

## BAB VI UKURAN SAMPEL

### A. Penentuan Sampel

Karena berbagai alasan, tidak semua hal yang ingin dijelaskan atau diramalkan atau dikendalikan dapat diteliti. Penelitian ilmiah boleh dikata hampir selalu hanya dilakukan terhadap sebagian saja dari hal-hal yang sebenarnya mau diteliti. Jadi penelitian hanya dilakukan terhadap sampel, tidak terhadap populasi. Namun kesimpulan-kesimpulan penelitian mengenai sampel itu akan dikenakan atau digeneralisasikan terhadap populasi. Generalisasi dari sampel ke populasi ini mengandung resiko bahwa akan terdapat kekeliruan atau ketidak-tepatan, karena sampel tidak akan mencerminkan secara tepat keadaan populasi. Makin tidak sama sampel itu dengan populasinya, maka makin besarlah kemungkinan kekeliruan dalam generalisasi itu. Karena hal yang demikian itulah maka teknik penentuan sampel itu menjadi sangat penting peranannya dalam penelitian. Berbagai teknik penentuan sampel itu pada hakekatnya adalah cara-cara untuk memperkecil kekeliruan generalisasi dari sampel ke populasi. Hal ini dapat dicapai kalau diperoleh sampel yang representatif, yaitu sampel yang benar-benar mencerminkan populasinya.

Diantara berbagai teknik penentuan sampel yang dianggap paling baik adalah penentuan sampel secara rambang (*random sampling*). Kebaikan teknik ini tidak hanya terletak pada teori yang mendasarinya, tetapi juga pada bukti-bukti empiris. Perkembangan teknologi komputer telah memungkinkan orang melakukan berbagai simulasi untuk membuktikan keunggulan teknik pengambilan sampel secara rambang itu. Di dalam penentuan sampel secara rambang semua anggota populasi, secara individual atau secara kolektif, diberi peluang yang sama untuk menjadi anggota sampel. Alat untuk mengambil sampel secara rambang ini yang paling praktis (dan dianggap paling valid juga) ialah dengan menggunakan tabel bilangan rambang atau kalkulator yang mempunyai program untuk bilangan rambang. Jika besarnya populasi terbatas, peluang rambang dapat diberikan kepada anggota-anggota secara individual. Tetapi kalau populasi itu sangat besar, sebaiknya peluang rambangnya diberikan kepada anggota-anggota populasi secara kelompok, dan kalau perlu dilanjutkan dengan rambang individual.

Walaupun teknik pengambilan sampel secara rambang itu merupakan teknik terbaik, namun tidak selalu dapat dilaksanakan, karena berbagai alasan. Kadang-kadang orang terpaksa puas dengan sampel

rumpun (*cluster sample*), karena rumpun-rumpun yang merupakan kelompok individu-individu itu yang tersedia sebagai unit-unit dalam populasi. Penelitian mengenai murid-murid sekolah biasanya tidak dapat menggunakan teknik pengambilan sampel secara rambang, melainkan harus secara rumpun. Yang mendapatkan peluang sama untuk menjadi sampel bukan murid secara individual, melainkan sekolah (jadi murid secara kelompok).

Seringkali terjadi pula sampel diambil dari rumpun-rumpun yang telah ditentukan atau tersedia. Teknik yang demikian itu disebut penentuan sampel secara bertingkat (*stratified sampling*). Kalau dari kelompok-kelompok yang tersedia itu diambil sampel-sampel yang sebanding dengan besarnya kelompok dan pengambilannya secara rambang, maka teknik itu disebut pengambilan sampel secara rambang proporsional (*proportional random sampling*).

Seperti telah disebutkan tujuan berbagai teknik penentuan sampel itu ialah untuk mendapatkan sampel yang paling mencerminkan populasinya, atau secara teknis disebut sampel yang paling representatif. Dalam penelitian terhadap sampel, ciri representativeness sampel itu tidak pernah dapat dibuktikan, melainkan hanya dapat didekati secara metodologis melalui parameter-parameter yang diketahui dan diakui baik secara teoritis maupun secara eksperimental. Ada empat parameter yang bisa dianggap menentukan representativeness sesuatu sampel, yaitu (a) variabilitas populasi, (b) besar sampel, (c) teknik penentuan sampel, dan (d) kecermatan memasukkan ciri-ciri populasi dalam sampel.

