SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943



Analisis Kecekapan Relatif dan Ketepatan Anggaran dalam Reka Bentuk Pensampelan: Perbandingan Teoretikal dan Empirikal antara Pensampelan Rawak Mudah dan Pensampelan Rawak Berstrata

Analysis of Relative Efficiency and Estimation Accuracy in Sampling Designs: A Theoretical and Empirical Comparison between Simple Random Sampling and Stratified Random Sampling

Muhammad Fadhil Marsani[1], Vikneswari Someetheram[1,*], Farid Zamani Che Rose[2,3], Mohd Shareduwan Mohd Kasihmuddin[1]

[1] School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia
[2] Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[3] Institut Penyelidikan Matematik (INSPEM), Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 20 April 2026
Received in revised form 24 August 2026
Accepted 13 October 2026
Available online 9 November 2026

ABSTRACT

Cabaran utama dalam inferens berasaskan reka bentuk ialah meminimumkan varians penganggar tanpa meningkatkan kos pensampelan secara berlebihan. Walaupun Pensampelan Rawak Mudah (SRS) merupakan pendekatan asas dalam teori pensampelan, kecekapan reka bentuknya berpotensi menurun dalam populasi yang bersifat heterogen. Walau bagaimanapun, bukti empirikal dalam konteks populasi pendidikan sebenar masih terhad. Kajian ini menilai kecekapan relatif antara SRS dan Pensampelan Rawak Berstrata (StRS) dalam menganggar purata gred akhir Matematik (G3) bagi populasi pelajar berjumlah N = 395. Pengstrataan dilakukan berdasarkan jantina dan dianalisis berasaskan prinsip alokasi optimum Neyman. Lima saiz sampel berbeza (n = 60, 100, 150, 200, dan saiz minimum bagi had ralat 5%) diuji melalui replikasi pensampelan bagi menilai bias, ralat piawai dan Ralat Kuasa Dua Min (MSE). Hasil kajian menunjukkan bahawa kedua-dua reka bentuk adalah tidak berat sebelah terhadap min populasi 10.4152; namun StRS menghasilkan varians yang lebih stabil serta mengurangkan turun naik anggaran melampau. Peruntukan Neyman menunjukkan saiz sampel optimum yang sedikit lebih kecil bagi StRS (n = 68) berbanding SRS (n = 70) untuk mencapai had ralat ±1.0. Kecekapan relatif StRS dianggarkan sekitar 1.42%, menunjukkan peningkatan marginal tetapi konsisten. Dapatan ini menegaskan bahawa keberkesanan pengstrataan bergantung kepada tahap heterogeniti populasi dan sejauh mana pembentukan strata selari dengan struktur varians sebenar. Kajian ini menyumbang bukti empirikal bahawa walaupun perbezaan saiz sampel optimum adalah kecil, pengstrataan mampu meningkatkan kestabilan penganggar dan mengurangkan risiko sisihan ekstrem dalam populasi pendidikan berskala sederhana.

The main challenge in design-based inference is to minimize estimator variance without excessively increasing sampling costs. Although Simple Random Sampling (SRS) is a fundamental approach in sampling theory, its design efficiency may decline in

* *Corresponding author.*
E-mail address: viknessomeetheram@student.usm.my

Keywords:
Pensampelan Rawak Mudah (SRS); Pensampelan Rawak Berstrata (StRS); inferens berasaskan reka bentuk; alokasi Neyman; kecekapan relatif; varians penganggar; populasi pendidikan; Simple Random Sampling (SRS); Stratified Random Sampling (StRS); Design-Based Inference; Neyman Allocation; Relative Efficiency; Estimator Variance; Educational Population

heterogeneous populations. However, empirical evidence in the context of actual educational populations remains limited. This study evaluates the relative efficiency of SRS and Stratified Random Sampling (StRS) in estimating the mean final Mathematics grade (G3) for a student population of N = 395. Stratification was performed based on gender and analyzed according to the principles of Neyman optimal allocation. Five different sample sizes (n = 60, 100, 150, 200, and the minimum sample size required for a 5% margin of error) were examined through repeated sampling to evaluate bias, standard error, and Mean Squared Error (MSE). The findings indicate that both sampling designs produce unbiased estimates of the population mean of 10.4152; however, StRS yields more stable variance and reduces extreme fluctuations in estimates. Neyman allocation indicates a slightly smaller optimal sample size for StRS (n = 68) compared with SRS (n = 70) to achieve a margin of error of ±1.0. The relative efficiency of StRS is estimated at approximately 1.42%, indicating a marginal but consistent improvement. These findings emphasize that the effectiveness of stratification depends on the degree of population heterogeneity and the extent to which the formation of strata aligns with the underlying variance structure. This study provides empirical evidence that, although the difference in optimal sample sizes is relatively small, stratification can improve estimator stability and reduce the risk of extreme deviations in moderately sized educational populations.

