Pak. J. Statist.
Saleh-Aly Special Edition
2000, Vol. 16(2), pp.99-109

OPTIMAL ALLOCATION OF OBSERVATIONS UNDER CONSTRAINED SUPERVISION

Bikas K. Sinha
Stat/Math Division
Indian Statistical Institute, Calcutta, India

Bimal K. Sinha
Department of Math/Stat
University of Maryland Baltimore County, USA

Abstract

Following Chatterjee and Chatterjee (Amer. Journal of Math. and Mgmt. Sciences, 1987, 7, 271-295), we address the problem of unbiased estimation of the success probability P of a coin in a fixed number (n) of throws when the entire coin-tossing experiment is performed under supervision of three observers. The observers are supposed to have noted different segments of the results of the tosses, each one not knowing what the other two have observed. We investigate the problem of most efficient unbiased estimation of P and related issue using the Best Linear Unbiased Estimator (BLUE) based on the summary statistics provided by the observers.

1. Introduction

The *key* reference to this article is Chatterjee & Chatterjee (1987). Henceforth we abbreviate it as *CC*. This nicely written and highly interesting paper is bound to provoke thoughts to a serious reader on certain problems and issues discussed in it.

We address one such problem viz., **Combining Expert Estimates Which are Dependent**. *CC* described a possible model for exhibiting this dependence in relation to a coin - tossing experiment.

The problem is to estimate the *success* probability (P) for a coin and several observers have noted different segments of the tosses - each one NOT knowing what others have observed. The ith observer reports the summary statistics : r_i (the number of heads) and n_i (the number of tosses), $1 \leq i \leq k$.

In *CC* the kind of dependence studied corresponds to the case where there are m common tosses among all the observers. This fact is *known only to the planner* and *not* to the observers. Furthermore, no specific information as to the outcome of these tosses is known even to the planner. Under this scenario, *CC* derived an expression for the *BLUE* of P based on the outcomes as reported by the individual observers.

CC derived the following conditions on the *plan parameters* (n_i, m) in order to ensure non-negativity while three observers are in action $(k = 3)$:

$$n_1 + n_2 + n_3 \leq min[n_2 n_3(2m + n_1), n_3 n_1(2m + n_2), n_1 n_2(2m + n_3)]/m^2 \quad (1.1)$$

It appears that there is a slight mistake in their computation. The conditions should actually read as (see Appendix for details):

$$\begin{aligned} n_2 n_3(2m - n_1) &\leq m^2(n_2 + n_3 - n_1) \\ n_1 n_2(2m - n_3) &\leq m^2(n_2 + n_1 - n_3) \\ n_1 n_3(2m - n_2) &\leq m^2(n_1 + n_3 - n_2) \end{aligned} \quad (1.2)$$

However, the conclusions drawn in Table 1 of *CC* in regard to the signs of the coefficients in the BLUE seem to be valid.

In a general scenario with k observers, it is easy to see that the individual estimates of P, namely, r_i/n_i, are dependent and that the dispersion matrix of the estimates is given by $\Sigma = P(1-P)W$ where the elements w_{ij} of the matrix W can be written as

$$w_{ij} = n_{ij}/n_i n_j \quad (1.3)$$

where n_{ij} is the number of *common* tosses shared by the ith and jth observers.

It is also well known that the *BLUE* of P corresponds to the coefficients of the individual estimators determined as $W^{-1}\mathbf{1}/\mathbf{1}'W^{-1}\mathbf{1}$. This, however, does *not* ensure non-negativity of the coefficents of the individual estimators!

The algebraic complexity of the problems of (i) obtaining an explicit expression of the BLUE of P, and (ii) examining the nonnegativity of the coefficients of the BLUE in the general case of k observers is quite deep since there are $(2^k - 1)$ possible cross-sectional studies.

In this paper we take up this investigation in the case of three observers while they are involved in the most general form of dependence with the following *plan parameters*: $(n_1, n_2, n_3, n_{12}, n_{13}, n_{23}, n_{123})$. Among other things, we examine the conditions on the *plan parameters* in order to extract maximum possible information on the parameter P, subject to a given total number of tosses n. It turns out that it is indeed possible to attain the highest possible precision (least possible variance) in a variety of situations.

