SUPPLEMENT TO " DESIGNING EXPERIMENTS FOR ESTIMATING AN APPROPRIATE OUTLET SIZE FOR A SILO TYPE PROBLEM""

BY JESUS LOPEZ-FIDALGO[1], CATERINA MAY[2] AND JOSE ANTONIO MOLER[3]

[1]*Instituto de Ciencia de los Datos e Inteligencia Artificial, Universidad de Navarra, Spain. fidalgo@unav.es*

[2]*Dipartimento DiSEI, Università degli Studi del Piemonte Orientale, Italy. caterina.may@uniupo.it*

[3]*Departamento de Estadística, Informática y Matemáticas, Universidad Pública de Navarra, Spain. jmoler@unavarra.es*

1. Proofs of the theoretical results in section 3. To prove the Proposition 3.1 a previous lemma is needed.

LEMMA 1.1. *The curve*

$$\begin{cases} x(\phi) = G(\phi,\,\theta)/C, \\ y(\phi) = -G(\phi,\,\theta)\,\phi^2, \\ \phi \in [a,b]. \end{cases} \tag{1}$$

and the points A_2,A_3 are always in the fourth quadrant of the Cartesian plane, while its reflection and the points A_1,A_4 are always in the second quadrant.

PROOF. It is enough to observe that $x(\phi) > 0$ and $y(\phi) < 0$ for any $\phi \in [a,\, b]$ because $L > 0$ and $e^{\phi^2 L} > C > 0$ for any ϕ. □

PROOF OF PROPOSITION 3.1. The main point is to prove that the curve (1) is always above the segment $\overline{A_3A_2}$ and below $\overline{A_1A_2}$.

The proof will be organized in the following steps:

1. From Lemma 1.1, the curve (1) and the points A_3,A_2 are always in the fourth quadrant of the Cartesian plane, while its reflection and the points A_4,A_1 are always in the second quadrant.
2. We have $x'(\phi) < 0$ for any $\phi \in [a,\, b]$; it follows that $x(b) \le x(\phi) \le x(a)$.
3. From the first equation of (1) we have $G(\phi,\,\theta) = Cx(\phi)$; moreover, by the definition of $G(\phi,\theta)$,

   $$\phi^2 = \frac{1}{L}\log\left(\frac{C^2 x}{Cx-1}\right);$$

   then plugging into the second equation of (1), we obtain the cartesian equation of the curve:

   $$y(x) = -\frac{C}{L}\,x\log\left(\frac{C^2 x}{Cx-1}\right), \quad x \in [x(b), x(a)]. \tag{2}$$

4. Notice that $y \in \mathscr{C}^2([x(b),x(a)])$ and that

   $$y''(x) = -\frac{C}{Lx(Cx-1)^2} < 0;$$

   it follows that (2) is concave and therefore (1) is above the segment $\overline{A_3A_2}$.

5. To prove that (1) is below the segment $\overline{A_1A_2}$ it is enough to prove that the tangent to the curve in A_2 is below $\overline{A_1A_2}$ (which has a negative slope m); this means that the slope of (2) in $x = x(a)$ is greater than the slope of $\overline{A_1A_2}$.
   We have

$$y'(x) = \frac{C}{L}\left(\frac{1}{Cx-1} - \log\frac{C^2x}{Cx-1}\right) \tag{3}$$

   and then

$$\begin{aligned} y'(x)|_{x=x(a)} &= \frac{C}{L}\left(\frac{1}{G(a,\theta)-1} - \log\frac{C\,G(a,\theta)}{G(a,\theta)-1}\right) \\ &= \frac{C}{L}\left(\frac{e^{a^2L}-C}{C} - a^2L\right) \\ &= \frac{1}{L}e^{a^2L} - \frac{C}{L} - a^2C. \end{aligned} \tag{4}$$

   At this point there are two cases:

   (a) If $C < e^{a^2L}/(1+a^2L)$ then $y'(x)|_{x=x(a)} > 0$ and it is straightforward that the slope of the curve is greater than the slope of $\overline{A_1A_2}$. Note that in this case we have $y'(x) > 0$ for any $x \in [x(b), x(a)]$ (as in Figure 2 in the paper).
   (b) If $e^{a^2L}/(1+a^2L) < C < e^{\phi^2L}$, then $y'(x)|_{x=x(a)} < 0$ (as in Figure 3) and we have to prove that

$$y'(x)|_{x=x(a)} > m = -C\frac{a^2G(a,\theta)+b^2G(b,\theta)}{G(a,\theta)+G(b,\theta)} \tag{5}$$

   Taking into account (4) can be written as

$$\frac{C}{L}\left(\frac{1}{G(a,\theta)-1} - a^2L\right),$$

   the inequality (5) is equivalent to

$$\frac{1}{L}\frac{1}{G(a,\theta)-1} - a^2 > -\frac{a^2G(a,\theta)+b^2G(b,\theta)}{G(a,\theta)+G(b,\theta)},$$

   giving

$$\frac{1}{L}\frac{1}{G(a,\theta)-1} + \frac{(b^2-a^2)G(b,\theta)}{G(a,\theta)+G(b,\theta)} > 0,$$

   which is always satisfied since the left term is a sum of two positive quantities.

