ON AN OPTIMAL SAMPLE SURVEY DESIGN FOR ESTIMATION OF SOLID WASTES IN SAUDI ARABIA

Munir Ahmad, T.A. Bader and H. Khararjian
University of Petroleum & Minerals
Dhahran, Saudi Arabia

Abstract

The purpose of this investigation is to estimate total solid wastes of various kinds generated by domestic and small commercial establishments in Saudi Arabia so as to plan for proper disposal and reutilization of some wastes. Three cities of Dammam, Dhahran and Al-Khobar are selected for the survey. Each city is divided into areas called districts, each district is then divided into blocks, block into structures and structure into houses. The districts, blocks, structures and houses are of unequal sizes. Various designs that include clusters, sub-clusters and stratifications are compared in one replicate factorial experiment for selection of a survey design that may be optimal in the sense of efficiency, minimum manpower and minimum time.

1. Introduction

Solid wastes are defined to be the discarded and unwanted solids from human, animal, industrial, agricultural and mineral wastes. The types of solid wastes could be in varying degrees and composition.

Solid waste problems are progressive in nature; it grows with the growth of an area; growth in the sense of industrial development and increase in human population. It creates tremendous environmental problems and health hazards. Various agencies are interested in reducing its impact by its (i) proper disposal, (ii) management, and (iii) reprocess. A vast amount of literature exists in many developed and developing countries and numerous surveys had been conducted to determine the magnitude of the problem and their possible solutions. A comprehensive study of the solid waste has been made by Tebobanoglous, et al. (1977). A national survey was conducted by US Department of Health (1968 a,b) which shows that the average amount of solid waste actually collected in 1968 was about 5.3 pounds per person per day. On the basis of 4% annual increase, the projected 1985 solid waste generation will be about 9 pounds per person per day. It is interesting to note that whereas the US population has increased 30% since 1950, the waste load has increased 60% and rose to another 50% by 1980.

There has been very little work done on the quantitative and qualitative estimation of the solid wastes in the Kingdom of Saudi Arabia. Khan (1979) made an isolated study in the city of Jeddah. Khan and Bayunus (1980,1981) made a very restrictive study in the city of Jeddah without using sampling technology and as such the results do not have wider scope of applicability to other cities.

Many cities in Saudi Arabia, in very recent years, have doubled and tripled in size. The high standard of living is also evident in many aspects of a typical family by looking at the annual imports of hundres of luxurious items. In order to determine, as accurately as possible, the total solid wastes and factors contributing to the generation of solid waste, it is essential to develop a proper sample design. It was decided to conduct a laboratory-controlled-type experiment in order to achieve the desired degree of success.

The aim of the experiment is to develop a proper design to provide reliable estimates of (i) total solid waste, (ii) proportions of various kinds of solid wastes, and (iii) standard errors of the estimates. The sample survey design should be easy to operate and should give a reasonably efficient estimates with a minimum number of personnels. The experiment may provide a good design for similar surveys in other city areas of the Kingdom.

2. Method

No basic statistical information is available for a sample survey of solid waste in the Kingdom. The first ever population census was conducted in 1974 which could not provide us any suitable frame for the survey. Municipal corporation called Baladia was contacted and maps of the areas of Eastern Province were collected. A pilot study was made in the city of Al-Khobar to determine the extent of problems, the team of investigators would face and to know if (i) blocks are reasonable units for estimation of properties of various kinds of solid wastes; (ii) blocks are efficient as primary units in the statistical sense; (iii) cost, labor force and overall organizational control is better in case of blocks as compared to other units like housing, structures or housing units. The pilot study was made on a simple random sample of 10 blocks (0.5%) out of 2183 city blocks.

The pilot study showed that blocks as shown on the maps which were of unequal sizes in areas vary from 57×14 to 60×255 sq.ft. The number of structures per block is 6.5 ranging from 2 to 12 structures in a block. In the absence of a proper frame, it was decided to conduct an experiment into two phases. Phase I consists of block survey and Phase II consists of structure-housing units survey.

