

## TEKNIK PENARIKAN SAMPEL

Karena polling ingin membuat generalisasi agar hasilnya dapat menggambarkan pendapat publik secara keseluruhan, maka sampel yang harus diambil adalah sampel acak (random sampling/probability sampling). Sampling probabilitas pada intinya berbicara tentang peluang terjadinya peristiwa kebetulan. Probabilitas diterapkan agar setiap anggota individu memiliki peluang yang sama besarnya untuk terpilih menjadi sampel. Di bawah ini akan diuraikan berbagai teknik pengambilan sampel.

**Sampel acak sederhana (*Simple Random Sampling*).** Teknik pengambilan sampel ini memastikan setiap unsur mempunyai peluang yang sama untuk dijadikan sampel. Peluang yang sama berarti setiap unsur mempunyai probabilitas yang sama untuk dijadikan sampel. Contohnya dalam populasi mahasiswa UGM sebanyak 24.600 mahasiswa berarti setiap mahasiswa mempunyai  $1/24.600$  kesempatan untuk terpilih sebagai sampel. Kalau kita mengambil sampel sebanyak 500 responden maka kesempatan seseorang untuk dipilih sebagai sampel adalah  $500/24.600 = 1/49$ . Angka ini sering disebut sebagai sampling fraction.

Pemakaian metode sampel acak sederhana perlu memenuhi beberapa syarat : (1) Harus tersedia kerangka sampel. Kalau kerangka sampel itu belum tersedia, harus dibuat terlebih dahulu; (2) Sifat populasi homogen dan keadaan populasi tidak terlalu tersebar secara geografis.

Ada dua cara untuk mengambil sampel acak sederhana. Cara pertama adalah dengan jalan mengundi. Misalnya seorang peneliti ingin mengetahui bagaimana tanggapan mahasiswa UGM terhadap keberadaan Dewan Mahasiswa. Peneliti ingin mengetahui apakah lahirnya Dewan Mahasiswa yang menyempal dari organisasi kemahasiswaan seperti SMPT yang sudah ditetapkan oleh Dikti Depdikbud itu mendapat persetujuan dari mahasiswa. Mahasiswa UGM dipilih karena di sinilah pertama kali ide terbentuknya Dewan Mahasiswa. Ada 24.600 mahasiswa UGM tingkat sarjana dan semuanya kita masukkan sebagai populasi yang akan kita ambil sampelnya, dan dari populasi itu dipilih 500 sampel. Pertama kali peneliti perlu membuat kerangka sampel yang terdiri atas daftar nama seluruh mahasiswa. Kemudian ke-24.600 mahasiswa tersebut diberi nomor masing-masing, setiap mahasiswa mendapat satu nomor. Kemudian 500 nomor dipilih secara acak dengan jalan diundi sehingga setiap individu mempunyai peluang yang sama untuk terpilih. Nama yang terundi dicatat sampai terpilih 500 orang mahasiswa yang menjadi sampel kita. Selain dengan undian dapat juga dengan mengaduk. Masing-masing nama mendapat satu nomor dan ditulis di atas lembaran kertas. Selanjutnya tiap nomor ditulis pada selembar kertas kecil yang kemudian dilipat-lipat. Semua, ke 24.600 lipatan kertas kecil tadi dimasukkan dalam satu wadah untuk diaduk, dan dari adukan tadi diambil secara buta 500 lipatan kertas. Nomor yang terpilih itu yang akan menjadi sampel penelitian. Cara apapun yang kita gunakan, 500 nama tadi dicocokkan dengan daftar nama seluruh mahasiswa UGM yang sudah kita buat.

Di lihat dari namanya, seharusnya metode ini yang paling sederhana, tetapi dalam kenyataannya rumit. Kesulitan utamanya terletak pada kerangka sampel, yang seringkali tidak tersedia dengan baik. Dalam contoh di atas kerangka sampel tersedia baik, sebab daftar nama mahasiswa ada di setiap Perguruan Tinggi. Tetapi bayangkan polling dengan populasi yang besar seperti masyarakat Yogya. Catatan-catatan penduduk seringkali tidak tersedia, walaupun tersedia peneliti harus menghimpunnya dan menyusunnya kembali dari data penduduk tiap kelurahan.

**Sampel acak sistematis (*systematic sampling*).** Sampel sistematis adalah cara yang lebih sederhana untuk mengambil sampel jikalau tersedia sebuah daftar populasi dengan urutan

tertentu. Pengambilan sampel sistematis adalah suatu metode dimana hanya unsur pertama saja dari sampel dipilih secara acak sedangkan unsur-unsur selanjutnya dipilih secara sistematis menurut suatu pola tertentu. Metode ini dijalankan apabila ada dua keadaan: (1) Apabila nama atau identifikasi dari individu dalam populasi itu terdapat dalam suatu daftar sehingga satuan-satuan tersebut dapat diberi nomor urut; (2) Apabila populasi tersebut mempunyai pola berurutan, seperti urut abjad dan sebagainya.

Kerangka sampel dari sampel sistematis harus sudah tersedia dengan baik. Daftar nama mahasiswa adalah daftar yang dapat dipakai sebagai kerangka sampel pengambilan acak sistematis. Daftar-daftar serupa ada berbagai macam yang dapat dipergunakan untuk berbagai keperluan seperti daftar tagihan pajak, daftar rekening telepon, pelanggan listrik, PDAM, daftar pemilih pemilu, daftar anggota organisasi, daftar tahanan dalam buku lembaga pemasyarakatan, daftar pegawai/pekerja dan sebagainya. Buku Petunjuk Telepon adalah suatu daftar yang dapat dipakai sebagai kerangka sampel untuk acak sistematis karena tersusun dalam suatu urutan tertentu, yakni abjad. Daftar ini aspek kebaruannya juga lebih terjamin karena diperbaharui tiap tahun.

