y, \theta) = \frac{M!}{x_1!\,x_2!\,y!}\theta^{x_1}(\theta\tau)^{x_2}(\tau^2)^{y} \qquad (1)$$

となる。ただし，$\tau = 1 - \theta$ で，$y = M - x_1 - x_2$ は2回とも見落された不良品の数を表わす。

ここで　$m = x_1 + x_2$　とし，$p(x\,;\,n, p)$　を２項分布の確率とすれば，(1)は

$$p(x_1, x_2\,;\,y, \theta) = p(x_2\,;\,m, \gamma_1)\,p(m\,;\,M, \gamma_2)$$

ただし，　$\gamma_1 = \tau/(1+\tau)$

$\gamma_2 = 1 - \tau^2$

と，mの分布と，mを与えたときのx_2の条件付分布に分解される。

これから，$\hat{\gamma}_1 = x_2/m$，したがって

$$\hat{\theta}_c = (x_1 - x_2)/x_1 \tag{2}$$

となり，さらに　$\hat{\gamma}_2 = (x_1^2 - x_2^2)/x_1^2$　から

$$\hat{M}_c = [\,m/\hat{\gamma}_2\,] = [\,x_1^2/(x_1 - x_2)\,]$$
$$\hat{y}_c = \hat{M}_c - m = [\,x_2^2/(x_1 - x_2)\,] \tag{3}$$

となる。ここに$[x]$はxを超えない最大の整数を示す。

この条件付最尤法はSanathanan (1972) による。

1.1　最尤推定

(1)をθで微分して0とおくと

$$\hat{\theta}' = m/(m + x_2 + 2y) \tag{4}$$

となる。これを(1)のθに代入するとyの尤度が得られるがそれは次の形となる。

$$l(y\,;\,x_1, x_2) = \log p(x_1, x_2\,;\,y, \hat{\theta}')$$

$$= \log((m+y)!) - \log(y!)$$
$$+ (x_2+2y)\log(x_2+2y)$$
$$-(m+x_2+2y)\log(m+x_2+2y) + \text{const.}$$

この値を最大にする整数$\hat{y}$がyの最尤推定量で，θの最尤推定量は$\hat{y}$を(4)に代入して得られる。

$\hat{y}$については次の定理が成立つ($m \geq 1$)。

定理　$x_2 \geq x_1+2$のとき，yには有限な最尤解が存在しない。$x_2 < x_1+2$では有限な解が存在して$\hat{y} \leq \max(0, y_0)$となる。ここに$y_0$は$(2x_2^2+x_1x_2-2x_1-4)/2(x_1-x_2+2)$を下まわらない最小の整数を意味する。

データが与えられたとき，$\hat{y}$を求めるのは幾分面倒であるが

$$\hat{y}^* = [(x_2-1)^2/(x_1-x_2+2)] \tag{5}$$

が$\hat{y}$に対する満足すべき近似値を与える。

1.2　推定量の性質の数値的な検討

最尤推定量と条件付最尤推定量はそれぞれ$x_2 \geq x_1+2$および$x_2 \geq x_1$で存在しなくなるが，このことのおこる確率はMとθがある程度大きければ小さい。たとえばM=20，θ=0.7では，それぞれ0.002および0.009であって，実際上無視できる大きさになる。

次に各推定量の，それらが有限になるという条件のもとでの平均と分散を，$M = 20, 50, 100$，$\theta = 0.1\ (0.1)\ 0.9$ の場合について求め，比較してみた。

M の推定量では，$\hat{M}$ の方が $\hat{M}_c$ よりも，θ がある程度大きいときには偏りが小さく，分散もまた小さい。一方，θ の推定については，$\hat{\theta}_c$ よりも $\hat{\theta}$ の方が幾分偏りが小さくなる。ただし，分散は $\hat{\theta}_c$ の方が多少大きい。

1.3 不良品の検出確率が変動する場合

不良品の検出される確率が，不良品ごとにことなるものとして，この確率 $\theta_i\ (i = 1, 2, \cdots, M)$ が，互いに独立に同一の分布に従う確率変数 Θ_i の実現値であるとする。

