

BAB I PENDAHULUAN

1.1 Latar Belakang

Informasi karakteristik populasi sering diperlukan untuk pengambilan kebijakan. Cara yang cepat dan murah adalah dengan memilih sebagian populasi untuk diukur. Levy & Lemeshow (2013) mendefinisikan proses tersebut sebagai survei sampel, lebih rinci mereka memaparkan survei sampel sebagai penelitian yang melibatkan beberapa individu dari satu populasi. Dan sebaliknya, penelitian yang melibatkan seluruh anggota populasi disebut sebagai survei sensus. Cochran (1977) menyatakan kelebihan survei sampel adalah hemat, cepat, jangkauan luas, dan akurat. Latpate, dkk. (2021) juga menyatakan pada era dewasa ini, keekonomisan pembiayaan pengukuran menjadi satu pertimbangan dalam memilih metode sampling. Sementara sensus membutuhkan waktu yang lama dan biaya yang besar (Lohr, 2022; Thompson, 2012; Hansen, Hurwitz, & Madow, 1953), sedangkan informasi dari populasi seringkali diperlukan cepat. Karena hal tersebut, survei sampel memiliki kelebihan dibandingkan dengan sensus karena relatif efektif dan efisien dari segi waktu dan biaya. Selain itu, Lohr (2022) menyatakan pengambilan data melalui survei survei lebih terkontrol dibandingkan dengan sensus yang biasanya melibatkan banyak orang sehingga sulit untuk diawasi. Walaupun demikian, penggunaan sampel bukan tanpa resiko, kesalahan penggunaan metode sampel akan menyedatkan pengambilan keputusan (Lohr, 2022). Di sisi lain, tidak setiap pengambilan karakteristik populasi harus menggunakan sampel, pada kondisi anggota populasi kecil, pengambilan karakteristik melalui sensus lebih masuk akal karena semua anggota populasi diukur. Selain itu, kelebihan penggunaan sensus adalah tidak adanya sampling eror yang selalu muncul pada sampel (Lohr, 2022).

Sebuah contoh klasik kesalahan dalam metode sampling terjadi pada survei poling pemilihan presiden AS oleh majalah *Literary Digest* pada tahun 1936. Survei tersebut memprediksi kemenangan Landon, akan tetapi pemilihan presiden ternyata dimenangkan oleh Roosevelt (Little, 2016). Studi lebih lanjut, seperti Lusinchi (2012) yang menjelaskan alasan utama kesalahan survei karena rendahnya tingkat respon survei, yaitu hanya 24%, sehingga mungkin saja mayoritas pemilih yang mendukung Roosevelt tidak terekam pada poling tersebut. Lohr & Brick (2017) mensimulasikan kembali data survei dengan menggunakan bobot sampel, hasilnya menunjukkan kemenangan Roosevelt, namun Lohr & Brick (2017) memberi catatan bahwa tingkat respon survei poling *Literary Digest* relatif rendah, sehingga bias estimasi survei tersebut masih tinggi.

Meskipun survei sampel mengandung resiko, metode ini tetap menjadi pilihan utama pada studi-studi internasional. Studi-studi internasional seperti PISA, TIMMS, dan PIRLS, menggunakan sampel survei dengan metode sampling acak bertingkat untuk pemilihan sampel (OECD, 2024; Almaskut, LaRoche, & Foy, 2023a; Mang, Küchenhoff, Meinck, & Prenzel, 2021; LaRoche, Joncas, & Foy, 2020). Di tingkat nasional, Asesmen Nasional (AN) yang diselenggarakan oleh Kemendikbudristek juga menggunakan metode sampel tertentu untuk pemilihan sampel. AN yang diselenggarakan sejak 2021 adalah salah satu bentuk evaluasi sistem pendidikan oleh Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) pada jenjang pendidikan dasar dan menengah jalur formal dan non formal (Mendikbudristek, Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi RI no. 71 th 2021 tentang Asesmen Nasional, 2021). Komponen AN terdiri dari Asesmen Kompetensi Minimal (AKM), Survei Karakter, dan Survei Lingkungan Belajar (Mendikbudristek, Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi RI no. 71 th 2021 tentang Asesmen Nasional, 2021).