2.1. Block Survey

The city of Al-Khobar was divided into 5 administrative areas. Table I shows the number of blocks (as counted from the map of the city) in each area.

North and Thugba comprise one-third of the city. West is the smallest area. Since pilot study showed a high variation in blocks, it was desirable to test 1/2% and more number of blocks. The four levels of the blocks namely 0.5%, 1%,

Table I: Number of Blocks in Area of the city of Al-Khobar

Districts		No. of Blocks
North		624
South		224
West		260
Agrabia		426
Thugba		649
	Total:	2183

1.5% and 2% are therefore considered. A simple random sample was then employed to select sample blocks from Al-Khobar. This selection resulted into the following no. of blocks in each area. The post stratification of the sample blocks by areas are given in Table II.

Table II: Post-Stratification of Sample Blocks by Areas of Al - Khobar

Area	.5%	1%	1.5%	2%
North	7	8	7	11
South	2	3	2	3
West	2	3	2	2
Agrabia	-	3	6	11
Thugba	-	5	16	17
Total:	11	22	33	44

Post-stratification is necessary for collection. The team will go to an area and collect waste from the assigned blocks.

Once the blocks are selected, the sample collection team began the data acquisition. The garbage is collected in the city twice a day; in the morning the collection started at 4:00 a.m. and in the afternoon at 1:30 p.m. The data collection includes taking the depth of garbage inside the containers to determine the volume (the size and type of each container was known). Card-boxes were counted and categorized separately. Other unusual items, such as mattresses, etc. are noted. Once the depth is measured, approximately 25% of garbage is sampled in plastic bags for laboratory analysis. This includes segregation into various categories, determining the bulk weight and volume. The density of the solid waste (without card-boxes) is then determined. After calculating the density, the weight of the garbage is determined by multiplying the density by the volume measured in the field. The bulk density is then calculated by summing the weights of all components divided by the total bulk volume.

2.2. Block Survey Analysis

Table III shows the summary results of the experiments on blocks only.

The most efficient sample seems to have a 1.5% size which has the smallest standard deviation and standard errors of means of weight and volume of solid waste. The estimated densities of solid waste are on the upper range of the values reported in the literature which is 60-120 kg/m^3. Only 1.5% sample size design renders a density of 122.9 kg/m^3 whereas the densities for other sample sizes lie outside the established range. The estimates of density from the sample density with .5% and 2% sizes are well above the upper limits. In order to study this marked variation in the densities, Table IV shows means, standard deviation by post-stratified areas and sample sizes.

It is observed from Table IV that the density estimated at 2% in Agrabia is 299.8 kg/m^3 with a standard deviation of 113.4. This value is quite high and when checking with the field group notes, it is found that on the sample collection date, one block in this group contains two 5m^3 containers full to the top with food waste (mostly cooked rice). This is an unusual event and is one of the reasons for the high density.

Table III: Means, standard deviations and estimated totals of solid waste weight, volume and density in Al-khobar by sample sizes (Values are in kg.)

Sample Sizes		Mean $\bar{x}$	Standard Deviations σ (of means)	Estimated Totals = $\hat{Y}(10)^5$
	Weight	402.56	154.01	4.39
0.5%	Volume	8.24	2.46 (1.56)	0.90
	Density	400.54	36.68 (-)	-
	Weight	315.64	89.77 (47.98)	3.45
1.0%	Volume	5.32	0.98 (0.52)	0.58
	Density	126.07	65.25 ( )	-
	Weight	258.48	42.59 (17.76)	2.82
1.5%	Volume	3.96	0.66 (0.28)	0.43
	Density	122.90	57.04 (-)	-
	Weight	224.84	60.06 (20.02)	2.45
2.0%	Volume	3.56	2.17 (0.72)	0.39
	Density	140.82	115.25 (-)	-

Table IV: Means and standard deviations of densities by post-stratified areas.