1. Pengenalan

Pensampelan dilaksanakan apabila kajian terhadap keseluruhan populasi tidak praktikal dari segi kos, masa dan sumber, khususnya bagi populasi yang besar atau tersebar secara geografi [11]. Pemilihan reka bentuk pensampelan penting kerana mempengaruhi ketepatan dan kecekapan anggaran. Pensampelan Rawak Mudah (SRS) menjadi asas kerana setiap elemen populasi mempunyai kebarangkalian pemilihan yang sama [15]. Namun, bagi populasi yang heterogen, seperti berdasarkan lokasi sekolah, prestasi atau latar belakang sosioekonomi, Pensampelan Rawak Berstrata (StRS) dapat memastikan subkumpulan diwakili dengan lebih baik serta berpotensi mengurangkan varians anggaran [14]. Kajian ini menilai kecekapan dan ketepatan SRS dan StRS melalui perbandingan varians, ralat min kuasa dua (MSE), kestabilan anggaran dan implikasi praktikal berdasarkan populasi pendidikan berskala sederhana [5]. Walaupun literatur moden menekankan inferens berasaskan reka bentuk serta strategi pemilihan dan pembobotan untuk mengurangkan ralat, bukti empirikal yang membandingkan SRS dan StRS secara langsung menggunakan populasi pendidikan sebenar bersaiz sederhana masih terhad [1]. SRS kekal sebagai reka bentuk asas kerana kesederhanaan dan ketelusan anggarannya dalam kerangka design-based [5]. Walau bagaimanapun, keberkesanannya boleh menurun dalam populasi heterogen akibat peningkatan varians dan risiko under-representation subkumpulan kecil [6,8]. Dalam konteks pendidikan, SRS juga mungkin gagal menangkap variasi antara subkumpulan melainkan saiz sampel yang besar digunakan. Namun, sebahagian besar bukti sedia ada berasaskan simulasi atau perbincangan metodologi umum, manakala kajian menggunakan populasi pelajar sebenar masih kurang [6].

StRS pula memanfaatkan ketidakseragaman populasi dengan membentuk strata yang lebih homogen bagi mengurangkan varians, manakala maklumat tambahan dan strategi alokasi dapat meningkatkan lagi ketepatan anggaran [1, 5]. Antaranya, alokasi berasaskan expected response rate dilaporkan dapat mengurangkan varians berbanding alokasi berkadar saiz strata [14]. Estimator log-type dan estimator berasaskan pembolehubah tambahan turut menunjukkan pengurangan MSE serta peningkatan kecekapan relatif [9, 17]. Penambahbaikan estimator varians berasaskan peminimuman MSE juga menunjukkan prestasi yang lebih baik berbanding estimator piawai. Dalam konteks pendidikan, estimator StRS yang menggunakan pembolehubah tambahan menunjukkan prestasi empirikal yang lebih baik dalam meramalkan skor peperiksaan serta mengurangkan MSE berdasarkan data pelajar sebenar [16]. Walaupun metodologi StRS semakin berkembang, penilaian

kos–faedah masih terhad, khususnya bagi populasi berskala sederhana dan keuntungan kecekapan yang kecil tetapi konsisten [1].

Oleh itu, kajian ini menyediakan bukti empirikal berasaskan populasi sebenar untuk membandingkan bias, varians, MSE dan kestabilan anggaran antara SRS dan StRS. Kajian turut menilai kecekapan relatif serta strategi alokasi proportional dan ERR dari aspek varians dan saiz sampel. Hasil kajian diharapkan dapat memberikan panduan praktikal dalam pemilihan reka bentuk pensampelan yang sesuai bagi kajian akademik berskala sederhana.

2. Metodologi Kajian

2.1 Kerangka Reka Bentuk dan Inferens

Kajian ini menggunakan pendekatan inferens berasaskan reka bentuk (design-based inference), dengan kebarangkalian pemilihan unit sampel sebagai asas pembentukan penganggar dan pengiraan varians. Tiada andaian model parametrik dikenakan terhadap taburan Gred Akhir Matematik (G3). Set data yang mengandungi N = 395 pemerhatian dianggap sebagai populasi penuh dengan parameter yang diketahui. Parameter utama ialah min populasi, μ. Pendekatan ini membolehkan penilaian langsung terhadap bias, varians dan Ralat Kuasa Dua Min (MSE) bagi setiap reka bentuk pensampelan. Dua reka bentuk kebarangkalian dibandingkan:

1. Pensampelan Rawak Mudah Tanpa Penggantian (SRSWOR)
2. Pensampelan Rawak Berstrata (StRS)

Perbandingan dilakukan untuk menilai sama ada pengstrataan dapat meningkatkan kecekapan penganggar bagi populasi yang menunjukkan heterogeniti berdasarkan ciri demografi tertentu.

Heterogeniti antara subkumpulan. Secara teoretikal, sifat tidak berat sebelah (unbiasedness) bagi penganggar min Pensampelan Rawak Mudah ($\bar{y}_{SRS}$) dan Pensampelan Rawak Berstrata ($\bar{y}_{ST}$) disandarkan sepenuhnya kepada pembuktian matematik melalui operator jangkaan iaitu $E(\bar{y}_{SRS}) = \mu$ dan $E(\bar{y}_{ST}) = \mu$ di bawah kerangka teori pensampelan klasik [2, 10]. Oleh itu, sifat tidak berat sebelah penganggar tidak bergantung pada dapatan empirikal atau andaian tentang bentuk taburan populasi, sebaliknya merupakan sifat yang terhasil daripada reka bentuk pensampelan dan kebarangkalian pemilihan unit. Dapatan empirikal dalam kajian ini, khususnya apabila nilai min sampel daripada replikasi berulang didapati menghampiri min populasi, tidak bertujuan untuk membuktikan sifat teoretikal tersebut secara mutlak. Sebaliknya, dapatan tersebut berfungsi sebagai ilustrasi praktikal dan sokongan empirikal terhadap kestabilan serta prestasi penganggar dalam keadaan sampel terhingga [7]. Pendekatan ini seterusnya membolehkan perbandingan yang lebih jelas antara SRSWOR dan StRS dari segi ketepatan dan kecekapan penganggar bagi populasi yang mempunyai tahap heterogeniti yang berbeza.

