Imputation with Item Nonrespondents

Hansheng Wang
Peking University, Beijing, People's Republic of China

Shein-Chung Chow
Millennium Pharmaceuticals, Inc., Cambridge, Massachusetts, U.S.A.

INTRODUCTION

In clinical research, an instrument (or questionnaire) consisting of a number of items (or questions) is usually used to quantitate a subject's behavior and the ability to function in day-to-day activities as perceived by the subject. For example, the EORTC QLQ-C30 questionnaire is commonly used as a reliable instrument to measure the quality of life of patients with cancer. On the other hand, the positive and negative syndrome scale (PANSS) is useful for assessing positive symptoms (e.g., delusions, hallucinatory behavior, and hostility), negative symptoms (e.g., emotional/passive social withdrawal and poverty of thought and spontaneous activity), and general psychopathology (e.g., anxiety and depression) in schizophrenic patients and patients with schizoeffective disorder. For each item, the patient (if self-administered) or the investigator is asked to provide a rating score from 0 (none) to k (severe). A self-administered instrument is often used because it can reflect how a patient feels and may have less bias because patients feel that their answers are more confidential. However, a self-administered instrument is expected to have a larger variability due to the unstable condition of the patient. Based on patient ratings, several statistical methods for assessment of drug effects have been proposed (see, e.g., Refs. [1–3]).

OVERVIEW

In practice, missing values are commonly encountered, which can be classified into two categories, namely, unit nonrespondent and item nonrespondent (see, e.g., Wang[4,5]). A unit nonrespondent refers to a subject who fails to answer all of the items (questions). An item nonrespondent refers to a subject who fails to answer some (not all) of the items. In practice, because unit nonrespondents do not provide any information regarding the treatment effect, they are usually excluded from the analysis under the assumption of missing completely at random. On the other hand, because item nonrespondents do provide partial information, it is not acceptable to exclude these subjects from analysis.

In clinical research, two approaches are commonly suggested for handling item nonrespondents. The first approach is to simply ignore all the item nonrespondents from the analysis. Although this method is statistically valid under the assumption of ''missing completely at random,'' it suffers from decreasing power/efficiency and from excluding too many evaluable patients from the analysis. The second approach is to impute the missing item and then calculate the total score based on the imputed data. As indicated in Wang,[6] imputation methods in clinical research commonly include last observation carry forward (LOCF) or endpoint analysis, mean/median imputation, regression with covariates, etc. To examine the statistical properties of imputation with item nonrespondents, for illustration purpose, we will focus only on mean imputation method.

In the next section, the mean imputation approach will be described under a commonly employed statistical model. Also included in this section are the statistical properties of the estimators derived from the model based on the imputed data. In the section ''Statistical Inference,'' two procedures including the so-called linearization procedure and the jackknife method proposed by Rao and Shao[7] for estimation of the asymptotic variance of the derived estimators are discussed. Also included are statistical tests based on the mean imputation approach. The section ''An Example'' gives an example concerning the study of agitated behavior in schizophrenic patients to illlustrate the use of the mean imputation method. A brief concluding remark is given in the last section.

STATISTICAL MODEL

Single-Imputation Class

In clinical research, the treatment groups are usually considered as imputation classes. The imputation will be carried out within each imputation class. For simplicity, we first consider the case of a single-imputation class. The results can be generalized to multiple-imputation classes as discussed at the end of this section.

Let $x_{ij} \in S_j$, $i = 1,\ldots, n$; $j = 1,\ldots, m$ be the score of the jth item from the ith subject. S_j is the finite sample

Encyclopedia of Biopharmaceutical Statistics
DOI: 10.1081/E-EBS 120007638

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

space for x_{ij}. Note that S_j need not to be the same for different j. The observations from the same subject (e.g., $x_{i1},\ldots, x_{im}$) are usually correlated. However, the vectors $(x_{i1},\ldots,x_{im})'_{i=1,\ldots,n}$ are assumed to be independent and identically distributed (i.i.d.) for different i. The total score of the ith subject is defined to be

$$z_i = \sum_{j=1}^{m} x_{ij}$$

The parameter of interest is $\mu = E(z_i) = \sum_j \mu_j$, where $\mu_j = E(x_{ij})$. Because there are missing scores, denote y_{ij} by missing scores of x_{ij}. That is, $y_{ij} = 0$ means missing and $y_{ij} = 1$ indicates not missing. As a result, $(y_{i1},\ldots,y_{im})'_{i=1,\ldots,n}$ are usually correlated for a fixed i but are i.i.d. random vectors for different i. It is assumed that $E(y_{ij}) = p_j$.