We conclude with some further open problems.

2. BLUE of P and Variance Inequality

Here we discuss the details of the derivation of the BLUE of P and establish the main variance inequality.

Write $a = n_{11}$, $b = n_{12}$, $c = n_{13}$, $d = n_{22}$, $e = n_{23}$, $f = n_{33}$, $g = n_{123}$. Define

$$n = a+b+c+d+e+f+g,\ \alpha = a+b+c+g,\ \beta = b+d+e+g,\ \gamma = c+e+f+g \tag{2.1}$$

With the above notations, it is easy to verify that the three individual estimators of P are given by $\hat{P}_1 = r_1/\alpha$, $\hat{P}_2 = r_2/\beta$, and $\hat{P}_3 = r_3/\gamma$, where r_1, r_2 and r_3 are the observed *outcomes* of the three observers. The dispersion matrix $\Sigma = (\sigma_{ij})$ of these estimators obviously has the following elements (apart from the multiplier $P(1-P)$): $\sigma_{11} = 1/\alpha$, $\sigma_{12} = (b+g)/\alpha\beta$, $\sigma_{13} = (c+g)/\alpha\gamma$, $\sigma_{22} = 1/\beta$, $\sigma_{23} = (e+g)/\beta\gamma$ and $\sigma_{33} = 1/\gamma$. Note that in order for Σ to be *pd*, it is *necessary* that $g < n$. Moreover, trivially, we must have: $\alpha > 0$, $\beta > 0$, $\gamma > 0$ and furthermore, $n > a + d + f$.

It is easy to show that

$$|\Sigma| = \frac{\alpha\beta\gamma + 2(b+g)(c+g)(e+g) - \alpha(e+g)^2 - \beta(c+g)^2 - \gamma(b+g)^2}{\alpha^2\beta^2\gamma^2} \tag{2.2}$$

We denote the numerator of (2.2) by A. Write $\Sigma^{-1} = (a_{ij})/|\Sigma|$. Routine calculations yield

$$a_{11} = \frac{\beta\gamma - (e+g)^2}{\beta^2\gamma^2}$$

$$
\begin{aligned}
a_{12} &= \frac{(c+g)(e+g)-(b+g)\gamma}{\alpha\beta\gamma^2} \\
a_{13} &= \frac{(b+g)(e+g)-(c+g)\beta}{\alpha\beta^2\gamma} \\
a_{22} &= \frac{\alpha\gamma-(c+g)^2}{\alpha^2\gamma^2} \\
a_{23} &= \frac{(b+g)(c+g)-\alpha(e+g)}{\alpha^2\beta\gamma}. \\
a_{33} &= \frac{\alpha\beta-(b+g)^2}{\alpha^2\beta^2}
\end{aligned} \tag{2.3}
$$

It is clear that

$$
\mathbf{1}'\Sigma^{-1}\mathbf{1} = \frac{a_{11}+a_{22}+a_{33}+2a_{12}+2a_{13}+2a_{23}}{|\Sigma|} \tag{2.4}
$$

The numerator B of (2.4) simplifies to

$$
\begin{aligned}
B = {} & [\alpha\beta\gamma(\alpha+\beta+\gamma)-\alpha^2(e+g)^2-\beta^2(c+g)^2-\gamma^2(b+g)^2 \\
& +2\alpha\beta(c+g)(e+g)+2\alpha\gamma(b+g)(e+g)+2\beta\gamma(b+g)(c+g) \\
& -2\alpha\beta\gamma(b+c+e+3g)]/\alpha^2\beta^2\gamma^2
\end{aligned} \tag{2.5}
$$

Using (2.2) and (2.5), we can compute the value of $\mathbf{1}'\Sigma^{-1}\mathbf{1}$. It is easy to verify that, as expected, this quantity is a symmetric function of (a, d, f) on one side and of (b, c, e) on the other. Recall that the variance of the BLUE of P is

$$
var(\hat{P}_{blue}) = \frac{1}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}. \tag{2.6}
$$

We claim that the following is true. Recall that $\alpha > 0$, $\beta > 0$, $\gamma > 0$, and Σ is *pd*.