Yang menjadi masalah adalah dalam topik polling apa seharusnya daftar tersebut dipakai? Bagaimana kita menggunakan daftar itu karena daftar itu amat spesifik. Bisakah kita memakai buku tagihan pajak atau langganan PDAM untuk menggambarkan masyarakat Yogyakarta? Apakah semua anggota populasi masyarakat Yogyakarta terdaftar dalam daftar tersebut? Daftar nama mahasiswa tentu saja adalah informasi yang valid dari semua mahasiswa. Tetapi daftar tagihan pajak Yogyakarta, atau langganan telepon/PDAM, tentu saja tidak menunjukkan populasi masyarakat Yogyakarta secara keseluruhan. Daftar-daftar tersebut memang bisa dipakai sebagai kerangka sampel tetapi penggunaannya harus hati-hati. Misalnya buku petunjuk telepon Yogyakarta tidak menunjukkan populasi masyarakat Yogyakarta, tetapi kelas menengah atau kalangan terdidik masyarakat Yogyakarta.

Cara penggunaan sampel acak sistematis adalah sebagai berikut: misalnya, jumlah satuan-satuan elementer dalam populasi itu adalah  $N$  dan besar sampel yang akan diambil adalah  $n$ , maka bagi hasil itu dinamakan interval sampling yang biasa diberi kode  $k$ . Unsur pertama dalam sampel lalu dipilih secara acak di antara individu bernomor urut  $I$ , dan satuan bernomor urut  $k$  dari populasi. Andaikan yang terpilih itu adalah individu bernomor urut  $s$ , maka unsur-unsur selanjutnya dalam sampel dapat ditentukan yaitu: Sampel pertama :  $s$  Sampel kedua :  $s + k$  Sampel ketiga :  $s + 2k$  dan seterusnya.

Sebagai contoh, kita menginginkan mengambil 300 nama dari populasi sejumlah 900 orang. Setelah menyeleksi secara acak pada langkah pertama, kita menyeleksi setiap tiga nama sampai 900 sehingga terpilih 300 nama. Interval sampel kita disini adalah 3, interval sampel mudah dihitung. Interval adalah kebalikan dari rasio sampel. Rasio sampel untuk 300 sampel adalah  $300/900 = 0,3$  (33%). Interval sampling adalah  $900/300 = 3$ .

Pemakaian sampel sistematis ini relatif sederhana. Tetapi pemakaian sampel ini membutuhkan syarat tertentu, apakah daftar tersebut dirancang dalam urutan tertentu ataukah tidak. Tetapi kelemahan utama sampel ini, sebagaimana dikatakan oleh Blalock justru karena sifat populasi yang berurutan itu. Menurut Blalock, apabila daftar unsur tersebut disusun berdasarkan abjad, akan terjadi penyimpangan yang berhubungan dengan representasi unsur dalam sampel yang berlebihan/kurang. Misalnya nama yang dimulai huruf A, M akan mempunyai kecenderungan untuk dimunculkan secara berlebihan dibandingkan nama yang dimulai huruf Z, atau Q. Kalau sebuah daftar nama mahasiswa berurutan menurut tahun masuk, bisa jadi mereka yang baru masuk (mahasiswa baru) mempunyai kesempatan lebih besar untuk

dipilih dibandingkan angkatan lama.

**Sampel acak stratifikasi proporsional.** Sampel acak sederhana dipakai apabila populasi homogen, tetapi tidak bisa dipakai untuk populasi heterogen berbeda dalam hal karakteristik populasi seperti tingkat pendidikan atau tingkat penghasilan. Makin heterogen suatu populasi makin besar pula perbedaan sifat antara lapisan-lapisan tersebut. Presisi dan hasil yang dapat dicapai dalam penggunaan suatu metode sampel antara lain dipengaruhi oleh derajat keseragaman populasi yang bersangkutan. Untuk dapat menggambarkan secara tepat mengenai sifat-sifat populasi yang heterogen, maka populasi yang bersangkutan harus dibagi terlebih dahulu dalam lapisan (strata) yang seragam dan dari setiap lapisan itu baru diambil sampel secara acak.

Ada tiga persyaratan yang harus dipenuhi apabila sampel acak stratifikasi proposional dipakai: (1) Harus ada kriteria yang jelas yang dipergunakan sebagai dasar untuk menstratifikasi populasi ini dalam lapisan-lapisan. Memang banyak sekali karakteristik populasi dalam pengambilan unsur sampel, peneliti harus menentukan dimensi yang terpenting dan relevan untuk tujuan polling.

(2) Harus ada data pendahuluan mengenai strata populasi. (3) Harus diketahui dengan tepat jumlah satuan-satuan elementer dari tiap lapisan (strata) dalam populasi itu. Misalnya kalau populasi mahasiswa UGM tadi mau distratifikasi menurut fakultas dengan asumsi fakultas sosial mempunyai derajat ketertarikan politik yang lebih tinggi dibandingkan fakultas eksakta. Harus ada data pendahuluan, berapa jumlah mahasiswa sosial, berapa jumlah mahasiswa eksakta. Peneliti juga perlu membuat kerangka nama-nama mahasiswa sosial dan eksakta sebagai dasar dimana sampel itu akan diambil.

Cara melakukan sampel acak stratifikasi proporsional ini sebagai berikut. Setelah daftar kerangka itu ditetapkan maka responden dibagi menurut stratanya masing-masing. Lalu responden diambil sesuai dengan proporsinya dalam populasi dengan perbandingan tertentu. Proporsi yang terbesar tentu saja mendapat sampel terbesar, sedangkan proporsi kecil akan mendapatkan sampel yang kecil juga. Dengan cara penarikan sampel ini lebih menjamin keadilan dan lebih mencerminkan representasi dari populasi. Tetapi sampel acak proporsional ini mempunyai kelemahan, diantaranya membutuhkan pengetahuan tentang komposisi populasi dan penyebaran karakteristik populasi sebelum diambil unsurunsur sampel. Karena itu metode ini membutuhkan waktu yang lama dalam menyusun kerangka sampel. Dalam contoh kasus di atas, peneliti tidak hanya perlu tahu daftar nama mahasiswa, tetapi juga perlu tahu mahasiswa menurut fakultasnya masing-masing sebagai kriteria dasar pengelompokkan.