2.2 Pensampelan Rawak Mudah (SRSWOR)

Dalam SRSWOR, sampel bersaiz n dipilih daripada populasi bersaiz N tanpa penggantian, di mana setiap kombinasi n unit mempunyai kebarangkalian yang sama untuk dipilih [12]. Penganggar min populasi di bawah reka bentuk ini ialah:

$$\bar{y}_{SRS} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Varians reka bentuk bagi penganggar min adalah:

$$Var(\bar{y}_{SRS}) = \left(1-\frac{n}{N}\right)\frac{S^2}{n}$$

di mana S^2 ialah varians populasi dan faktor $\left(1-\frac{n}{N}\right)$ merupakan Pembetulan Populasi Terhingga (Finite Population Correction, FPC).
Saiz sampel optimum bagi mencapai had ralat anggaran B ditentukan menggunakan:

$$n = \frac{NS^2}{(N-1)D + S^2}, D = \frac{B^2}{4}$$

dengan had ralat ditetapkan pada tahap 5%.

2.3 Pensampelan Rawak Berstrata (StRS)

Dalam reka bentuk ini, populasi dibahagikan kepada L strata yang saling eksklusif dan menyeluruh. Setiap strata ke-h mengandungi N_h unit dengan $\sum_{h=1}^{L} N_h = N$. Sampel bersaiz n_h dipilih secara SRSWOR dalam setiap strata.
Penganggar min populasi berstrata ialah:

$$\bar{y}_{ST} = \sum_{h=1}^{L} W_h \bar{y}_h$$

di mana $W_h = \frac{N_h}{N}$ ialah wajaran strata, $\bar{y}_h$ ialah min sampel dalam strata ke-h. Varians reka bentuk bagi penganggar berstrata diberikan oleh:

$$Var(\bar{y}_{ST}) = \sum_{h=1}^{L} W_h^2 \left(1-\frac{n_h}{N_h}\right)\frac{S_h^2}{n_h}$$

Dengan S_h^2 ialah varians populasi dalam strata ke-h. Bagi meminimumkan jumlah varians, alokasi Neyman digunakan:

$$n_h = n\frac{N_h S_h}{\sum_{h=1}^{L} N_h S_h}$$

Saiz sampel keseluruhan optimum ditentukan melalui:

$$n = \frac{\sum_{h=1}^{L} N_h^2 S_h^2}{N^2 D + \sum_{h=1}^{L} N_h S_h^2}$$

Pendekatan ini memastikan strata dengan varians yang lebih besar menerima peruntukan sampel yang lebih tinggi bagi memaksimumkan kecekapan. Dalam kajian ini, pemilihan jantina sebagai pemboleh ubah pengstrataan dalam kajian ini adalah berdasarkan faktor ketersediaan maklumat demografi asas tersebut dalam pangkalan data sekunder sekolah yang boleh diakses dengan mudah tanpa melibatkan kos tambahan. Walau bagaimanapun, keberkesanan reka bentuk berstrata amat bergantung kepada kewujudan varians antara strata yang tinggi dan varians dalam strata yang rendah.

2.4 Reka Bentuk Perbandingan dan Replikasi

Bagi menilai kestabilan anggaran, lima saiz sampel ditetapkan:

n = 60, 100, 150, 200, dan satu saiz optimum yang dikira berdasarkan had ralat. Bagi setiap saiz sampel, pensampelan dijalankan secara berulang bagi kedua-dua reka bentuk untuk memerhati variasi anggaran antara replikasi. Pendekatan ini membolehkan analisis terhadap turun naik anggaran yang berpunca daripada proses pemilihan rawak, bukan semata-mata perbezaan reka bentuk.

2.5 Ukuran Prestasi Penganggar

Prestasi kedua-dua reka bentuk pensampelan dinilai menggunakan ukuran statistik seperti yang dirumuskan dalam Jadual 1.

Jadual 1
Ukuran Statistik bagi Penilaian Prestasi Penganggar di bawah Reka Bentuk Pensampelan

Ukuran	Notasi / Formula
Bias	$Bias(\bar{y}) = E(\bar{y}) - \mu$
Varians Reka Bentuk	$Var(\bar{y})$
Ralat Piawai (SE)	$SE = \sqrt{Var(\bar{y})}$
Ralat Kuasa Dua Min (MSE)	$MSE = Bias^2 + Var(\bar{y})$
Kecekapan Relatif (RE)	$RE = \frac{Var(\bar{y}_{SRS})}{Var(\bar{y}_{ST})}$
Kesan Reka Bentuk (Deff)	$Deff = \frac{Var_{design}}{Var_{SRS}}$
Lebar Selang Keyakinan (CI Width)	$2 \times 1.96 \times SE$

Berdasarkan Jadual 1, prestasi kedua-dua reka bentuk dinilai berdasarkan ketepatan, kejituan dan kecekapan penganggar. Bias menunjukkan sejauh mana anggaran menyimpang daripada parameter populasi, manakala varians dan ralat piawai (SE) menunjukkan kestabilan anggaran. Ralat Kuasa Dua Min (MSE) menggabungkan bias dan varians untuk memberikan ukuran keseluruhan prestasi penganggar. Kecekapan Relatif (RE) dan Kesan Reka Bentuk (Deff) digunakan untuk membandingkan kecekapan kedua-dua reka bentuk dengan SRS sebagai rujukan. Lebar selang keyakinan 95% turut digunakan untuk menilai ketepatan anggaran. Selang yang lebih sempit menunjukkan anggaran yang lebih tepat. Selang keyakinan 95% bagi min populasi dibina seperti berikut:

$$\bar{y} \pm 1.96\sqrt{Var(\bar{y})}$$

Lebar selang keyakinan digunakan sebagai indikator tambahan kejituan penganggar.

