Chapter 14
Combining Optimization and Randomization Approaches for the Design of Clinical Trials

Victor Fossaluza, Marcelo de Souza Lauretto, Carlos Alberto de Bragança Pereira and Julio Michael Stern

Abstract Intentional sampling methods are non-randomized procedures that select a group of individuals for a sample with the purpose of meeting specific prescribed criteria. In this paper we extend previous works related to intentional sampling, and address the problem of sequential allocation for clinical trials with few patients. Roughly speaking, patients are enrolled sequentially, according to the order in which they start the treatment at the clinic or hospital. The allocation problem consists in assigning each new patient to one, and only one, of the alternative treatment arms. The main requisite is that the profiles in the alternative arms remain similar with respect to some relevant patients' attributes (age, gender, disease, symptom severity and others). We perform numerical experiments based on a real case study and discuss how to conveniently set up perturbation parameters, in order to yield a suitable balance between *optimality* – the similarity among the relative frequencies of patients in the several categories for both arms, and *decoupling* – the absence of a tendency to allocate each pair of patients consistently to the same arm.

> We describe a possible allocation that the experimenter judges to be free of covariate interference as haphazard. Randomization may be a convenient way of producing a haphazard design. We argue that it is the haphazard nature, and not the randomization, that is important. It seems therefore that a reasonable approximation to an optimal design would be to select a haphazard design. ... a detailed Bayesian consideration of possible covariates would almost certainly not be robust in that the analysis might be sensitive to small changes in judgments about covariates.
> Lindley (1982, p.438-439) - The Role of Randomization in Inference.

Victor Fossaluza
IME – USP, São Paulo, Brazil, e-mail: victorf@ime.usp.br

Marcelo de Souza Lauretto
EACH – USP, São Paulo, Brazil, e-mail: marcelolauretto@usp.br

Carlos Alberto de Bragança Pereira
IME – USP, São Paulo, Brazil, e-mail: cpereira@ime.usp.br

Julio Michael Stern
IME – USP, São Paulo, Brazil, e-mail: jstern@ims.usp.br

14.1 Introduction

Lindley (1991, p.47-48) illustrates the celebrated Simpson's paradox with a medical trial example. In such example, the association between two variables, Treatment and Recovery from a given illness, is reversed if the data is aggregated or disaggregated over a confounding variable, Sex, see also Yule (1903). Randomized and double-blind (masked) clinical trials are designed to shield the experiment from undesired bias effects caused by undue interference or deliberate manipulation of confounding variables. The introduction of randomized tests, first proposed in Peirce and Jastrow (1885) and popularized by Fisher (1935), has established a new paradigm for statistically valid empirical research. However, the standard sampling methods by randomized design are not always appropriate; for example, they have limited application when cost, ethical or inherent rarity constraints only allow the use of very small samples.

Intentional sampling methods are non-randomized procedures that select or allocate groups of individuals with the purpose of meeting specific prescribed criteria. Such methods can overcome some of the aforementioned limitations of standard randomized designs for statistical experiments. However, intentional or purposive sampling methods pose several interesting questions concerning statistical inference, as extensively discussed in Basu and Ghosh (1988), see also Schreuder et al. (1993, Sec.6.2), Brewer and Särndal (1983) and following discussions in Madow et al. (1983).

This paper focus on sequential allocation methods, following previous research in the field of intentional sampling presented in Fossaluza et al. (2009) and Lauretto et al. (2012). Particularly, we discuss an allocation scheme that combines aspects of intentional and randomized sampling methods.

The paper is organized as follows. Section 14.2 provides a brief discussion of sampling design under the perspective of linear regression superpopulation model. Section 14.3 elucidates how to use linear regression models to handle compositional data – which is of direct interest for our case study. Section 14.4 illustrates our approach of combining purposive sampling with random perturbation techniques for "*providing samples which are approximately balanced*", as stated by Royall and Pfeffermann (1982, p.20), in an application case concerning a clinical trial allocation. Section 14.5 presents and discusses our numerical experiments and results.