Peserta AN mencakup seluruh kepala sekolah, seluruh guru, dan perwakilan siswa kelas 5, 8, dan 11 dari setiap sekolah seluruh Indonesia yang dipilih secara acak (Kepala-BSKAP, 2024). Berdasarkan Pedoman Penyelenggaraan AN tahun 2024, jumlah siswa yang dipilih secara acak adalah 30 siswa untuk jenjang SD dan 45 siswa untuk jenjang SMP, SMA, dan SMK (Kepala-BSKAP, 2024). Pemilihan peserta AN berada dibawah tanggung jawab Pusat Asesmen Pendidikan (Pusmendik), salah satu unit kerja di Kemendikbudristek. Berdasarkan Petunjuk Teknis AN 2024 (Pusmendik, 2024c) Pusmendik berwenang mengembangkan sistem pendataan yang salah satu fungsinya adalah melakukan sampling peserta AN. Satu alasan pengambilan sampel pada AN sebagaimana dinyatakan oleh Kemendikbudristek adalah untuk menghemat biaya penyelenggaraan (Putra, 2020).

Pusmendik mengembangkan metode sampel AN berdasarkan teori sampling acak bertingkat (*stratified random sampling*) (Pusmendik, 2024a). Metode sampling acak bertingkat adalah salah satu tipe probabilitas sampling dimana populasi dibagi menjadi lapisan-lapisan yang disebut strata. Dari setiap strata, sampel dipilih secara acak menggunakan metode sampling acak sederhana (*simple random sampling*) (Lohr, 2022; Wu & Thompson, 2020; Taherdoost, 2016; Levy & Lemeshow, 2013). Dengan penggunaan metode probabilitas sampling memungkinkan semua anggota populasi memiliki probabilitas yang diketahui untuk terpilih sebagai sampel (Lohr, 2022; Wu & Thompson, 2020; Taherdoost, 2016; Levy & Lemeshow, 2013). Lebih jauh, Wu & Thompson (2020) menegaskan pentingnya probabilitas sampling yang telah menjadi pondasi teori dan praktek sampling modern.

Pemilihan metode sampling acak bertingkat oleh Pusmendik memiliki beberapa alasan mendasar yang dijabarkan melalui pernyataan Lohr (2022) dan Levy & Lemeshow (2013) terkait kelebihan metode sampling acak bertingkat:

- Tingkat keakuratan estimasi yang lebih baik dibandingkan dengan metode lain.
- Semua subpopulasi yang dikenali terwakili dalam sampel.
- Kemampuan menghitung tingkat keakuratan pengukuran subpopulasi.
- Pengadministrasian yang efisien karena memungkinkan pengambilan sampel sedikit atau pengadministrasian sampel yang berbeda untuk setiap strata.

Walaupun demikian, metode ini juga memiliki beberapa kelemahan. Berndt (2020) dan Sharma (2017) menjelaskan beberapa tantangan utama dari penggunaan metode sampling acak bertingkat:

- pemilihan dan pendefinisian strata memegang peranan yang penting karena mengasumsikan pengelompokan populasi tersebut penting untuk dikenali. Kesalahan pemilihan strata dapat menyebabkan karakteristik penting populasi tidak terukur.
- Tidak dapat diterapkan untuk semua populasi, terutama pada populasi yang memiliki perbedaan yang samar diantara subpopulasinya.
- Ketidakseimbangan varian antar strata dapat menyebabkan permasalahan tersendiri.
- Tidak dapat diterapkan pada data yang telah diambil yang diambil tidak menggunakan metode acak bertingkat.