For any given j, the observed sample mean for the jth item is given by

$$x_j^* = \frac{\sum_{i=1}^{n} x_{ij}y_{ij}}{\sum_{i=1}^{n} y_{ij}}$$

As a result, if x_{ij} is missing (i.e., $y_{ij} = 0$), it is replaced by $x_{\cdot j}^*$. Let x_{ij}^* be the value of the jth item of the ith subject after imputation. It can be verified that

$$x_{ij}^* = y_{ij}x_{ij} + (1 - y_{ij})x_j^*$$

Hence, the total score of the ith subject after imputation is given by

$$z_i^* = \sum_{j=1}^{m} x_{ij}^* = \sum_{j=1}^{m} (y_{ij}x_{ij} + (1 - y_{ij})x_j^*)$$

It is common practice to treat the imputed values as the observed values and to obtain estimators designed based on complete data sets for estimation of the parameter of interest. In the situation where there are no missing values, the sample mean of z_i is a natural estimator for μ. Consequently, the sample mean of z^* becomes our estimator for μ, which is given by

$$\hat{\mu}^I = \frac{1}{N}\sum_{i=1}^{n} z_i^* = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}(y_{ij}x_{ij} + (1 - y_{ij})x_j^*)$$

$$= \sum_{j=1}^{m}\left(\frac{\frac{1}{n}\sum_{i=1}^{n} x_{ij}y_{ij}}{\frac{1}{n}\sum_{i=1}^{n} y_{ij}}\right)$$

The following theorem can be obtained.

Theorem 1. Assuming single-imputation class, based on the proposed mean imputation procedure,

$$\sqrt{n}(\hat{\mu}^I - \mu) \rightarrow_{\mathrm{d}} N(0, \sigma^2)$$

where ''$\rightarrow_{\mathrm{d}}$'' denotes ''converge in distribution'' and σ^2 is given by

$$\sigma^2 = \mathrm{var}\left(\sum_{j=1}^{m}\frac{1}{p_j}(x_{ij}y_{ij} - \mu_j y_{ij})\right)$$

Proof:

By Taylor's expansion and Slusky's theorem, it follows that

$$\sqrt{n}(\hat{\mu}^I - \mu) = \sum_{j=1}^{m}\sqrt{n}\left(\frac{\frac{1}{n}\sum_{i=1}^{n} x_{ij}y_{ij}}{\frac{1}{n}\sum_{i=1}^{n} y_{ij}} - \mu j\right)$$

$$= \sum_{j=1}^{m}\sqrt{n}\left(\frac{1}{p_j n}\sum_{i=1}^{n}[(x_{ij}y_{ij} - \mu_j p_j) - \mu_j(y_{ij} - p_j)]\right)$$

$$+ o_p(1) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\sum_{j=1}^{m}\frac{1}{p_j}(x_{ij}y_{ij} - \mu_j y_{ij})\right)$$

$$+ o_p(1) \rightarrow_d N(0, \sigma^2)$$

where σ^2 is defined in Theorem 1.

Multiple-Imputation Classes

Assuming there are H imputation classes, let $x_{h,ij}$ be the score of the jth item from the ith subject in the hth imputation class. The quantities μ_h, $y_{h,ij}$, $\mu_{h,j}$, $z_{h,i}$, $p_{h,j}$ are similarly defined.

For multiple-imputation classes, the parameter of interest is usually a linear combination of μ_h, where $\mu_h = {}_jE(x_{h,ij})$. Let w_h denote the coefficient. The parameter of interest is given by

$$\mu = \sum_{h=1}^{H} w_h\mu_h$$

In clinical research study, μ usually denotes the treatment effect of interest and w_hs are usually the coefficients of contrast satisfying $\sum_h w_h = 0$. For example, in a placebo-controlled clinical trial, the number of imputation classes equals the number of treatment groups, which is two [treatment ($h = 1$) and placebo ($h = 2$)]. w_hs are then chosen to be $w_1 = 1$ and $w_h = -1$. The parameter of interest is $\mu = \mu_1 - \mu_2$, which is the treatment effect.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

For multiple-imputation classes, mean imputation can be carried out within each imputation class. The sample mean of the total score after imputation ($\hat{\mu}_h^I$) is used as an estimator for μ_h. The estimator of μ is given by

$$\hat{\mu}^I = \sum_h w_h \hat{\mu}_h^I \quad (1)$$

Similarly, the following results can be obtained.

