

PENGGUNAAN STATISTIK DALAM PENELITIAN SOSIOLOGI

Oleh : Partini¹

Abstract

The usage of statistic in social research, particularly in sociology is very uncomplicated and straightforward as long as the related researcher understands the basic argument from the chosen statistic formula. The choosing of specific formula is based on the type and data quality in order to get a more precise and accurate formula. The precision and accuracy in choosing the technique and statistic formula strongly define the quality of conducted research so that the formulizing of the summary is more precise, too.

Statistic is nothing but a mean to create a more "meaningful" data. The usage itself is varies and the right understanding is not only based on the number resulted manually or computerize which usually used on the SPSS program but also have to be supported with the concept and theory. Furthermore, the understanding can also be get from research result conducted by other and from a balance battle of concept and empiric fact. Simply by using this method, the results can be accepted by others and can be more meaningful particularly for the sake of science development and policy makers.

One of the excellences of by using the quantitative research method that usually applies the statistic formula can be used to predict. Moreover, the results and the summary can be generalized to the condition of population so that those can be utilized by a lot of people.

Keywords: Statistic, Quantitative Method, Sociological Research

A. PENDAHULUAN

Para peneliti sosial sering merasa enggan berbicara masalah statistik, karena dalam bayangan mereka kata statistik memiliki

makna rumus-rumus yang sangat rumit dan angka-angka yang menjemukan, sehingga dapat membuat peneliti lebih pusing. Pada awalnya kata statistik diartikan sebagai keterangan-keterangan yang dibutuhkan oleh negara, untuk kepentingan negara. Secara definitif statistik merupakan angka yang diperoleh dengan menggunakan metode statistika untuk mendeskripsikan sampel. Kemudian arti tersebut mengalami pergeseran makna karena statistik lebih diartikan sebagai data kuantitatif baik yang sudah diolah, yang berupa data-data mendasar. Padahal data kuantitatif dengan statistik adalah berbeda. Data kuantitatif hanyalah merupakan kumpulan angka-angka, dan kumpulan angka-angka ini tidak dapat disebut sebagai statistik. Statistik lebih merupakan metode atau azas-azas guna "mengerjakan" atau menampilkan data kuantitatif agar supaya angka-angka tersebut dapat berbicara (Anto Dayan, 1973). Di dalam ilmu sosial dikenal ada 2 macam jenis statistik yang biasa dipergunakan dalam menganalisis data, yaitu : (1) Statistik deskriptif, (2) Statistik eksplanatif (*inference*).

Dalam penelitian sosiologi, ada banyak metode penelitian yang dapat dipakai, salah satu dari sekian banyak jenis penelitian itu adalah survei. Penggunaan jenis penelitian survei membutuhkan tahapan yang jelas, dengan melakukan interaksi secara simbiosis mutualistik antara teori yang dipergunakan dengan data empiris yang ada di lapangan. Berangkat dari teori yang menjadi dasar berpijak dan frame work peneliti, dengan menggunakan logika formal yang rasional, teori tersebut kemudian diturunkan di dalam konsep-konsep yang lebih mendasar, yang lebih mendekati realitas sosial, melalui bekerjanya logika deduktif. Konsep merupakan abstraksi dari fenomena sosial yang sifatnya general. Secara definitif, pengertian konsep seperti yang ditulis oleh Singarimbun dan Sofian Effendi (1987), konsep adalah abstraksi mengenai suatu fenomena yang dirumuskan atas dasar generalisasi dari sejumlah karakteristik, kejadian, keadaan, kelompok atau individu tertentu. Konsep dalam pandangan Turner (dalam Babbie, 1989) bahwa konsep merupakan kumpulan bangunan dasar dari teori. Konsep adalah elemen abstrak yang menampilkan sekelompok fenomena dalam suatu bidang

¹ Dosen Jurusan Sosiologi, Fakultas Ilmu Sosial dan Ilmu Politik, Universitas Gadjah Mada

kajian. Konsep yang di dalam penelitian sosial biasanya bersifat abstrak, harus didefinisikan dan kemudian diturunkan menjadi variabel penelitian, karena dengan variabel ini dapat dicari indikasi-indikasi atau gejala-gejalanya, yang merupakan kenyataan riil yang terjadi di dalam kehidupan masyarakat. Dari beberapa variabel yang akan diukur, dan hubungan di antara beberapa variabel tersebut sampailah peneliti pada satu tahapan yang disebut sebagai hipotesis. Hipotesis ini merupakan jawaban sementara yang mencerminkan hubungan dua variabel atau lebih yang harus dibuktikan kebenarannya di lapangan.

B. PENGGUNAAN STATISTIK DESKRIPTIF

Statistik deskriptif merupakan suatu metode untuk mengumpulkan, mengolah dan menyajikan serta menganalisa data kuantitatif agar dapat memberikan gambaran yang teratur tentang suatu peristiwa. Statistik jenis ini hanya dapat dipakai untuk memberikan gambaran dan menganalisa data dari variabel tunggal atau hanya dipergunakan untuk menganalisis dan mengukur satu variabel saja. Statistik deskriptif biasanya hanya bisa dilakukan untuk jenis penelitian yang sifatnya deskriptif juga. Penelitian deskriptif ini dilakukan untuk melakukan pengukuran terhadap fenomena sosial tertentu, seperti misalnya, banyaknya pengangguran intelektual di suatu wilayah tertentu atau banyaknya kasus perceraian yang terjadi pada kurun waktu tertentu. Dalam penerapan statistik deskriptif ini, peneliti melakukan pengumpulan data, namun tidak melakukan pengujian hipotesis. Analisis datanya lebih bersifat deskriptif, dan bukan analisis yang bisa menunjukkan ada tidaknya hubungan antara beberapa variabel yang dipakai di dalam penelitian tersebut.

Biasanya jenis statistik ini melihat pada jenis ukuran nilai-nilai tengah atau central tendency, ukuran variabilitas, dan ukuran posisi relatif. Ukuran tendensi sentral digunakan untuk menentukan nilai yang umum (kecenderungan) dari suatu kelompok nilai. Ukuran variabilitas menunjukkan bagaimanakah penyebaran nilai dalam suatu kelompok. Data-data yang sudah dikumpulkan melalui

penelitian tersebut akan lebih menarik pembaca jikalau disajikan di dalam bentuk grafik atau tabel, sehingga cepat difahami oleh orang lain (pembaca) untuk berkomunikasi melalui angka dan bukan dalam bentuk angka mentah atau "Raw data".