**Variabilitas populasi.** Dari keempat parameter tersebut diatas itu variabilitas populasi merupakan hal yang sudah *given*, artinya penelitian harus menerima sebagaimana adanya, dan tidak dapat mengatur atau memanipulasikannya. Ketiga parameter yang lain tidak demikian halnya; penelitian dapat atau memanipulasikannya untuk meningkatkan taraf *representativeness* sampel.

**Besar sampel.** Makin besar sampel yang diambil akan makin tinggi taraf *representativeness* sampelnya. Ketentuan ini berlaku selama populasinya tidak homogen secara sempurna. Jika populasinya homogen secara sempurna besar sampel tidak mempengaruhi taraf representatifnya sampel. Untuk populasi yang demikian itu sampel cukup kecil saja.

**Teknik penentuan sampel.** Makin tinggi tingkat rambang dalam penentuan sampel, akan makin tinggilah tingkat representatif sampel. Ketentuan ini juga hanya berlaku selama populasinya tidak homogen secara sempurna. Jika populasinya homogen secara sempurna rambang sama sekali tidak diperlukan.

**Kecermatan memasukkan ciri-ciri populasi** ke dalam sampel, akan makin tinggi tingkat representatifnya sampel. Dengan mempertimbangkan parameter-parameter tersebut di atas, peneliti diharapkan dapat

mengusahakan, dapat menentukan sampel yang paling tinggi tingkat representatifnya yang mungkin dicapai. Kecepatan untuk ini, seperti untuk melakukan langkah-langkah yang lain dalam penelitian, sangat tergantung kepada latihan dan pengalaman.

Walau berbagai teknik pengambilan sampel telah dikembangkan dan parameter-parameter untuk perkiraan telah diidentifikasi, namun hampir tidak pernah penelitian dapat menentukan sampel yang mencerminkan populasi secara sempurna. Hal yang demikian itu terjadi terutama dalam lapangan ilmu-ilmu sosial dan kemanusiaan. Karena itu kekeliruan yang timbul karena penggunaan sampel (*sampling error*) hampir selalu ada. Keadaan yang demikian itu lalu menimbulkan kebutuhan untuk dapat memperhitungkan atau setidaknya memperkirakan, besar kecilnya kekeliruan itu. Dalam analisis kekeliruan dalam generalisasi dari sampel ke populasi itu disebut kekeliruan baku (*standard error*). Dasar teoritis yang digunakan untuk memperkirakan kekeliruan baku itu ialah teori probabilitas. Sampel-sampel tunduk kepada hukum probabilitas, demikian juga harga-harga yang diperoleh dari sampel. Interpretasi kekeliruan baku itu adalah sama dengan interpretasi harga-harga lain yang menggunakan tabel probabilitas.

## B. Rencana Sampling

### 1. Unit Sampling

Unit sampling adalah sesuatu yang berdasarkan kriteria tertentu, dijadikan sebuah ketentuan yang karakteristiknya akan diukur. Unit sampling ini dapat merupakan sebuah individu yang berdiri sendiri, kumpulan individu, sebuah daerah tertentu, atau sebuah satuan waktu tertentu. Yang penting dalam menentukan unit sampling adalah terdapatnya kriteria yang secara ketat menentukan sesuatu sebagai sebuah kesatuan yang jelas batas-batasnya.

Menentukan unit-unit sampling secara keseluruhan sama dengan menentukan *sampled population* (populasi yang akan disampel). Populasi yang akan disampel ini hendaknya identik dengan target *population* (populasi sasaran), yaitu populasi yang akan merupakan generalisasi dari fakta empirik yang terdapat dalam sampel.

### 2. Kerangka Sampling (*Sampling Frame*)

Kerangka sampling adalah sebuah daftar berisi unit-unit sampling yang ada dalam populasi. Unit-unit ini masing-masing diberi nomor urut yang satu sama lain berbeda, sehingga dengan menunjuk sebuah nomor unit, kita dapat mengidentifikasinya di lapangan. Nomor-nomor ini terdiri

dari digit (angka) yang sama banyaknya. Kerangka Sampling ini mutlak harus ada, karena dari kerangka sampling inilah kita akan memilih unit ke dalam sampel.