Theorem 2. Assuming multiple-imputation classes and $n_h/n \to \rho_h$ for some $\rho_h > 0$, based on the mean imputation procedure described above, we have

$$\sqrt{n}(\hat{\mu}^I - \mu) \to_d N(0, \sigma^2)$$

where σ^2 is given by

$$\sigma^2 = \sum_{h=1}^{H} \frac{w_h^2}{\rho_h} \text{var}\left(\sum_{j=1}^{m} \frac{1}{p_{h,j}} (x_{h,ij} y_{h,ij} - \mu_{h,j} y_{h,ij}) \right)$$

Proof

Similar to the proof of Theorem 1, we have

$$\begin{aligned} \sqrt{n}(\hat{\mu}^I - \mu) &= \sqrt{n} \sum_h w_h (\hat{\mu}_h^I - \mu_h) \\ &= \sum_h \frac{w_h}{\sqrt{n_h/h}} \sqrt{n_h} (\hat{\mu}_h^I - \mu_h) \\ &= \sum_h \frac{w_h}{\sqrt{\rho_h}} \sqrt{n_h} (\hat{\mu}_h^I - \mu_h) + o_p(1) \\ &\to N\left(0, \sum_h \frac{w_h^2}{\rho_h} \sigma_h^2\right) \end{aligned}$$

where σ_h^2 is the asymptotic variance of $\sqrt{n_h}(\hat{\mu}_h^I - \mu_h)$, which can be estimated according to Theorem 1.

STATISTICAL INFERENCE

Variance Estimates

Although the sample variance of $\sum_{j=1}^{m} (1/p_{h,j})(x_{h,ij} y_{h,ij} - \mu_{h,j} y_{h,ij})$ can be used to estimate its true variance, $p_{h,j}$ and $\mu_{h,j}$ are unknown. One way to overcome this problem is to substitute $p_{h,j}$ and $\mu_{h,j}$ with their consistent estimators (e.g., $x_{h,\cdot j}^*$ and $1/n\sum_i y_{h,ij}$, respectively. Since this procedure is based on Taylor's expansion or the linearization of $\hat{\mu}^I$, we denote this estimator by $\hat{\sigma}_{\text{linear}}^2$.

The above procedure is relatively complicated although it requires less computation. As an alternative, the jackknife procedure proposed by Rao and Shao[7] can be used to estimate the asymptotic variance. Referring to the proofs of Theorem 1 and Theorem 2 in the Appendix, we note that $\hat{\mu}^I$ is the smooth function of $1/n\sum x_{h,ij} y_{h,ij}$ and $1/n\sum y_{h,ij}$. Both of them are sample means of i.i.d. random variables. As a result, Rao and Shao's procedure can be applied here to estimate the asymptotic variance. More specifically, for any given h and i, let $\hat{\mu}_{h,-i}^I$ be an estimate of μ. It can be obtained by first constructing a reduced data set by dropping the ith subject in the hth stratum. Then, perform the proposed mean imputation procedure to the reduced data set. Finally, $\hat{\mu}_{h,-i}^I$ is obtained by (1) but based on the reduced data set. Then, the jackknife estimate of σ^2 is given by

$$\hat{\sigma}_{\text{jack}}^2 = \sum_{h=1}^{H} \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} (\hat{\mu}_{h,-i}^I - \hat{\mu}_h^I)^2$$

Testing Hypothesis

In clinical research, the statistical hypotheses of interest can usually be written as

$$H_0: \mu = \mu_0 \text{ and } H_1: \mu \neq \mu_0$$

where μ_0 is a prespecified value. One simple example is $\mu_0 = 0$ for a two-group parallel design comparing treatment effects (recall $H = 2$, $w_1 = 1$, and $w_2 = -1$). The test statistic is given by

$$Z = \frac{\hat{\mu}^I - \mu_0}{\hat{\sigma}}$$

where $\hat{\sigma}$ is an estimate for σ. It could be $\hat{\sigma}_{\text{jack}}$ or $\hat{\sigma}_{\text{linear}}$. Refer to the proofs of Theorem 1 and Theorem 2 given in the Appendix. It can be noted that $\hat{\mu}^I$ is a smooth function of $1/n\sum x_{h,ij} y_{h,ij}$ and $1/n\sum y_{h,ij}$, which are sample means of i.i.d. random variables. As a result, according to central limit theorem and Slusky's theorem, Z is asymptotically normally distributed with mean 0 and variance 1 under the null hypothesis. Therefore, for a given significance level α, the null hypothesis can be rejected if $|Z| > z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the $(1-\alpha/2) \times 100\%$ percentile of a standard normal distribution.