14.2 Linear Regression Superpopulation Model

We will introduce the basic ideas for our approach in the context of the Linear Regression Superpopulation Model, as presented by Royall and Pfeffermann (1982), Pereira and Rodrigues (1983) and Tam (1986).

$$E\left(\begin{bmatrix} y_s \\ y_r \end{bmatrix}\right) = \begin{bmatrix} X_s \\ X_r \end{bmatrix} \beta , \ \ \mathrm{Cov}\left(\begin{bmatrix} y_s \\ y_r \end{bmatrix}\right) = \begin{bmatrix} V_{s,s} & V_{s,r} \\ V'_{s,r} & V_{r,r} \end{bmatrix} .$$

Row i of the $n \times m$ matrix X contains the explanatory variables for individual i, and is known for the entire population. The response variable, y_i, is observed for the

individuals in a given sample, S, indexed by i in $s = [1, 2, \ldots, m]$, and unobserved for the remaining individuals, indexed by i in $r = [m+1, m+2, \ldots, n]$. (Whatever the sample, $s = [i_1, i_2, \ldots, i_m]$, we can always reorder the indices so to place them first.) We partition all vectors and matrices of the model accordingly and assume that $V_{s,s} > 0$.

We seek an optimal linear predictor, $p'y_s$, for a quantity of interest, $\kappa = q'y$, and define the auxiliary matrices: $t' = [t'_s, t'_r] = [p' - q'_s, -q'_r]$ and $M = \left(X'_s V_{s,s}^{-1} X_s\right)^{-1}$.

Since probability expectation is a linear operator, that is, for any random (vector) variable, z, $\mathrm{E}(Az+b) = A\mathrm{E}(z) + b$, one can compute the expected value of the prediction error,

$$\mathrm{E}\left(t'y\right) = \left(t'X\right)\beta = \left(t'_s X_s - q'_r X_r\right)\beta .$$

Hence, for a general parameter β, the predictor p is unbiased if and only if it obeys the *balance* constraint,

$$t'X = t'_s X_s - q'_r X_r = p'X_s - q'X = 0 .$$

Solving the normal linear system for the minimum variance unbiased estimator problem at hand yields the solution

$$t^*_s = (p^* - q_s) = V_{s,s}^{-1}\left(V_{s,r} + X_s U\left(X'_r - X'_s V_{s,s}^{-1} V_{s,r}\right)\right)q_r .$$

Finally, we can write the optimal (minimum-variance unbiased) prediction for the quantity of interest, $\kappa = q'y$, as

$$\hat{\kappa} = q'_s y_s + q'_r \left(X_r \hat{\beta} + V_{r,s} V_{s,s}^{-1}\left(y_s - X_s \hat{\beta}\right)\right) , \quad \text{where} \quad \hat{\beta} = M X_s V_{s,s}^{-1} y_s , \quad \text{and}$$

$$\mathrm{Var}\left(\hat{\kappa}\right) = q'_r \left(V_{r,r} - V_{r,s} V_{s,s}^{-1} V'_{r,s}\right)q_r + q'_r\left(X_r - V_{r,s} V_{s,s}^{-1} X_s\right)M\left(X_r - V_{r,s} V_{s,s}^{-1} X_s\right)' q_r .$$

The balance and optimality conditions obtained above can be used for choosing a good predictor p, but they can also be used to select a "good" sample $s = [i_1, i_2, \ldots, i_m]$. Many survey sampling studies are interested in population totals, where $q = \mathbf{1}$, that is, $\kappa = \mathbf{1}'y$ where $\mathbf{1}$ is the column vector of ones of appropriate dimension in the context. In this case, the balance equation takes the form $p'X_s = \mathbf{1}'X$.

Robustness is also a desirable characteristic of a survey design. Imagine our model is misspecified, say by omission of important covariates, Z, in the expanded covariate matrix $\widetilde{X} = [X, Z]$. Without loss of generality, assume that we use "orthogonal" covariates, for witch $X'Z = 0$ and $Z'Z = I$. We would like to still be able to make a useful prediction. In general, this desire is just wishful thinking if we know nothing about the ignored covariates in Z. Now, assume that we fix p as the *expansion* predictor, $p = (N/n)\mathbf{1}$, and choose a *representative* sample, X_s, for which the sample totals are (approximately) balanced, that is, $p'X_s \approx \mathbf{1}'X$. If we are lucky enough, the balance equation will also (approximately) hold for the omitted covariates, that is, $(N/n)\mathbf{1}'Z_s \approx \mathbf{1}'Z$. For further developments of this idea, see Royall and Pfeffermann (1982), Pereira and Rodrigues (1983) and Tam (1986).

But how do we manage to get lucky? According to Lindley (1982, p.438-439), that seems to be the role of randomization in experimental design. For complementary and mutually supportive views of the role of randomization in statistical design

of experiments, see Bonassi et al. (2009) and Stern (2008, 2011a). There is a vast literature in design-based random sampling that aims to achieve this goal. In this paper, we explore the use of purposive sampling. For a more extensive discussion of this approach and a complete application case, see Lauretto et al. (2012).