Banyak macam ragamnya ukuran tendensi sentral (*central tendency*) antara lain ukuran ini memberi kemudahan bagi peneliti untuk membuat deskripsi dari sekelompok data dengan suatu angka. Angka hasil perhitungan ukuran tendensi sentral menyajikan nilai rata-rata atau nilai umum dari subyek atau kelompok. Jenis atau macam ukuran tendensi sentral tersebut antara lain adalah : Mean (rerata); Median, Modus, Kuartil, Desil, dan Persentil. Selain ukuran-ukuran tersebut, orang akan lebih suka melihat tampilan grafik, karena grafik yang dikenal di dalam statistik deskriptif yaitu grafik Poligon, grafik Histogram atau sering kali disebut dengan chart dan yang terakhir adalah grafik ogive (Bandingkan dengan Anto Dayan, Pengantar metode Statistik Deskriptif, 1973).

Sedangkan untuk ukuran variabilitas dalam statistik deskriptif yang sering dijumpai adalah : range, deviasi dan standar deviasi. Deviasi standar atau standar deviasi ini yang paling sering dipergunakan dibandingkan dengan ukuran variabilitas yang lain. *Range* adalah beda antara nilai tertinggi dengan nilai terendah pada suatu distribusi dan ditentukan dengan cara mengurangi nilai-nilai ekstrim tersebut (nilai tertinggi dikurangi dengan nilai terendah). Sayangnya *range* bukan merupakan ukuran variabilitas yang akurat, tetapi merupakan ukuran variabilitas yang sifatnya kasar dan tingkat akurasi rendah, sehingga jarang dipergunakan di dalam analisis sosial. Deviasi merupakan perbedaan ukuran dari nilai mentah (*raw data*) atau dari sebuah ukuran tertentu dan kemudian dikurangi dengan nilai reratanya, sedangkan standar deviasi merupakan ukuran dispersi yang memiliki standar tertentu.

1. Mean (rerata)

Mean adalah rata-rata aritmatik nilai-nilai, dihitung dengan cara menjumlahkan semua nilai dan membagi total tersebut dengan sejumlah data (subyek penelitian) baik sebagai individu maupun kelompok.

Diantara sekian banyak jenis nilai tengah yang sering dipakai di dalam analisis sosial adalah Mean atau rerata. Sayangnya nilai mean ini sangat dipengaruhi oleh adanya nilai ekstrem, baik ekstrem kiri maupun ekstrem kanan. Nilai ekstrem kiri adalah nilai yang rendah atau sangat rendah jika dibandingkan dengan nilai lain pada umumnya, demikian juga nilai ekstrem kanan adalah nilai yang terlalu tinggi dibandingkan dengan nilai-nilai lainnya dari hasil pengumpulan data tersebut. Nilai-nilai ekstrem baik ekstrem kiri maupun ekstrem kanan ini seringkali menyesatkan hasil nilai rata-rata yang dihitung, sehingga tidak memberikan gambaran yang sesungguhnya terhadap sekumpulan data hasil penelitian. Oleh karena itu penggunaan nilai tengah, seperti mean (rerata) sebaiknya dikontrol dengan menghitung nilai deviasinya. Jika nilai deviasinya tinggi, maka pembaca atau pemakai hasil penelitian tersebut harus hati-hati, karena penyimpangan dari hasil perhitungan datanya cukup besar.

Tingginya nilai penyimpangan ini dapat menunjukkan sampai sejauh mana persebaran data yang dikumpulkan oleh peneliti. Semakin tinggi nilai deviasinya, maka persebaran datanya semakin lebar, dan tentunya hal tersebut berlaku sebaliknya. Setidaknya dengan melakukan kontrol pada nilai rerata tersebut pengguna hasil penelitian atau pembaca dapat lebih memahami hasil yang disajikan oleh peneliti.

Jika peneliti tidak melakukan kontrol terhadap nilai rerata, maka sebaiknya nilai-nilai ekstrem tersebut dikeluarkan terlebih dahulu sebelum reratanya dihitung. Dengan begitu, maka hasil mean atau rerata lebih menunjukkan realitas yang sesungguhnya, dan jikalau data-data tersebut digambarkan ke dalam bentuk curve, akan mendekati normal (seperti lonceng). Curve ini pada langkah selanjutnya akan sangat bermanfaat tatkala peneliti melakukan analisis yang tidak lagi bersifat deskriptif, tetapi yang lebih bersifat eksplanatif. Pada penelitian eksplanatif curve normal menjadi persyaratan penting untuk dapat memilih jenis statistik apa yang akan dipergunakan di dalam menganalisis hasil penelitiannya. Karena jika peneliti akan mengoperasikan jenis statistik yang

parametrik, maka normalitas ini menjadi syarat yang harus dipenuhi agar hasil perhitungannya dapat dipertanggungjawabkan.

2. Median

Median, Kuartil, Desil dan Persentil lebih merupakan ukuran nilai tengah yang serumpun artinya bahwa ukuran-ukuran ini dipergunakan dalam memberikan pemahaman tentang pembagian distribusi saja, sehingga perbedaan dari pengertian ini hampir tidak berbeda. Median adalah bilangan pada suatu distribusi yang menjadi batas separo frekuensi atau data, sehingga dapat membagi suatu distribusi menjadi dua bagian yang sama besarnya, sedangkan kuartil adalah bilangan yang membagi suatu distribusi menjadi 4 bagian yang sama besarnya. Desil adalah bilangan yang membagi suatu distribusi menjadi 10 bagian yang sama, dan Persentil membagi menjadi 100 bagian yang sama. Oleh karena itu jika seseorang sudah dapat menghitung median, dengan sendirinya Kuartil kedua, Desil ke 5 dan Persentil ke 50 adalah sama persis. Rumus yang dipakai untuk menghitung nilai tengah yang serumpun ini hampir sama, yang berbeda hanyalah pembagiannya saja. Kalau median pembagiannya 2; Kuartil pembagiannya 4, Desil pembagiannya 10 dan persentil pembagiannya 100. Meskipun nilai tengah ini jarang diaplikasikan di dalam analisis sosial, namun setidaknya pemahaman tentang macam-macam nilai tengah perlu dilakukan, agar dalam pemilihan teknik-teknik statistik menjadi lebih tepat.