### 3. Ukuran Sampel

Ukuran sampel adalah banyaknya unit sampling yang ada dalam sampel. Ukuran sampel biasanya diberi simbol *n* (simbol ukuran populasi adalah *N*) dalam urutan langkah rencana sampling, menentukan ukuran sampel inilah yang paling sulit, karena banyak sekali faktor yang melandasinya. Ukuran sampel tergantung kepada:

- (1). Variabel karakteristik yang akan diukur yang dimiliki unit-unit sampling. Variabelitas ini diperlihatkan oleh varians populasi yang besarnya jarang sekali diketahui.
- (2). Kekeliruan sampling yang dapat ditolerir. Kekeliruan ini ditentukan oleh peneliti, yang besarnya tergantung kepada tujuan penggunaan hasil penelitian.
- (3). Biaya penelitian per unit sampling.
- (4). Waktu.

Biasanya biaya dan waktulah yang dijadikan kendala untuk menentukan ukuran sampel ini. Secara umum ukuran sampel dapat dipakai rumus Cochran (1991), yaitu:

$$n = \frac{\frac{Z^2 P.Q}{d^2}}{1 + \frac{1}{N} \left[ \frac{Z^2 P.Q}{d^2} - 1 \right]}$$

Keterangan: *n* adalah ukuran sampel; *P* merupakan proporsi dari masing-masing kelompok sampel pada kelas yang terpilih; sedangkan *Q* = 1 - *P*. *N* adalah ukuran populasi; *Z* adalah nilai deviasi normal terhadap probabilitas keyakinan yang diinginkan, dan *d* = standar error.

### 4. Memilih Unit Sampling

Unit sampling dipilih ke dalam sampel melalui prosedur acak, artinya pemilihan unit sampling dilakukan sedemikian rupa, sehingga setiap unit sampling yang ada dalam populasi mempunyai peluang terpilih yang diketahui besarnya, dan peluang ini tidak boleh sama dengan nol. Untuk memilih sampling secara acak, diperlukan kerangka sampling, dan tabel angka random.

Pemilihan unit sampling secara acak ini mutlak harus dilakukan, agar supaya kita dapat mempergunakan metode statistik sebagai alat analisis. Apabila persyaratan acak di atas tidak dipenuhi, maka rumus-rumus yang disediakan oleh metode estimasi tidak sah untuk dipergunakan.

## 5. Macam-macam Rencana Sampling

Dari sekian banyak rencana sampling yang ada, hanya beberapa buah saja yang akan kita bicarakan di bawah ini.

### 5.1. Sampling Acak Sederhana (SAS)

Ciri utama SAS adalah bahwa peluang setiap unit sampling untuk terpilih ke dalam sampel adalah sama. Kerangka sampling berisi semua unit sampling yang ada dalam populasi, dan pemilihan dilakukan langsung dari kerangka sampling dengan mempergunakan angka random (angka/bilangan acak).

Rencana sampling ini dipergunakan apabila kita mempunyai keterangan/petunjuk bahwa keadaan karakteristik yang akan diukur tersebar meluas diseluruh populasi. Keuntungan rencana ini terletak pada kesederhanaan analisisnya (estimasi parameternya), karena tidak memerlukan pembobotan (*weighting*). Tetapi kesederhanaan ini harus dibayar oleh prestasi yang rendah dan biaya yang cukup mahal, sebagai akibat dari harus tersediannya kerangka sampling yang lengkap, yang harus disusun sekaligus.

Ukuran sampelnya dapat dipakai dengan menggunakan rumus dibawah ini.

Untuk Estimasi Mean: 
$$n = \frac{N.r^2}{(N-1)D + r^2}$$

dimana, 
$$D = \frac{B^2}{4}$$

Untuk Estimasi Total: 
$$n = \frac{N.r^2}{(N-1)D + r^2}$$

dimana, 
$$D = \frac{B^2}{4N^2}$$

Untuk Estimasi Proporsi: 
$$n = \frac{N \cdot p(1-p)}{(N-1)D + p(1-p)}$$

dimana, 
$$D = \frac{B^2}{4}$$

Keterangan: N = Ukuran populasi; n = Ukuran sampel; p = Proporsi populasi; B = Bound of error; r<sup>2</sup> = Variance populasi. Apabila p tidak diketahui, biasanya dapat diambil dari hasil survei sebelumnya.