Sampel stratifikasi adalah modifikasi dari sampel acak sederhana dan sampel sistematis yang didesain untuk menghasilkan sampel yang lebih representatif dan lebih akurat. Kenapa? Untuk mendapatkan sampel representatif, proporsi representatif dari sejumlah kelompok dalam sampel haruslah sama proporsinya dalam populasi. Sebagai contoh kalau kita mengambil sampel acak sederhana atau sistematis, kita barangkali akan mendapatkan lebih banyak orang dari kelas menengah, atau lebih banyak orang yang berumur muda. Untuk polling dengan tema tertentu, jika karakteristik dalam sampel tidak representatif bisa jadi akan menimbulkan kesalahan. Sebagai contoh, dalam polling mengenai perilaku pemilih, kalau pengambilan sampel tidak diambil berdasarkan proporsi dari populasi, kita mungkin akan mendapatkan sampel orang berumur muda dibandingkan tua. Akibatnya, kita tidak cukup punya gambaran yang lebih akurat mengenai perilaku pemilih dari semua karakteristik umur. Sampel stratifikasi menolong mengatasi masalah ini. Dalam sampel stratifikasi peneliti pertama kali membagi populasi ke dalam sub-populasi (strata) sebagai dasar informasi. Setelah populasi terbagi ke dalam strata-

strata, peneliti mengambil sampel dengan cara acak, dapat acak sederhana atau acak sistematis. Sebagai ilustrasi kita akan menggunakan sampel acak stratifikasi untuk mengambil sampel dari populasi mahasiswa Universitas Gadjah Mada.

Salah satu syarat utama dari penarikan sampel acak stratifikasi adalah kita harus mengetahui proporsi dari strata pada populasi yang menjadi dasar dalam menentukan strata pada sampel. Sehingga strata sampel akan sama dengan strata populasi. Kalau pengetahuan mengenai populasi sedikit, pemakaian sampel stratifikasi justru akan menyebabkan kekeliruan. Pengetahuan mengenai strata populasi ini bisa didapatkan dari data-data sekunder.

**Sampel acak stratifikasi tidak proporsional.** Sampel acak stratifikasi tidak proporsional idenya hampir sama dengan proporsional. Perbedaannya terletak pada strata dalam sampel tidak sama dengan proporsi strata dalam populasi. Bisa saja dalam strata disajikan lebih, sementara strata lain disajikan kurang. Dengan menggunakan strata Status Sosial Ekonomi (SSE) sebagai contoh, peneliti misalnya membagi populasi dalam tiga strata yaitu atas, menengah, bawah. Pembagian persentase yang akan diteliti: kelas atas 30%, menengah 30% dan bawah 40%. Di sini peneliti memperbesar sampel kelas atas dan memperkecil sampel kelas bawah. Pengambilan sampel acak tidak proporsional berarti peneliti akan memberikan bobot yang sama untuk setiap strata dan ia akan memberikan bobot yang lebih untuk beberapa strata dan mengurangi bobot terhadap strata yang lain sesuai dengan penyebaran yang tidak proporsional dalam populasi.

Teknik pengambilan sampel ini kita pakai jikalau salah satu dari strata itu jumlahnya teramat kecil (sedikit) sehingga apabila dipakai strata proporsional, ada strata yang tidak terwakili dalam sampel. Banyak sekali proporsi populasi tidak seimbang dimana salah satu strata jumlahnya besar dan strata lain jumlahnya sangat kecil kepangkatan dalam militer, eselon pegawai negeri dan sebagainya.

Misalkan polling yang menanyakan pandangan anggota ABRI terhadap Dwi Fungsi ABRI yang memasukkan semua perwira ABRI dari perwira rendah sampai jenderal sebagai populasi. Karena jumlah jenderal dalam populasi sedikit, peneliti dapat memutuskan untuk memakai semua jenderal dalam sampel. Kalau memakai sistem proporsional, para jenderal tersebut bisa jadi hanya mendapat 1 sampel atau bahkan tidak sama sekali. Seperti juga pada contoh penarikan sampel anggota DPR hasil Pemilu 1997. Ada 500 anggota DPR, kalau sampel acak proporsional yang dipakai, maka FPDI mungkin tidak akan mendapat wakil. Hal ini disebabkan begitu sedikitnya jumlah anggota FPDI di DPR RI yaitu 11 orang, jumlah ini amat timpang dibandingkan dengan FKP yang mencapai 325 orang. Dengan demikian peneliti memberi bobot kelompok tertentu yang tidak muncul dalam populasi supaya menjadi lebih seimbang, terutama apabila terdapat perbedaan proporsi yang mencolok di antara berbagai strata dalam populasi asal.

**Sampel klaster.** Dalam berbagai teknik pengambilan sampel yang sudah diuraikan terdahulu, unit analisisnya adalah individu dan membutuhkan tersedianya kerangka sampel sebagai dasar pengambilan sampel. Padahal kita seringkali dihadapkan dengan kenyataan dimana kerangka sampel tidak tersedia, atau tidak memungkinkan dibuat karena membutuhkan waktu yang lama atau biaya yang sangat besar untuk membuatnya. Kalau kita meneliti pendapat mahasiswa UGM terhadap Dewan Mahasiswa, kerangka sampel tersedia dengan baik. Tetapi bagaimana kalau kita ingin meneliti pendapat mahasiswa Yogyakarta? Kerangka sampel yang berisi daftar nama seluruh mahasiswa Yogyakarta pasti tidak ada. Untuk menyusunnya dibutuhkan waktu yang sangat lama mengingat di Yogya ada 40 Perguruan Tinggi dengan 150.000 mahasiswa, belum termasuk akademi. Dibutuhkan waktu berbulan-bulan untuk menyusun kerangka sampel mahasiswa sebanyak itu. Untuk mengatasi masalah itu diterapkan teknik sampel klaster.