Theorem 2.1. $var(\hat{P}_{blue}) \geq \frac{P(1-P)}{n}$, with equality if and only if either $n = \alpha$, or $n = \beta$, or $n = \gamma$, unless $n = a + d + f$.

Proof. See Appendix.

Remark 2.1. It is indeed quite interesting to clearly point out the practical significane of the conditions for equality as mentioned above. The condition $n = \alpha$ indicates that the first observer makes a note of the outcomes of all the n tosses. From (2.1), another implication of this condition is that $d = e = f = 0$ which in turn implies $\beta = b + g$ and

$\gamma = c + g$. In other words, when the first observer notes the outcomes of all the tosses, observer 2 notes the results of outcomes in conjunction with observer 1 alone ($b > 0$), or in conjunction with both the observers 1 and 3 ($g > 0$). Of course, a similar interpretation holds in respect of observer 3 and in the other two cases ($n = \beta$ and $n = \gamma$) as well.

Appendix A
Proof of (1.2)

For $k = 3$, write $a = n_1 - m$, $b = n_2 - m$, $c = n_3 - m$. It is easy to show that the dispersion matrix Σ of $(\hat{P}_1, \hat{P}_2, \hat{P}_3)'$, apart from $P(1-P)$, can be written as

$$\Sigma = D + \alpha\alpha \tag{A.1}$$

where

$$D = diag[a/(a+m)^2, b/(b+m)^2, c/(c+m)^2] \tag{A.2}$$

and

$$\alpha = [\sqrt{m}/(a+m), \sqrt{m}/(b+m), \sqrt{m}/(c+m)]' \tag{A.3}$$

Recalling that (see Rao, 1973)

$$\Sigma^{-1} = D^{-1} - \frac{D^{-1}\alpha\alpha' D^{-1}}{1 + \alpha' D^{-1}\alpha} \tag{A.4}$$

we readily get

$$\Sigma^{-1}\mathbf{1} = D^{-1}\mathbf{1} - \frac{D^{-1}\alpha\alpha' D^{-1}\mathbf{1}}{1 + \alpha' D^{-1}\alpha} \tag{A.5}$$

Straightforward computations yield

$$\begin{aligned} D^{-1}\mathbf{1} &= [(a+m)^2/a, (b+m)^2/b, (c+m)^2/c] \\ \alpha' D^{-1}\mathbf{1} &= \sqrt{m}[\frac{a+m}{a} + \frac{b+m}{b} + \frac{c+m}{c}] \\ \alpha' D^{-1}\alpha &= m[\frac{1}{a} + \frac{1}{b} + \frac{1}{c}] \\ D^{-1}\alpha &= \sqrt{m}[\frac{a+m}{a}, \frac{b+m}{b}, \frac{c+m}{c}]' \end{aligned} \tag{A.6}$$

The condition that the first component of $\Sigma^{-1}\mathbf{1}$ is nonnegative can be expressed as

$$\frac{(a+m)^2}{a} \geq \frac{\frac{m(a+m)}{a}[\frac{a+m}{a} + \frac{b+m}{b} + \frac{c+m}{c}]}{1 + m(\frac{1}{a} + \frac{1}{b} + \frac{1}{c})} \tag{A.7}$$

The above condition in turn simplifies to

$$abc \geq m(bc - ab - ac) \tag{A.8}$$

which is equivalent to

$$m^2(n_2 + n_3 - n_1) \geq n_2 n_3 (2m - n_1) \tag{A.9}$$

Appendix B
Proof of Theorem 2.1

It is enough to show that $\frac{A}{B} \geq \frac{1}{n}$ which is the same as proving $C = nA - B \geq 0$. There are many steps in the proof. We show that C, which can be expressed as a polynomial in g of degree four, is essentially *linear* in g with nonnegative coefficients!