14.3 Compositional Models and Simplex Geometry

We begin this section reviewing some basic notions of Compositional Models and Simplex Geometry, as presented in Aitchison (1986) and Aitchison (2008). The open (m-1)-Simplex is the set $S^{m-1} = \{x \in R^m \mid x > 0 \wedge \mathbf{1}'x = 1\}$, where $\mathbf{1}$ in the vector of ones of appropriate dimension. The *closure-to-unity* transformation, from R_+^m to S^{m-1}, the *additive logratio transformation*, from S^{m-1} to the unrestricted R^{m-1} space, and the *centered logratio transformation*, from S^{m-1} to a hyperplane through the origin of R^m, are defined as

$$\text{clu}(x) = (1/\mathbf{1}'x)x \ , \quad \text{alr}(x) = \log\big((1/x_m)[x_1, \ldots x_{m-1}]\big) \ ,$$
$$\text{and } \text{clr}(x) = \log\big((1/g(x))[x_1 \ldots x_m]\big) \ , \quad g(x) = (x_1 x_2 \ldots x_m)^{1/m} \ .$$

It is easy to check the normalization conditions stating that $\mathbf{1}'\text{clu}(x) = 1$ and $\mathbf{1}'\text{clr}(x) = 0$, as well the following expressions of the inverse transformations

$$\text{alr}^{-1}(z) = \text{clu}\big(\exp([z_1, \ldots z_{m-1}, 0])\big) \ , \quad \text{clr}^{-1}(z) = \text{clu}\big(\exp([z_1, \ldots z_m])\big) \ .$$

We can introduce the *power* (scalar multiplication) operator, $\star$, and the *perturbation* (vector summation) operation, $\oplus$, providing a vector space structure for the Simplex, namely, $\alpha \star x = \text{clu}([x_1^\alpha, \ldots x_m^\alpha])$ and $x \oplus y = \text{clu}([x_1 y_1, \ldots x_m y_m])$. The perturbation operation can be interpreted as the effect of proportional decay rates in y over the fractional composition in x. The power operation can be interpreted as the α-times repeated effect of proportional decay rates. The perturbation operation defines an Abelian (commutative) group, where the identity element is $e = (1/m)\mathbf{1}$, and the inverse of a perturbation is given by $x^{-1} = \text{clu}([1/x_1, \ldots 1/x_m])$. Hence, we define the difference $x \ominus y = \text{clu}([x_1/y_1, \ldots x_m/y_m])$.

Next, we equip the Simplex with a metric structure. More specifically, we search for a distance function, $D_S(x, y)$, that exhibits the invariance properties that are most adequate for the purpose of compositional analysis. The most important of these invariance properties are:
- Perturbation invariance: For any perturbation, z, $D_S(x \oplus z, y \oplus z) = D_S(x, y)$.
- Permutation invariance: For any permutation matrix, P, $D_S(Px, Py) = D_S(x, y)$.
- Power scaling: For any $\alpha > 0$, $(1/\alpha)D_S(\alpha \star x, \alpha \star y) = D_S(x, y)$.

The following distance function exhibits all of these desirable invariance properties, as well as the standard properties required from a distance function, like positivity, symmetry and the triangular inequality.

$$D_S^2(x, y) = [\text{clr}(x) - \text{clr}(y)]' I [\text{clr}(x) - \text{clr}(y)] =$$

$$= [\text{alr}(x) - \text{alr}(y)]' H^{-1} [\text{alr}(x) - \text{alr}(y)] \ , \quad H_{i,j} = 2\delta_{i,j} + 1(1 - \delta_{i,j}) \ .$$

We can further extend the mathematical structure over the Simplex to a vector (finite Hilbert) space, defining the inner product

$$\langle x,y \rangle_S = \mathrm{clr}(x)' \, I \, \mathrm{clr}(y) = \mathrm{alr}(x)' H^{-1} \, \mathrm{alr}(y) \; .$$