## 5.2. Sampling Acak dengan Stratifikasi (SASTRA)

Langkah pertama yang harus dilakukan dalam SASTRA adalah membagi populasi ke dalam strata (subpopulasi). Penstratifikasi ini gunanya untuk menghomogenkan karakteristik yang diperkirakan keadaannya heterogen.

Setelah kita menentukan ukuran sampel keseluruhan, misalnya saja n, maka setiap stratum kita harus memilih unit secara acak sedemikian rupa, sehingga banyaknya unit yang diambil dari stratum-stratum itu secara keseluruhan sama dengan n. Langkah ini dinamakan alokasi unit ke dalam strata. Ada beberapa metode alokasi:

### (1). Alokasi Sembarang

Dari setiap stratum dipilih unit yang banyaknya sembarang, dengan syarat bahwa banyaknya secara keseluruhan adalah n. Alokasi ini kurang baik untuk dipergunakan, karena mungkin akan memberikan prestasi yang lebih rendah dari rencana SAS.

### (2). Alokasi Sama Banyak

Dari setiap stratum diambil unit yang sama banyaknya, sehingga secara gabungan akan sama dengan n. Prestasi yang diberikan mungkin lebih buruk dari yang diberikan oleh SAS.

Untuk ukuran sampelnya dipakai rumus sebagai berikut:

$$n = \frac{L \cdot \sum N_i^2 \cdot p_i(1-p_i)}{N^2 \cdot D + \sum N_i \cdot p_i(1-p_i)} \longrightarrow n_i = \frac{n}{L}$$

L adalah banyaknya strata.

## (3). Alokasi Proporsional

Alokasi ini berpegang pada kriteria, bahwa makin besar ukuran stratum, makin banyak ukuran unit yang diambil dari stratum itu. Dalam praktek, alokasi inilah yang banyak digunakan, karena adanya beberapa keuntungan:

- Sederhana
- Analisisnya tidak memerlukan bobot karena sampel yang diperoleh merupakan "self-weighting".
- Presisi yang diberikan tinggi, paling rendah sama dengan presisi yang dapat diberikan oleh SAS.

$$n = \frac{N \cdot \sum N_i \cdot p_i (1 - p_i)}{N^2 \cdot D + \sum N_i \cdot p_i (1 - p_i)}$$

dimana

$$n_i = \frac{N_i}{N} \cdot n$$

## (4). Alokasi Optimum

Kriteria yang dipergunakan disini adalah :

- makin besar variasi karakteristik dalam sebuah stratum, makin banyak unit yang diambil dari stratum itu
- makin murah biaya per unit sampling dalam sesuatu stratum, makin banyak yang diambil
- makin besar ukuran stratum, makin banyak unit yang diambil dari stratum yang bersangkutan.

Sekalipun alokasi ini dapat memberikan presisi yang tinggi, tetapi alokasi ini melibatkan varians stratum (sebagai ukuran variasi karakteristik dalam stratum) yang jarang sekali diketahui besarnya.

Ukuran sampel untuk alokasi optimum dapat dipakai rumus di bawah ini:

$$n = \frac{\sum N_i \cdot \sqrt{p_i (1 - p_i)} \cdot \sqrt{C_i} \cdot \sum N_i \cdot \sqrt{p_i (1 - p_i)}}{\sqrt{C_i} \cdot (N^2 \cdot D + \sum N_i \cdot p_i (1 - p_i))}$$

$$n_i = \frac{\frac{N_i \cdot \sqrt{p_i (1 - p_i)}}{\sqrt{C_i}}}{\frac{\sum N_i \cdot \sqrt{p_i (1 - p_i)}}{\sqrt{C_i}}} \cdot n$$

Nilai  $C_i$  adalah unit biaya per stratum

(5). Alokasi Neyman

Alokasi ini tergantung kepada besarnya ukuran stratum dan besarnya varians stratum. Jika sekiranya kita mengetahui besarnya varians untuk tiap-tiap stratum, maka alokasi ini merupakan yang baik sekali karena dapat memberikan presisi yang tinggi.

