ALBUKHARY INTERNATIONAL UNIVERSITY
albukhary socialbusiness journal
Vol. 1, Issue 2, 2020

Determination of Sample Size in Early Childcare Centre (TASKA) Service Project in Malaysia: Classification and Analytical Approach

Azrul Fazwan Kharuddin [a], Norazura Azid [b], Zaida Mustafa [c], Muhamad Naim Kamari [d], Ku Faridah Ku Ibrahim [c] Darvinatasya Kharuddin [e]

[a] *Graduate School of Business, Universiti Tun Abdul Razak (UNIRAZAK), Malaysia*
[b] *SMK Jeram, Selangor, Malaysia*
[c] *School of Education & Humanities, Universiti Tun Abdul Razak (UNIRAZAK), Malaysia*
[d] *Faculty of Hospitality & Tourism, City International University, Malaysia*
[e] *School of Mathematical Science, Universiti Sains Malaysia (USM), Malaysia*

[a] *azrulfazwan@unirazak.edu.my*
[b] *zurazid@gmail.com*
[c] *zaida@unirazak.edu.my*
kufaridah@unirazak.edu.my
[d] *naimkamari62@gmail.com*
[e] *tasya.darvina@gmail.com*

ABSTRACT

Determining sample size for large-scale projects such as TASKA services throughout Malaysia is important and requires research from all aspects. The difficulty in determining the size of these samples seems to be scary and tempting researchers to use formal calculation rule-of-thumb based on goals from past trials. We consider a wide range of situations and experimental designs commonly used in determining the sample size of the study, including discussions on procedures for controlling the rates of false findings and adjustments for small sample situations. Here we present theoretical formula to determine the sample size to accomplish the research objective, including class clarification and analytical approach. The results obtained explain the impact of minimum approaches, technical replication and strata arrangement based on proportional to random size requirements. These results are derived from collective discussion of dependency studies on the relative size of various variables. We are discussing the procedures for controlling the rates of false discovery. Our calculations are based on a fairly simple but realistic statistical model for data, and provide a simple sample size calculation formula.

Keywords: Analytical approach, class clarification, sampling technique, stratified random sampling

1.0 INTRODUCTION

Early childhood care and education is a term that applies to all types of child care services and education (Phillips & Lowenstein, 2011). The term illustrates that early education and childcare is an essential service. However, early education and child care programs are often linked to support services to working mothers and termed as child education services (Rohaiza Rokis, 2014). In Malaysia, this centre is called Child Care Centre (TASKA) which accepts children aged 0 to 4 years. This education center was set up to provide opportunities for working mothers to obtain educational services that promote a comprehensive, balanced and integrated child growth and development (Rohaiza Rokis, 2014).

We consider a wide range of situations and experimental designs commonly used in determining the sample size of the study, including discussions on procedures for controlling the rates of false findings and adjustments for small sample situations. This calculation is based on a relatively simple yet realistic statistical model for data, and provides a simple sample size calculation formula. There are few researches published in the TASKA service literature to determine the number of samples needed for class comparison problems, or analytical justification. We focus here on two statistical goals:

i) class clarification, and
ii) analytical justification,

We present the equation for the sample size formula as it gives insight into the effect that variance parameters and experimental designs have an algorithmic sample size requirement or a 'black box' calculation. Often in practice, especially in small sample situations, statistical software is used to determine sample size. The result will be better for estimation because the software package typically uses a calculation algorithm to achieve a better estimate for the true power and level. Hence, the sample size estimation that resulting is better. In class clarification, one is interested in identifying the number of sample sizes according to the different classes specified. Thus, generating a study by focusing on the conclusions for the appropriate selected TASKA is the prioritized. When constructing the analytical justification, the main goal is to build a multivariate indicator; the first step is usually to identify the TASKA with interest. Once again, the research needs to be empowered based on the conclusions for the TASKA by category (Kharuddin *et al,* 2019). The multivariate approach to sample size does not seem to be appropriate for this purpose. All the sample size formulas presented here are based on the assumption of the typical linear model for each TASKA category. We think this assumption is close enough to the truth for log intensity data as many educational studies using this approach have identified free expression of independence certified by other technologies. The calculation of power based on the assumption of a typical linear model will be empirically tested and sufficient.