Unit tempat pertama kali klaster diambil adalah PSU (*Primary Sampling Unit*) dapat berupa organisasi, asosiasi, batas geografis dengan batasan yang jelas. Dalam sampel klaster, unit analisis dalam populasi digolongkan dalam gugusgugus yang disebut klaster yang merupakan satuan-satuan dari mana sampel akan diambil. Jumlah gugus yang diambil sebagai sampel harus acak. Lalu dari gugus terpilih, individu dalam gugus itu diambil secara acak. Dengan kata lain, peneliti secara acak mengambil sampel klaster, kemudian secara acak pula mengambil elemen dari dalam klaster yang telah diseleksi. Dalam sampel klaster tidak diperlukan daftar dari individu untuk kerangka sampel, tetapi cukup daftar gugus saja. Misalnya dalam polling dengan populasi mahasiswa Yogyakarta, semua Perguruan Tinggi merupakan kelompok/gugus yang disebut *Primary Sampling Unit* (PSU). Karena di Yogyakarta ada 40 Perguruan Tinggi maka ada 40 PSU. Kemudian dari PSU tersebut, ditarik sampel fraction tingkat pertama yang besarnya misalnya:  $10/40 \times 100\% = 25\%$ . Dengan perkataan lain jika diinginkan sebuah sampel yang ditarik secara acak dengan sampel fraction sebesar 25% maka besarnya sampel adalah 10 PSU atau 10 Perguruan Tinggi. Untuk menarik sampel Perguruan Tinggi ini dapat dipakai Daftar Perguruan Tinggi Yogyakarta yang telah tersedia di Kopertis Yogyakarta. Lalu dipilih secara acak 10 Perguruan Tinggi dari 40 Perguruan Tinggi tadi. Dalam memilih Perguruan Tinggi secara acak tersebut dapat diterapkan stratifikasi, misalnya membagi terlebih dahulu Perguruan Tinggi negeri dan swasta atau antara Universitas dan Institut dan sebagainya.

Tiap Perguruan Tinggi mempunyai mahasiswa yang jumlahnya berbedabeda. Mahasiswa yang merupakan sumber informasi dari polling adalah unit elementer dari PSU. Tidak semua mahasiswa dari 10 Perguruan Tinggi terpilih itu otomatis menjadi sampel. Artinya harus ditarik lagi sampel mahasiswa dari Perguruan Tinggi terpilih. Peneliti membuat kerangka sampel, sebuah daftar mahasiswa dari 10 Perguruan Tinggi tersebut ini jauh lebih ringan dibandingkan harus membuat kerangka sampel 40 Perguruan Tinggi. Lalu ditarik sampel dari tiap-tiap PSU terpilih dengan sampel fraction yang berimbang dengan jumlah anggota/unit elementer dalam tiap PSU. Dengan demikian ada dua tahap sampling yaitu sampling untuk memilih Perguruan Tinggi dan sampling untuk memilih mahasiswa. Teknik menarik sampel untuk memilih mahasiswa dari PSU terpilih sama dengan prinsip sampel acak yang lain.

Ada dua situasi dimana sampel klaster dipakai. Pertama, wilayah/area sampel tersebar amat luas sehingga untuk menyusun kerangka sampel amat sulit. Kalau kita ingin mengadakan polling pendapat umum masyarakat Jakarta, kerangka sampel berupa daftar nama penduduk tidak tersedia, walaupun ada membutuhkan waktu berbulan-bulan untuk mengumpulkannya dari tiap kelurahan. Dengan sampel klaster, kerangka sampel yang dipakai adalah daftar kelurahan wilayah Jakarta. Sampel yang kita ambil pertamakali adalah kelurahan, dari kelurahan diambil Rumah Tangga dan dari Rumah Tangga inilah individu diambil. Kedua, peneliti tidak mempunyai kerangka sampel yang baik dari populasi ataupun kalau ada harus dibuat dengan biaya yang mahal. Sebagai contoh tidak ada daftar kerangka sampel nama-nama dosen di Jakarta. Kita bisa membuat kerangka sampel secara akurat dengan melakukan survei, tetapi hal ini jelas membutuhkan biaya yang mahal. Biasanya populasi di atas 50.000 orang sudah sukar untuk menyusun kerangka sampel. Dengan sampel klaster tidak dibutuhkan survei seperti itu asalkan daftar nama perguruan tinggi negeri ataupun swasta seluruh Jakarta tersedia.

Seorang peneliti yang menggunakan sampel klaster harus memutuskan jumlah klaster dan jumlah individu (elemen) yang akan diambil dari setiap klaster. Sebagai contoh, dalam sampel klaster dua tahap yang mengambil 240 orang, apakah peneliti mengambil secara acak 120 klaster dan memilih 2 orang dari setiap klaster, ataukah secara acak mengambil 2 klaster dan memilih 120 orang dari tiap klaster? Yang terbaik adalah mengambil jumlah klaster lebih banyak, karena

elemen dalam kluster (seperti orang yang tinggal dalam satu kelurahan/RW yang sama, pelajar dalam satu sekolah, pekerja yang bekerja dalam perusahaan yang sama) relatif mempunyai karakteristik sama dibandingkan lainnya (seperti orang yang tinggal di lain RW, pelajar dalam SMA yang berbeda, pegawai di dua perusahaan yang berbeda). Jika hanya beberapa kluster saja yang dipilih, akan banyak elemen yang sama sehingga representasi dari total populasi secara keseluruhan menjadi berkurang.

Sampel kluster mempunyai kelebihan dalam hal efisiensi terutama menghemat waktu dan biaya. Biaya perjalanan dan waktu wawancara dapat dihemat sekecil mungkin. Pemakaian sampel kluster amat cocok dipakai dalam wilayah penelitian dimana sampel tersebar luas, tidak mengumpul dalam satu tempat. Tetapi biaya yang lebih murah ini diikuti oleh akurasi sampel yang lebih rendah jika dibandingkan dengan teknik sampel acak sederhana. Dalam sampel kluster unit elementer yang dipilih adakalanya berdekatan, sehingga informasi yang diberikan tidak cukup representatif dibandingkan dengan informasi dari unit elementer yang cukup berpencar seperti pada sampel acak. Dalam sampel acak sederhana, pengambilan sampel dari populasi akan menyebabkan sampling error sekali, tetapi sampel kluster dua tahap akan menyebabkan sampling error dua kali. Pertama, sampling error yang terjadi pada saat pengambilan PSU untuk menggambarkan populasi kluster. Kedua, sampling error yang terjadi ketika individu diseleksi dari kluster untuk menghasilkan individu yang representatif dari kluster.