このとき，事象

不良品 $1, \cdots, x_1$ は第1回目に検出

不良品 $x_1 + 1, \cdots, m$ は第2回目に検出

不良品 $m + 1, \cdots, M$ は検出されない

の，$\Theta_i = \theta_i\ (i = 1, \cdots, M)$ という条件のもとでの確率は

$$\prod_{j=1}^{x_1} \theta_j \cdot \prod_{j=x_1+1}^{m} \theta_j (1-\theta_j) \prod_{m+1}^{M} (1-\theta_j)^2$$

となる。

ゆえに，この事象の無条件の確率は

$$\{E(\Theta)\}^{x_1}\{E(\Theta(1-\Theta))\}^{x_2}\{E(1-\Theta)^2\}^{y}$$

となる。

このことから，１回目，２回目の検査で検出される不良品の個数 x_1，x_2 の分布は

$$p(x_1,\ x_2)$$
$$=\frac{M!}{x_1!\ x_2!\ y!}\{E(\Theta)\}^{x_1}\{E(\Theta(1-\Theta))\}^{x_2}\{E(1-\Theta)^2\}^{y} \qquad (6)$$

で与えられる。

ここで

$$E(\Theta)=\theta \quad , \quad Var(\Theta)=\sigma^2$$

とすれば，(6)は

$$p(x_1,\ x_2)$$
$$=\frac{M!}{x_1!\ x_2!\ y!}\ \theta^{x_1}\ (\theta\tau-\sigma^2)^{x_2}\ (\tau^2+\sigma^2)^{y} \qquad (7)$$

と書かれる。こうして，$\sigma^2=0$ の場合とくらべ，x_2 の周辺分布が小さい方に偏ることが知られる。

(7)を

$$p(x_1,\ x_2)=p(x_2\ ;\ m,\ \gamma_1^*)\,p(m\ ;\ M,\ \gamma_2^*)$$

$$\gamma_1^*=\frac{\theta\tau-\sigma^2}{\theta(1+\tau)-\sigma^2}$$

$$\gamma_2^*=1-\tau^2-\sigma^2$$

と書きなおすと，

$$\gamma_1^* \leqq \gamma_1 \quad , \quad \gamma_2^* \geqq \gamma_2$$

となるから，$\sigma^2=0$ の場合にくらべ，mはふさくなり易く，Mの推定量はふさくなり易く，θ の推定量は大きくなり易いことがわかる。

1.4　抜取検査の場合

ロットサイズNがあまりに大きければ，そのうちのn個を抜取って上述の検査をする。

この場合，抜取られたn個のサンプルの中の不良品の個数をM^*とすれば，x_1，x_2の分布は，$h(x; n, M, N)$を超幾何分布の確率として

$$p(x_1, x_2)$$
$$= h(M^*; n, M, N)\, p(x_1 : M^*, \theta)\, p(x_2 : M^* - x_1, \theta)$$

となる。

条件付最尤法では

$$\hat{M}^*_c = [\, x_1^2 / (x_1 - x_2)\,]$$
$$\hat{\theta}_c = (x_1 - x_2)/x_1 \qquad (8)$$
$$\hat{M}_c = [\, \hat{M}^*_c (N+1)/n\,]$$

となる。

これらの推定量の性質の数値的な検討はまだなされていない。

2. 200％検査の場合

次に，1回目に不良品とされたものを含め，すべての製品を再検査する場合を考える。

不良品の総数をM，そのぞれぞれの検出される確率が一定値θとし，1回不良品とされたものの個数を z_1，2回不良品とされたものの個数を z_2 とすれば，それらの同時分布は

$$p(z_1, z_2; y, \theta) = \frac{M!}{z_1!\, z_2!\, y!} (\theta^2)^{z_2} (2\theta\tau)^{z_1} (\tau^2)^{y} \qquad (9)$$