A Simulation Study

A simulation study was conducted to evaluate the finite sample performance of the proposed method. The objective of the simulation study is multifold. It includes the following.

1. Note that the estimators derived by both imputation and complete units are consistent for μ. As a result, it is of interest to compare their relative efficiency in terms of standard deviation.

2. Two methods (linearization and jackknife) are proposed for estimation of the standard deviation of $\hat{\mu}^I$. Although they are asymptotically equivalent to each other, it is of interest to compare their finite sample performance by comparing them with the true standard deviation.
3. The proposed Z-test for the hypothesis of interest is based on large sample theory. As a result, the empirical significance level of the test may not be exactly the same as the nominal level. Therefore, it is of interest to evaluate the finite sample performance of the proposed Z-test based on both $\hat{\sigma}_{\text{linear}}$ and $\hat{\sigma}_{\text{jack}}$.

Consider a two-arm parallel randomized clinical trial comparing a test treatment with a placebo. There are two imputation classes ($H=2$), $w_1 = 1$, $w_h = -1$, and the parameter of interest is the treatment effect $\mu = w_1\mu_1 + w_2\mu_2 = \mu_1 - \mu_2$.

For illustration purposes and for the sake of convenience, we fix $m = 5$ and $n = 20$, which means there are a total of 20 items (questions), and for each question the rating score ranges from 1 to 5. We also assume the score of each item is marginally distributed with probabilities of 0.3, 0.1, 0.3, 0.1, and 0.2. Note that it is not necessary for each item to have the same sample space and the same distribution.

Let n_1 and n_2 denote the sample sizes for group 1 and group 2, respectively. They range from 20 to 40. For a given subject, we first generate a continuous standard uniform random number U_0. For any $j \leq n$, we generate another standard uniform random number R_j independent of U_0. Then, define $U_j = U_0$ if $R_j \leq 0.5$. Otherwise, let U_j be another standard uniform random number

Table 1 Simulation results (m = 5 and n = 20)