3. Mode (Modus)

Mode adalah nilai atau skor yang paling sering muncul dalam suatu distribusi. Atau nilai yang paling besar atau paling tinggi frekuensinya. Mode merupakan alat yang cepat untuk menaksir tendensi sentral suatu distribusi. Meskipun demikian mode merupakan alat penaksir yang kasar, kemungkinan dalam suatu distribusi dapat memiliki 2 atau bahkan 3 mode, karena mode tidak terpengaruh sama sekali oleh nilai yang ekstrem, baik ekstrem kanan maupun ekstrem kiri. Modus, di dalam penggunaannya biasanya dipakai untuk melihat harga yang sering muncul. Di dalam penyajian data dengan grafik, apakah poligon, ataukah histogram, nilai modus menjadi puncak dari grafik tersebut. Jika data yang diperoleh peneliti

cenderung normal, maka dalam distribusi tersebut hanya memiliki satu nilai modus. Jika dalam sebuah distribusi frekuensi ada nilai modus yang jumlahnya lebih dari satu, sudah bisa diduga bahwa data tersebut tidak mungkin bersifat normal (lonceng).

4. Ukuran Penyimpangan

Ada banyak jenis ukuran penyimpangan, mulai dari penyimpangan Mean, Median, dan penyimpangan Kuartil, namun yang paling sering dipergunakan di dalam menganalisa data di dalam penelitian sosial adalah "penyimpangan standar". Harga penyimpangan standar ini akan sangat bermanfaat jika dikaitkan dengan harga dari Meannya. Semakin besar atau tinggi harga penyimpangannya dari mean, maka data yang diperoleh semakin menyebar atau heterogen dan itu berlaku sebaliknya. Harga penyimpangan yang semakin kecil dapat menunjukkan bahwa nilai Mean atau reratanya semakin mendekati kenyataan dari sekelompok data yang sedang diukur atau dihitung. Dengan kata lain, data tersebut dapat menggambarkan keadaan populasinya dan kemungkinan besar distribusi tersebut jika digambar curvenya akan semakin mendekati curve normal (bentuk lonceng).

Sebuah curve dikatakan normal (seperti bentuk lonceng) apabila nilai Meannya = 1 dan nilai penyimpangan standarnya = 0 → ini curve yang dikatakan normal sempurna. Curve normal sempurna atau yang mendekati normal menjadi salah satu syarat penting dalam memilih dan menentukan teknik statistik yang akan dipergunakan untuk mengolah dan menganalisis data. Curve dapat mendekati normal apabila jumlah sampel yang diambil dalam suatu penelitian cukup besar jumlahnya, sehingga dapat merepresentasikan keadaan populasinya. Dengan kata lain, semakin besar sampel penelitiannya, maka akan semakin mendekati ciri-ciri populasinya dan sampel tersebut dapat dikategorikan semakin representatif.

C. STATISTIK EKSPLANATIF/INFERENSIAL

Statistik jenis ini merupakan metode yang sering dipakai dalam rangka pengambilan kesimpulan tentang populasi, berdasarkan sampel yang diambil. Melalui uji hipotesa yang telah

dirumuskan sebelumnya, penelitian ini berusaha untuk membuat kesimpulan. Untuk itu diperlukan parameter atau ukuran yang dapat dipakai untuk mengukur variabel sosial, sesuai dengan persyaratan dari masing-masing rumus yang ada harus dipenuhi sebelum mengoperasikan rumus tersebut. Adapun persyaratan yang harus dipenuhi dapat meliputi skala pengukurannya (nominal, ordinal, interval dan rasio), jenis datanya (diskrit atau continuum), bentuk curvenya (normal atau tidak) dan tentu saja juga jumlah sampelnya (kecil atau besar). Di dalam penelitian sosial, sampel penelitian dikelompokkan sebagai sampel kecil apabila jumlahnya kurang dari 30 orang, dan apabila jumlahnya telah melebihi dari jumlah ambang batas tersebut dikelompokkan sebagai sampel besar. Sudah barang tentu semakin besar jumlah sampelnya, maka tingkat presisinya semakin tinggi, dan tingkat kesalahannya (error) semakin kecil, dan kondisi tersebut berlaku sebaliknya.

Biasanya statistik yang akan dipakai dalam rangka menguji hipotesis penelitian, dan hipotesa tersebut terlebih dahulu harus dirumuskan dalam bentuk hipotesa Nol atau hipotesa nihil. Hipotesa Nol dirumuskan dengan kata "*tidak ada*" hubungan antara parameter pada populasi, atau "*tidak ada*" perbedaan atau "*tidak ada*" hubungan antara variabel bebas dengan variabel terganggunanya. Hal inimenunjukkan bahwa hubungan atau perbedaan yang diperoleh dari sampel itu adalah akibat dari kesalahan sampling. Oleh karena itu hipotesis Nol/Nihil ini sering juga disebut sebagai *hipotesa statistik*, karena munculnya hipotesis ini untuk keperluan uji statistik dan keberadaannya hanya untuk ditolak melalui uji statistik.

Peneliti yang dapat menolak hipotesa nol secara meyakinkan (*significant*), secara otomatis akan menerima hipotesa kerja. Kedua hipotesis tersebut "*saling berlawanan*" atau bertolak belakang, karena apabila peneliti dapat menolak hipotesis nol penolakan ini dapat memberikan dukungan terhadap konseptualisasi yang dibuat oleh peneliti. Sebaliknya peneliti yang gagal menolak hipotesa Nol/Nihil (*failed to rejection nol hypothesis*), ada beberapa kemungkinan yang dapat dipertimbangkan dan dapat dirunut kembali di mana letak

kesalahannya. Apakah kesalahan tersebut terletak pada teori dan konseptualisasinya, dalam pengambilan dan penentuan jumlah sampelnya, dalam pengukurannya, dalam analisisnya ataukah dalam perhitungannya yang tidak atau kurang teliti? Persoalan lain yang muncul adalah, apakah peneliti yang tidak mampu menolak hipotesa Nol/Nihil itu adalah peneliti yang gagal. Peneliti yang dapat membuktikan bahwa "*penelitiannya atau hipotesisnya tidak terbukti*" itu adalah sebuah hasil dari penelitian. Hanya saja apabila dalam menguji hipotesis peneliti tidak dapat menolak hipotesis Nol/Nihil ada faktor penting yang berpengaruh tetapi tidak termasuk atau tidak diperhitungkan di dalam diagram matriksnya.