2.0 LITERATURE REVIEW

The primary history of Early Childhood Education began in the 1900s in neighboring countries such as Philippines, Singapore and Thailand. While the development of Early Childhood Education in Malaysia started in the 1940s. In Malaysia, beginning in 1950 to 1960, Christian socialist church welfare works have opened the kindergarten (TADIKA) and are in charge of providing Early Childhood Education only for parents who can afford to pay only fees. The implication is that only a few children aged 4 to 6 years have the opportunity to follow the TADIKA class. However, many children from poor and rural families are unable to receive early education at TADIKA. The Malaysian Social Welfare Department (JKMM) or KEMAS has opened a Child Care Centre (TASKA) which is in line with the new philosophy of rural development. The program provides child care services to urban, rural and suburban communities, especially for the poor and underprivileged income.

National policy on child placement in care center varies depending on the public or private sector. An example of the implementation of parenting and education of children using a full civil responsibility system is as in Denmark. Children aged 26 weeks up to the age of compulsory admission are entitled to a place at the child care center (Danish Ministry for Children and Education, 2012). This is the same as in Sweden, where children have the legal right to be placed in care center from 1 year old to the age of compulsory schooling (Naumann, 2005). As discussed in this study, many aspects of custodial care include the items not discussed above have been linked to quality assessments. Although this aspect influences the decision on quality of care, it should be used with caution (Kharuddin *et al,* 2020). The environment in which the childcare system is located will affect the philosophy of the system and will determine the characteristics of the service as well as manipulate its quality control (Tayler et al., 2006). Therefore, comparisons of standards and quality assertions need to be interpreted through the number of sample sizes considered adequate by assuming the objective criteria of child care are met.

3.0 SAMPLING METHOD

Generally, there are several methods of classification in determining the sample size of the study for known or unknown populations. There are at least four methods of sample size determination commonly used, namely:

i) Yamane Method (Burgess et al., 2014),
ii) Krejcie and Morgan Methods (Anderson et al., 2016),
iii) Cochran method (Guo & Fraser, 2014) and
iv) G*Power Software (Kelly, 2015).

Sampling is required to obtain the sample value or size to estimate the population value from which the sample was taken. This is to save time, energy and finance because the samples have a smaller degree of variability than the degree of population. So often the variability of the population is scattered to produce a normal distribution graph. Therefore the determination of the sample size should be given attention so that the inferential statistics can be done. In particular, the total number of TASKA population in Malaysia whether registered or unregistered is obtained with the assistance of JKMM. According to the latest report dated November 30, 2016, TASKA numbers in each state in Malaysia are sorted according to TASKA types reported as in Table 1.

As reported, the total number of TASKA population in Malaysia is 4660. However, after investigation, the TASKA number that has complete information, which is legally registered with JKMM, is 3730 and the rest of 930 are not or not yet registered. The following is a sample size determination technique commonly used for the determination of TASKA sample size registered with JKMM and unregistered.

i) Yamane Method

This study first discusses Yamane's technique in determining the size of the sample in TASKA selection to be carried out throughout Malaysia. This technique was created in 1967 where a simple formula was introduced in determining the size of the sample if the size of the population was known (Levy & Lemeshow, 2013).

$$n = \frac{N}{(1 + Ne^2)}$$

where,
N = limited population is known
e = margin of error

In this situation, a 95% confidence interval is used to describe a 5% margin of error (Westland, 2010). For the calculation of the minimum sample size under the limited population method, stratified sampling, n is:

$$n = \frac{4660}{(1 + 4660(0.05^2))}$$
$$n \approx 368$$

Through this Yamane method, 368 TASKA will be identified for research purposes.

ii) Krejcie and Morgan method

The second sample size calculation techniques were Krejcie and Morgan method, introduced in 1970. If population size is known (target or accessibility), the Krejcie and Morgan tables can be used. The calculations in this table do not take into account statistical power and effect size (Collins, 2016). The sample size calculation for this study can be determined based on Table Size Determination of Krejcie and Morgan Samples. This sample size determination table is adopted by academics and researchers and facilitates the researcher to determine the sample size for research purposes. In the table, there is no exact number of population sizes 4660 but it is estimated that the nearest population is 4500. If the population is known to be equal to 4500, it is estimated that the number of samples to be chosen is 354 (Table 2). After obtaining sample sizes based on the population of the study, a sample size check patterned will be implemented. Reviews can be made online at surveysystem.com. Researchers only need to include population size and standard deviation so the total sample size will be estimated at 354 (Table 3).