Sample Size %	Mean/Standard Deviation	North	South	West	Agrabia	Thugba	Total
0.5%	$\bar{x}$	97.6	82.4	107.6	-	-	400.54
	s	35.9	19.8	83.93	-	-	36.68
1%	$\bar{x}$	86.8	82.6	65.7	187.5	163.2	126.07
	s	11.7	33.8	-	32.5	76.4	65.25
1.5%	$\bar{x}$	69.6	118.1	85.4	197.8	129.0	122.90
	s	16.1	77.8	31.3	50.9	49.6	57.04
2%	$\bar{x}$	60.8	83.4	128.5	299.8	119.9	140.82
	s	34.5	63.1	21.9	113.4	79.5	115.25
	$\bar{x}$	77.5	89.9	106.9	255.6	129.5	-
	s	32.1	45.6	38.7	102.9	69.4	-

In Table IV, one also observes that as areas become more commercialized, the density decreases. An example is the comparison between North (commercial) and Agrabia (mostly residential) areas where the overall densities are 77.5 and 255.6 kg/m^3 respectively. The other three areas follow the same pattern of higher densities for residential areas. The solid waste generation rate based on 1½% sample with an estimated population* of Al-Khobar of 1.03×10^5 is 2.489 kg/person/day. The values reported in the literature are in the range of 0.91 - 2.268 kg/m^3 with 1.588 kg/m^3 most commonly mentioned. Our estimates are at one of the extremes of this range. It should be noted that this estimate is very crude because (i) there was no daily replicate of solid waste, (ii) population estimate is very crude (and could not be checked with the official figures of Central Department of Statistics, Kingdom of Saudi Arabia). However, it shows that a sample design with 1.5% sample size does provide a well balanced sample design.

3. Housing Units - Structure Survey

After completing the block survey, an experiment on a sample design for housing unit structure is conducted. It was decided not to conduct the experiment in a larger area like the whole city of Al-Khobar because (i) the amount of work involved was monumental, (ii) area cannot be made laboratory-controllable, (iii) manpower was limited, (iv) period of waste collection during each day was limited, (v) of large variation in the waste collection of various areas/blocks, and (vi) frame was not available. It was decided to conduct the experiment in a smaller area of Thughba where there were 649 blocks with estimated totals of 4550 structures and 12,8000 housing units. If a list of housing units had been available, a simple random sample of housing units would have been selected for the survey. The list was not available, so various levels of structures and housing units need to be determined for the final sample design. A pilot study made at an early stage showed that a three-stage sample survey design was proper. A three-stage-2x2 factorial experiment is conducted. In this experiment, the selection procedure is as follows.

In the first stage, 1.5% blocks are independently selected from Thughba; from each selected block, 2 levels of structures (i.e., 2 structures and 4 structures) at the second stage are selected and from each selected structure, 2 levels of housing units (i.e., 2 housing units and 4 housing units) are sampled at the third stage. The notations followed are given below:

- (1) : Structures at lower level and housing units at lower level, i.e., 2 structures and 2 housing units.
- h : Structures at lower level and housing units at upper level, i.e, 2 structures and 4 housing units.
- s : Structures at upper level and housing unit at lower level (i.e., 4 structures and 2 housing units).
- hs : Structures at upper level and housing units at upper level (i.e., 4 structures and 4 housing units.

3.1 A General Method:

A general method of estimation for a multi-stage-2^n factorial experiment design is developed. In this section, a design is proposed to test two factors each at two levels. In this section, a 3-stage-2^3 factorial experiment design is proposed along with estimation procedure. Table V shows an example of 3-stage 2^3 factorial experiment in one replicate.

Table V: Three-Stage 2 x 2 x 2 Factorial Experiment with one Replicate

Stages	Combinations							
	(1)	h	s	hs	b	bh	bs	bhs
First Stage (# Blocks)	2	2	2	2	4	4	4	4
Second Stage (# Structures)	2	2	4	4	2	2	4	4
Third Stage (# Housing Units)	2	4	2	4	2	4	2	4
Total of Housing Units in one Replicate	8	16	16	32	16	32	32	64

From each selected block, 2 levels of structures (2 structures and 4 structures) are randomly selected. Similarly, 2 levels of housing units are randomly selected from relevant structures. The scheme is shown in Table V. Each column in Table V is a three-stage scheme, i.e., each (treatment)combination is an outcome of a 3-stage random sampling scheme.