Mengapa bisa demikian ? Karena perilaku sosial atau fenomena sosial itu sangat kompleks sifatnya. Perilaku manusia tertentu dipengaruhi oleh banyak faktor yang tidak mudah diduga dan tidak mudah diukur dengan menggunakan ukuran angka seperti yang dilakukan oleh para peneliti ilmuwan eksakta. Dalam penelitian survei, biasanya ada keterbatasan peneliti di dalam menguraikan dan mengukur secara keseluruhan faktor-faktor sosial tersebut. Ada beberapa fenomena atau perilaku manusia yang tidak mudah atau bahkan tidak dapat ditanyakan dengan menggunakan kuesioner (angket) dan kemudian diukur dengan ukuran angka. Ada kalanya data bisa berupa uraian responden yang diidentifikasi dan kemudian diklasifikasikan, baru bisa diangkakan (diberi score), tetapi hanya bisa dituangkan melalui cerita atau penggambaran peneliti terhadap fenomena yang menjadi perhatian peneliti. Yang jelas, cerita responden (informan) tidak dapat "*ditelan mentah*" sebagai realitas, namun perlu dimaknai dan diinterpretasikan terlebih dahulu agar menjadi sebuah pemahaman penting untuk menganalisis fenomena yang sedang dicari jawabnya. Dengan kata lain, cerita responden merupakan data mentah yang membutuhkan interpretasi dari peneliti agar data tersebut menjadi lebih bermakna.

Apakah semua penelitian survei memerlukan uji statistik ? Jawabnya bisa "**ya**" tetapi juga bisa "**tidak**". "Ya", apabila model penelitiannya menggunakan beberapa macam variabel penelitian yang satu sama lain saling berhubungan. Hubungan tersebut perlu

dicari kebenarannya di dalam kenyataan empiris. "Tidak", apabila jenis penelitian yang dilakukan sifatnya deskriptif, sehingga jika harus menggunakan statistik, cukup dengan statistik yang deskriptif pula sifatnya (statistik jenis ini telah diuraikan di depan) dan atau peneliti cukup membuat deskriptif terhadap data yang dikumpulkan dengan menggunakan prosentase "*one way*" yaitu prosentase tunggal yang tidak mencerminkan hubungan antar variabel penelitian. Misalnya, ada 60 % responden remaja/mahasiswa yang senang "*ngabuburit*". Untuk memberikan penjelasannya membutuhkan data kualitatif yang akurat, sehingga data prosentase yang dibuat lebih dapat menjelaskan realitas sosial yang sedang menjadi perhatian peneliti.

Dalam upaya untuk menguji hipotesis dan menolak hipotesis nol/nihil tersebut, peneliti memerlukan suatu tes atau nilai untuk mengetahui signifikannya tidaknya sebuah hasil perhitungan, yaitu dengan membandingkan antara hasil statistik yang diperoleh dari perhitungan dengan hasil angka yang ada di dalam tabel yang sudah disediakan oleh para pakar dalam bidang ini yang biasanya ada di dalam buku-buku statistik. Angka di dalam tabel itu sering disebut dengan "nilai kritis" (critic value). Hasil perhitungan sebaik dan secanggih apapun, tidak akan ada artinya jika tidak dikonsultasikan dengan tabel statistik. Agar supaya peneliti dapat mengkonsultasikan dan mencocokkan hasil perhitungan yang diperoleh dengan angka yang ada di dalam tabel statistik, peneliti harus menentukan terlebih dahulu tingkat signifikansi yang diinginkan, yang menunjukkan seberapa besar peneliti membuat kesalahan. Pada umumnya para peneliti menggunakan tingkat kesalahan sebesar 5 % atau $< 0,05$ sebagai standar penolakan dan 1 % atau $< 0,01$. Itu artinya kesalahan yang dilakukan peneliti sama dengan atau kurang dari 5 % dan sama dengan dan atau kurang dari 1 %. Khusus untuk penelitian sosial ada toleransi kesalahan maksimum sampai dengan 10 % atau $< 0,10$, sedangkan hal ini tidak bisa diterapkan di dalam ilmu kedokteran atau ilmu eksakta yang lain. Selain menentukan tingkat kesalahannya, peneliti juga harus mengetahui derajat kebebasannya (*degrees of freedom*) agar dapat menemukan angka yang pas (cocok)

untuk mengetahui mampu tidaknya peneliti menolak hipotesa Nol/Nihilnya.

Apabila peneliti mampu menolak hipotesis nol pada derajat kesalahan sebesar 5 % (0,05), maka hipotesis kerja diterima pada batas kepastian benar (significant) sebesar 95 %. Hal ini berarti bahwa setiap 100 kali peneliti mengambil sampel, maka sebanyak 5 kali melakukan kesalahan terhadap sampel yang diambilnya. Sebagai contoh : hasil perhitungan korelasi Product Moment (R_{xy}) sebesar 0,625, dengan kesalahan 5 % dan derajat kebebasan (D.B) $n=48$ (diperoleh dari jumlah sampel yang diteliti dikurangi dengan jumlah variabelnya, misalnya sampelnya sebanyak 50 responden dan 2 variabel yang digunakan) maka besarnya angka kritis (critic value) atau besarnya angka korelasi yang ada dalam tabel adalah : 0,273. Secara statistis, peneliti dapat menolak hipotesa Nol/Nihil, karena hasil hitung ($R_{xy} = 0,625$) . dari hasil korelasi yang ada di dalam tabel ($R_{xy} 0,05$, d.b.=48=0,273). Berarti hasil perhitungan tersebut "significant" artinya, ada korelasi atau hubungan antara 2 variabel yang sedang dihitung. Jikalau peneliti tidak mampu menolak hipotesa Nol/Nihil maka hubungan yang ada hanya terjadi secara kebetulan saja.