In contrast to the Yamane method, this technique illustrates where, the first table is only an estimate of 354. The calculator is then used to verify the actual sample size based on the confined population size of 355. Nevertheless, the Krejcie & Morgan approach does not discuss directly on the significant level and sampling error compared to Cohen et al. (2001).

iii) Cochran (1977)

Cochran technique based on marginal error formula for researchers to obtain item errors such as 5% margin for categorical data and 3% margin for continuous data. These are errors that are deliberately held and accepted by researchers as each measurement must have an error. This includes receiving error with α = 0.05 (Type I error). Therefore, Cochran has provided two formulas namely continuous data and discrete data. The population size is not compulsory to know when using the Cochran formula (Singh & Masuku, 2014).

Cochran sample size formula for categorical data with confidence level at 0.05 (5% error):

$$n_0 = \frac{t^2(p)(1-p)}{d^2}$$
$$n_0 = 384$$

Where,
n_0 is the sample size,
t is the chosen value based on alpha level such as 1.96 for 25% per two-way test,
p is an estimate of the probability of a successful test,
1 - p is an estimate of the probability of a test failure and
d is the marginal error received for proportional budgeting.

However, the size of 384 is not an end to this technique; Cochran also creates a correction formula where if the population is estimated to be less than 50,000 known then the proposed number of samples is:

$$n_1 = \frac{384}{\left(1 + \frac{384}{4660}\right)}$$
$$n_1 = 354$$

Through this Cochran method, it is estimated that 354 TASKA will be identified for research purposes.

Moreover, if we do not know the total population of the study, the G*Power software is more appropriate as it helps us to set the size of the sample taking into account the size of the effect, the sampling error and the significant level of the study. Cohen's technique also discusses the size of this effect in his study (Gao et al., 2014).

iv) G*Power (2011 - version 3.1)

Based on statistical power and size effects, this free software is unfamiliar. G*Power is very useful because its output is not only based on the type of statistical analysis, but also includes graphical views. By making 80% statistical power estimates (prevalence rate) and the effect size you expect first, you will know the required 0- (Table 4) (Hair et al., 2016).

If the population is known, it is easy for researchers to determine sample size through three option; Yamane method, Krejcie and Morgan schedule or Cochran formula calculation technique. However, if population size is unknown, G*Power software is the solution (Button et al., 2013). This software will be used if the actual number of TASKA in Malaysia is unknown and the researchers are experiencing difficulties in determining the actual number of population. Discussion on this software method will be dismissed if any information on TASKA registered and unregistered is obtained from JKMM.

Based on the three techniques described above, the sample sizes to be observed are slightly the same as 368 (Yamane), 355 (Krejcie & Morgan) and 354 (Cochran). The researcher agreed to choose the sample size of 368 which is by Yamane method because the ideal number to represent the size of the TASKA population in Malaysia. The Yamane method was given priority as the researchers managed to collect and know the actual amount of TASKA population size and it was limited to 4660. In addition to the highest number of the three techniques discussed, if the sample size was not reached, at least the lowest sample size option can be achieved and it's still at a 95% confidence interval estimate with a marginal error of 0.05.