**Sampel kluster proporsional (*Probability Proportionate to Size/PPS*).** Asumsi yang dipakai dalam penarikan sampel kluster adalah tiap kluster mempunyai elemen (individu) yang sama banyaknya dan sama homogenya. Padahal dalam kenyataannya tidak demikian. Dalam ilustrasi penarikan sampel kluster SMA, anggota kluster (jumlah pelajar dalam kluster satu SMA) relatif agak sama satu SMA rata-rata mempunyai jumlah siswa 400 orang. Tetapi seringkali terjadi kelompok kluster mempunyai elemen/jumlah elemen yang berbeda. Kalau hal ini terjadi, peneliti harus membuat perlakuan, agar probabilitas atau rasio sampling seimbang dalam beberapa langkah dalam sampling. Kita misalnya mensampel kelurahan. Pada satu kelurahan ada yang memiliki 60 RW (kelurahan besar) dan ada kelurahan lain yang hanya memiliki 6 RW (kelurahan kecil). Dalam sampel kluster, baik kelurahan yang memiliki RW banyak maupun sedikit diperlakukan sama. Ini menjadi masalah karena dengan demikian kesempatan seseorang untuk terpilih sebagai sampel berbeda, untuk kelurahan dengan jumlah RW kecil mempunyai kesempatan lebih besar. Contoh sama polling dengan populasi pekerja yang diambil secara kluster menurut perusahaan. Padahal antara satu perusahaan dengan perusahaan lain mempunyai jumlah pekerja yang berlainan tergantung pada besar kecilnya perusahaan.

Misalnya kita mengambil sampel 300 Perguruan Tinggi dari populasi 3.000 Perguruan Tinggi. Dengan metode ini berarti setiap kluster (Perguruan Tinggi) mempunyai kesempatan yang sama untuk terpilih sebagai sampel  $300/3.000$  atau 10 %. Tetapi tiap Perguruan Tinggi mempunyai jumlah mahasiswa yang berbeda, yang mengakibatkan tiap mahasiswa tidak mempunyai kesempatan yang sama terpilih sebagai sampel. Pada penarikan sampel tahap pertama (memilih Perguruan Tinggi), Perguruan Tinggi besar dengan jumlah mahasiswa 40.000 orang dan Perguruan Tinggi kecil dengan 400 mahasiswa mempunyai kesempatan yang sama terpilih sebagai sampel. Tetapi pada penarikan sampel tahap kedua (memilih mahasiswa dari PSU Perguruan Tinggi terpilih), kesempatan seorang mahasiswa untuk terpilih sebagai sampel berbeda. Untuk Perguruan Tinggi besar, kesempatan seorang mahasiswa untuk terpilih sebagai sampel adalah  $5/40.000 = 0,0125\%$ , sementara mahasiswa pada Perguruan Tinggi kecil mempunyai kesempatan  $5/4000 = 1,25\%$  terpilih sebagai sampel. Mahasiswa dari Perguruan Tinggi kecil memiliki kesempatan 100 kali lebih besar menjadi sampel. Dengan memakai sampel

klaster, tiap mahasiswa tidak mempunyai kesempatan yang sama untuk terpilih sebagai sampel.

Untuk mengatasi masalah ini, ada satu metode pengembangan dari sampel klaster yang dikenal sebagai *Probability Proportionate to Size* (PPS). Dengan pemakaian sampel ini setiap mahasiswa mempunyai kesempatan yang sama besarnya untuk terpilih sebagai sampel. Sesuai dengan namanya, setiap klaster mempunyai kesempatan yang proporsional sesuai dengan ukuran dari klaster itu. Inti dari metode ini adalah mengatur pemberian kesempatan lebih besar atau lebih kecil pada langkah pertama pengambilan sampel. Kita memberikan Perguruan Tinggi yang besar dengan lebih banyak mahasiswa kesempatan yang lebih besar untuk diseleksi sebagai sampel. Dasar dari pengaturan itu adalah proporsi dari semua mahasiswa dalam populasi yang akan diambil.

Teknik penarikan sampel mana yang dipilih? Teknik penarikan sampel mana yang akan kita pilih dalam polling tergantung dalam dua hal. Pertama, kemungkinan ada atau tidaknya kerangka sampel. Masalah utama dalam penarikan sampel acak adalah tidak semua populasi mempunyai kerangka sampel. Di samping itu, apabila populasi polling bukan masyarakat umum sehingga menyulitkan kita dalam menyusun kerangka sampel. Populasi yang khusus ini misalnya para veteran ABRI, janda pahlawan nasional, olahragawan, korban perkosaan, anak-anak gelandangan, penggemar otomotif, dan sebagainya. Populasi-populasi itu tidak ada catatan anggota populasi, sehingga peneliti perlu membuat kerangka sampel sendiri, dan pekerjaan ini membutuhkan waktu panjang dan dana yang besar.

Pertimbangan kedua, bagaimana karakteristik populasi tersebut. Apakah populasi cukup homogen atau heterogen, apakah populasi itu menyebar atau mengumpul. Untuk populasi yang homogen, cara yang terbaik adalah sampel acak sederhana, apabila tidak homogen teknik sampel terbaik adalah sampel acak stratifikasi. Apabila populasi mengumpul, cara terbaik untuk menarik sampel adalah dengan sampel acak sederhana/sistematis. Tetapi apabila populasi menyebar cara yang lebih efisien adalah dengan menggunakan sampel klaster.

## **B. BESAR SAMPEL**

Kita kembali dengan pengandaian Benyamin S. Memang benar kita tidak perlu memakan seluruh daging sapi, mencicipi sedikit saja daging sapi sudah dapat mengukur dengan tepat bagaimana rasa daging sapi. Karakteristik daging sapi sama sehingga dengan mengambil daging untuk sekedar mencicipi saja sudah dapat mewakili populasi tubuh sapi secara keseluruhan. Tetapi tidak demikian halnya dengan pendapat umum. Pendapat orang berbeda-beda, bahkan ada adagium setiap orang pada dasarnya mempunyai pendapat yang berbeda-beda. Kalau ada 24.600 orang itu berarti ada 24.600 pendapat. Pertanyaannya, berapa jumlah sampel yang harus kita ambil sehingga dengan jumlah sampel itu dapat menggambarkan pendapat 24.600 orang? Dalam contoh daging sapi, unsur dalam populasi betul-betul seragam sehingga pengujian sedikit saja sudah dapat mengukur semua populasi. Sementara unsur-unsur manusia amat beragam karakteristik etnik, latar belakang sosial, demografik amat beragam. Di sinilah kita bertemu dengan masalah jumlah sampel yang dibutuhkan dalam polling.