We further define some notations:

N = the number of blocks in the area.

n = the number of sample blocks (first-stage units).

M_i = the number of structures in the i-th block.

m_i = the number of sample structures from i-th block (second-stage units).

Q_{ij} = the number of houses in j-th structure of i-th block.

q_{ij} = the number of sample houses from j-th structure and i-th block.

y_{ijk} = the information from k-th house of j-th structure and of i-th block.

The above notations relate to any one combination. N, M_i and Q_{ij} are known and are fixed numbers.

The number of blocks in the sample is either 2 or 4, i.e., $n = 2$ or $n = 4$. The number of structures in the example is either 2 or 4. Similarly, the two levels of housing units is 2 or 4. The notation m denotes 2 structures (lower level of structures) and m_1 denotes 4 structures (upper level of structures). Similarly q_{ij} is defined for levels at the third-stage.

3.2 Estimation of Total and its Variance

Let $y_{ijk}(t)$ denote an observation from the k-th housing unit in j-th structure and the i-th block at t-th combination (upper level or lower level of block, structure or housing unit). t stands for the treatment combination, i.e., $t = 1, 2, \ldots, 8$ where

(1) = 1: (1) is designed by the number 1 and stands for lower levels of the 3 factors.

h = 2: denotes lower level of block and structures and upper level of housing units.

Similarly, s = 3, hs = 4, b = 5, bh = 6, bs = 7, bhs = 8.

Observations as recorded under each treatment combination will be denoted by the following notations:

$(1): y_{ijk(1)},\ i = 1, 2, \ldots, n_{(1)},\ j = 1, 2, \ldots, m_{i(1)},\ k = 1, 2, \ldots, q_{ij(1)}$

$h: y_{ijk(2)},\ i = 1, 2, \ldots, n_{(2)},\ j = 1, 2, \ldots, m_{i(2)},\ k = 1, 2, \ldots, q_{ij(2)}$

and so on,

where number in the parentheses stands for a particular combination. The numbers of observations in rows are schematically different from each other. The totals in rows are not comparable and as such averages may be computed and analysis made. However, analysis of variance can be performed as if 2^3 factorial experiment had been designed in a completely randomized design.

The following formulae are obtained for estimating totals and their variances for each combination (scheme of 3-stage sampling):

$$y'_{(t)} = \frac{N}{n_{(t)}} \sum_{i=1}^{n} \frac{M_i}{m_{i(t)}} \sum_{j=1}^{m_{i(t)}} \frac{Q_{ij}}{q_{ij(t)}} \sum_{k=1}^{q_{ij(t)}} y_{ijk(t)},\quad t = 1, 2, \ldots, 8$$

$$Var(y'_{(t)}) = \frac{N(N-n_{(t)})}{n_{(t)}} S^2_{1y_{(t)}} + \frac{N}{n_{(t)}} \sum_{i=1}^{N} \frac{M_i(M_i-m_{i(t)})}{m_{i(t)}} S^2_{2iy_{(t)}} + \frac{N}{n_{(t)}} \sum_{i=1}^{N} \frac{M_i}{m_{i(t)}} \frac{Q_{ij}(Q_{ij}-q_{ij(t)})}{q_{ij(t)}} S^2_{2ijy_{(t)}},\quad t = 1, 2, \ldots, 8$$

where

$$S^2_{1y_{(t)}} = \sum_{i=1}^{N} (y_{i(t)} - \bar{Y}_{(t)})^2/(N-1),$$

$$S^2_{2iy_{(t)}} = \sum_{j=1}^{M_i} (Y_{ij(t)} - \bar{Y}_{i(t)})^2/(M_i - 1),\ \text{and}$$

$$S^2_{2ijy_{(t)}} = \sum_{k=1}^{Q_{ij}} (Y_{ijk(t)} - \bar{Y}_{ij(t)})^2/(Q_{ij} - 1)$$