3. Keputusan dan Analisis

Penerangan Populasi dan Parameter

Set data Prestasi Pelajar (student-mat.csv) daripada Kaggle (Student Performance Data) digunakan sebagai populasi kajian. Populasi terdiri daripada 395 orang pelajar, dengan skor akhir Matematik (G3) sebagai pemboleh ubah utama. Parameter populasi G3 dikira berdasarkan

keseluruhan set data dan dilaporkan dalam Jadual 1. Nilai ini digunakan sebagai penanda aras untuk menilai ketepatan dan kejituan anggaran daripada kaedah pensampelan yang berbeza.

Jadual 1
Parameter Populasi bagi Skor Matematik (G3)

Statistic	Value
Saiz populasi (N)	395
Min populasi (G3)	10.4152
Varians populasi (G3)	20.9896
Sisihan piawai populasi (G3)	4.5814

Sebelum analisis dijalankan, Analisis Data Penerokaan (*Exploratory Data Analysis*, EDA) ringkas dilakukan dan jenis data bagi setiap pemboleh ubah disesuaikan untuk memudahkan analisis seterusnya. Berdasarkan set data, terdapat 208 pelajar perempuan dan 187 pelajar lelaki. Maklumat ini digunakan dalam Pensampelan Rawak Berstrata (StRS). Saiz sampel bagi kedua-dua kaedah iaitu Pensampelan Rawak Mudah (SRS) dan StRS turut ditentukan. Bagi memudahkan penulisan dan notasi, SRS1 merujuk kepada sampel pertama yang dipilih menggunakan SRS, SRS2 kepada sampel kedua, dan seterusnya. STR1 merujuk kepada sampel pertama yang dipilih menggunakan StRS, STR2 kepada sampel kedua, dan seterusnya. Jadual 2 menunjukkan anggaran statistik bagi gred akhir Matematik (G3), dengan min dan sisihan piawai populasi sebagai rujukan. Tiada ralat piawai (SE) atau selang keyakinan 95% dilaporkan bagi populasi kerana nilai tersebut merujuk kepada anggaran sampel, bukan parameter populasi.

Jadual 2
Anggaran Statistik

Sampel	n	Min	Sisihan Piawai	Ralat Piawai (SE)	Selang Keyakinan 95% bagi Min
SRS1	60	11.2667	3.9525	0.4705	(10.3445, 12.1889)
SRS2	100	10.6800	4.3155	0.3734	(9.9481, 11.4119)
SRS3	150	10.2400	4.3914	0.2827	(9.6859, 10.7941)
SRS4	100	10.2600	4.7703	0.2373	(9.7949, 10.7251)
SRS5	174	10.5632	4.6176	0.2622	(10.0493, 11.0771)
STR1	60	10.4667	3.7618	0.4478	(9.5890, 11.3444)
STR2	100	10.7800	4.6244	0.4001	(9.9958, 11.5642)
STR3	150	10.0533	4.4281	0.2851	(9.4945, 10.6121)
STR4	200	10.2450	4.5990	0.2288	(9.7966, 10.6934)
STR5	173	10.3391	4.4430	0.2523	(9.8446, 10.8336)
Populasi		**10.4152**	**4.5814**	–	–

*Ralat Piawai (SE) Dibetulkan dengan FPC

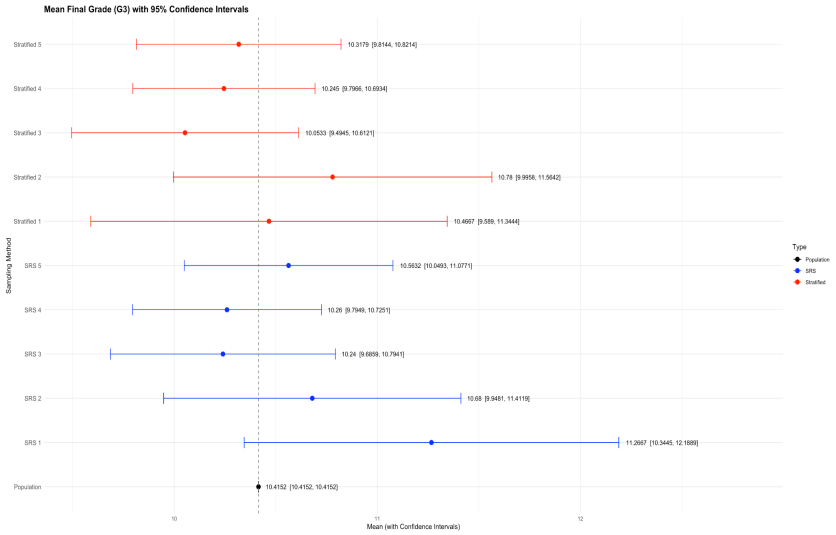
Berdasarkan Jadual 2, min populasi G3 ialah 10.4152. Secara umum, min sampel SRS dan StRS tertumpu di sekitar nilai ini, dengan kebanyakan anggaran berada dalam julat ±0.3. Sebagai contoh, SRS1 (11.2667) terlebih anggar min populasi sebanyak kira-kira 0.8, manakala STR3 (10.0533)