Statistik inferensial dapat dikelompokkan menjadi 2 kelompok, yaitu statistik "non parametrik dan statistik yang parametrik". Statistik non parametrik tidak membutuhkan parameter atau ukuran yang tinggi, yang canggih, yang *sophisticated*, tetapi parameter yang dipersyaratkan hanya sederhana dan mudah dipenuhi oleh para peneliti sosial. Penggunaan statistik non parametrik harus memenuhi syarat seperti misalnya : skala pengukurannya bisa nominal dan atau ordinal, jenis datanya diskrit, curvenya tidak dituntut harus normal dan lain-lain. Sebaliknya, penggunaan jenis statistik parametrik membuuthkan persyaratan yang lebih rumit, lebih canggih, lebih *sophisticated*, dengan persyaratan : skala pengukurannya harus interval dan atau rasio, jenis datanya harus continuum dan linear, serta curvenya harus normal (berbentuk lonceng). Di bawah ini kedua jenis statistik tersebut akan diuraikan secara lebih terinci.

1. STATISTIK NON PARAMETRIK

Ada banyak macam ragamnya statistik jenis ini antara lain adalah: *Chi Square*, *Wilcoxon Test*, *Sign Test* (uji tanda), *Cochran Q Test*, *Korelasi Spearman*, *Kedall Rank Correlation Coefficient*, dll (lihat tabel yang ada dalam bukunya Sidney Siegel, atau buku-buku lain yang sejenisnya). Dari beberapa jenis dan rumus-rumus statistik yang tersedia, yang paling sederhana tetapi populer dan sering dipergunakan oleh peneliti sosial maupun peneliti eksakta adalah X^2 (Chi-Square atau Kai-Kuadrat). Chi-Square biasanya dipergunakan untuk melihat ada tidaknya perbedaan dari dua kelompok sampel. Apabila dengan menggunakan Chi-Square tersebut peneliti dapat menolak hipotesa Nol/ Nihil, maka perhitungan tersebut sebaiknya diteruskan dengan menggunakan *Contingency Coefficient*. *Contingency Coefficient* ini dimaksudkan untuk mengetahui ada tidaknya hubungan yang sifatnya kualitatif pada kedua variabel yang sedang dihitung dan ingin dibuktikan kebenarannya secara empiris.

Cara menghitung Chi Square sangatlah mudah, dan persyaratan yang dibutuhkan pun sangat sederhana. Agar supaya data-data yang diperoleh, dapat dibuat ke dalam bentuk tabel silang dan kemudian dapat dihitung dengan menggunakan rumus Chi-Square diperlukan syarat:

- Skalanya pengukuran boleh Nominal, maupun Ordinal, atau yang satu Nominal dan yang lainnya Ordinal dan keduanya mempunyai skala ordinal
- Distribusi datanya kalau dibuat dalam bentuk kurve tidak harus normal
- Setiap sel dalam tabel silang jumlah frekuensi yang diharapkan (F_e) tidak boleh kurang dari 5

Apabila ternyata tabel silangnya tidak memenuhi syarat-syarat seperti tersebut di atas maka Chi-Square tidak perlu dihitung, sebaiknya data tersebut diinterpretasikan dengan menggunakan prosentase tabel silang (cross tabulation), di mana setiap baris atau kolom yang dipilih diberi bobot 100 % (pemilihan arah prosentase tergantung pada apa yang akan diinterpretasikan). Biasanya tabel silang ini untuk melihat ada tidaknya perbedaan options dari

kelompok sampel. Misalnya: apakah ada perbedaan antara kelompok laki-laki dengan kelompok perempuan terhadap gaya hidup konsumtif? Contoh dapat dikembangkan dalam kasus apapun, sesuai dengan apa yang sedang menjadi perhatian peneliti. Meskipun data-data tersebut berwujud angka, namun data tersebut lebih bersifat kualitatif, oleh karena itu perlu penjelasan lebih lanjut dengan ditopang data-data kualitatif.

Jika peneliti dapat menggunakan rumus Chi-Square, agar supaya dapat menolak dan atau menerima hipotesis Nol/Nihil, maka hasil perhitungan tersebut harus dikonsultasikan dengan tabel yang telah disediakan oleh pembuat rumus (biasanya para ahli statistik dan matematika, karena di dalam mengukur signifikansi tabel menggunakan perhitungan probabilitas yang sangat rumit). Jika perhitungannya menggunakan komputer, maka secara otomatis komputer tersebut akan menunjukkan tingkat signifikansinya. Biasanya di dalam komputer ada tanda bintang di belakang hasil perhitungannya. Bintang 1 menunjukkan tingkat signifikansi 95 %, dengan kesalahan 5 %, bintang 2 dengan tingkat signifikansi 99 % dengan kesalahan 1 % dan bintang 3 dengan signifikansi 99,9 % dengan kesalahan 0,001 % (jika kesalahannya mencapai 10 % biasanya tidak terdapat tanda bintangnya). Jadi para peneliti sosial tinggal mencocokkan hasil perhitungannya dengan tabel yang ada. Ingat: sebelum mencocokkan dengan tabel, peneliti harus menghitung derajat kebebasannya (degrees of freedom)-nya terlebih dahulu untuk melihat sampai batas mana hasil perhitungannya dapat dikatakan signifikan (berarti). Adapun untuk menghitung derajat kebebasan Chi-Square dengan rumus sebagai berikut:

$$DB = (b-1)(k-1)$$

b adalah baris atau *k* adalah kolom dalam tabel silang tersebut

Biasanya dalam mengukur taraf signifikansi tersebut dengan kriteria sebagai berikut:

- a. Jika hasil hitungannya lebih tinggi dari hasil yang ada di dalam tabel, maka hasil perhitungan Chi-Square signifikan atau hasil tersebut mempunyai arti

- b. Sebaliknya apabila hasil hitungannya lebih kecil dari hasil tabelnya, maka kesimpulannya adalah tidak signifikan

Jika hasil hitungannya signifikan, maka hipotesa Nol/ Nihil yang menyatakan "tidak ada perbedaan" atau "tidak ada hubungan" **ditolak**, dan secara otomatis hipotesa kerjanya **diterima**. Sebaliknya jika hasil hitungannya tidak signifikan, maka hipotesa Nol/ Nihil diterima dan hipotesa kerjanya ditolak.