4.0 SAMPLING CORRECTION

Based on table 5, the highest number of TASKA types to be visited is the institutional TASKA of 172. Compared to community TASKA, there are 3 centres only to be selected in the sample list. Through the minimum approach, the researcher feels that the number of community TASKA needs to be amplified because the feedback from this group is also important even as a minority group. Similarly, the total number for OKU TASKA (for disabled person) is only 6 throughout Malaysia. The researcher felt that all OKU TASKA should be visited because the contribution of this group should be given attention and may contribute to the findings of the study. As previously mentioned, unregistered TASKA numbers are fluctuate time to time. Thus, through technical replication the researcher feels the number of previous 73 samples should be increased to 84 as the percentage proportion of the group. The researcher chose to place a reasonable sampling size in order for this distribution to be fulfilled within the scheduled period to make the whole sample a total of 450. The focus of the TASKA selection is also directed towards purposive sampling techniques where a few criteria are taken into account:

i) homogeneous sampling from various locations such as urban, suburban, rural and inland
ii) respondents consist of different genders for parent and guardian categories
iii) the involvement of all OKU TASKA and
iv) minimum number of sample sizes is 20 for all TASKA types except OKU TASKA.

5.0 DISCUSSION

In this section we examine the context of child care 0 to 4 years old and the initial setting of education in relation to some aspects of the

quality reported by the owner / caretaker / child / parent or guardian. These include: structural features, which focus on the demographics of all respondents; positive practices in relation to child-caring interactions and approaches to teaching and learning in early education; space allocation and resources that enhance game play, learning and development of infants and children; as well as the characteristics of an environment organization that supports a positive environment for community.

Despite the limitations of the sample size and the respondents selected, the data presented in this section provide early support for the validity and robustness of TASKA service quality in Malaysia in early childhood education. This section also supports the suitability including selection of measures in multivariate analysis to evaluate the effect of early childcare education on the dependency on sample size being obtained.

We have presented a sample size formula for class clarifications and analytical justification for different types of TASKA and proportion stratified random sampling technique. We have also presented a sample size formula for a theoretical and practical study. In general, determining the number of TASKA is needed for the study does not require a researcher to 're-invent the wheel' in the sense that many reasons for the size of the classic statistical sample may be used in this new context.

On the contrary, the context of sampling correction attempts is fairly novel that it rises interesting questions as sources of relative variation are expected to affect sample size requirements. It will impact on how design decisions are made such as minimal approach, technical replication, and structural sampling on cost associated with the procedure of determining sample size by category. We have addressed these issues in our intuitive and hopeful ways so that this study will help as well as guide researchers in designing the determination in the future.

TABLE 1: Distribution of TASKA Population by Category in Each State in Malaysia

State	Institutional	Work Place	Community	Home	Unregistered	KEMAS	PAPN	JPNIN	YPKT	Population
Perlis	27	0	2	6	6	14	4	2	0	61
Kedah	183	7	3	9	50	48	3	3	0	306
P. Pinang	106	12	2	18	20	15	3	3	0	179
Perak	193	6	1	36	47	38	11	3	0	335
Selangor	596	43	7	258	283	28	2	2	0	1219
W.P	109	54	4	95	100	25	7	3	0	397
N. Sembilan	79	7	1	27	104	35	9	2	0	264
Melaka	72	12	2	10	12	20	10	4	0	142
Johor	237	10	3	43	32	53	6	2	0	386
Pahang	90	16	2	70	45	24	9	3	0	259
Kelantan	125	3	2	0	54	23	6	4	0	217
Terengganu	125	11	3	9	46	23	2	2	65	286
Sarawak	80	6	6	34	101	80	3	4	0	314
Sabah	143	13	4	21	30	50	13	2	0	276
Labuan	10	3	2	0	0	2	0	2	0	19
Total	**2175**	**203**	**44**	**636**	**930**	**478**	**88**	**41**	**65**	**4660**

TABLE 2: Determination of Sample Size from Known Population by Krejcie & Morgan Table