terkurang anggar sebanyak kira-kira 0.36. Secara keseluruhan, kedua-dua kaedah menghasilkan anggaran yang hampir dengan parameter populasi. Dari segi variabiliti, sisihan piawai sampel bagi kedua-dua kaedah adalah hampir dengan sisihan piawai populasi, iaitu 4.5814. StRS secara umum menghasilkan ralat piawai yang sedikit lebih rendah berbanding SRS pada saiz sampel yang sama. Sebagai contoh, bagi n ≈ 60, SRS1 mencatatkan ralat piawai 0.4705, manakala STR1 ialah 0.4478. Walau bagaimanapun, StRS tidak sentiasa mengatasi SRS kerana ralat piawai StRS bagi Sampel 2 dan Sampel 3 didapati lebih tinggi. Bagi kedua-dua kaedah, ralat piawai berkurang apabila saiz sampel meningkat. Perlu diingatkan, di bawah reka bentuk pensampelan rawak, kedua-dua SRS dan StRS adalah reka bentuk yang tidak bias (*unbiased*) secara purata jangkaan. Walau bagaimanapun, varians reka bentuk SRS yang lebih besar bermaksud ia mempunyai variabiliti pensampelan yang lebih tinggi. Implikasi praktikal bagi situasi ini ialah, dalam pengambilan sampel tunggal, SRS mempunyai kebarangkalian rawak yang lebih tinggi berbanding StRS untuk menghasilkan sampel pencilan ekstrem (lihat sampel SRS1) yang menyimpang jauh daripada keadaan populasi sebenar dan bukannya disebabkan oleh bias sistematik dalam kaedah tersebut. Bagi selang keyakinan 95%, semua selang SRS dan StRS merangkumi min populasi sebenar. Selang paling sempit diperoleh bagi STR4 (0.8968 unit), manakala selang paling lebar ialah SRS1 (1.8444 unit). Dapatan ini menunjukkan bahawa peningkatan saiz sampel meningkatkan kejituan dan mengurangkan lebar selang keyakinan, selaras dengan teori pensampelan. Seterusnya, kecekapan relatif digunakan untuk membandingkan prestasi StRS dengan SRS, dengan nisbah varians min sebagai ukuran kuantitatif kelebihan StRS.

$$RE = \frac{\sum_{i=1}^{5} SE_{SRS}^2}{\sum_{i=5}^{5} SE_{STR}^2} = \frac{0.5658}{0.5579} = 1.0142$$

Secara keseluruhan, pensampelan berstrata didapati kira-kira 1.42% lebih cekap berbanding Pensampelan Rawak Mudah (SRS). Walaupun peningkatan ini kecil, dapatan tersebut menunjukkan bahawa pengstrataan berdasarkan jantina memberikan peningkatan kecekapan sederhana dengan mengambil kira sebahagian varians anggaran. Walaubagaimanapun, dalam kajian pendidikan sebenar, kos pentadbiran untuk membina rangka pensampelan berstrata berdasarkan jantina, merekodkan wajaran strata dan melaksanakan alokasi Neyman mungkin tidak menghasilkan pulangan kecekapan yang berbaloi berbanding kesederhanaan reka bentuk SRS. Melainkan maklumat strata demografi ini sudah tersedia dalam pangkalan data sekolah tanpa kos tambahan, penggunaan SRS adalah lebih praktikal dan cekap kos bagi populasi berskala sederhana ini. Dalam kajian ini, analisis data kami menunjukkan varians gred Matematik (G3) adalah hampir setara antara pelajar lelaki dan perempuan yang membuktikan bahawa jantina bukan pemboleh ubah pembeza yang kuat bagi prestasi Matematik dalam populasi ini. Ini menjelaskan mengapa peningkatan kecekapan StRS yang diperoleh adalah sangat marginal.

Rajah 1 menunjukkan plot garis dengan bar ralat bagi min gred akhir (G3) dan selang keyakinan 95%. Min sampel serta selang keyakinan dipaparkan di sebelah kanan garis. Visualisasi yang lebih jelas boleh dirujuk dalam Lampiran kerana saiz garis dan fon dalam rajah adalah kecil.



Rajah 1. Plot Garis dengan Bar Ralat (Min Gred Akhir dengan Selang Keyakinan 95%)

Sesuatu selang keyakinan dianggap berjaya apabila merangkumi min sebenar populasi, iaitu apabila garis selang keyakinan bersilang dengan garis putus-putus dalam plot. Berdasarkan plot, empat daripada lima sampel SRS mempunyai min yang hampir dengan min populasi dan selang keyakinan 95% yang merangkumi parameter tersebut. Hal ini menunjukkan bahawa SRS tidak menunjukkan bias sistematik, walaupun terdapat variabiliti rawak antara sampel. SRS1 menunjukkan penyimpangan paling ketara dengan min sampel yang terlebih anggar dan selang keyakinan yang hampir tidak merangkumi min populasi. Keadaan ini menunjukkan bahawa sampel tertentu masih boleh menghasilkan anggaran yang jauh daripada parameter sebenar. Sebaliknya, semua sampel StRS merangkumi min populasi dan secara umum mempunyai selang keyakinan yang lebih sempit berbanding SRS pada saiz sampel yang hampir sama. Sebagai contoh, STR4 mempunyai selang yang lebih sempit berbanding SRS4, menunjukkan peningkatan kejituan melalui pengstrataan berdasarkan jantina. Min sampel StRS juga lebih tertumpu di sekitar min populasi, manakala SRS1 menjadi pencilan yang jelas. Plot turut menunjukkan bahawa saiz sampel yang lebih besar mengurangkan ketidakpastian, dengan SRS4 dan STR4 menunjukkan selang yang lebih sempit berbanding SRS1 dan STR1. Dapatan ini selaras dengan prinsip bahawa saiz sampel yang lebih besar menghasilkan anggaran yang lebih tepat.