□

For proving Proposition 3.2 the following lemma is needed.

LEMMA 1.2. *From Proposition 3.1, ϕ_i can be equal to a or equal to b.
Then, depending on the fixed value of T_0, the convex hull is crossed by $\mathbf{c}$ in the fourth quadrant through $\overline{A_iA_{i+1}}$, $i = 1,2,3$:*

i. *if $T_0 \in (max\{0,\ (1-C_0)/C_0\},\ \frac{1}{C_0}\exp(-y_2L_0/x_2C_0) - 1]$ then the crossing point is in $\overline{A_1A_2}$;*

ii. *if* $T_0 \in (\frac{1}{C_0}\exp(-y_2L_0/x_2C_0)-1,\ \frac{1}{C_0}\exp(-y_3L_0/x_3C_0)-1]$ *then the crossing point is in* $\overline{A_2A_3}$*;*
iii. *if* $T_0 > \frac{1}{C_0}\exp(-y_3L_0/x_3C_0)-1$ *then the crossing point is in* $\overline{A_3A_4}$.

PROOF. Taking into account that $\mathbf{c}$ is given by

$$\mathbf{c}(\theta) = \frac{1}{2\sqrt{L}}\left(\frac{1}{C\sqrt{\log(C(T_0+1))}}, -\frac{\sqrt{\log(C(T_0+1))}}{L}\right)^T;$$

as $\log(C(T_0+1))>0$, then

$$T_0 > \frac{1-C}{C}.$$

Moreover, since $\partial g(\theta;T_0)/\partial L<0$ and $\partial g(\theta;T_0)/\partial C>0$, $\mathbf{c}$ always moves into the fourth quadrant. As only the vertices A_2 and A_3 can be in the fourth quadrant, as stated by Lemma 1.1, then $P_0=A_i$, $i=2,3$, are the only two situations where the optimal design reduces to one point. In such a case, $y_i/x_i=K$, and then if

$$T_{0i} = \frac{1}{C}\exp\left(-\frac{y_iL}{x_iC}\right)-1, \quad \text{for} \quad i=2,3,$$

we have that:

i. for $T_0 \in (max(0,\ (1-C)/C),\ T_{02}]$ the crossing point is in $\overline{A_1A_2}$;
ii. for $T_0 \in (T_{02},\ T_{03}]$ the crossing point is in $\overline{A_2A_3}$;
iii. for $T_0 > T_{03}$ the crossing point is in $\overline{A_3A_4}$,

which gives the thesis. □

PROOF OF PROPOSITION 3.2. The intersection between the line of $\mathbf{c}$ and the line containing the segment $\overline{A_iA_{i+1}}$ is

$$\begin{cases} y-y_i = \dfrac{y_{i+1}-y_i}{x_{i+1}-x_i}(x-x_i) \\ y = Kx, \end{cases}$$

hence the solution of the crossing point P_0 follows straightforwardly after some algebra. From the Elfving method we have that if the crossing point P_0 is in the side $\overline{A_iA_{i+1}}$, then the c-optimal design is given by

$$\left\{\begin{array}{cc} \phi_i & \phi_{i+1} \\ 1-p_i & p_i \end{array}\right\}$$

with $p_i = \|\overline{A_iP_0}\|/\|\overline{A_iA_{i+1}}\|$, where $\|\cdot\|$ is the euclidean norm. The result hence follows. □

2. Details for the sensitivity analysis. The following steps describe the procedure to perform a sensitivity analysis for the choice of the nominal values of the parameters.

Step 1: Consider $\phi\in[a,b]$ and the nominal values (C_0,L_0). The c-optimal design is obtained from Proposition 3.2,

$$\xi_c^{(0)} = \left\{\begin{array}{cc} \phi_i^{(0)} & \phi_{i+1}^{(0)} \\ 1-p_i^{(0)} & p_i^{(0)} \end{array}\right\}$$

Step 2: We consider a grid where the parameters C and L take potential actual values (C^*, L^*) in a neighborhood of the nominal values (C_0, L_0). Thus, we obtain the c-optimal design

$$\xi_c^* = \left\{ \begin{array}{cc} \phi_i^* & \phi_{i+1}^* \\ 1-p_i^* & p_i^* \end{array} \right\},$$

when the true values of the parameters is a pair (C^*, L^*) in the grid.