Jika Chi Square signifikan, maka peneliti dapat melanjutkan untuk menghitung Assosiasi (Cass). Assosiasi ini yang akan menunjukkan ada tidaknya hubungan dari kedua variabel, tetapi hubungan yang ada sifatnya kualitatif. Hal tersebut karena persyaratan yang dibutuhkan oleh Chi Square hanya skala Nominal dan atau Ordinal, dan ukuran kurvenya tidak harus normal, sehingga hubungan yang dihitung dengan assosiasi tersebut lebih bersifat kualitatif. Mengapa demikian? Karena hasil assosiasi hanya dapat menyatakan ada tidaknya hubungan dalam tingkatan atau gradasi tertentu. Gradasi hubungan tersebut hanya dapat dinyatakan tinggi, sedang atau rendah saja dan tidak dapat menunjukkan hubungan dengan koefisien korelasinya. Untuk mengetahui tinggi rendahnya hubungan assosiasi diperlukan perhitungan C maximum, yang rumusnya sebagai berikut:

$$C_{\max} = \sqrt{\frac{M-1}{M}} \quad \text{dimana } M \text{ adalah baris atau kolom dalam tabulasi silang yang jumlahnya kecil}$$

Jika hasil Cass lebih tinggi dari setengah C maximum maka hubungannya tinggi

Jika hasil Cass sama dengan setengahnya C maximum maka hubungannya sedang.

Jika hasil Cass lebih rendah dari setengah C maximum maka hubungannya rendah

Dengan kata lain, Cass dan C maximum adalah untuk mengetahui tinggi rendahnya gradasi dari hubungan kedua variabel yang sedang dihitung, apakah hubungan tersebut berada pada gradasi tinggi, sedang, ataukah rendah.

Teknik statistik non parametrik lain yang biasa dipergunakan oleh peneliti sosial, khususnya Sosiologi, adalah jenis Korelasi Rank Spearman (Korelasi Berjenjang). Korelasi ini biasanya dipakai untuk menghitung korelasi dengan menggunakan sampel yang kecil (< 30 responden). Persyaratan untuk menghitung korelasi ini hampir sama dengan persyaratan untuk menghitung Chi-Square, hanya saja rumus korelasi ini setidaknya harus memiliki skala pengukuran yang ordinal dan tidak bisa dengan skala nimal. Cara menghitungnya pun mudah, yaitu dengan mengurutkan skor dari urutan yang paling rendah sampai ke urutan yang paling tinggi, atau sebaliknya. Ada 2 macam rumus yang dapat dipakai, yaitu rumus yang nilainya atau skor tidak ada *ties* (nilai kembar) dan yang ada nilai atau skor yang sama (*ties*-nya).

$$R_s = 1 - \sqrt{\frac{6.d^2}{N^3 - N}}$$

Rumus² = (Rumus R_s tanpa *ties* atau tanpa nilai kembar)

Keterangan

R_s = korelasi Rank Spearman

d^2 = deviasi, yaitu perbedaan antara nilai dari variabel X dengan nilai variabel Y

N = jumlah sampel (subyek penelitian)

6 = nilai konstan (rumus)

Untuk mengetahui apakah hasil korelasi Rank Spearman tersebut signifikan atau tidak (dapat menolak hipotesa Nol/ Nihil atau tidak), sama seperti pada penggunaan Chi-Square, juga harus berkonsultasi dengan tabel korelasi yang disediakan di dalam setiap buku statistik. Agar dapat mencocokkan hasil perhitungan dengan hasil tabelnya harus diketahui terlebih dahulu derajat kebebasannya

² Rumus R_s yang ada nilai kembarnya (*ties*) ada di dalam buku-buku statistik non parametrik, karena ruang dalam tulisan ini sangat terbatas sehingga tidak memungkinkan untuk memuat semua rumus yang bisa diaplikasikan di dalam penelitian sosial. Dan bagi yang tertarik dengan rumus ini bisa membaca buku.

(db). Untuk menghitung db dari korelasi adalah: $n-2$, dimana n adalah jumlah subyek yang diteliti dan 2 adalah jumlah variabel yang ada dalam perhitungan korelasi tersebut. Karena dalam contoh ini hanya ada 2 variabel (*independent dan dependent variable*) maka dikurangi 2. Tetapi jika jumlah variabel yang dihitung dalam korelasi itu ada 3 variabel, maka db-nya adalah = $n-3$ (karena korelasi tersebut terdiri dari 3 variabel) dst.

Bagi mahasiswa atau peneliti yang berminat menerapkan teknik statistik dalam memperkuat kesimpulannya dapat mempelajari teknik-teknik lain yang ada di dalam banyak buku statistik sosial, baik yang ditulis dan diterbitkan oleh penerbit internasional, nasional maupun lokal. Pada statistik non parametrik ini dapat dipelajari juga rumus Kendall Tau. Rumus-rumus statistik non parametrik tidak hanya dapat digunakan untuk menghitung 2 variabel saja, tetapi juga bisa untuk menghitung dengan 3 variabel. Agar dapat memilih rumus mana yang paling cocok dengan penelitiannya perlu dipertimbangkan beberapa hal antara lain:

- Apa jenis data yang diperoleh dari penelitian
- Perlu dikenali apa skala pengukurannya
- Bagaimana komposisi persebaran datanya, apakah datanya mengelompok atautkah tersebar
- Bagaimana komposisi jawaban responden jika dilihat dari tabel frekuensi *one-way* (jawaban yang baik, jika komposisinya seimbang). Jika jawaban responden mengelompok pada salah satu opsi jawaban, sebaiknya dihindari, karena datanya ada yang bias

Dengan mengetahui hal tersebut, maka peneliti dapat dengan mudah memilih jenis statistik yang paling cocok dengan data yang diperoleh. Dengan begitu kesalahan dapat dieliminir sedemikian rupa, agar hasil penelitiannya menjadi lebih baik (akurat).