Dari segi taburan, median kebanyakan sampel adalah hampir dengan median populasi. Namun, SRS menunjukkan perubahan median dan julat antara kuartil (IQR) yang lebih ketara antara sampel. SRS1 mempunyai median yang lebih tinggi, selaras dengan dapatan bahawa sampel tersebut terlebih menganggar min populasi. Sebaliknya, StRS menunjukkan IQR yang lebih konsisten serta taburan yang lebih hampir dengan populasi, khususnya bagi STR3, STR4 dan STR5. Nilai pencilan pada skor rendah terdapat dalam kedua-dua kaedah dan mencerminkan ciri semula jadi data prestasi pelajar. Secara keseluruhan, plot kotak menyokong dapatan berangka bahawa StRS menghasilkan taburan yang lebih konsisten dan representatif, manakala SRS lebih terdedah kepada variabiliti pensampelan antara replikasi.

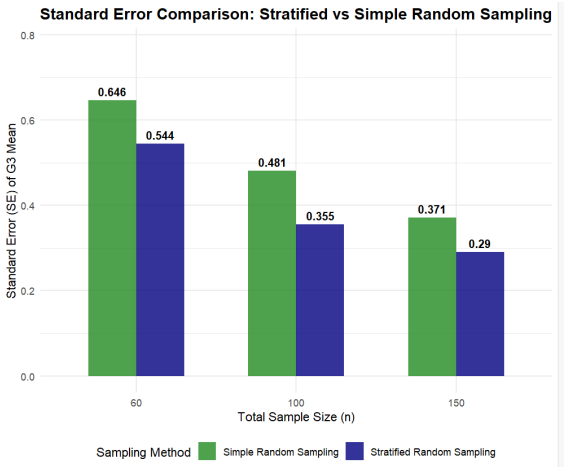
Jadual 3
Metrik Penilaian

Sampel	Ketepatan (Bias)	Kejituan (SE)	Ralat Kuasa Dua Min (MSE)
SRS1	0.8515	0.4705	0.9464
SRS2	0.2648	0.3734	0.2095
SRS3	-0.1752	0.2827	0.1106
SRS4	-0.1552	0.2373	0.0804
SRS5	0.1480	0.2622	0.0907
Purata (SRS)	**0.18678**	**0.32522**	**0.28752**
STR1	0.0515	0.4478	0.2032
STR2	0.3648	0.4001	0.2932
STR3	-0.3619	0.2851	0.2123
STR4	-0.1702	0.2288	0.0813
STR5	-0.0973	0.2569	0.0755
Purata (StRS)	**-0.04262**	**0.32374**	**0.1731**

Jadual 3 membandingkan prestasi Pensampelan Rawak Mudah (SRS) dan Pensampelan Rawak Berstrata (StRS) berdasarkan bias, ralat piawai dan Ralat Kuasa Dua Min (MSE). Dari segi bias, SRS menunjukkan variasi yang lebih besar. SRS1 terlebih menganggar min populasi (0.8515), manakala SRS3 dan SRS4 menunjukkan terkurang anggar masing-masing sebanyak −0.1752 dan −0.1552. StRS tidak menunjukkan penyimpangan ekstrem seperti SRS1, walaupun STR2 dan STR3 mempunyai bias yang lebih tinggi berbanding sampel SRS yang sepadan. Ini menunjukkan bahawa StRS dapat mengurangkan penyimpangan ekstrem, tetapi tidak semestinya menghasilkan bias yang lebih rendah bagi setiap sampel.

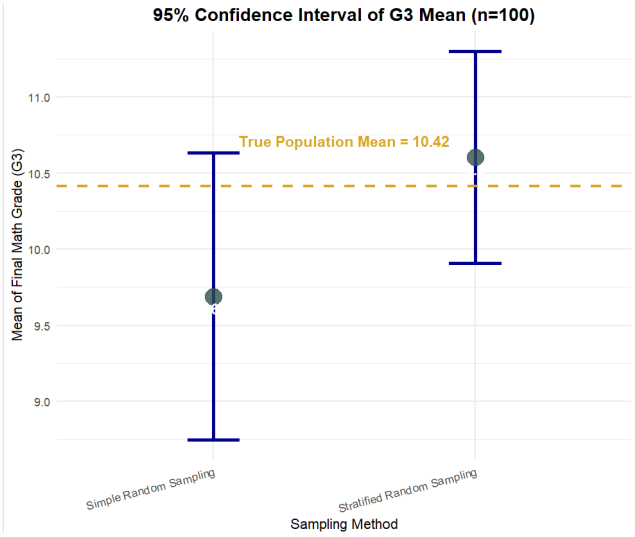
Dari segi ralat piawai, SRS berada antara 0.2373 hingga 0.4705, manakala StRS antara 0.2288 hingga 0.4478. Perbezaan yang kecil menunjukkan bahawa pengstrataan berdasarkan jantina tidak banyak mengurangkan varians. Ralat piawai terendah diperoleh bagi SRS4 (0.2373) dan STR4 (0.2288), menunjukkan bahawa saiz sampel yang lebih besar meningkatkan kejituan.