$y'_{(t)}$ is an unbiased estimate of the population total $Y_{(t)}$. The unbiased variance estimator of $Var(y'_{(t)})$ is given by

$$var(y'_{(t)}) = \frac{N(N-n_{(t)})}{n_t} s^2_{1y(t)} + \frac{N}{n_{(t)}} \sum_{i=1}^{n_{(t)}} \frac{M_i(M_i-m_{i(t)})}{m_{i(t)}} s^2_{2iy_{(t)}} + \frac{N}{n_{(t)}} \sum_{i=1}^{n_{(t)}} \sum_{j=1}^{m_{i(t)}} \frac{M_i}{m_{i(t)}} \frac{Q_{ij}(Q_{ij}-q_{ij(t)})}{q_{ij(t)}} s^2_{2ijy_{(t)}}$$

where $s^2_{1y} = \sum_{i=1}^{n_{(t)}} (y_{i(t)} - \bar{Y}_{(t)})^2/(n_{(t)}-1)$

$$s^2_{2iy} = \sum_{k=1}^{m_{i(t)}} (y_{ij(t)} - \bar{Y}_{i(t)})^2/(m_{i(t)}-1),\ \text{and}$$

$$s^2_{2ijy} = \sum_{k=1}^{q_{ij(t)}} (y_{ijk(t)} - \bar{Y}_{ij(t)})^2/(q_{ij(t)}-1)$$

These formulae become simpler as all the units in a stage are equal.

If a proportion $p_{(t)}$ of solid waste is collected from each housing unit of t-th combination, the estimated totals and variances become $p^{-1}_{(t)}(y'_{(t)})$ and $p^{-2}_{(t)}\ Var(y'_{(t)})$ respectively.

A comparison of the variances may lead to the selection of a particular 3-stage sample design.

A further analysis using the analysis of variance technique is made to test if there is any difference between using different levels of the three factors. Table V shows that these numbers of observations in combinations are schematically different. There are 8 observations in (1), 16 in h, s and b, etc. The analysis can be made on two schemes (we employ (a) here):

a) 2^3 factorial experiment in a completely randomized design, and
b) 2^3 factorial experiment with one replicate with average considered as one observation per cell.

Here we shall assume that a random sample of size 8 is assigned to the combination (1), a random sample of size 16 is assigned to treatment h, etc. The size of the random sample is not assumed to be constant for all treatments. The notations that will be used for computation of means and variances are outlined in Table VI.

The estimates of variation due to error and combination effects are obtained by an additive model $y_{ijk(t)} = \mu + \tau_{(i)} + \varepsilon_{ijk(t)}$ with usual conditions.

Least-squares estimates of the parameters μ and τ_i are obtained by minimizing $\Sigma_t\ \varepsilon^2_{(t)}$ under the constraints $\Sigma_t\ n_{(t)}\ \hat{\tau}_{(t)} = 0$ and analysis of variance is made.

The combination degrees of freedom can be split up into single degrees of freedom for testing various treatment contrasts(comparisons).

The summary results based on a 3-stage 2 x 2 factorial experiment are given from Tables VII to X for weight and volume.

Table VI: Computation of Mean and Variance for the 3-Stage 2x2x2 Factorial Experiment Design