2. PENGGUNAAN STATISTIK PARAMETRIK

Yang termasuk dengan statistik parametrik adalah statistik yang membutuhkan ukuran atau parameter tertentu. Yang termasuk di dalam jenis statistik ini macamnya banyak, rumus dan cara

penghitungannya lebih canggih, karena itu peneliti yang ingin menggunakan statistik jenis ini membutuhkan persyaratan yang lebih canggih pula. Yang termasuk di dalam jenis statistik ini antara lain: Korelasi Product Moment, Korelasi Partial, Korelasi Ganda, Regresi (*step wise* dan *multiple* regresi), Faktor Analysis, Path Analysis, Analisis Varian maupun Covarian, dan masih banyak lagi jenis yang lainnya. Yang telah disebutkan ini paling tidak yang sering dioperasikan di dalam penelitian sosial, khususnya Sosiologi. Beberapa rumus statistik di atas biasanya ada di dalam komputer dan sudah sering di *setting* di dalam program SPSS serta sudah banyak dikenal secara umum.

a. KORELASI PRODUCT MOMENT

Sebagai gambaran, peneliti yang ingin menggunakan rumus Korelasi Product Moment misalnya, harus dapat memenuhi syarat yang dibutuhkan oleh rumus tersebut, yang antara lain adalah:

- 1) Skalanya harus interval atau rasio
- 2) Memiliki distribusi normal
- 3) Datanya linear

Semua jenis korelasi, apapun rumusnya, apakah itu Korelasi Rs (yang termasuk di dalam kelompok statistik non parametrik), Korelasi Product Moment atau Korelasi Pearsons, sampai pada Korelasi Partial, maupun Korelasi Ganda, semuanya memiliki fungsi yang sama, karena sama-sama untuk menghitung hubungan antara 2 variabel (*independent* dan *dependent variable*) dan atau lebih dari 2 variabel (*independent*, *intervening*, dan *dependent variable*). Hubungan tersebut dapat dicerminkan dari besar kecilnya koefisien korelasi hasil perhitungannya, apakah perhitungannya dengan manual atautkah dengan menggunakan komputer seharusnya hasilnya sama. Sama seperti pada jenis statistik yang lain, untuk dapat menyimpulkan hubungan kedua variabel tersebut, hasil perhitungannya harus dikonsultasikan dengan tabel yang ada di dalam buku statistik. Jika hasil hitungnya lebih besar dari

harga kritik, maka hubungan tersebut signifikan, dengan demikian peneliti dapat menolak hipotesa nol dan menerima hipotesa kerja.

Namun, perlu diingat bahwa ada perbedaan di antara beberapa jenis statistik tersebut dan perbedaan itu terletak pada persyaratan yang harus dipenuhi, karena adanya perbedaan jenis dan pengelompokannya. Oleh karena itu, dalam memilih rumus mana yang akan dipergunakan tergantung dari data seperti apa yang dimiliki oleh peneliti. Sudah barang tentu semakin banyak (canggih) persyaratan yang dapat dipenuhi, maka pengukurannya akan menjadi semakin baik, peneliti lebih leluasa untuk memilih jenis statistik yang diinginkan, karena ukuran yang dimiliki semakin teliti.

Hubungan yang ada di dalam korelasi biasanya bersifat asimetris. Di dalam penelitian sosial, sebaiknya menghindari hubungan variabel yang sifatnya simetris, karena hubungan seperti ini tidak dapat mengetahui dengan pasti mana variabel yang mempengaruhi dan mana variabel yang dipengaruhi. Hubungan simetris atau timbal balik ini menimbulkan kebingungan peneliti maupun pembacanya. Dari hubungan tersebut juga sulit diketahui variabel mana yang mempunyai pengaruh lebih besar terhadap yang lain. Jika peneliti ingin mengetahui hal ini, maka harus menghitung lagi dengan regresi atau dengan Path Analysis.³

b. KORELASI PARTIAL DAN KORELASI GANDA

Seorang peneliti yang ingin menghitung korelasi dengan variabel yang jumlahnya lebih dari 2, dapat dipergunakan Korelasi Partial dan atau Korelasi Ganda. Perbedaan kedua korelasi tersebut terletak pada fungsi dari masing-masing yang

³ ada beberapa rumus yang dapat dipakai untuk menghitung korelasi untuk lebih jelasnya dapat dilihat di beberapa buku statistik yang ada, agar peneliti dapat memilih yang paling pas bagi penelitiannya

akan dicapai. Korelasi Partial berfungsi untuk mengetahui apakah hubungan antara variabel bebas dengan variabel terikat itu bersifat murni atautkah semu. Hadirnya variabel antara atau *intervening variable* dimaksudkan untuk mengontrol hubungan antara variabel bebas dengan variabel terikat. Dengan dimasukkannya variabel ketiga ini apakah akan mempengaruhi tingkat signifikansi dari perhitungan Korelasi Product Moment (2 variabel). Sedangkan Korelasi Ganda lebih banyak melihat hubungan antara variabel terikat dengan variabel lain secara bersama-sama (variabel bebas dan variabel antara atau *intervening variable*).

Dalam menerapkan rumus Korelasi Partial dibutuhkan persyaratan seperti yang berlaku di dalam Korelasi Product Moment, karena Korelasi Partial hanya dapat dihitung apabila peneliti terlebih dahulu harus sudah menghitung Korelasi Product Moment. Korelasi Product Moment ini merupakan korelasi tahap awal yang pada akhirnya dapat dikembangkan dalam penghitungan korelasi yang lebih canggih, yang bersifat multivariat sedangkan Korelasi Product Moment adalah korelasi yang bersifat bivariat. Dengan lain perkataan, bahwa koefisien Korelasi Product Moment merupakan prasyarat untuk dapat menghitung Korelasi Partial dan Korelasi Ganda. Selain itu, dapat juga dijadikan dasar di dalam menghitung Regresi maupun untuk menghitung Path Analysis. Adapun rumus dari Korelasi Product Moment adalah sebagai berikut:⁴

$$\text{Rumus Korelasi Partial } R_{12.3} = R_{12.3} = \frac{r_{12} - r_{13} \cdot r_{23}}{\sqrt{\{1 - r_{13}\}\{1 - r_{23}\}}} \quad (\text{untuk 3 variabel})$$

⁴ Untuk rumus Korelasi Partial yang lebih dari 4 variabel, dapat dilihat dalam buku-buku statistik dan langsung saja dihitung dengan bantuan komputer (karena hitungannya semakin rumit). Demikian juga untuk menghitung Korelasi Ganda dan Residunya, Regresi dan Path Analysis.

$$\text{Rumus Korelasi Partial } R_{12.34} = R_{12.34} = \frac{r_{12.3} - r_{14.3} \cdot r_{24.3}}{\sqrt{\{1 - r_{14.3}\}\{1 - r_{24.3}\}}}$$

(untuk 4 variabel)

Hasil perhitungan dari Korelasi Partial ini hanya akan dapat bermakna apabila hasil koefisiennya kemudian dikonsultasikan dengan tabel statistik yang ada. Setelah itu barulah peneliti melakukan perbandingan hasil akhir dari kedua koefisien korelasi tersebut untuk mengetahui apakah hubungan yang terjadi antara variabel bebas dengan variabel terikat bersifat murni atautkah semu.