Total	*Sample*	*Total*	*Sample*	*Total*	*Sample*
10 ⇒	10	220 ⇒	140	1200 ⇒	291
15 ⇒	14	230 ⇒	144	1300 ⇒	297
20 ⇒	19	240 ⇒	148	1400 ⇒	302
25 ⇒	24	250 ⇒	152	1500 ⇒	306
30 ⇒	28	260 ⇒	155	1600 ⇒	310
35 ⇒	32	270 ⇒	159	1700 ⇒	313
40 ⇒	36	280 ⇒	162	1800 ⇒	317
45 ⇒	40	290 ⇒	165	1900 ⇒	320
50 ⇒	44	300 ⇒	169	2000 ⇒	322
55 ⇒	48	320 ⇒	175	2200 ⇒	327
60 ⇒	52	340 ⇒	181	2400 ⇒	331
65 ⇒	56	360 ⇒	186	2600 ⇒	335
70 ⇒	59	380 ⇒	191	2800 ⇒	338
75 ⇒	63	400 ⇒	196	3000 ⇒	341
80 ⇒	66	420 ⇒	201	3500 ⇒	346
85 ⇒	70	440 ⇒	205	4000 ⇒	351
90 ⇒	73	460 ⇒	210	4500 ⇒	354
95 ⇒	76	480 ⇒	214	5000 ⇒	357
100 ⇒	80	500 ⇒	217	6000 ⇒	361
110 ⇒	86	550 ⇒	226	7000 ⇒	364
120 ⇒	92	600 ⇒	234	8000 ⇒	367
130 ⇒	97	650 ⇒	242	9000 ⇒	368
140 ⇒	103	700 ⇒	248	10000 ⇒	370
150 ⇒	108	750 ⇒	254	15000 ⇒	375
160 ⇒	113	800 ⇒	260	20000 ⇒	377
170 ⇒	118	850 ⇒	265	30000 ⇒	379
180 ⇒	123	900 ⇒	269	40000 ⇒	380
190 ⇒	127	950 ⇒	274	50000 ⇒	381
200 ⇒	132	1000 ⇒	278	75000 ⇒	382
210 ⇒	136	1100 ⇒	285	100000 ⇒	384

TABLE 3: Calculator to Calculate Sample Size by Krejcie & Morgan Approach

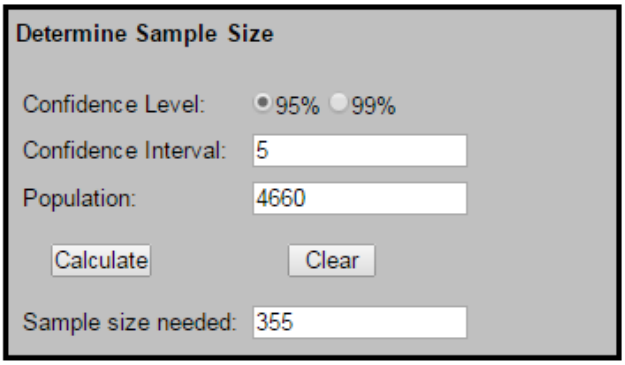

TABLE 4: G*Power Statistical Software to Calculate Estimation Sample Size Based on Expected Prevalence Rate and Level of Accuracy

Calculate Your Own Sample Size Here!		
Expected Prevalence (P)	**0.20**	(Between 0.01 till 0.99)
Level of Accuracy (d)	**0.05**	(Usually between 0.03 till 0.05)
Sample Size Required	**246**	
Confidence level 95%		

TABLE 5: Distribution of Population TASKA Placement All Over Malaysia by Registered and Unregistered Category

State	Institutional TASKA	Work Place TASKA	Community TASKA	Home TASKA	Unregistered TASKA	KEMAS	PAPN	JPNIN	OKU	YPKT
Perlis	2	1	1	0	1	2	1	1		0
Kedah	13	1	1	0	2	5	2	1		0
Pulau Pinang	7	1	1	2	2	3	2	1	1	0
Perak	16	0	1	3	2	5	1	2	1	0
Selangor	46	4	3	20	4	19	2	1	1	0
Wilayah Persekutuan	16	4	2	7	6	6	2	2	1	0
Negeri Sembilan	5	1	1	3	3	7	1	2		0
Melaka	2	1	1	0	1	3	1	2		0
Johor	15	1	1	4	2	3	2	1		0
Pahang	7	1	1	6	2	8	2	2		0
Kelantan	11	1	1	0	3	4	1	1		0
Terengganu	12	1	1	0	2	3	1	1		20
Sarawak	7	1	2	3	4	13	1	2	1	0
Sabah	12	2	2	2	4	3	1	1	1	0
Labuan	1	0	1	0	0	0	0	0		0
TOTAL	**172**	**20**	**20**	**50**	**84**	**38**	**20**	**20**	**6**	**20**

6.0 ACKNOWLEDGEMENT

The authors thanked for the comments and suggestions received from the editor in the making of this article. The authors will also admit UNIRAZAK staff and students for their assistance in completing the fieldwork of research.