### **Ukuran Populasi Tidak Menentukan**

Perlu dipahami bahwa besarnya sampel tidak tergantung pada besar/ukuran populasi. Kita mengambil sampel 500 orang mahasiswa UGM yang jumlahnya 36.000 mahasiswa sama tingkat teliti dan akuratnya dengan sampel 500 orang dari seluruh penduduk Yogya yang jumlahnya 3,2 juta jiwa. Ketelitian kedua sampel tersebut sama yakni menggunakan sampel 500 orang, dan memperbandingkan penduduk Yogya yang berjumlah 3,2 juta jiwa dan mahasiswa UGM

berjumlah 24.600 tidaklah relevan. Karena tingkat ketelitian sebuah sampel diukur dari besar sampel dan bukan besar populasi. Dengan perkataan lain, populasi pada sebuah kota, sebuah daerah, sebuah propinsi mempunyai ukuran yang sama dan diperlakukan sama dalam sampling.

Ukuran dari suatu populasi tidak mempunyai hubungan dengan besar sampel, barangkali suatu hal yang susah dimengerti. Untuk mengilustrasikan hal ini dapat digambarkan demikian.1 Kita misalnya mempunyai sekotak kelereng, lalu kita ambil dari kotak itu sebuah kelereng yang berwarna biru. Kita akan membuat kesalahan apabila menarik kesimpulan bahwa semua kelereng dalam kotak itu berwarna biru. Jika kita mengambil segenggam kelereng sejumlah 10 butir kelereng, yang terdiri atas empat biru dan enam merah, kita akan lebih teliti dalam menebak bahwa kotak itu berisi lebih banyak kelereng berwarna merah daripada kelereng berwarna biru. Apabila kita menggunakan kedua genggam tangan secara bersama-sama dan kita mendapatkan 75 kelereng berwarna merah dan 25 kelereng berwarna biru, kita dapat lebih khusus lagi berkata bahwa kotak itu berisi 3 kelereng merah untuk setiap kelereng biru. Dengan kata lain, tidaklah penting kita mengambil kelereng dari kotak kecil atau besar, yang membuat estimasi warna kelereng jauh lebih teliti/presisi adalah ukuran tangan (besar sampel) dan tidak bergantung pada ukuran kotak dan banyaknya kelereng (ukuran populasi). Tetapi dengan asumsi bahwa kelereng dalam kotak itu bercampur dengan baik dan tangan itu merupakan sampel acak yang benar.

Sampel probabilitas sejumlah 1.000 orang yang diambil dari seluruh Indonesia tidak lebih baik dari sampel 1.000 orang yang diambil dari populasi masyarakat Yogyakarta. Kedua sampel itu mempunyai tingkat ketelitian sama, yakni 1.000 sampel, meskipun diambil dari populasi yang ukurannya berbeda. Perbedaannya adalah populasi yang besar (seperti Indonesia) boleh jadi mempunyai “kelereng” yang berbeda warna dan ukuran. Populasi yang besar menjadi lebih kompleks dalam subgroup, tingkat sosial ekonomi, jenjang pendidikan dan menjadi lebih sulit dalam menggambarkan perbedaan dalam berbagai macam subgroup. Kesulitannya bukan pada jumlah sampel tetapi teknik pengambilan sampel yang dapat menjamin keragaman karakteristik populasi. Sementara populasi yang kecil relatif lebih mudah menarik sampel yang dapat menggambarkan populasi.

### **Presisi sebagai Estimasi Besar Sampel**

Ada dua bagian yang penting dalam sampel yaitu bagaimana caranya mengeliminasi bias dan bagaimana meningkatkan ketelitian/presisi. Mengeliminasi bias berhubungan dengan teknik pengambilan sampel sementara presisi berhubungan dengan besar sampel. Presisi adalah suatu ukuran yang berhubungan dengan pertanyaan, bagaimana kita dapat mengestimasi nilai populasi (parameter) dari sampel dan bagaimana hasil dari sampel dapat digeneralisasikan untuk populasi yang lebih luas. Tingkat presisi ini harus diperhitungkan dan direncanakan sebelum polling dijalankan. Kalau kita ingin presisi yang lebih baik, konsekuensinya adalah penambahan jumlah sampel. Ada tiga faktor yang diperlukan dalam menghitung besar sampel yang diambil: variasi dalam populasi, tingkat kesalahan yang ditoleransi dan tingkat kepercayaan.

Variasi dalam populasi yang disampel. Andaikan kita membuat sampel dari semangkok kolak. Jika kolak itu terdiri atas berbagai ramuan termasuk pisang, kelapa, dan kacang hijau, akan diperlukan beberapa sendok sebelum kita menyimpulkan rasa seluruh isinya. Akan tetapi, jika kita memakan es krim dengan ramuan dan komposisi yang tepat, barangkali satu sendok sudah cukup untuk mengetahui rasanya. Demikian pula jika peneliti membuat sampel dari populasi, semakin beragam (dalam hal latar belakang sosial, karakteristik demografis, nilai, kepercayaan, tingkat pendidikan dan sebagainya), semakin besar jumlah yang diperlukan untuk memperoleh pukulan rata dari campuran yang heterogen tersebut.

Semakin besar keragaman dalam populasi, maka semakin besar sampel yang diperlukan.



Bila unsur dalam populasi betul-betul seragam, maka sampel dengan ukuran satu saja sudah cukup representatif. Pengujian mutu obat hasil produksi sebuah pabrik obat modern barangkali cukup satu sampel untuk setiap

100.000 pil, karena keragaman pil-pil tersebut hampir 100% seragam. Tetapi pengujian mutu hasil rokok yang dikerjakan manual dengan ribuan buruh pabrik membutuhkan jumlah sampel yang lebih besar agar tingkat keragamannya tercakup.

Besar sampel ditentukan oleh keragaman dari populasi. Populasi yang heterogen, mempunyai derajat perbedaan amat besar, sehingga dibutuhkan sampel yang besar untuk mencakup perbedaan yang ada. Umpamakan ada populasi 1.000 kelereng dalam satu kotak. Apabila populasi kotak itu terdiri atas 950 kelereng berwarna putih dan 50 kelereng berwarna hitam, kita hanya membutuhkan 50 sampel untuk menyimpulkan bahwa populasi kelereng dalam kotak itu lebih banyak kelereng berwarna putih. Hal ini karena populasi homogen (kelereng berwarna putih) sehingga sedikit saja kita mengambil sampel, cukup dapat menggambarkan populasi. Tetapi bagaimana kalau populasi kelereng itu terdiri atas 500 kelereng berwarna putih dan 500 lainnya berwarna hitam? Kita tidak cukup hanya mengambil 50 sampel kelereng, karena populasi amat heterogen, sehingga dibutuhkan jumlah sampel yang lebih besar. Kalau sampel yang diambil sedikit sementara tingkat heterogenitas populasi tinggi, maka akan terjadi peristiwa kebetulan: kita mendapat sampel kelereng lebih banyak berwarna putih karena kita kebetulan mendapat kelereng berwarna putih. Peristiwa kebetulan lebih besar daripada yang terjadi sebenarnya.