Defining the norm, $\|x\|_S^2 = \langle x,x \rangle_S$, it is easy to check that $D_S(x,y) = \|x \ominus y\|_S$ and that $\|x\|_S = D_S(x,e)$. Finally, we can compose the additive logratio transformation with an orthogonalization operation that translates Aitchison's inner product for the Simplex to the standard inner product for the unrestricted Euclidean space. For example, we can use the Cholesky factorization $L'L = H$ to define the *isometric logratio transformation* $\mathrm{ilr}(x) = L^{-t} \mathrm{alr}(x)$, where L^{-t} denotes the transpose of L^{-1}. In this case, since $H^{-1} = L^{-1}L^{-t}$, we have

$$\langle x,y \rangle_S = \mathrm{ilr}(x)' L (L^{-1}L^{-t}) L' \mathrm{ilr}(y) = \mathrm{ilr}(x)' \, I \, \mathrm{ilr}(y) = \langle \mathrm{ilr}(x), \mathrm{ilr}(y) \rangle_E \; .$$

All these compatible vector space structures make it easy to develop linear regression models for compositional data, see Egozcue et al. (2000) and Pawlowsky-Glahn and Egozcue (2001). In the most simple terms, in the Euclidean space we can use the standard linear properties of the expectation operator, and consequent transformation properties for the covariance operator, as reviewed in the previous section. These properties suffice to prove Gauss-Markov theorem, allowing the computation of optimal unbiased estimators via least-squares linear algebra, see Whittle (2000, Sec.14.4). Hence, we can map compositional data, naturally presented in S^{m-1}, into R^m or R^{m-1}, analyse the data with linear regression models and, if so desired, map the models back to the Simplex.

The centered logratio transformation from S^{m-1} to R^m requires regression models under linear equality constraints. Meanwhile, the additive logratio transformation allows the use of standard (unconstrained) regression models in R^{m-1}. Moreover, in many practical applications, the last coordinate in the Simplex, x_m, is a significant proportion that may represent the preponderant component, an aggregate of many residual or indiscriminated components, a dispersion medium, etc. In this case, the coordinates generated by the additive logratio transformation have an intuitive interpretation. Furthermore, under appropriate conditions, the random variates corresponding to these coordinates in the statistical model have probability distributions with interesting statistical properties, as analysed in Aitchison and Shen (1980) and Pereira and Stern (2008, Sec.7). Finally, the isometric logratio transformation provides orthogonal coordinates in the unrestricted Euclidean space. However, in many practical applications, these orthogonal coordinates have a less intuitive interpretation than the oblique coordinates given by the additive logratio transformation.

Nevertheless, all these approaches will define compatible statistical models, render coherent statistical inferences, and mutually support each other for rich interpretations. In particular, it is easy to translate to the context of compositional data the results obtained in the last section concerning best unbiased predictors and well balanced samples. This will be the starting point of the next section.

14.4 Haphazard Intentional Allocation for Clinical Trials

In the setting discussed in Section 14.2, we choose purposively a sample S that represents well the covariates X, that is, so that $(N/n)\mathbf{1}'X_s \approx \mathbf{1}'X$. (We could even seek a sample that, simultaneously, also approximately minimizes the variance of our prediction $\hat{\kappa}$ for $\kappa = \mathbf{1}'y$.) However, at the same time, we would like to use some sort of randomization technique in order to obtain a sample S that is haphazard with respect to the omitted covariates, so that it is (probably) balanced, that is, $(N/n)\mathbf{1}'Z \approx \mathbf{1}'Z$. In this section, we introduce a technique for conciliating these goals, in a clinical trial case study.

The case study discussed in this work is the allocation of patients with Obsessive-compulsive disorder (OCD) between two treatment arms, see Fossaluza et al. (2009). Patients are enrolled sequentially, according to the order in which they start the treatment at the clinic or hospital. The allocation problem consists in assigning each new patient to one, and only one, of two alternative treatments (arms). A requisite stated by the trial coordinators is that profiles in the alternative arms remained similar with respect to some relevant patients' factors. In other words, it was expected that the compositional vectors (i.e. relative frequencies of patients in each variable category) remained similar each other as new patients were allocated. The available clinical trial dataset consists of $T = 277$ patients.

Roughly speaking, the factors and respective number of classes considered are: 1. Current patient's *age* (a): three classes; 2. Treatment *history* (h): three classes; 3. OCD symptom *severity* (v): nine classes; 4. *Gender* (g): two classes. A more detailed description on these factors and respective categories may be found in Fossaluza et al. (2009).

After some patients are already in treatment, we denote by n_i^a, n_i^h, n_i^v and n_i^g the quantities of patients already allocated to arm i belonging to each category of factors *age*, *history*, *severity* and *gender*. For example, $n_1^a = [n_{1,1}^a, n_{1,2}^a, n_{1,3}^a]$ denotes the quantity vector of patients in arm 1 belonging to the three age classes.