Step 1. We first simplify C as

$$\begin{aligned} C = {} & (a+b+c+d+e+f+g)[\alpha\beta\gamma + 2(b+g)(c+g)(e+g) \\ & -\alpha(e+g)^2 - \beta(c+g)^2 - \gamma(b+g)^2] - \alpha\beta\gamma(\alpha+\beta+\gamma) \\ & +\alpha^2(e+g)^2 + \beta^2(c+g)^2 + \gamma^2(b+g)^2 + 2\alpha\beta\gamma(b+c+e+3g) \\ & -2\alpha\beta(c+g)(e+g) - 2\alpha\gamma(b+g)(e+g) \\ & -2\beta\gamma(b+g)(c+g) \end{aligned} \tag{B.1}$$

The coefficient of $\alpha\beta\gamma$ in the above can be simplified as

$$\begin{aligned} & a+b+c+d+e+f+g \\ & \quad -a-b-c-g-b-d-e-g-c-e-f-g \\ & \quad +2b+2c+2e+6g \\ & \quad = b+c+e+4g \end{aligned} \tag{B.2}$$

Thus, C can be written as

$$C = term1 + term2 - term3 - term4 - term5 - term6 - term7 - term8 \tag{B.3}$$

where term 1 (= terms involving $\alpha\beta\gamma$) is given by

$$\begin{aligned} term1 = {} & (b+c+e+4g)(a+b+c+g)(b+d+e+g) \\ & \times(c+e+f+g) \\ & = 4g^4 + g^3[(b+c+e) + 4(a+b+c) + 4(b+d+e) \\ & +4(c+e+f)] \\ & +g^2[(b+c+e)\{(a+b+c) + (b+d+e) + (c+e+f)\} \end{aligned}$$

$$
\begin{aligned}
&+4\{(a+b+c)(b+d+e)+(a+b+c)(c+e+f)\\
&\quad +(b+d+e)(c+e+f)\}]\\
&+g[(b+c+e)\{(a+b+c)(b+d+e)+(a+b+c)(c+e+f)\\
&+(b+d+e)(c+e+f)\}+4(a+b+c)(b+d+e)(c+e+f)]\\
&+(b+c+e)(a+b+c)(b+d+e)(c+e+f)\\
&\quad = 4g^4+g^3[4a+4d+4f+9b+9c+9e]\\
&+g^2[(b+c+e)(a+2b+2c+d+2e+f)+4\{(a+b+c)(b+d+e)\\
&+(a+b+c)(c+e+f)+(b+d+e)(c+e+f)\}]\\
&+g[(b+c+e)\{(a+b+c)(b+d+e)+(a+b+c)(c+e+f)\\
&+(b+d+e)(c+e+f)\}+4(a+b+c)(b+d+e)(c+e+f)]\\
&+(b+c+e)(a+b+c)(b+d+e)(c+e+f) \qquad (B.4)
\end{aligned}
$$

term 2 (= terms independent of α, β, γ) is given by

$$
\begin{aligned}
& term2 = 2(b+g)(c+g)(e+g)(a+b+c+d+e+f+g)\\
= \;& 2g^4+2g^3[(a+b+c+d+e+f)+(b+c+e)]\\
& +2g^2[(b+c+e)(a+b+c+d+e+f)+(bc+be+ce)]\\
& +2g[(bc+be+ce)(a+b+c+d+e+f)+bce]\\
& +2bce(a+b+c+d+e+f)\\
= \;& 2g^4+2g^3[a+d+f+2b+2c+2e]\\
& +2g^2[(b+c+e)(a+b+c+d+e+f)+(bc+be+ce)]\\
& +2g[(bc+be+ce)(a+b+c+d+e+f)+bce]\\
& +2bce(a+b+c+d+e+f) \qquad (B.5)
\end{aligned}
$$

term 3 (= terms involving α, α^2) is given by

$$
\begin{aligned}
term3 &= (a+b+c+g)(d+e+f)(e^2+2eg+g^2)\\
&= g^3[d+e+f]+g^2[(a+b+c+2e)(d+e+f)]\\
&\quad +g[\{e^2+2e(a+b+c)\}(d+e+f)]\\
&\quad +e^2(a+b+c)(d+e+f) \qquad (B.6)
\end{aligned}
$$

term 4 (= terms involving β, β^2) is given by

$$
\begin{aligned}
term4 &= (b+d+e+g)(a+c+f)(c^2+2cg+g^2)\\
&= g^3[a+c+f]+g^2[(b+d+e+2c)((a+c+f)]\\
&\quad +g[\{c^2+2c(b+d+e)\}(a+c+f)]\\
&\quad +c^2(b+d+e)(a+c+f) \qquad (B.7)
\end{aligned}
$$