Secara keseluruhan, SASTRA merupakan rencana sampling yang superior dibandingkan dengan SAS, tetapi superioritas ini harus dibayar dengan biaya yang sangat mahal. Ukuran sampelnya adalah:

$$n = \frac{\left[ \sum N_i \cdot \sqrt{p_i(1-p_i)} \right]^2}{N^2 D + \sum N_i \cdot p_i(1-p_i)}$$

$$n_i = \frac{N_i \cdot \sqrt{p_i(1-p_i)}}{\sum N_i \cdot \sqrt{p_i(1-p_i)}} \cdot n$$

### 5.3. Sampling Acak Klaster (SAK)

SAK merupakan rencana yang banyak sekali dipergunakan dalam penelitian survai karena biayanya yang relatif murah. Tetapi sayangnya, presisi yang diberikan oleh rencana ini rendah, dan rumus-rumus estimasinya merupakan estimator yang bias.

Untuk memperkecil biasnya harus diambil dengan ukuran relatif besar. Langkah pertama yang dilakukan adalah membentuk kluster-kluster, yang dalam bahasa samplingnya disebut juga unit sampling primer (USP). USP merupakan unit terbesar, di dalamnya terdapat unit-unit terkecil, yang disebut unit sub sampling (USS). USS dapat dibagi lagi kedalam USS yang lebih kecil, yang disebut USS-kedua, USS-ketiga, dan seterusnya sampai kepada unit terkecil (unit elementer) yang tidak dapat dibagi-bagi lagi. Langkah selanjutnya adalah melakukan pemilihan tingkat pertama (*first stage selection*), yaitu memilih USP secara acak, dari kerangka USP yang tersedia.

Dari USP yang terpilih pada pemilihan tingkat pertama, dilakukan pemilihan USS. USP yang tidak terpilih dibiarkan. Pemilihan USS dari USP yang terpilih itu merupakan pemilihan tingkat kedua (*second stage selection*). Pemilihan unit-unit tersebut dapat berhenti pada pemilihan tingkat pertama (disebut sampling acak kluster satu tingkat), dapat berhenti pada pemilihan tingkat kedua (disebut sampling acak kluster dua tahap), dapat pula terus sampai unit terakhir yang terkecil (disebut Sampling Acak Kluster Multi Tingkat).



## 5.4. Sampling Sistematis

Sampling sistematis dapat memberikan presisi yang tinggi dengan prosedur pemilihan yang sangat sederhana. Pertama-tama, setelah ukuran sampel ditentukan, kita tentukan besarnya interval pemilihan:

$$I = \frac{N}{n}$$

Kemudian pilih sebuah angka acak, awal pemilihan (AP), yang memenuhi batas  $1 < AP < I$ . AP yang diperoleh secara acak itu merupakan unit sampling pertama yang dipilih ke dalam sampel, sedangkan unit-unit selanjutnya dipilih dengan cara menambahkan  $I$  kepada nomor unit yang sebelumnya (AP, AP + I, AP + I + I, dan seterusnya).

Apabila dalam kerangka sampling terdapat pengulangan karakteristik yang sama secara priodik, maka pemilihan sistematis ini dapat membahayakan, karena dalam sampel nantinya terdapat banyak unit yang mempunyai karakteristik yang sama, sehingga informasi yang diberikan oleh sampel adalah kecil.

Keempat rencana sampling di atas merupakan rencana sampling dasar. Rencana sampling lainnya diperoleh dengan cara mengkombinasikan rencana sampling dasar tersebut disertai dengan modifikasi-modifikasi tertentu. Para peneliti memilih rencana sampling yang tepat, yang sesuai dengan situasi "lapangan" dan tujuan yang digunakannya hasil penelitian. Rencana sampling yang didasarkan kepada prosedur acak seperti di atas, termasuk ke dalam kategori sampling peluang (probabilitas sampling). Untuk data yang terkumpul melalui rencana ini dapat digunakan analisis statistika.