	1	2	3	4	5	6	7	8	Totals
Number of Observations	$n_{(1)}$	$n_{(2)}$	...	...	...	...	...	$n_{(8)}$	$M = \Sigma_{t=1} n_{(t)}$
Sum of Observations	$T_{(1)}$	$T_{(2)}$	...	...	...	...	...	$T_{(8)}$	$G = T_{(i)}$
Mean of Observations	$\bar{T}_{(1)}$	$\bar{T}_{(2)}$	...	...	...	...	...	$\bar{T}_{(8)}$	$\bar{G} = \frac{G}{n_{(t)}}$
Sum of Squares of Observations	$\Sigma Y_{(1)}$	$\Sigma Y^2_{(t)}$	...	...	...	...	...	$\Sigma Y^2_{(8)}$	$\Sigma\Sigma Y^2_{(t)}$
T_i/n_i	$T^2_{(1)}/n_{(2)}$	$T^2_{(2)}/n_{(2)}$	...	...	...	...	...	$T_{(8)}/n_{(8)}$	
Within-class Variation	$SS^2_{(1)} = \Sigma Y^2_{(1)} - \frac{T^2_{(1)}}{8}$		...	...	...	...	...	$SS_{(8)}$	
Within-class Variance	$s^2_{(1)} = \frac{SS^2_{(1)}}{n_{(t)}-1}$,	$s^2_{(2)}$						$S^2_{(8)} = \frac{SS^2_{(8)}}{n_{(t)}-1}$	

Table VII: Summary Results of Weight of Solid Waste in a 2 x 2 Factorial Experiment

	Treatments (Structure-Housing Unit)				Combination Total
	(1)	H	S	HS	
n	6	12	22	23	63
$\bar{X}$	9.98	5.78	5.35	7.14	7.04
S	4.22	5.52	3.14	6.21	

Table VIII: Analysis of Variance of Weight

Sources of Variation	df	ss	ms	F	Remarks
Between combination	3	120.94	4.31	<1	Insignificant
Within combination	59	1635.98	27.73	-	-
Total	62	1756.92			

Table IX: Summary Results of Volume of Solid Waste in a 2x2 Factorial Experiment

	Treatments (Structure-Housing Unit)				Combination Total
	(1)	H	S	SH	
n	6	12	22	23	63
$\bar{x}$	0.0393	0.033	0.0436	0.0606	0.4737
s	0.01309	0.0157	0.01508	0.08665	

Table X: Analysis of Variance for Volume

Sources of Variation	df	ss	ms	F	Remarks
Between combination	3	.00772	.00241	<1	Insignificant
Within combination	59	.17355	.002042		
Total	62	.18077			

The analysis variance technique shows that there is no significant difference between structures and between housing units with regards to weight and volume of solid waste. The built-in variation in structures necessitates to eliminate structure as a factor. The results further show that a sample of 2 housing units from each selected block will suffice and as such, the final design for the survey would be to select a simple random sample of 1.5% blocks from each area and to select 2 housing units from each sample block.

Acknowledgements

The work is supported by the Saudi Arabian Center for Science and Technology under Project No. AR-04-049.

REFERENCES

1. Bayunus, O and Khan, M.Z., (1981), "Forecasting Automobile Rate of Abandonment - A Case Study in Solid Waste Management, City of Jeddah", 10th Pakistan Statistical Conference, Islamabad.
2. Khan, M.Z. Ali, (1979), "Community Housing and Refuse Disposal in Developing Countries", Edited by Oktay Ural, Vol. I, Pergamon Press, New York.

3. Khan, M.Z. and Bayunus O., (1980), "Energy Conservation by Resource Recovery and Reuse in Solid Waste Management", Published in the book "Energy Resources and Conservation Related to Built Environemt", Edited by Oktal Ural Vol. I, Pergamon Press, New York.

4. Khan, M.Z. and Bayunus, O., (1981), "Factors Affecting the Solid Waste Generation and Composition in Semi Arid Coastal Zones", Proceedings Institute of Endironmental Science, Energy and the Environment, Volume II, Los Angeles, California.

* Estimation done during the project.

5. Office of Solid Waste Management, Environmental Health Service, Department of Health, Education and Welfare, (1968a), Massive Statistics Public Health Service Publication No. 1867.

6. Office of Solid Waste Management, Environmental Health Service, Department of Health Education and Welfare, Interim Report, (1968b) National Survey of Community Solid Waste Practices.

7. Tebobanoglous, G., Theisen, H. and Elliasen, R., (1977), "Solid Wastes Engineering; Principles and Management Issues", McGraw-Hill New York.