term 5 (= terms involving γ, γ^2) is given by

$$\begin{aligned} term5 &= (c+e+f+g)(a+b+d)(b^2+2bg+g^2) \\ &= g^3[a+b+d]+g^2[(2b+c+e+f)(a+b+d)] \\ &\quad +g[\{b^2+2b(c+e+f)\}(a+b+d)] \\ &\quad +b^2(a+b+d)(c+e+f) \end{aligned} \tag{B.8}$$

term 6 (= terms involving $\alpha\beta$) is given by

$$\begin{aligned} & term6 = 2(c+g)(e+g)(a+b+c+g)(b+d+e+g) \\ = \;& 2g^4+2g^3[c+e+a+b+c+b+d+e] \\ & +2g^2[ce+(a+b+c)(b+d+e)+(c+e)(a+b+c+b+d+e)] \\ & +2g[ce(a+b+c+b+d+e)+(c+e)(a+b+c)(b+d+e)] \\ & +2[ce(a+b+c)(b+d+e)] \\ = \;& 2g^4+2g^3[a+2b+2c+d+2e] \\ & +2g^2[ce+(a+b+c)(b+d+e)+(c+e)(a+2b+c+d+e)] \\ & +2g[ce(a+2b+c+d+e)+(c+e)(a+b+c)(b+d+e)] \\ & +2[ce(a+b+c)(b+d+e)] \end{aligned} \tag{B.9}$$

term 7 (= terms involving $\alpha\gamma$) is given by

$$\begin{aligned} & term7 = 2(b+g)(e+g)(a+b+c+g)(c+e+f+g) \\ = \;& 2g^4+2g^3[b+e+a+b+c+c+e+f] \\ & +2g^2[be+(a+b+c)(c+e+f)+(b+e)(a+b+c+c+e+f)] \\ & +2g[be(a+b+c+c+e+f)+(b+e)(a+b+c)(c+e+f)] \\ & +2be(a+b+c)(c+e+f) \\ = \;& 2g^4+2g^3[a+2b+2c+2e+f] \\ & +2g^2[be+(a+b+c)(c+e+f)+(b+e)(a+b+2c+e+f)] \\ & +2g[be(a+b+2c+e+f)+(b+e)(a+b+c)(c+e+f)] \\ & +2be(a+b+c)(c+e+f) \end{aligned} \tag{B.10}$$

term 8 (= terms involving $\beta\gamma$) is given by

$$\begin{aligned} & term8 = 2(b+g)(c+g)(b+d+e+g)(c+e+f+g) \\ = \;& 2g^4+2g^3[b+c+b+d+e+c+e+f] \\ & +2g^2[bc+(b+d+e)(c+e+f)+(b+c)(b+d+e+c+e+f)] \\ & +2g[bc(b+d+e+c+e+f)+(b+c)(b+d+e)(c+e+f)] \end{aligned}$$

$$
\begin{aligned}
&+2bc(b+d+e)(c+e+f) \\
= \;& 2g^4 + 2g^3[2b+2c+d+2e+f] \\
&+2g^2[bc+(b+d+e)(c+e+f)+(b+c)(b+c+d+2e+f)] \\
&+2g[bc(b+c+d+2e+f)+(b+c)(b+d+e)(c+e+f)] \\
&+2bc(b+d+e)(c+e+f) \qquad \text{(B.11)}
\end{aligned}
$$

We now search for the coefficients of various powers of g in C. Our choice of the *factor* g rather than the others is a matter of convenience along with the fact that g is the *odd* term! All other *factors* appear in a symmetric fashion (as mentioned after (2.5)).

Step 2. Coefficient of g^4 in $C = 0$.

Step 3. Coefficient of g^3 in $C = 0$ because the coefficient of each *factor* a, b, c, d, e and f is 0.

Step 4. To determine the coefficient of g^2, first note that this is a quadratic in the rest of the factors $(a - f)$. We now separate this into two parts: terms involving a and those independent of a. For the first type, we simplify the coefficients of terms like ab, ac,...,af, and each of the coefficients is 0. The various terms independent of a are now collected and arranged in different groups. First, terms involving b^2, bc, bd, be and bf are simplified, resulting in 0. Next, terms involving c^2, cd, ce and cf are simplified, and again these are all 0. Finally, we have terms involving e^2, de, ef and df. All these are 0.