Selain tantangan metodologis, laporan monitoring pelaksanaan AN menunjukkan adanya ketidakpuasan atas penggunaan metode sampling AN dari pihak sekolah. Beberapa sekolah merasa dirugikan karena siswa yang terpilih terdiri dari siswa yang memiliki kemampuan akademik rendah sehingga dianggap tidak merepresentasikan sekolah (Pusmendik, 2024b)

Setelah tiga tahun pelaksanaan, metode sampling AN perlu dikaji ulang untuk menilai keakuratan metode sampling tersebut dengan membandingkannya dengan metode sampling acak sederhana karena berdasarkan data dari Pusmendik, proporsi sampel AN th 2022 terhadap populasi relatif besar. Untuk jenjang SD proporsi sampel terhadap populasi sebesar 73,65%, didapat dari jumlah sampel AN jenjang SD sebesar 3.416.583 dibagi dengan 4.639.078 populasi siswa kelas 5. Sedangkan untuk SMP, proporsi sampel terhadap populasi sebesar 44,74%, didapat dari sampel siswa SMP sebanyak 1.908.413 siswa dibagi dengan populasi 4.262.615 siswa kelas 8. Persentase ini jauh lebih besar dibandingkan dengan persentase sampling Indonesia pada PISA 2022 yang hanya sekitar 0,4% (OECD, 2023) dengan rata-rata proporsi sampel – populasi sekitar 16%. Dimana pada Studi PISA 2022, sampel siswa

Indonesia yang diambil sebanyak 13.439 siswa dari populasi 3.790.846 siswa PISA. Pada survei PISA 2022, proporsi sampling terhadap populasi yang terkecil adalah Amerika, yaitu sebesar 0,1%, dengan sampel siswa sebanyak 4.552 siswa dari populasi 3.661.328 siswa. Dan proporsi terbesar adalah Makao dengan sampel siswa sebesar 4.384 siswa dan populasi sebesar 4.423 siswa.

Studi tentang keakuratan pada metode sampling AN ini penting karena sebagaimana dipaparkan oleh Lohr (2022) yang menjelaskan bahwa konsep keakuratan berkaitan dengan konsistensi dan ketepatan metode sampling. Konsistensi dan ketepatan hasil pengukuran hasil metode sampling sangat penting untuk digunakan bersamaan karena metode sampling yang konsisten atau tidak bias belum tentu tepat sasaran, dan sebaliknya metode sampling yang tepat mungkin saja mengandung bias yang besar (Lohr, 2022). Untuk menyelidiki keakuratan hasil sampel, penelitian ini menggunakan dua pendekatan, yaitu dengan menggunakan sampling eror dan kuadrat rerata eror atau MSE (*Mean Square Error*). Sampling eror menghitung kadar kesalahan sampling dan kuadrat rerata eror mengestimasi seberapa jauh ukuran yang diperoleh menyimpang dari varian populasi (Lohr, 2022; Levy & Lemeshow, 2013). Pendekatan ini memungkinkan evaluasi yang komprehensif terhadap metode sampling AN dan memastikan hasil AN akan tetap representatif dan dapat diandalkan. Sehingga lebih jauh, penelitian ini dapat memberi manfaat berupa metode sampling yang akurat yang digunakan dan diterapkan secara nasional pada sampling AN.

1.2 Pembatasan Penelitian

Penelitian ini berfokus pada metode sampling yang digunakan untuk pemilihan sampel dan tidak melebar pada diskusi mengenai jumlah sampel per sekolah yang telah ditetapkan melalui Pedoman Penyelenggaraan AN (Kepala-BSKAP, 2024). Penelitian ini membatasi diri pada sampel jenjang SD dan SMP pada pelaksanaan AN tahun 2022.

1.3 Pertanyaan penelitian

Pertanyaan penelitian ini adalah sebagai berikut:

- Apakah terdapat perbedaan ketepatan metode sampling acak bertingkat dan metode sampling acak sederhana dalam memprediksi parameter populasi AN?
- Apakah terdapat perbedaan konsistensi metode sampling acak bertingkat dan metode sampling acak sederhana dalam memprediksi parameter populasi AN?

1.4 Tujuan penelitian

Tujuan Penelitian ini adalah mengevaluasi ketepatan dan konsistensi metode sampling acak bertingkat dibandingkan dengan metode sampling acak sederhana.

1.5 State of The Art

Metode sampling acak bertingkat telah digunakan secara luas pada asesmen berskala besar seperti PISA, TIMMS, dan PIRLS (OECD, PISA 2022 Technical Report, 2024; Almaskut, LaRoche, & Foy, 2023a; Mang, Küchenhoff, Meinck, & Prenzel, 2021; LaRoche, Joncas, & Foy, 2020). Namun dalam konteks AN, belum pernah dilakukan pengujian keakuratan penggunaan sampling acak bertingkat dalam proses pemilihan sampel.