In order to yield allocations with approximately the same number of patients in each arm, we also consider, besides the previous factors, the sample *size* (z) in each arm. With that purpose we define q_i as the total number of patients allocated to arm i, and the vector of total allocation to arm i and its complement, $n_i^z = [q_i, (q_1 + q_2 - q_i)]$.

The complete profile of arm i, $i = 1,2$, is stored in the concatenated vector $n_i = [n_i^a, n_i^h, n_i^v, n_i^g, n_i^z]$. In order to avoid empty categories in the allocation process, we may add to vector n a *ground-state* or *weak-prior*, see Pereira and Stern (2008), in the form of vector $w = [w^a, w^h, w^v, w^g, w^z]$. For any character ξ in the set $\{a,h,v,g,z\}$, where factor w^ξ has $\kappa(\xi)$ categories, we take $w^\xi = [1/\kappa(\xi), \ldots, 1/\kappa(\xi)]$. From vectors n and w we obtain the *regularized proportions* vector: $p_i = [p_i^a, p_i^h, p_i^v, p_i^g, p_i^z]$, where $p_i^\xi = clu(n_i^\xi + w_i^\xi)$, $\xi \in \{a,h,v,g,z\}$.

We define the heterogeneity measure between arms 1 and 2 by the function:

$$\Delta(p_1, p_2) = \frac{1}{5}\left(D_s(p_1^a, p_2^a) + D_s(p_1^h, p_2^h) + D_s(p_1^v, p_2^v) + D_s(p_1^g, p_2^g) + D_s(p_1^z, p_2^z)\right).$$

$$(14.1)$$

Let us consider a new patient that enrolls the study and must be allocated to one of arms 1 or 2. We denote by x^a, x^h, x^v, x^g and x^z the binary vectors indicating to which categories the new patient belongs. For example, in vector $x^a = [x_1^a, x_2^a, x_3^a]$, $x_k^a = 1$ if and only if the patient belongs to age category k, $k \in [1, \ldots, \kappa(a)]$. Vector x^z is set as $x^z = [1, 0]$. So, the relevant information about the new patient is carried by the vector $x = [x^a, x^h, x^v, x^g, x^z]$.

The arm allocation decision for the new patient is taken as follows.

1. For $j = 1, 2$, consider the allocation of the new patient, x, in arm j, that is, for $i = 1, 2$, make $m_i = n_i + \delta(i, j)x$ and perform the following steps:

   a) For $i = 1, 2$ and $\xi \in \{a, h, v, g, z\}$, compute the regularized proportions

   $$p_i^\xi = clu(m_i^\xi + w_i^\xi) \; ;$$

   b) For $i = 1, 2$, set $p_i = [p_i^a, p_i^h, p_i^v, p_i^g, p_i^z]$;
   c) For $i = 1, 2$, set $b_i = [u_i, 1 - u_i]$, where u_i are independently generated from $Uniform(0, 1)$ distribution;
   d) For $\varepsilon \in [0, 1]$, compute the ε-perturbed distance

   $$d_\varepsilon(j) = (1 - \varepsilon)\Delta(p_1, p_2) + \varepsilon D_s(b_1, b_2).$$

2. Choose the allocation j that minimizes $d_\varepsilon(j)$, assign the new patient to the corresponding arm, and update vector n accordingly.

The perturbation parameter ε introduces a random component in the allocation method. The higher the value of ε, the higher the proportion of randomness in the allocation. For $\varepsilon = 0$, we have a deterministic intentional allocation scheme, as described in Fossaluza et al. (2009), and for $\varepsilon = 1$, we have the pure random allocation method, which consists in assign each patient randomly (with probability 0.5) to one of the two arms.

14.5 Numerical Experiments

We analyse the performance of our haphazard intentional allocation procedure, for $\varepsilon \in \{0, 0.005, 0.01, 0.05, 0.25, 1\}$. The experiments were conducted as follows. We generated $P = 300$ random permutations of the original data, each one representing a possible sequence of patients arriving to the hospital or clinic. For each permutation, we ran the pure random method and the haphazard intentional allocation method $H = 300$ times. For $\varepsilon = 0$, the (deterministic) procedure was executed only once.

Two criteria were used to analyse the performance of the haphazard intentional allocation method: *Optimality* and *Decoupling*. The first criterion, *Optimality*, is based on the distance Δ defined in Equation 14.1 and concerns the difference among the relative frequencies of patients in the several categories for both arms.