Hence the coefficient of g^2 is 0.

Step 5. We now determine the coefficient of g. Clearly this is a cubic in the factors $(a, \cdots, f)$. We start by collecting terms which are various powers of b such as b^3, b^2, b, and lastly terms independent of b.

Coefficient of $b^3 = 0$.

Coefficient of $b^2 = 0$.

Coefficient of b involves many terms. We arrange them in powers of c.

Coefficient of $c^2 = 0$.

Coefficient of c simplifies to $d + e + f$.

Constant terms simplify to $a(d + e + f) + f(d + e)$.

Finally, the terms independent of b are simplified as follows. We express them in powers of c such as c^3, c^2, c and lastly terms independent of c.

Coefficient of $c^3 = 0$.

Coefficient of $c^2 = 0$.

Coefficient of c has many terms. We express them in powers of d and get the coefficient of d^2 as 0, coefficient of d as $a+e+f$, and finally the term independent of d as $a(e + f)$. Thus, the total contribution from this part is: $cd(a + e + f) + ca(e + f)$.

We now collect and simplify terms which do not involve c. There are many terms here, and we express them in powers of e such as e^3, e^2, e and lastly terms independent of e.
Coefficient of $e^3 = 0$.
Coefficient of $e^2 = 0$.
Coefficient of e simplifies to $ad + af + df$.
Terms independent of e simplify to $4adf$.
Collecting all the above terms, we get the coefficient δ_1 of g as

$$\begin{aligned}\delta_1 &= bc(d+e+f) + ba(d+e+f) + bf(d+e) + ac(e+f) \\ &\quad +cd(a+e+f) + e(ad+af+df) + 4adf \qquad \text{(B.12)}\end{aligned}$$

Step 6. We now determine the terms in C which are independent of g. We do this by collecting terms in powers of a.
(*i*) Coefficient of $a^2 = 0$.

(*ii*) Coefficient of $a = (b+c+e)(b+d+e)(c+e+f) + 2bce - e^2(d+e+f) - c^2(b+d+e) - b^2(c+e+f) - 2ce(b+d+e) - 2be(c+e+f)$.
In the above,
coefficient of $b^2 = 0$.

coefficient of $b = (c+e+f)(c+d+2e) + 2ce - c^2 - 2ce - 2e(c+e+f) = cd + ce + cf + de + df$.

term independent of $b = f(cd + ce + de)$.

(*iii*) terms independent of a: we simplify them in powers of b.

coefficient of $b^3 = 0$.

coefficient of $b^2 = 0$.

coefficient of $b = cde + cdf + cef$.

term independent of $b = cdef$.

Combining all the above terms, we get δ_2 (= terms independent of g) as

$$\begin{aligned}\delta_2 &= ab(cd + ce + cf + de + df) + af(dc + de + ce) \\ &\quad +b(cde + def + cef) + cdef \qquad \text{(B.13)}\end{aligned}$$

Thus, C eventually simplifies to

$$C = g\delta_1 + \delta_2 \qquad \text{(B.14)}$$

which is ≥ 0. This completes the proof of the first part of the Theorem.

For the conditions for equality, we carefully examine the quantities δ_1 and δ_2. For C to be 0, we must have both $g\delta_1$ and δ_2 to be 0. It is not difficult to verify that, under the conditions implying nonsingularity of Σ, this happens only when one of the conditions mentioned in the Theorem holds, and this is also sufficient. Incidentally, we exclude the possibility of $n = a + d + f$ since this would totally nullify the spirit of constrained supervision! This completes the proof of the Theorem.

3. Concluding Remarks

A simple proof of Theorem 3.1 (along with a study of '=') has so far eluded us. For $k > 3$ supervisors, the algebra will be extremely cumbersome and a direct algebraic proof of an appropriate version of Theorem 2.1 would be almost impossible. Cost consideration (on the part of the supervisors) and variance estimation would be other interesting directions for future study.

References

Chatterjee, S. and Chatterjee, S. (1987). On Combining Expert Opinions. *American Journal of Mathematical and Management Sciences*, **7**, 271-295.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications.* John Wiley, New York.