Analogi kelereng tersebut dapat kita pakai dalam kasus nyata. Kalau kita meneliti berapa rata-rata umur siswa SMU di Jakarta, tidak dibutuhkan jumlah sampel yang besar. Hal ini karena rata-rata umur siswa SMU relatif homogen, yaitu berkisar antara umur 15-20 tahun. Tetapi apabila kita meneliti rata-rata umur masyarakat Jakarta, maka dibutuhkan sampel dalam jumlah besar. Hal ini karena populasi amat heterogen dari bayi sampai kakek berusia 100 tahun. Demikian juga kalau kita membuat polling mengenai suatu isu. Polling yang menanyakan kepada responden apakah perlu para koruptor diadili tidak dibutuhkan sampel besar. Hal ini karena secara teoritis akan banyak suara homogen, kemungkinan 90% setuju dan mungkin hanya 10% saja yang tidak setuju. Tetapi sebaliknya untuk isu-isu kontroversial dimana pendapat masyarakat akan suatu masalah amat heterogen, maka dibutuhkan sampel dalam jumlah besar. Misalnya polling yang menanyakan apakah setuju dengan pelaksanaan UULL proporsi populasi yang setuju barangkali seimbang dengan yang tidak setuju.

Memperhitungkan proporsi populasi dalam mengambil sampel, berarti kita harus mempunyai pengetahuan yang cukup mengenai proporsi populasi. Pengetahuan itu bisa kita dapatkan dari data sekunder dalam kasus polling mengenai umur siswa atau masyarakat Jakarta tadi adalah buku statistik. Atau dengan melihat hasil penelitian atau polling yang pernah dilakukan sebelumnya. Tetapi umumnya kita tidak tahu secara pasti proporsi populasi. Ahli-ahli statistik mengatakan kita sebaiknya memakai proporsi populasi seimbang, dalam hal ini 50% : 50%. Ini artinya kita mengasumsikan populasi heterogen, masyarakat terbagi dalam aneka pendapat. Kita memakai proporsi populasi 50% : 50%, karena proporsi populasi ini adalah angka maksimal, dengan mengasumsikan populasi heterogen, kita tidak akan mengalami kekeliruan apabila ternyata populasi itu homogen.

### **Tingkat kesalahan yang ditoleransi (*sampling error*).**

Berapa banyak kesalahan yang dapat kita toleransi umumnya tergantung pada estimasi sebelum suatu polling dijalankan. Sampel berbeda dari populasi, karenanya peneliti harus memandang hasil dari suatu sampel bukanlah hasil yang pasti, tetapi lebih sebagai kira-kira.

Akurasi dari suatu polling diukur diantaranya dari sejauh mana ketepatan sampel kita dalam menggambarkan populasi. Presisi merupakan pernyataan sejauh mana perbedaan antara nilai statistik dengan nilai parameter. Parameter adalah ciri-ciri yang menjelaskan populasi, sedangkan statistik adalah ciri-ciri yang menjelaskan sampel. Peneliti polling selalu mengharapkan bahwa nilai statistik sama persis dengan nilai parameter, tetapi dalam kenyataannya selalu saja ada perbedaan antara nilai statistik dan nilai parameter.

### **Presisi tergantung pada besar sampel.**

Dalam sampel probabilitas, jumlah sampel yang besar akan memberikan presisi yang lebih besar oleh karena dapat menurunkan kesalahan kesempatan dalam acak. Akan lebih mudah memahami hal ini dengan ilustrasi demikian. Misalkan kita membuat polling pendapat mahasiswa UGM terhadap keberadaan Dewan Mahasiswa. Dari sampel sebanyak 400 mahasiswa yang dipilih secara acak dari seluruh mahasiswa UGM, ditemukan hasil bahwa 60% responden menyetujui keberadaan Dewan Mahasiswa. Hasil polling memakai sampel ini kita perbandingkan dengan sensus: kita mewawancarai semua mahasiswa UGM yang berjumlah 24.600 orang, hasilnya pasti berbeda. Misalnya, setelah dirata-rata ternyata yang menyetujui keberadaan Dewan Mahasiswa sebanyak 66%. Dengan demikian ada kesalahan (error) sebesar 6% dalam memproyeksikan populasi. Kesalahan sebesar 6% itu merupakan kesalahan yang terjadi sebagai efek pengambilan sampel: kita mewawancarai 500 orang dan bukan 25 ribu orang. Dalam sensus kesalahan adalah 0% karena semua mahasiswa UGM diwawancarai sehingga tidak ada kesalahan. Kesalahan 6% itu dapat diperkecil dengan jalan menambah sampel.

Masalahnya dalam melakukan polling kita tidak mungkin tahu nilai parameter. Kita tidak tau berapa persentase seluruh mahasiswa UGM yang setuju adanya Dewan Mahasiswa. Yang kita teliti dan kita tahu hanyalah nilai statistik, dalam hal ini persentase dari sampel yang mendukung Dewan Mahasiswa. Pertanyaannya berapa persentase sesungguhnya mahasiswa UGM yang mendukung Dewan Mahasiswa? Pertanyaan ini dengan kata lain menanyakan nilai parameter dari hasil sampel yang diperoleh dalam polling. Kita harus menentukan terlebih dahulu berapa tingkat kesalahan yang kita inginkan, baru kita dapat memprediksikan nilai populasi. Dengan mengetahui Sampling Error, kita dapat memprediksikan nilai parameter dari hasil sampel. Misalnya Sampling Error yang kita ambil dalam polling mengenai Dewan Mahasiswa tadi adalah 6% maka hasil 60% mahasiswa yang mendukung Dewan Mahasiswa harus dibaca  $60 \pm 6\%$ . Artinya, nilai sebenarnya dari mahasiswa UGM yang mendukung Dewan Mahasiswa berada diantara 54-66%.