Beberapa penelitian telah membandingkan metode sampling untuk menyelidiki tingkat akurasi, seperti studi yang dilakukan oleh Saini & Kumar (2018). Mereka menggunakan data simulasi untuk membandingkan keakuratan sampling acak bertingkat dan sampling set peringkat bertingkat (*stratified ranked set sampling*) dalam memprediksi parameter populasi. Hasilnya menunjukkan bahwa sampling set peringkat bertingkat lebih akurat. Secara ringkas sampling set peringkat bertingkat adalah penerapan sampling set peringkat (*ranked set sampling*) pada sampling acak bertingkat. Sampling set peringkat adalah metode sampling yang dikenalkan oleh McIntyre (dalam (Saini & Kumar, 2018)). Pada metode ini, kerangka sampel dikelompokkan dalam beberapa set berdasarkan kriteria pengukuran yang mudah diperoleh. selanjutnya pada setiap set, sampel dipilih menggunakan metode sampling acak sederhana. Walaupun demikian, metode ini belum populer dan tidak digunakan pada survei-survei internasional seperti PISA, TIMMS, dan PIRLS kemungkinan karena konsepnya yang mirip dengan strata implisit yang telah digunakan pada survei-survei tersebut (OECD, 2024; LaRoche, Joncas, & Foy, 2020; Almaskut, LaRoche, & Foy, 2023a). Pada studi lain yang dilakukan oleh Bujang dkk. (2018), menggunakan data populasi empiris untuk meneliti ukuran sampel terkecil yang menghasilkan ukuran statistik yang mirip dengan parameter populasi. Penelitian Bujang dkk. (2018) tersebut mencoba menunjukkan berapa besar ukuran sampel terkecil yang diperlukan dari sebuah studi empiris. Mereka menyimpulkan ukuran sampel minimal 500 subyek diperlukan untuk menggambarkan populasi berukuran 1.500 subyek. Berbeda dengan penelitian-penelitian sejenis yang telah dijabarkan, penelitian ini memiliki keunikan karena berfokus pada penerapan metode sampling acak bertingkat pada AN dan menggunakan data empiris yang belum pernah digunakan pada penelitian-penelitian lainnya.

1.6 Road Map Penelitian

Penelitian ini memerlukan tahapan-tahapan pelaksanaan yang dijabarkan pada tabel 1.1. Dari tabel tersebut terlihat terdapat tujuh tahapan penelitian ini, yaitu studi kepustakaan, penyiapan data, analisa data, intepretasi analisa, penarikan kesimpulan, implikasi keilmuan dan kebijakan, dan terakhir penulisan laporan. Pada tahap studi kepustakaan, dilakukan penelusuran dan pengkajian penelitian-penelitian sebelumnya pada topik yang sama. Pada tahap penyiapan data, peneliti meminta data kepada pihak ketiga, yaitu Pusmendik, hal ini karena penelitian ini menggunakan data sekunder yang dikumpulkan oleh Pusmendik. Pada tahap ketiga, yaitu analisa data, dilakuan analisa seperti analisa deskriptif dan uji hipotesis. Pada tahap keempat dilakukan pengintepretasian hasil analisa, kemudian dilakukan penarikan kesimpulan yang merupakan tahap selanjutnya dari rangkaian tahapan penelitian. Pada tahap keenam dirumuskan implikasi keilmuan dan kebijakan, kemudian pada tahap terakhir disusun laporan penelitian yang merangkum semua langkah-langkah penelitian.

s

Tabel 1.1 Peta Jalan Penelitian

No	Tahapan Penelitian	Minggu																					
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	Studi kepustakaan	■	■	■	■																		
2	Penyiapan data					■	■	■	■														
3	Analisa data									■	■	■	■										
4	Intepretasi analisa												■	■	■	■	■						
5	Penarikan kesimpulan																	■	■	■	■	■	
6	Implikasi keilmuan dan kebijakan																					■	■
7	Penulisan laporan																						■