Step 3: For each (C^*, L^*) in the grid, the following values are obtained:

$$M_1 = (1-p_i^*)I(\phi_i^*, C^*, L^*) + p_i^* I(\phi_{i+1}^*, C^*, L^*)$$

$$Var_1(g) = \nabla g(C^*, L^*; T_0)^T M_1^{-1} \nabla g(C^*, L^*; T_0)$$

- Consider the nominal values (C_0, L_0), where $p_i^{(0)}$ and $\phi_i^{(0)}$ were obtained in Step 1, then, using the actual values for computing its FIM,

$$M_0 = (1-p_i^{(0)})I(\phi_i^{(0)}, C^*, L^*) + p_i^{(0)} I(\phi_{i+1}^{(0)}, C^*, L^*)$$

$$Var_0(g) = \nabla g(C^*, L^*; T_0)^T M_0^{-1} \nabla g(C^*, L^*; T_0)$$

- Compute the relative efficiency given by $Var_1(g)/Var_0(g)$.

3. Simulation study to check the goodness of the approximations. The simulations are performed in the following steps:

Step 1: For the nominal values (C_0, L_0) the optimal design from Proposition 3.2 for a fixed value T_0 is obtained. Following the notation in Proposition 3.2, n_i observations are randomly allocated at ϕ_i and $n_{i+1} = n - n_i$ at ϕ_{i+1}.

Step 2: The MLEs of C, L and $g(\theta)$ are computed. As the responses follow an exponential distribution with mean

$$\eta(\phi;\theta) = \frac{1}{C}\exp(L\phi^2) - 1,\ \phi \in \mathscr{X} = [a,b],$$

there are n_i responses from an exponential distribution with parameter λ_i that are denoted by $t_k^{(i)}$, $k = 1, \cdots, n_i$ and n_{i+1} with parameter λ_{i+1}, which are denoted by $t_k^{(i+1)}$, $k = 1, \cdots, n_{i+1}$ where

$$\lambda_j = \frac{C}{e^{L\phi_j^2} - C}, \qquad j = i, i+1.$$

The likelihood function depends on the sample obtained, $\mathbf{t}$, and the parameter values C and L,

$$\mathscr{L}_n = \mathscr{L}_n(\mathbf{t}, \theta) = \lambda_i^{n_i} e^{-\lambda_i \sum_{k=1}^{n_i} t_k^{(i)}} \lambda_{i+1}^{n_{i+1}} e^{-\lambda_{i+1} \sum_{k=1}^{n_{i+1}} t_k^{(i+1)}}.$$

By solving the equations $\partial \mathscr{L}_n / \partial C = 0$ and $\partial \mathscr{L}_n / \partial L = 0$ we have that:

$$\hat{\lambda}_j = \frac{n_j}{\sum_{k=1}^{n_j} t_k^j} = \frac{1}{\overline{T}_j}; \qquad j = i, i+1$$

Solving this system of equations we finally obtain the MLEs of C and L:

$$(6) \qquad \hat{C} = \left(\frac{(1+\overline{T}_{i+1})^{\phi_i^2}}{(1+\overline{T}_i)^{\phi_{i+1}^2}} \right)^{\dfrac{1}{\phi_{i+1}^2 - \phi_i^2}}, \qquad \hat{L} = \log \left(\frac{(1+\overline{T}_{i+1})}{(1+\overline{T}_i)} \right)^{\dfrac{1}{\phi_{i+1}^2 - \phi_i^2}}$$

The MLE of $g(\theta)$ is given by $g(\hat{\theta})$, where $\hat{\theta}^T = (\hat{C}, \hat{L})$.

Step 2 is repeated m times obtaining three m-vectors, $\hat{\mathbf{C}}$, $\hat{\mathbf{L}}$, $\hat{\mathbf{g}}$, which contain, respectively, the MLEs of C, L and $g(\theta)$ computed at each step.

Step 3: To analyze the goodness of the approximation, the covariance matrix of $\hat{\theta}$ is approximated by the empirical covariance matrix of $(\hat{C}, \hat{L})$. Because the MLE is asymptotically efficient, the covariance matrix of $\hat{\theta}$ should be similar to the Frechet-Cramer-Rao bound for n sufficiently large. In the multiparameter case, this bound is equal to $\Im = \partial\Psi/\partial\theta^T \, I(\phi, \theta) \, \partial\Psi^T/\partial\theta$, where $\Psi(\theta) = E(\hat{\theta})$ and

$$I(\phi, \theta) = \frac{e^{2\phi^2 L}}{C(e^{\phi^2 L} - C)^2} \begin{pmatrix} \dfrac{1}{C} & -\phi^2 \\ -\phi^2 & C\phi^4 \end{pmatrix}.$$