### **Tingkat kepercayaan.**

Dalam menentukan besar sampel, peneliti juga memperhatikan tingkat kepercayaan yang harus diberikan dalam menyimpulkan bahwa jika seseorang menarik sampel lain dari populasi itu, sampel tersebut harus memberikan hasil yang kira-kira sama dengan yang pertama. Diulang berapapun pengambilan sampel, akan memberikan hasil yang sama. Tingkat kepercayaan ini erat hubungannya dengan Sampling Error. Sampling Error mengacu kepada bagaimana keakuratan taksiran yang diinginkan oleh peneliti, sedangkan tingkat kepercayaan mengacu kepada bagaimana kepastian yang diinginkan bahwa taksiran itu sendiri akurat.

Kita bisa membuat analogi demikian. Andaikan kita meneliti kolak dalam satu dandang besar (populasi), peneliti kemudian mengambil segelas kolak sebagai sampel. Misalkan peneliti menemukan bahwa 20% dari sampel kolak terdiri atas pisang. Apakah angka itu mewakili proporsi pisang dalam seluruh kolak dandang besar? Dalam setiap sampel, betapapun telitinya sampel itu ditarik, biasanya ada saja kesalahan. Misalnya bisa jadi 25% dari kolak itu terdiri atas

pisang meskipun pencicip hanya menemukan 20%. Orang yang halus pengecapnya mungkin berpendapat bahwa ketidakakuratan seperti itu tidak dapat ditoleransi, tetapi orang yang kelaparan tidak akan begitu menghiraukannya. Tingkat kesalahan itulah yang disebut Sampling Error. Sehingga ia selalu mempertanyakan: Apakah 20% dari kolak itu terdiri dari pisang sementara tingkat kepercayaan selalu mempertanyakan bahwa 20% dari setiap mangkok kolak akan terdiri dari pisang.

Bila satu mangkok berisi 20% pisang, dan diambil semangkok lagi kolak dari dandang, peneliti ingin yakin bahwa hasilnya sama yaitu 20% terdiri dari pisang. Tetapi tidak mungkin 100%, karena pengambilan sampel sebarangpun telitinya selalu ada kesalahan. Tingkat kepercayaan yang sering dipakai adalah 90%, 95% dan 99%. Disini peneliti merasa yakin 90% atau 95% bahwa komposisi sampel bisa diulang serta tetap identik jika dipilih sampel lain dari populasi yang sama. Semakin tinggi tingkat kepercayaan yang diinginkan, semakin besar ukuran sampel yang diperlukan. Tingkat kepercayaan dapat memberi keyakinan kepada kita bahwa temuan-temuan dalam sampel dapat digeneralisasikan kepada populasi. Bila digunakan tingkat kepercayaan 90 atau 95%, ini berarti jika terdapat 100 sampel, maka perbedaan yang diamati akan muncul dalam 90 atau 99 dari sampel itu.

Berapa besar sampel yang dibutuhkan? Ketiga faktor yang menentukan besar sampel itu dapat dirangkum dalam rumus sebagai berikut:

$$N = (p \times q) \cdot Z^2/E^2$$

Dimana  $(p \times q)$  adalah variasi proporsi populasi,  $Z$  adalah ukuran tingkat kepercayaan dan  $E$  adalah Sampling Error / kesalahan yang dapat ditoleransi. Sebagai misal, kita ingin membuat sampel dimana populasinya adalah mahasiswa UGM. Sebagai taksiran tentang keragaman populasi biasanya diambil proporsi seimbang 50 : 50. Misalnya proporsi antara laki-laki dan perempuan 50% : 50%. Untuk menghitung variasi dari populasi, kita mengalikan proporsi laki-laki ( $p$ ) dan proporsi wanita ( $q$ ) sehingga  $(p \cdot q) = 0,5 (0,5) = 0,25$ . Ukuran tingkat kepercayaan adalah suatu skor  $Z$ . Kita cukup saja dengan mengatakan bahwa pada tingkat kepercayaan 99%, skor  $Z$  adalah 2,58; pada tingkat kepercayaan 95% skor  $Z$  adalah 1,96 dan pada tingkat kepercayaan 90% skor  $Z$  adalah 1,65. Misalkan dalam contoh ini tingkat kepercayaan adalah 95%, sehingga skor  $Z$  adalah 1,96 masuk ke dalam rumus. Ukuran kesalahan yang dapat ditoleransi sederhana. Andaikan kita menginginkan taksiran berdasarkan sampel dengan kesalahan tidak lebih dari  $\pm 5\%$ ,. Jadi jika sampel itu menunjukkan bahwa 45% setuju Dewan Mahasiswa, maka diantara 40 dan 50% dari populasi seluruh mahasiswa UGM pasti menyetujui keberadaan Dewan Mahasiswa. Dengan demikian  $E = 0,05$ . Oleh karena  $N = (0,25) \cdot 1,96^2/0,05^2$ , atau  $N = (0,25) \cdot 3,842/0,0025 = 384$ .

Besar sampel yang diperlukan adalah 384 orang. Jika kita naikan tingkat kepercayaan kita menjadi 99%, maka akan diperlukan besar sampel berikut bagi setiap Sampling Error/tingkat kesalahan:  $\pm 1\% = 9.000$ ,  $\pm 2\% = 2.400$ ,  $\pm 3\% = 1.067$ ,  $\pm 4\% = 600$ ,  $\pm 6\% = 267$ ; dan  $\pm 7\% = 196$  dan seterusnya.

Dengan mempertimbangkan proporsi populasi, Sampling Error yang dipakai dan tingkat kepercayaan, kita dapat memprediksikan jumlah sampel polling yang kita butuhkan dalam polling. Ketiga hal tersebut menunjuk kepada presisi tingkat ketelitian yang kita perlukan dalam polling. Sebagaimana dikatakan Bradburn and Sudman, jika kita membutuhkan hasil polling yang teliti kita akan mengambil Sampling Error sekecil mungkin dan tingkat kepercayaan sebesar mungkin itu berarti jumlah sampel yang besar. Semua ini dapat diperhitungkan peneliti sebelum polling dilakukan.