					Imputation based					
			Complete units based				Jackknife		Linearization	
n_1	n_2	p_h	$\hat{\mu}^I$	σ	$\hat{\mu}$	σ	$\hat{\sigma}_{\text{jack}}$	p	$\hat{\sigma}_{\text{linear}}$	p
20	20	0.8000	0.7619	14.9606	0.0284	5.1977	5.1614	0.0597	5.1391	0.0609
		0.8500	0.1983	12.9868	0.0709	5.0866	5.1062	0.0557	5.0931	0.0559
		0.9000	−0.0739	9.3859	−0.0263	5.0755	5.0546	0.0546	5.0485	0.0552
		0.9500	0.1001	6.7447	0.0668	5.0279	5.0095	0.0587	5.0062	0.0587
		1.0000	0.0887	5.0813	0.0887	5.0813	4.9653	0.0635	4.9653	0.0635
20	30	0.8000	0.5605	14.6527	−0.0474	4.7282	4.7090	0.0567	5.1506	0.0367
		0.8500	0.1758	11.9008	0.0561	4.6332	4.6568	0.0564	5.1000	0.0371
		0.9000	0.0922	8.3996	0.0432	4.6271	4.6099	0.0586	5.0513	0.0373
		0.9500	0.0027	6.0743	0.0302	4.5251	4.5757	0.0533	5.0140	0.0343
		1.0000	−0.0289	4.5519	−0.0289	4.5519	4.5375	0.0569	4.9724	0.0360
20	40	0.8000	0.3131	14.0286	0.0088	4.5208	4.4748	0.0600	5.1697	0.0305
		0.8500	0.1110	11.1785	−0.0310	4.4393	4.4214	0.0571	5.1095	0.0300
		0.9000	0.0685	8.1120	0.0021	4.4150	4.3736	0.0573	5.0576	0.0310
		0.9500	−0.0792	5.7790	−0.0777	4.3477	4.3306	0.0606	5.0081	0.0306
		1.0000	0.0778	4.3305	0.0778	4.3305	4.3007	0.0598	4.9721	0.0291
30	30	0.8000	0.2106	13.9243	−0.0540	4.2101	4.2314	0.0502	4.2209	0.0508
		0.8500	0.1861	10.6147	−0.0280	4.1717	4.1854	0.0526	4.1797	0.0527
		0.9000	−0.0367	7.4941	0.0073	4.1271	4.1335	0.0525	4.1309	0.0530
		0.9500	0.0389	5.4454	−0.0039	4.1258	4.0919	0.0553	4.0911	0.0553
		1.0000	0.0283	4.0690	0.0284	4.0690	4.0602	0.0553	4.0602	0.0553
30	40	0.8000	0.4015	13.3119	−0.0112	3.9533	3.9597	0.0536	4.2282	0.0396
		0.8500	−0.0337	9.9220	−0.0091	3.9083	3.9156	0.0548	4.1834	0.0408
		0.9000	−0.0291	7.0364	−0.0068	3.8735	3.8665	0.0541	4.1342	0.0408
		0.9500	0.0070	5.1251	−0.0164	3.8366	3.8333	0.0549	4.0983	0.0391
		1.0000	0.0486	3.7901	0.0486	3.7901	3.7975	0.0548	4.0601	0.0399
40	40	0.8000	0.0155	12.5925	−0.0331	3.6945	3.6690	0.0577	3.6627	0.0586
		0.8500	0.0626	9.0309	−0.0264	3.6302	3.6248	0.0525	3.6216	0.0528
		0.9000	0.0285	6.3868	0.0412	3.5928	3.5842	0.0534	3.5823	0.0535
		0.9500	0.0573	4.6978	0.0412	3.5562	3.5499	0.0556	3.5493	0.0552
		1.0000	0.0413	3.5004	0.0414	3.5004	3.5129	0.0535	3.5129	0.0535