7.0 REFERENCES

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Cochran, J. J. (2016). *Statistics for business & economics*. Nelson Education.

Burgess, Scott C., Kerry J. Nickols, Chris D. Griesemer, Lewis AK Barnett, Allison G. Dedrick, Erin V. Satterthwaite, Lauren Yamane, Steven G. Morgan, J. Wilson White, and Louis W. Botsford (2014). "Beyond connectivity: how empirical methods can quantify population persistence to improve marine protected-area design." *Ecological Applications* 24, no. 2: 257-270.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365-376.

Cohen, A., Vigoda, E., & Samorly, A. (2001). Analysis of the mediating effect of personal-psychological variables on the relationship between socioeconomic status and political participation: A structural equations framework. *Political Psychology*, *22*(4), 727-757.

Collins, K. M. (2016). Sampling Decisions in Educational Research. *The BERA/SAGE Handbook of Educational Research*, 280.

Danish Ministry of Children and Education, 2012. Ny Nordisk Skole [New Nordic School]. Retrieved March 24, 2014 from http:// nynordiskskole.dk/

Gao, D., Grunwald, G. K., & Xub, S. (2014). A Simple Approach to Sample Size Calculation for Count Data in Matched Cohort Studies. *International Journal of Statistics in Medical Research*, *3*(3), 321-330.

Guo, S., & Fraser, M. W. (2014). *Propensity score analysis: Statistical methods and applications* (Vol. 11). Sage Publications.

Hair Jr, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications http://www.surveysystem.com/sscalc.htm

Kharuddin, A. F., Mustafa, Z., Azid, N., Kharuddin, D., Ibrahim, K. F. K., Hong, L. Y., ... & Zahri, Z. A. (2019). Comparison of registration status of Institutional Taska services in East Coast Malaysia. *Southeast Asia Early Childhood Journal*, *8*(2), 57-62.

Kharuddin, A. F., Azid, N., Mustafa, Z., Ibrahim, K. F. K., & Kharuddin, D. (2020). Application of Structural Equation Modeling (SEM) in Estimating the Contributing Factors to Satisfaction of TASKA Services in East Coast Malaysia. *Asian Journal of Assessment in Teaching and Learning*, *10*(1), 68-76.

Kelly, D. (2015). Statistical power analysis for sample size estimation in information retrieval experiments with users. In *European Conference on Information Retrieval* (pp. 822-825). Springer International Publishing.

Levy, P. S., & Lemeshow, S. (2013). *Sampling of populations: methods and applications*. John Wiley & Sons.

Naumann, I. K. (2005). Child care and feminism in West Germany and Sweden in the 1960s. *Journal of European Social Policy*, *15*(1), 47-63.

Phillips, D. A., & Lowenstein, A. E. (2011). Early care, education, and child development. *Annual review of psychology*, *62*, 483-500.

Rokis, R. (2014). Work-care Balance among Parents-workers in Malaysian Urban Organizations: Role and Quality of Children's Daycare Centers. *American International Journal of Social Science*, *3*, 109-117.

Singh, A. S., & Masuku, M. B. (2014). Sampling techniques & determination of sample size in applied statistics research: An overview. *International Journal of Economics, Commerce and Management*, *2*(11), 1-22.

Tayler, C. P., Wills, M., Hayden, J., & Wilson, C. (2006). A review of the approach to setting national standards and assuring the quality of care in Australian child care services. Centre for Learning Innovation, Faculty of Education, QUT, Australia, Brisbane, QUT Kelvin Grove.

Westland, J. Christopher (2010), Lower Bounds on sample size in structural equation modelling. Electron. Comm. Res. Appl. 9 (6): 476-487