Bagi MSE, nilai tertinggi dicatatkan oleh SRS1 (0.9464), manakala nilai terendah ialah STR5 (0.0755), diikuti SRS4 (0.0804) dan STR4 (0.0813). Secara umum, StRS menghasilkan MSE yang lebih rendah dan lebih konsisten. StRS tidak memberikan peningkatan kejituan yang besar berbanding SRS, tetapi menunjukkan kestabilan dan konsistensi yang lebih baik. Oleh itu, StRS lebih sesuai digunakan untuk menganggar min populasi dalam kajian ini.



Rajah 2. Perbandingan Ralat Piawai (Standard Error Comparison)

Rajah 2 menunjukkan perbandingan ralat piawai bagi SRS dan Pensampelan Rawak Berstrata (StRS) pada saiz sampel yang berbeza. Secara umum, StRS mempunyai ralat piawai yang lebih rendah berbanding SRS bagi setiap saiz sampel. Walaupun perbezaan semakin kecil apabila saiz sampel meningkat, StRS masih menunjukkan kelebihan dari segi kejituan. Apabila saiz sampel meningkat daripada 60 kepada 150, ralat piawai StRS menurun daripada 0.544 kepada 0.290, iaitu kira-kira 46.7%. Bagi SRS, ralat piawai menurun daripada 0.646 kepada 0.371, iaitu kira-kira 42.6%. Ini menunjukkan bahawa StRS bukan sahaja menghasilkan ralat piawai yang lebih rendah, malah menunjukkan peningkatan kejituan yang lebih baik apabila saiz sampel bertambah.



Rajah 3. Selang Keyakinan 95% (n = 100)

Rajah 3 menunjukkan selang keyakinan 95% bagi min G3 apabila saiz sampel adalah 100. Selang keyakinan bagi Pensampelan Rawak Berstrata (9.91–11.31) merangkumi min populasi sebenar (10.42), dengan lebar selang sebanyak 1.40. Sebaliknya, selang keyakinan bagi Pensampelan Rawak Mudah (8.75–10.63) mempunyai lebar 1.88, iaitu lebih luas berbanding pensampelan berstrata. Selain itu, min anggaran bagi pensampelan berstrata (10.61) adalah lebih hampir kepada min populasi sebenar berbanding anggaran SRS (9.69), sekali gus membuktikan kelebihan pensampelan berstrata dari segi ketepatan anggaran. Selaras dengan kerangka inferens berasaskan reka bentuk (design-based inference), skop kajian ini adalah dihadkan secara eksklusif kepada penilaian min sampel ($\hat{y}$) sebagai penganggar tidak bias (unbiased estimator) bagi min populasi (μ). Kajian ini tidak

bertujuan untuk membuat inferens terhadap parameter lokasi lain seperti median, mahupun mencirikan keseluruhan bentuk taburan populasi.

4. Kesimpulan

Analisis data pelajar Matematik menunjukkan bahawa kedua-dua Pensampelan Rawak Mudah (SRS) dan Pensampelan Rawak Berstrata (StRS) menghasilkan anggaran min populasi yang hampir tidak berat sebelah, dengan min sebenar 10.4152. Namun, StRS memberikan peningkatan kecekapan sederhana, iaitu kira-kira 1.42% berbanding SRS apabila jantina digunakan sebagai strata. SRS menunjukkan variabiliti rawak yang lebih tinggi, seperti SRS1 yang terlebih menganggar min populasi sebanyak 0.8515 dan mencatatkan MSE tertinggi (0.9464). Sebaliknya, StRS secara umum menunjukkan bias mutlak yang lebih kecil serta taburan yang lebih konsisten dengan populasi. Saiz sampel turut mempunyai hubungan langsung dengan kejituan anggaran. Sampel yang lebih besar seperti SRS4 dan STR4 menghasilkan ralat piawai lebih rendah serta selang keyakinan lebih sempit. Pada saiz sampel yang sama, StRS menghasilkan ralat piawai yang lebih rendah berbanding SRS. Bagi saiz sampel 100, ralat piawai StRS adalah 26% lebih rendah berbanding SRS dan min anggaran lebih hampir kepada min populasi, iaitu 10.42. Apabila saiz sampel meningkat daripada 60 kepada 150, ralat piawai StRS menurun sebanyak 46.7%, berbanding 42.6% bagi SRS, menunjukkan kelebihan StRS dalam mengurangkan ralat apabila saiz sampel meningkat.