Di samping rencana sampling peluang dikenal pula rencana sampling non-peluang (non-probabilitas sampling) atau disebut juga model sampling. Untuk model sampling, tidak berlaku analisis statistika.

## C. Metode Estimasi (Analisis)

Estimasi parameter atau analisis yang dipergunakan, tergantung kepada dua hal, rencana sampling dan tingkat pengukuran (skala) yang dimiliki oleh data.

Untuk data yang berskala nominal dan/atau ordinal digunakan metode statistik non-parametrik, sedang untuk data dengan skala interval, dan/atau ratio dapat dipergunakan metode statistika parametrik.

## D. Macam-macam Disain Sampling

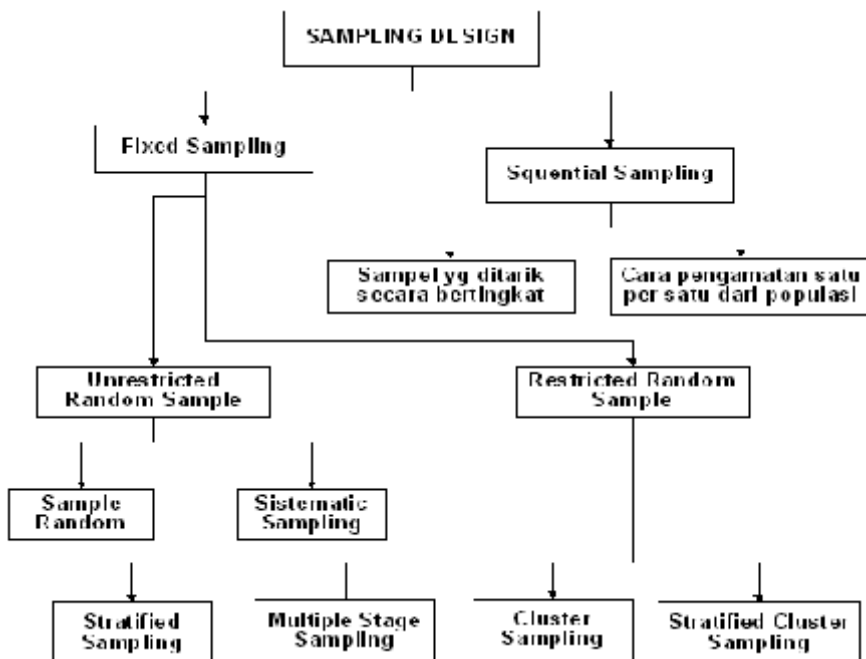
*Unrestricted random sample*: sampel ditarik secara langsung dari populasi, tanpa dibagi-bagi terlebih dahulu.

*Restricted random sample*: sampel ditarik dari populasi yang telah dikelompokkan lebih dahulu.

- ❑ *Stratified sampling*: populasi dibagi dalam kelompok yang homogen terlebih dahulu. Anggota sample ditarik dari setiap strata. Jika tidak semua strata ditarik sampelnya, maka ia menjadi;
- ❑ *Multiple stage sample*: Sampel ditarik dari kelompok populasi, tetapi tidak semua anggota kelompok populasi menjadi anggota sampel. Hanya sebagian dari anggota subpopulasi menjadi anggota sampel.
- ❑ *Cluster sampling*: populasi dibagi dulu atas kelompok berdasarkan area atau cluster. Anggota subpopulasi tiap cluster tidak perlu homogen. Beberapa cluster dipilih dulu sebagai sampel, kemudian dipilih lagi anggota unit dari sampel cluster di atas.
- ❑ *Stratified cluster sample*: sampel ditarik dengan kombinasi antara stratified sampling dan cluster sampling.

*Simple random sample* (sampel random sederhana): Sampel yang diinginkan ditarik secara random, baik menggunakan random numbers ataupun dengan undian biasa.

*Sistimatic sample* (sampel sistematik): Unit populasi diberi nomor urut, kemudian ditentukan nomor awal star secara acak. Untuk nomor berikutnya ditentukan secara sistematik.



Gambar 4. Macam-macam Disain Sampling (Menurut C.W. Churchman)