The second criterion, *Decoupling*, concerns the absence of a tendency to allocate each pair patients to the same arm. In this work, we use the Yule's coefficient of

association (Q), see Yule (1912), in the following way. After each batch of H runs of haphazard allocations, for each pair of patients, A and B, we build a 2×2 contingency table z where z_{ij} denotes the number of runs patient A was assigned to arm i and patient B was assigned to arm j ($i, j = 1, 2$).

The Yule' coefficient, given by $Q = (z_{11}z_{22} - z_{21}z_{12})/(z_{11}z_{22} + z_{21}z_{12})$, measures the balance among the number of pairs in agreement and disagreement. It ranges in the interval $[-1, 1]$; equals zero when the numbers of agreement and disagreement pairs is equal; and is maximum (-1 or $+1$) in the presence of total negative (complete disagreement) or positive (complete agreement) association.

Figures 14.1 and 14.2 present, respectively, the $5\%, 25\%, 50\%, 75\%, 95\%$ empirical percentiles of Δ and Q. In Figure 14.1 the quantiles for Δ span the H haphazard allocations. In Figure 14.2 the quantiles for Q span the $T(T-1)/2$ pairs of patients, where the Q for each pair is computed over the H haphazard allocations. Each bar height corresponds to the median over the P random permutations, and the vertical line in each bar represent the corresponding $(5\%, 95\%)$ percentiles. Continuous and dashed horizontal lines in Figure 14.1 represent, respectively, the median of distance Δ for the deterministic intentional allocation method, $\varepsilon = 0$, and the $(5\%, 95\%)$ percentiles over P random permutations. Figure 14.2 omits the percentile 50%, since the Yule's coefficient medians were close to zero for all allocation methods.

Figure 14.1 shows a clear difference between the optimality (Δ) achieved by the haphazard intentional allocation method and the pure random method. Notice that, for $\varepsilon \leq 0.01$, even the 5% percentile of Δ for the pure random method is larger than the 95% percentile of the haphazard intentional method. We also notice that, for the same range of ε, the optimality achieved by the haphazard intentional method comes close to the optimality achieved by the deterministic method, only showing moderate degradation in the 95% percentile.

Figure 14.2 shows that, for the lower range of ε, the absolute values of Yule's association coefficient (Q) tend to be high, indicating that the haphazard intentional allocation method, with too small an ε, tends to allocate the same pairs of patients in the same arms, that is, it fails to achieve the desired decoupling property. On the other hand, moderate values of ε attenuate these dependencies, making the haphazard intentional allocation method perform in the decoupling criterion almost as well as the the pure random method. Indeed, for $\varepsilon \geq 0.05$, the distribution of Yule's

Fig. 14.1 The 5%, 25%, 50%, 75% and 95% percentiles for Δ optimality, with $\varepsilon \in \{0, 0.005, 0.01, 0.05, 0.25, 1\}$.
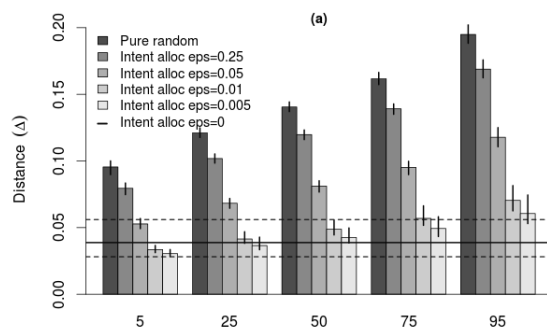
Q is very close to the pure random method – which provides our benchmark for decoupling performance.

It is worth mentioning that, for the haphazard intentional allocation method, the percentiles intervals (represented by vertical lines in each bar) are, in general, similar to the pure random method (except for the 5% and 95% percentiles of *Q* for $\varepsilon \leq 0.01$). This result suggests that the proposed method is highly adaptive.

The results of these numerical experiments suggest that, under an appropriate calibration of the perturbation parameter ε, the haphazard intentional allocation method proposed in this paper has the remarkable property of being able to conciliate the performance on optimality achieved by the deterministic intentional allocation with the performance on decoupling achieved by the pure random allocation method.

14.6 Acknowledgements and Final Remarks

The present article does not include Bayesian models incorporating prior information although, formally, such models only generalize the implicit uninformative priors. Despite our focus in this paper have been on the conceptual and practical discussions concerning randomization and intentional sampling, it is perfectly possible to extend our analysis to more general Bayesian models, a work we intend to do following Bolfarine et al. (1987). The article *Six Approaches to Enumerative Survey Sampling*, by Brewer and Särndal (1983), see also Schreuder et al. (1993, Sec.6.2), has been used for the last 30 years as a classification scheme concerning, among other things, the role of randomization in survey sampling. However, it is not straightforward to fit the allocation method we have just presented in that classification scheme. We hope to explore this theme in following articles.