Hubungan antara variabel bebas dengan variabel terikat dikatakan murni apabila hasil koefisien korelasi dari hubungan antara 2 variabel (R_{xy}) dibandingkan dengan koefisien korelasi dari hubungan 3 variabel ($R_{xy.z}$) mengalami penurunan. Terjadinya penurunan koefisien dari hasil perhitungan hubungan 3 variabel tersebut menjadi tidak signifikan, maka hubungan tersebut "murni". Misalnya hasil koefisien korelasi 2 variabel sebesar 0,57 signifikan pada taraf 99 % dibandingkan dengan hasil koefisien korelasi dari 3 variabel dengan koefisien 0,53 dan setelah dikonsultasikan dengan tabel korelasimasih tetap signifikan, maka hubungan yang ada pada kedua variabel tersebut yaitu variabel bebas dan terikat adalah "murni". Hal tersebut berarti bahwa kehadiran variabel ke 3 atau variabel antara (*intervening variable*) tidak mempunyai peran yang signifikan di dalam merubah hubungan kedua variabel tersebut.

Sebaliknya apabila hasil koefisien korelasi 2 variabel dibandingkan dengan hasil koefisien korelasi 3 variabel dan ternyata hasilnya mengalami penurunan sehingga menjadi tidak signifikan, maka hubungan antara variabel bebas dengan variabel terikat adalah "semu". Misalnya hasil $R_{xy} = 0,57$ dan hasil $R_{xy.z} = 0,23$. Kedua koefisien korelasi tersebut dikonsultasikan dengan tabel, ternyata $R_{xy} = 0,57$ signifikan pada taraf 95 % dan $R_{xy.z} = 0,23$ hasilnya menjadi tidak signifikan → hubungan antara R_{xy} adalah "semu". Dalam hal

ini variable intervening berperan sebagai variabel kontrol terhadap hubungan antara variabel bebas dengan variabel terikat.

D. Kesimpulan

Penggunaan statistik di dalam penelitian sosial, khususnya Sosiologi sangatlah mudah dan sederhana, sepanjang peneliti yang bersangkutan memahami anggapan dasar dari rumus statistik yang dipilih. Pemilihan rumus tertentu didasarkan pada jenis dan kualitas data yang telah dimiliki, agar pemilihan rumusnya menjadi lebih tepat dan akurat. Ketepatan dan akurasi di dalam pemilihan teknik dan rumus statistik sangat menentukan kualitas penelitian yang sedang dilakukan, sehingga kesimpulan yang dirumuskan menjadi lebih akurat (tepat) juga.

Statistik hanyalah alat untuk membuat data menjadi lebih "bermakna". Pemaknaan itu sendiri banyak caranya, dan pemaknaan yang benar tidak hanya dari besar kecilnya angka yang diperoleh dari hasil perhitungan, baik secara manual, maupun dengan komputer yang biasanya menggunakan program SPSS, tetapi juga harus ditopang dengan konsep dan teori yang dipergunakan. Selain itu dapat juga dengan hasil-hasil penelitian orang lain yang senada yang pernah dilakukan, dan dengan pergulatan antara konsep dengan kenyataan empiris yang seimbang. Hanya dengan cara seperti ini, maka hasil penelitian tersebut dapat diterima oleh orang lain, akan menjadi bermakna bagi orang lain, terutama dalam rangka pengembangan ilmu pengetahuan dan bagi perumus kebijakan.

Adalah benar bahwa survei dengan menggunakan analisis kuantitatif hanyalah salah satu alat dari sekian banyak alat yang dapat diakses, dipelajari, dan diimplementasikan oleh peneliti sosial, khususnya Sosiologi, masih ada metode atau teknik lain yang dapat dipakai untuk mengungkapkan realitas sosial yang terjadi. Peneliti dapat saja memilih, teknik mana yang paling pas bagi penelitiannya.

Namun, perlu diingat bahwa salah satu keunggulan dari penggunaan metode penelitian kuantitatif, yang biasanya menerapkan rumus-rumus statistik, dapat dipergunakan untuk

melakukan prediksi. Selain itu, hasil penelitian dan kesimpulannya dapat digeneralisasikan terhadap kondisi populasinya. Dengan demikian, manfaat dari hasil penelitian ini dapat dinikmati oleh banyak orang, karena dapat diperuntukkan pada masyarakat lain dengan kondisi populasi yang hampir sama.

Daftar Pustaka

- Anto, Dayan: *Pengantar Metode Statistik Deskriptif*, 1973, Jakarta, LP3ES.
- Babbie, Earl: *The Practice of Social Research*, 1989, Wadsworth Publishing Company, Belmont, California, A Division of Wadsworth, Inc.
- Guilford, Joy Paul and Fruchter, Benjamin,: *Fundamental Statistic in Psychology and Education*, 1968, Tokyo, McGraw - Hill, Kogakusha Ltd.
- Neuman, Lawrence.W: *Social Research Methods, Qualitative an Quantitative Approaches*, 2003, SA, Library of Congress Cataloging-in-Publication Data.
- Siegel, Sydney: *Nonparametric Statistics, For Behavioral Sciences*, 1968, New York, Toronto, London, McGraw-Hill, Book Company.
- Singarimbun, Masri dan Effendi, Sofian: *Metode Penelitian Survei*, 1987, Jakarta, LP3ES.