Observe that $(\partial\Psi/\partial\theta)_{ij} = \partial\Psi_i/\partial\theta_j = Cov(\hat{\theta}_j, \, \partial log(\mathscr{L}_n)/\partial\theta_j)$, where $\mathscr{L}_n$ is the likelihood function. In order to approximate this matrix, in step 2 we will also obtain, in each run, the bidimensional vector:

$$\partial log(\mathscr{L}_n)/(\partial\theta) = \begin{pmatrix} \partial log(\mathscr{L}_n)/\partial C \\ \partial log(\mathscr{L}_n)/\partial L \end{pmatrix}$$

$$(7) \qquad = \sum_{j=i}^{i+1} \sum_{k=1}^{n_j} \begin{pmatrix} \dfrac{1}{C_0} + \dfrac{1}{e^{L_0\phi_j^2} - C_0} + \dfrac{e^{L_0\phi_j^2}}{(e^{L_0\phi_j^2} - C_0)^2} t_k^{(j)} \\ \dfrac{-\phi_j^2 e^{L_0\phi_j^2}}{e^{L_0\phi_j^2} - C_0} \left[1 - \dfrac{C_0}{e^{L_0\phi_j^2} - C_0} t_k^{(j)} \right] \end{pmatrix}$$

then, we approximate $(\partial\Psi/(\partial\theta))_{ij}$ with the corresponding sample covariance.

EXAMPLE 1 (Example 1 in the paper revisited). We apply the simulation study in the framework of Example 1.

Step 1: Consider several values of T_0 within the three intervals giving the three different crossing situations (see Lemma 1.2 and Table 1).

Step 2: Allocate randomly $n = 1{,}000$ experimental points following the optimal design obtained from Proposition 3.2. The MLE values of C, L and $g(\theta)$ are obtained jointly with the pair of values of the vector (7) that we denote, respectively, $\mathbf{f}_n^{(1)}$ and $\mathbf{f}_n^{(2)}$. Step 2 is repeated $m = 1{,}000$ times and the 1,000-dimensional vectors $\hat{C}$, $\hat{L}$, $g(\hat{\theta})$, $\mathbf{f}_n^1$ and $\mathbf{f}_n^2$ are stored.

Step 3: Table 1 shows a high similitude between the target value $g(\theta)$ and its MLE $\hat{g}$. Also, between the variance obtained with the simulated $Cov(\hat{C}, \hat{L})$ denoted in the table as $\hat{Var}(\hat{g})$ and the variance obtained with $\Im$, which is denoted in the table as $Var(\hat{g})$.

Observe that for $T_0 = 0.5$ neither the estimator, nor the variance are similar. As T_0 is in the interval $(0.4887, \infty)$, values close to the boundary carry out a slower convergence of the estimators. In Table 2 we study the approach for $T_0 = 0.5$ of $g = 0.1426$ and $\hat{g}$ and $Var[\hat{g}]$ and $\hat{Var}[\hat{g}]$ for increasing values of the sample size n. The decreasing rate is smaller for the bias than it is for the variance. In summary, these results show in detail that the approximations give good results, unless in some particular cases. For the examples considered in this paper the approximations are good enough.

	$T_0 < T_{02}$		$T_{02} < T_0 < T_{03}$					$T_0 > T_{03}$		
T_0	0.5	2	20	200	2,000	2×10^4	2×10^5	3×10^5	6×10^6	10^8
g	0.14	1.37	2.66	3.62	4.39	5.05	5.63	5.72	6.38	6.95
$\hat{g}$	0.23	1.37	2.66	3.62	4.39	5.05	5.63	5.72	6.38	6.95
p	0.91	0.98	0.21	0.45	0.66	0.84	0.999	0.02	0.15	0.21
$\hat{Var}(g)^*$	87	5.8	1.4	0.9	0.8	0.6	0.6	0.6	1.0	1.4
$Var(\hat{g})^*$	562	6.1	1.4	0.9	0.8	0.6	0.6	0.6	1.0	1.4

* The variances must be multiplied by 10^{-4}

TABLE 1

Simulation performance for Example 1 in the paper using several values of T_0

n	1,000	5,000	10,000	100,000	1,000,000
$bias = \hat{g} - g$	0.0874	0.0421	0.0218	0.0014	-0.0021
$Var[\hat{g}]^*$	562.6	166.0	66.6	7.8	0.5
$\hat{Var}[\hat{g}]^*$	87.3	44.8	25.5	7.6	0.6

* The variances must be multiplied by 10^{-4}

TABLE 2

Goodness of the approximations for different values of n when $T_0 = 0.5$ in Example 1.