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Table 2 Data listing

No	TRT	Complete	Itemized scores (1–20)																			
1	1	1	1	1	3	1	2	1	1	1	2	1	1	5	1	1	5	1	1	1	1	5
2	1	0	4	1	3	1	1	3	1	1	1	4	1	1	*	1	1	1	1	1	3	3
3	1	0	5	5	1	1	1	1	1	1	1	1	2	5	4	*	5	1	1	1	2	1
4	1	1	3	3	3	3	2	3	3	3	3	4	2	3	3	2	3	3	3	4	3	3
5	1	0	*	1	1	2	1	3	1	1	1	5	3	3	*	1	1	1	1	5	1	1
6	1	0	*	3	3	*	*	3	3	*	3	*	*	*	3	*	3	3	*	*	*	*
7	1	1	1	1	1	3	1	3	3	1	2	1	1	1	5	1	1	1	3	1	1	1
8	1	1	3	3	5	3	3	1	3	1	3	3	4	1	3	3	3	1	5	3	3	3
9	1	0	1	3	2	2	2	*	3	2	2	2	2	2	*	2	2	*	2	3	2	2
10	1	0	2	3	3	2	2	2	2	1	3	2	2	2	3	2	*	2	1	3	2	2
11	1	0	5	1	2	2	*	2	4	2	2	5	3	3	*	2	2	3	2	5	5	*
12	1	1	1	4	2	5	3	2	1	1	1	1	3	1	3	1	1	1	3	5	3	5
13	1	1	1	1	5	1	1	3	1	1	1	1	1	1	3	3	1	3	3	1	1	1
14	1	1	3	2	3	3	3	5	3	3	3	1	3	2	3	3	3	3	5	3	3	3
15	1	1	4	4	4	2	4	4	4	5	4	1	1	1	1	3	4	4	3	1	4	3
16	1	0	1	4	1	1	2	1	1	1	1	*	1	4	3	3	4	1	2	1	1	5
17	1	0	5	2	1	5	5	3	5	5	5	5	4	1	*	5	5	5	5	5	5	5
18	1	0	3	5	3	3	5	3	3	3	3	*	3	1	*	1	3	1	3	4	*	3
19	1	0	5	3	1	1	1	3	1	5	1	1	3	4	1	1	5	2	*	1	1	1
20	1	0	1	1	1	1	1	2	3	3	1	4	1	*	1	1	1	3	1	3	1	1
21	2	1	3	5	3	3	3	1	1	3	3	3	3	3	3	3	3	3	1	4	3	3
22	2	0	3	2	2	2	1	1	5	1	*	*	1	5	1	2	1	3	1	*	1	3
23	2	0	1	3	5	*	4	3	3	*	3	3	3	3	5	5	3	3	4	4	3	3
24	2	1	3	4	3	3	4	2	3	1	4	4	2	3	3	3	3	5	1	3	2	4
25	2	1	2	4	2	2	2	4	2	2	3	2	4	3	1	1	2	2	2	2	5	2
26	2	1	4	2	4	5	5	2	2	2	4	4	1	4	1	4	4	4	4	3	4	4
27	2	0	4	2	2	2	*	2	2	2	4	2	2	3	1	5	1	2	1	5	3	5
28	2	0	4	4	2	*	4	*	3	4	4	5	4	4	4	2	4	4	4	3	2	4
29	2	1	3	3	3	3	3	3	3	3	1	3	3	3	2	3	2	3	3	5	1	1
30	2	0	2	1	4	1	3	3	3	4	3	*	3	*	2	1	3	3	3	3	4	3
31	2	0	5	5	4	*	1	5	4	5	3	3	1	5	5	5	*	5	5	5	5	5
32	2	1	5	2	2	2	3	2	4	2	2	2	2	2	2	2	2	3	4	2	4	4
33	2	1	1	1	1	1	1	1	1	2	3	5	1	5	2	5	4	3	2	1	1	3
34	2	0	1	3	2	1	1	5	1	*	1	1	1	1	1	3	3	4	1	*	1	1
35	2	0	4	5	3	4	4	1	4	4	1	4	*	2	4	4	1	2	4	4	4	4
36	2	0	4	5	5	5	5	5	5	5	5	5	5	5	5	5	1	5	5	5	5	*
37	2	0	*	1	1	5	1	1	1	1	2	1	1	*	3	1	2	1	2	5	1	3
38	2	0	5	5	5	3	3	5	5	5	5	5	2	5	*	2	5	5	5	5	4	5
39	2	1	1	4	2	2	1	1	3	1	1	1	1	1	1	1	3	2	1	1	5	1
40	2	0	3	3	*	3	3	3	3	1	3	3	3	3	5	4	3	3	3	5	3	3

*indicates missing value.
No = subject number.
TRT = treatment group.
Complete denotes whether the unit is complete (1, complete and 0, not complete).

independent of both U_0 and R_j. Compare U_j with the cumulative distribution function of the score of the item to determine the value of the jth item. More specifically,

$$x_{h,ij} \doteq \begin{cases} 1 & \text{if } U_j \le 0.3 \\ 2 & \text{if } 0.3 < U_j \le 0.4 \\ 3 & \text{if } 0.4 < U_j \le 0.7 \\ 4 & \text{if } 0.7 < U_j \le 0.8 \\ 5 & \text{if } 0.8 < U_j \end{cases}$$

where ''ė'' means ''defines to be.'' There are two important properties of $x_{h,ij}$ generated by this procedure. First, U_j is still a marginally standard uniform random variable. As a result, $x_{h,ij}$ is distributed with probabilities of 0.3, 0.1, 0.3, 0.1, and 0.2. Second, all the U_j from the same subject have probability 0.5 to share the same U_0. Consequently, they are correlated. The purpose of this is to reflect the fact that the observations from the same subject should be correlated in some way.

In order to determine if the score of an item is missing or not, we generate a standard uniform random variable U_0 for each subject. Generate another standard uniform random variable R_j for each $j \leq n$. Then, define $U_j = U_0$ if $R_j \leq 0.5$. Otherwise, let U_j be another standard uniform random number independent of both U_0 and R_j. Then, the score of the jth item is set to be missing if $U_j > p_h$ otherwise it is observed, where p_h is a prespecified value. As a result, for a given item, it has a marginal probability p_h to be observed. In the simulation, we consider p_h to be 0.80, 0.85, 0.90, 0.95, and 1.00. As a result, the missing of all items from the same subject are correlated with each other but independent of the value of the item.