As duly noted by an anonymous referee, an important topic for further research concerns a comparative analysis of the logical status of all the aforementioned randomization, intentional and mixed-randomization sampling methods according to several possible theoretical and epistemological frameworks. In the standard

Fig. 14.2 The 5%, 25%, 50%, 75% and 95% percentiles for Yule's Q decoupling, with $\varepsilon \in \{0, 0.005, 0.01, 0.05, 0.25, 1\}$.
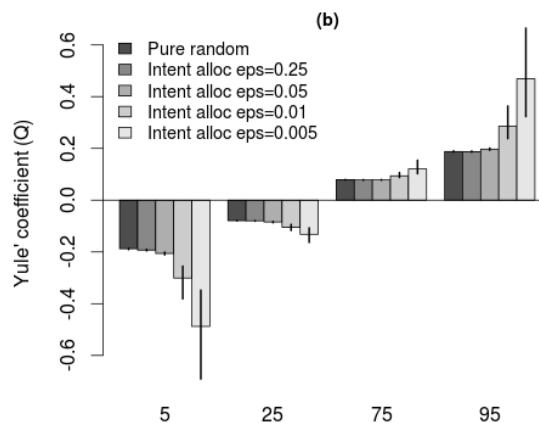
Bayesian decision theoretical framework, randomization methods may not be incompatible with optimal decisions, and may even be able to address some extra-theoretical demands. However, they can never find a direct intra-theoretical justification; see DeGroot (1970, Sec.8.5, p.128-130). Nevertheless, the standard decision theoretical framework can be expanded by taking into account games with multiple adversarial opponents, see Morgenstern (2008), Morgenstern and Neumann (1947) and Stern (2008b, Sec.6.8). Bonassi et al. (2009) explore this expanded decision-theoretical framework, survey the pertinent literature, and suggest interesting approaches for further development. Finally, the function and logical status of randomization methods can be analyzed in the framework of systems theory, see for example Pearl (2000, p.16-20, 340-348) and Stern (2008).

All the aforementioned theoretical frameworks offer alternative ways to deal with concepts related to decoupling, separation, shielding from undue interference, or defensive strategies against players with hostile objectives. Hence, logical analyses conducted in these frameworks should be able to provide guidelines for the coherent development and application of intentional but haphazard sampling methods in the scope of Bayesian statistics.

The anonymous referee also stresses the importance of carefully analyzing the ethical consequences of using alternative sampling methods. In the context of clinical trials, the experimenter must always consider at least two competing (if not conflicting) objectives: On one hand, the primary objective of any clinical trial is the acquisition of of valid or objective knowledge, see Stern (2014). On the other hand, providing appropriate health care for all patients participating in the trial is a second goal that should never be neglected, see Kadane and Sedransk (1980) and Kadane (1996). The use of intentional sampling methods is a technical solution that has the potential of facilitating the reconciliation of such multiple objectives. Moreover, these considerations can be extended to multi-phase and adaptive trials. All these are topics that certainly deserve further research.

References

1. Aitchison, J., Shen, S. M.: Logistic-normal Distributions: Some Properties and Uses. Biometrika, **67**, 261-272 (1980)
2. Aitchison, J.: The Statistical Analysis of Compositional Data. Chapman & Hall (1986)
3. Aitchison, J.: The single principle of compositional data analysis, continuing fallacies, confusions and misunderstandings and some suggested remedies. (2008)
4. Basu, D., Ghosh, J. K. (eds.): Statistical Information and Likelihood, A Collection of Essays by Dr.Debabrata Basu, Lecture Notes in Statistics, **45**, Springer (1988)
5. Bolfarine, H., Pereira, C. A. B., Rodrigues, J.: Robust Linear Prediction in Finite Populations: A Bayesian Perspective. Sankhya, B, **49(1)**, 23-35 (1987)