Secara keseluruhan, walaupun SRS kukuh dari segi teori, StRS dapat mengurangkan risiko sampel yang kurang mewakili populasi dengan memastikan perwakilan mengikut jantina, iaitu 208 perempuan dan 187 lelaki. Dapatan ini disokong oleh nilai MSE yang lebih rendah serta kestabilan anggaran yang lebih baik merentasi pengulangan sampel. Oleh itu, StRS wajar dipertimbangkan dalam kajian akademik apabila maklumat demografi populasi tersedia. Penggunaan saiz sampel yang lebih besar juga disarankan bagi meningkatkan kejituan anggaran. Kajian lanjut boleh meneroka pemboleh ubah pengstrataan lain seperti gred terdahulu (G1 dan G2) atau tempoh pembelajaran (studytime) bagi mencapai peningkatan kecekapan yang lebih tinggi. Walaubagaimanapun, cadangan untuk memberi keutamaan kepada StRS berbanding SRS tidak seharusnya digeneralisasikan secara mutlak kepada semua jenis kajian akademik. Kebolehan StRS mengatasi SRS adalah bersyarat, ia hanya disyorkan apabila pemboleh ubah strata mempunyai korelasi teoretikal atau empirikal yang kuat dengan pemboleh ubah bersandar (pencapaian akademik) dan apabila maklumat strata tersebut boleh diperoleh secara kos-efektif (Lohr, 2022). Untuk populasi pelajar berskala sederhana dengan sifat homogeniti gred yang tinggi merentasi gender, SRS kekal sebagai alternatif yang paling praktikal dan kukuh.

Kajian ini mempunyai beberapa keterbatasan penting yang perlu diambil kira semasa mentafsirkan hasil dapatan. Pertama, analisis hanya tertumpu kepada saiz populasi yang sederhana (N = 395) dalam satu institusi pendidikan tertentu yang mengehadkan kebolehgeneralisasian dapatan kepada populasi pendidikan berskala makro. Kedua, pengstrataan dihadkan kepada satu pemboleh ubah demografi sahaja iaitu jantina yang secara empirikal didapati kurang mempunyai perkaitan varians yang kuat dengan pencapaian gred Matematik pelajar. Akhir sekali, fokus kajian ini adalah terhad kepada gred pencapaian bagi mata pelajaran Matematik sahaja, di mana struktur varians prestasinya mungkin berbeza berbanding subjek akademik yang lain. Berdasarkan prinsip teori pensampelan, pengstrataan hanya akan memberikan peningkatan kecekapan yang ketara jika terdapat perbezaan min yang besar antara strata. Dalam kes gred Matematik (G3) populasi ini, didapati min dan varians antara pelajar lelaki dan perempuan adalah hampir homogen. Oleh itu, jantina didapati kurang sesuai untuk dijadikan sebagai pemboleh ubah strata utama sekiranya matlamat penyelidikan adalah untuk memaksimumkan kecekapan anggaran. Bagi kajian masa

hadapan, pemboleh ubah yang lebih sensitif terhadap prestasi akademik seperti gred pentaksiran awal (G1 atau G2) atau pemboleh ubah masa pembelajaran mingguan (*studytime*) adalah dicadangkan kerana ia mempunyai korelasi langsung yang lebih kuat dengan gred akhir, sekali gus berpotensi menghasilkan kecekapan relatif (RE) yang jauh lebih tinggi.

Rujukan

[1] Chaudhuri, A., and S. Pal. 2022. A Comprehensive Textbook on Sample Surveys. Springer.
[2] Cochran, W. G. 1977. Sampling Techniques. 3rd ed. John Wiley & Sons.
[3] Food and Agriculture Organization of the United Nations (FAO). 2020. Use of Sampling in the Census. FAO Statistics Division.
[4] Kaggle. n.d. "Student Performance Dataset." Kaggle.
[5] Kim, J. K. 2025. Statistics in Survey Sampling. Chapman & Hall/CRC. https://doi.org/10.1201/9781003606147.
[6] Lakhre, A., and A. Mishra. 2024. "A Review on Statistical Importance of Simple Random Sampling Techniques." International Journal of Multidisciplinary Innovation and Research Methodology.
[7] Lohr, S. L. 2022. Sampling: Design and Analysis. 3rd ed. CRC Press. https://doi.org/10.1201/9780429298899.
[8] Noor, S., O. Tajik, and J. Golzar. 2022. "Simple Random Sampling." International Journal of Education & Language Studies 1 (2): 78–82. https://doi.org/10.22034/ijels.2022.162982.
[9] Ramya, G. R. V., et al. 2024. "Optimizing Variance Estimation in Stratified Random Sampling through a Log-Type Estimator for Finite Populations." Symmetry 16 (5): 540.
[10] Särndal, C.-E., B. Swensson, and J. Wretman. 2003. Model Assisted Survey Sampling. Springer.
[11] Sarker, M., and M. A. Al-Muaalemi. 2022. "Sampling Techniques for Quantitative Research." In Principles of Social Research Methodology, edited by M. R. Islam et al., 221–234. Springer. https://doi.org/10.1007/978-981-19-5441-2_15.
[12] Scheaffer, R. L., W. Mendenhall, L. Ott, and K. Gerow. 1990. Elementary Survey Sampling. Boston, MA: PWS-Kent.
[13] Simkus, J. 2023. "What Is Simple Random Sampling?" Simply Psychology.
[14] Szeitl, B., and T. Rudas. 2022. "Reducing Variance with Sample Allocation Based on Expected Response Rates in Stratified Sample Designs." Journal of Survey Statistics and Methodology 10 (4): 1107–1120. https://doi.org/10.1093/jssam/smab021.
[15] Thomas, L. 2020. "Simple Random Sampling: Definition, Steps & Examples." Scribbr.
[16] Triveni, G. R. V., F. Danish, and O. Albalawi. 2024. "Advancing Survey Sampling Efficiency under Stratified Random Sampling and Post Stratification." Symmetry 16 (5): 604.
[17] Triveni, G. R. V., F. Danish, and K. Tawiah. 2024. "Evolving Techniques for Enhanced Estimation: A Comprehensive Survey of Stratified Sampling and Post Stratification Methods." AIP Advances 14 (10): 100703.