For a given set of parameters (n_1, n_2, p_h), 10,000 simulation runs were carried out. For each simulation, a data set mimics a two-arm parallel trial, and is generated according to the specified parameters. For the generated data set, μ is estimated by both imputation method ($\hat{\mu}^I$) and complete units only ($\hat{\mu}$). Their sample means are reported. Their sample standard deviations are used as estimates for their true standard deviation. For the same data set, both jackknife and linearization methods are used to estimate the standard deviation of $\hat{\mu}^I$. Their sample means are used to estimate the true values of those two estimates. For the same data set, the Z-tests are performed by using both $\hat{\sigma}_{\text{jack}}$ and $\hat{\sigma}_{\text{linear}}$. The empirical p value is estimated by the proportion of the data sets rejecting the null hypothesis. Table 1 summarizes the simulation results.

Table 1 findings are summarized below.

1. The efficiency (in terms of standard deviation) of the point estimate based on imputed data is improved as compared with the method of using the completers only.
2. Both jackknife and linearization provide reasonably good estimates for the standard deviation of $\hat{\mu}^I$. The jackknife estimator seems to perform slightly better than the estimate obtained by linearization.
3. The empirical level of both Z-tests are reasonably close to the nominal level 5%. However, the Z-test using $\hat{\sigma}_{\text{jack}}$ is slightly better than the Z-test using $\hat{\sigma}_{\text{linear}}$.

AN EXAMPLE

To illustrate the application of the proposed mean imputation with item nonrespondents, consider an example concerning the study of treatment effect on controlling agitated behavior of a drug product in schizophrenic patients and patients with schizoeffective disorders. A randomized, parallel-group clinical trial comparing two treatments was conducted in patients with mild to moderate schizophrenia or schizoeffective disorder to help control their agitated behavior. Twenty patients were randomized and completed the study of 12 weeks. An instrument consisting of 20 items was used to quantitate the patient's agitated behavior. The scale of each item ranges from 1 (none) to 5 (severe). The results at endpoint are given in Table 2.

Although, the trial recruited 20 patients for each treatment group, there were only eight unit respondents per group. The analysis based on complete units estimates the treatment effect $\hat{\mu} = -3.875$ with standard deviation $\hat{\sigma} = 5.662$. The corresponding Z statistic is given by $Z = -3.875/5.662 = -0.684$. The p value is given by 0.494, which is not significant (< 0.05).

On the other hand, the analysis based on the proposed mean imputation approach estimates the treatment effect $\hat{\mu}^I = -10.767$ with standard deviation $\hat{\sigma}_{\text{jack}} = 4.862$ and $\hat{\sigma}_{\text{linear}} = 4.859$. The corresponding Z statistics are given by 2.214 and 2.216, respectively, with p values given by 0.027 and 0.027, respectively, which are significant.

In this example, the analysis based on unit respondents only may lose too much power/efficiency and consequently may be misleading. The mean imputation approach, on the other hand, is more efficient than the analysis using complete units.

CONCLUSION

In summary, we propose a mean imputation approach to handle item nonrespondents when the total score is the parameter of interest. The derived estimator is consistent and asymptotically normal. The jackknife method is useful for estimation of the standard deviation of the proposed estimator. The proposed method is not only statistically valid but also more efficient as compared with the method based on complete units only.

REFERENCES

1. Tandon, P.K. Application of global statistics in analysing quality of life. Stat. Med. **1990**, *9*, 819–827.
2. Olschewski, M.; Schumacher, M. Statistical analysis of quality of life data in cancer clinical trials. Stat. Med. **1990**, *9*, 749–763.
3. Chow, S.C.; Ki, F. On statistical characteristics of quality of life assessment. J. Biopharm. Stat. **1994**, *4*, 1–17.
4. Wang, H. Two-way contingency tables with conditionally and marginally imputed data; Ph.D. Thesis; University of Wisconsin–Madison, 2001.
5. Wang, H. *Imputation in Clinical Research*; Encyclopedia of Biopharmaceutical Statistics, Chow, S.C., Ed.; Marcel Dekker, Inc.: New York, 2003.
6. Wang, H.; Shao. Sample correlation coefficients based on survey data under regression imputation. J. Am. Stat. Assoc. **2002**, *97*, 544–552.
7. Rao, J.N.K.; Shao, J. Jackknife variance estimation with survey data under hot deck imputation. Biometrika **1992**, *79*, 811–822.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016