6. Bonassi, F. V., Nishimura, R., Stern, R. B.: In defense of randomization: A subjectivist Bayesian approach. AIP Conf. Proc., **1193**, 32-39 (2009)
7. Brewer, K. R. W., Särndal, C. E.: Six approaches to enumerative survey sampling. Incomplete Data in Sample Surveys, **3**, 363-368 (1983)
8. Brewer, K. R. W.: Combined Survey Sampling Inference: Weighing of Basu's Elephants. A Hodder Arnold Publication (2002)
9. DeGroot, M. H.: Optimal Statistical Decisions. McGraw-Hill (1970).
10. Egozcue, J. J., Daunis-i-Estadella, J., Pawlowsky-Glahn, V., Hron, K., Filzmoser, P.: Simplicial regression: The normal model. J.of Applied Probability and Statistics, **6**, 87-108 (2001)
11. Fisher, R. A.: The Design of Experiments, 8ed (1966). Oliver and Boyd, London (1935)
12. Fossaluza, V., Diniz, J. B., Pereira, B. B., Miguel, E. C., Pereira, C. A. B.: Sequential Allocation to Balance Prognostic Factors in a Psychiatric Clinical Trial. Clinics, **64**, 511-518 (2009)
13. Kadane, J.B..: Bayesian Methods and Ethics in a Clinical Trial Design. Wiley (1996).
14. Kadane, J.B., Sedransk, N.: Toward a more ethical clinical trial. Trabajos de Estadistica Y de Investigacion Operativa **31**, 1, 329-346 (1980).
15. Lauretto, M. S., Nakano, F., Pereira, C. A. B., Stern, J. M.: Intentional Sampling by Goal Optimization with Decoupling by Stochastic Perturbation. AIP Conf.Proc., **1490**, 189-201 (2012)
16. Lindley, D. V.: The Role of Randomization in Inference. In: Asquith, P., Nickles, T. (eds.) Proceedings of the Biennial Meeting of the Philosophy of Science Association, V.2, 431-446. Univ. of Chicago Press (1982)
17. Lindley, D. V.: Making Decisions. John Wiley, New York (1991)
18. Madow, W. G., Olkin, E., Rubin, D. B.: Incomplete Data in Sample Surveys Vol. 3 (1983)
19. Morgenstern, O.: Game Theory. Dictionary of the History of Ideas Vol.2 p.264-275 (2008)
20. Morgenstern, O., Neumann, J.: The Theory of Games and Economic Behavior. Princeton University Press (1947)
21. Pawlowsky-Glahn, V., Egozcue, J. J.: Geometric approach to statistical analysis on the simplex. Stochastic Environmental Research and Risk Assessment, **15(5)**, 384-398 (2001)
22. Pearl, J.: Causality: Models, Reasoning, and Inference. Cambridge Univ. Press (2000)
23. Peirce, C. S., Jastrow, J.: On small differences of sensation. den Memoirs of the National Academy of Sciences, 3, 75-83 (1884)
24. Pereira, C. A. B., Rodrigues, J.: Robust Linear Prediction in Finite Populations. International Statistical Review, **3**, 293-300 (1983)
25. Pereira, C. A. B., Stern, J. M.: Special Characterizations of Standard Discrete Models. RevStat - Statistical Journal, **6**, 199-230 (2008)
26. Royall, R. M., Pfeffermann, D.: Balanced samples and robust Bayesian inference in finite population sampling. Biometrika, **69(2)**, 401-409 (1982)
27. Schreuder, H. T., Gregoire, T. G., Wood, G. B.: Sampling Methods for Multiresource Forest Inventory. Wiley, New York (1993)
28. Stern, J. M.: Decoupling, Sparsity, Randomization, and Objective Bayesian Inference. Cybernetics and Human Knowing, **15**, 49-68 (2008)
29. Stern, J. M.: Cognitive Constructivism and the Epistemic Significance of Sharp Statistical Hypotheses in Natural Sciences. arXiv:1006.5471 (2008).
30. Stern, J. M.: Cognitive-Constructivism, Quine, Dogmas of Empiricism, and Münchhausens Trilemma. To appear in this volume (2014).
31. Tam, S. M.: Characterization of best model-based predictors in survey sampling. Biometrika, **73(1)**, 232-235 (1986)
32. Valliant, R., Dorfman, A. H., Royall, R. M.: Finite Population Sampling and Inference: A Prediction Approach. Wiley, New York (2000)
33. Whittle, P.: Probability via Expectation. Springer, New York (2000)
34. Yule, G. U.: Notes on the Theory of Association of Attributes in Statistics. Biometrika, **2(2)**, 121-134 (1903)
35. Yule, G. U.: On the Methods of Measuring Association Between Two Attributes. Journal of the Royal Statistical Society, **75(6)**, 579-652 (1912)