となる。ここに，$\tau = 1 - \theta$，$y = M - z_1 - z_2$。

ここで $m = z_1 + z_2$ とすれば (9) は

$$p(z_1, z_2; y, \theta) = p(z_2; m, \gamma_1)\, p(m; M, \gamma_2)$$

ただし $\gamma_1 = \theta / (1 + \tau)$

$\gamma_2 = 1 - \tau^2$

と書かれる。

これから，条件付最尤推定量として，

$$\hat{\theta}_c = 2z_2 / (z_1 + 2z_2)$$

$$\hat{M}_c = m + [z_1^2 / 4z_2] \qquad (10)$$

$$\hat{y}_c = [z_1^2 / 4z_2]$$

が得られる。

これらは $z_2 > 0$ で存在する。

2.1 最大推定

y と θ の対数尤度は，

$$l(y,\theta;z_1,z_2)$$

$$=\log((m+y)!)-\log(y!)$$

$$+(z_1+2z_2)\log\theta+(z_1+2y)\log\tau+\text{const} \qquad (11)$$

となる。$\partial l/\partial\theta=0$ から

$$\hat{\theta}'=(z_1+2z_2)/2(m+y) \qquad (12)$$

を得る。これを(11)の θ に代入して，y のみの対数尤度は

$$l(y;z_1,z_2)$$

$$=\log((m+y)!)-\log(y!)$$

$$+(z_1+2y)\log(z_1+2y)$$

$$-2(m+y)\log\{2(m+y)\}+\text{const} \qquad (13)$$

となる。

$$l'(y;z_1,z_2)$$

$$=\frac{1}{y+1}+\cdots+\frac{1}{y+m}-2\log\left\{\frac{2(m+y)}{z_2+2y}\right\}$$

$$=-\frac{z_2}{y}-\frac{z_1^2-2m(m-1)}{4y^2}+O(y^{-3})$$

だから，$z_2>0$ なら y の有限な最尤解が存在する。

$z_2=0$ のとき，$z_1=0$ なら解は不定，$z_1=1$ なら $\hat{y}=0$，$z_1\geq 2$ では有限な解が存在しない。

最尤解は

$$\hat{y}^* = [z_1 (z_1 - 2) / 4z_2]$$

できわめてよく近似される。

2.2　推定量の性質の数値的な検討

$z_2 = 0$ のときには条件付最尤推定量が存在しないが，この確率 $(1-\theta^2)^M$ は，Mがある程度大きければきわめてちいさい。

次に，$z_2 > 0$ という条件のもとでの各推定量の平均と分散を，§1.2 と同じ M と θ の組合せについて調べてみた。

M の推定量，θ の推定量とも，最尤推定量と条件付最尤推定量の間に大きな差異はみられない。いずれにせよ，さきの一旦良品とされたもののみを再検査する場合にくらべ，偏りもちいさく，分散も少ない推定量が得られることは注目に値する。

2.3　不良品の検出確率が変動する場合

§1.3 と同じモデルのもとで，z_1，z_2 の同時分布は

$$p(z_1, z_2) = p(z_2 ; m, \gamma_1^*)\, p(m, M, \gamma_2^*)$$

ただし
$$\gamma_1^* = \frac{\theta^2 + \sigma^2}{\theta(1+\tau) - \sigma^2}$$

$$\gamma_2^* = 1 - \tau^2 - \sigma^2$$

となる。

$$\gamma_1^* \geqq \gamma_1 \quad , \quad \gamma_2^* \leqq \gamma_2$$

であるから，今度もまた，$\sigma^2=0$ の場合にくらべて，M の推定量は小さくなり易く，θ の推定量は大きくなり易い。

2.4 抜取検査の場合

ロットサイズ N があまりに大きくて，n 個のサンプルを抜取って，それらを全数2回検査すれば，サンプル中の不良品の数を M^* とするとき，条件付最尤法では

$$\hat{M}^*_c = m + [\, z_1^2 / 4z_2 \,]$$

$$\hat{\theta}_c = 2z_2 / (z_1 + 2z_2)$$

$$\hat{M}_c = [\, \hat{M}^*_c \, (N+1) / n \,]$$

となる。

3. 結論

製品ロットの検査で，検査方法が不完全で不良品の見落しがおこるとき，見落しの確率とロット中の不良品の総数を推定するには，良品のみを再検査するより，200％検査を実施する方がずっとよい推定値を得やすい。