

Dari Redaksi

Selamat berjumpa dengan edisi khusus kali ini, meskipun keterlambatan terbit untuk sementara masih merupakan norma daripada satu kasus pengecualian. Buletin edisi ini menyajikan makalah-makalah mengenai uji signifikansi hipotesis nihil, satu topik yang seringkali menjadi bahan perdebatan di kalangan ahli psikologi. Tulisan-tulisan tersebut sesungguhnya pernah dimuat dalam *proceeding* seminar statistika "Kontroversi penerapan prinsip-prinsip statistika" di tahun 2000. Tulisan-tulisan dalam *proceeding* itu dimuat kembali dengan harapan akan lebih banyak orang yang mendapatkan informasi seputar perdebatan tentang uji signifikansi hipotesis nihil.

Redaksi

SEPUTAR KONTROVERSI UJI SIGNIFIKANSI HIPOTESIS NIHIL

Dicky Hastjarjo

PERDEBATAN TENTANG PERLU-TIDAKNYA UJI SIGNIFIKANSI HIPOTESIS NIHIL

Sejumlah pakar menyatakan bahwa psikologi, ilmu perilaku, dan ilmu sosial mengalami rintangan dan hambatan untuk maju. Salah satu kambing hitamnya adalah kebiasaan berat psikologi dan ilmu sosial dalam menggunakan uji signifikansi hipotesis nihil (Cohen, 1994). Kebiasaan yang sulit dihilangkan ini membuat psikologi seperti serangkaian kudesa atau lorong buntu metodologis (Lofus, 1996) sehingga psikologi hanya sibuk memutar rodanya tanpa banyak kemajuan (Cohen, 1994), atau maju selambat jalannya gletser (Hunter, 1997).

Semenjak diperkenalkan 75 tahun yang lalu oleh Fisher, banyak kritikan telah dialamakan kepada pengujian signifikansi hipotesis nihil. Kritikan-kritikan tersebut diantaranya adalah (a) penerapan yang keliru penalaran silogistik deduktif, dan (b) hipotesis nihil selalu salah (Cohen, 1994), serta (a) tingkat kesalahan uji signifikansi bukan sebesar 5% melainkan 60%, dan (b) hipotesis nihil hampir selalu salah (Hunter, 1997). Salah satu alternatif yang disodorkan untuk menggantikan uji signifikansi adalah penggunaan interval kepercayaan (*confidence interval*) (Cohen, 1994; Hunter, 1997). Pada kenyataannya, uji signifikansi hipotesis nihil tetap menjadi metode utama analisis data dalam ilmu-ilmu sosial meskipun telah mendapat kritikan bertubi-tubi (Tyron, 1998). Cohen (1994) menggaribarkan uji signifikansi hipotesis nihil bak seorang maharaja tanpa busana yang tanpa rasa malu berlari berputar-putar dalam waktu lama. Situasi ini terjadi oleh karena buku teks menekankan pengujian hipotesis nihil, para dosen mensosialisasikan kepada mahasiswanya penggunaan (*use*) dan penyalahgunaan (*abuse*) uji signifikansi, serta para editor jurnal meminta pengarang naskah untuk melaporkan lebih banyak uji signifikansi (Shrout, 1997).

BULETIN PSIKOLOGI

Fakultas Psikologi Universitas Gadjah Mada

Bulaksumur, Yogyakarta 55281 - Telp.: (0274) 550435, Fax. 550436

- *Peindung*: Dekan Fakultas Psikologi UGM • *Pemangungjawab*: Mohamad Bachroni
- *Pemimpin Umum*: Sugyanto • *Sekretaris*: Sywi Dewajani • *Bendahara*: Esti Hayu Purnamaningsih
- *Dewan Redaksi*: Th. Dicky Hastjarjo (ketua), Johana Endang Pratiwiasari, Avin Fadilla Helmi, Falurochman, Subandi, Budi Andyanie • *Penerbitan*: Ira Paramasri
- *Sirkulasi*: Aisah Indati.

Buletin Psikologi diterbitkan dua kali dalam setahun, yang memuat artikel-artikel ilmiah populer, resensi buku, analisis dan diskusi, pembahasan isu mengenai psikologi dan bidang-bidang yang relevan baik dari segi teori maupun aplikasi praktisnya. Artikel untuk dimuat harus diketik dalam spasi rangkap disertai riwayat pendidikan serta pekerjaan.

Harga langganan Rp. 20.000,-/tahun untuk 2 kali penerbitan

Menyadari situasi yang demikian ini dalam konvensi *American Psychological Society (APS)* tahun 1996 terdapat simposium yang menyangkut kontroversi uji signifikansi hipotesis nihil. Makalah-makalah, baik yang pro maupun kontra uji signifikansi, diterbitkan dalam rubrik khusus di jurnal *Psychological Science* edisi Januari 1997. Sejumlah pengarang makalah menyarankan untuk melarang sama sekali pengujian signifikansi hipotesis nihil (Hunter, 1997, Shrout, 1997). Saran untuk melarang uji signifikansi hipotesis nihil ini juga telah banyak dikemukakan diantaranya oleh Cohen (1994), Hubbard (1995), dan McGrath (1998). Hubbard (1995) menegaskan bahwa lonceng kematian pengujian signifikansi hanya akan mungkin terjadi jika ada perubahan kurikulum di program pascasarjana dengan lebih menekankan penelitian replikasi serta menghilangkan perilaku tergiat-giat kepada nilai $p < 0.05$. Judul makalah Hunter (1997) misalnya juga sangat provokatif yakni "Dibutuhkan: satu pelajaran terhadap uji signifikansi". Seorang editor jurnal psikologi pernah menghimbau kepada para pengirim naskah untuk melakukan uji signifikansi (Loftus, 1996, catatan kaki). Seorang editor sebuah jurnal ilmu kesehatan masyarakat bahkan telah meminta calon pengirim naskah hasil penelitian untuk menghilangkan semua acuan mengenai pengujian hipotesis dan signifikansi atau jika tidak mau melakukan hal itu penulis dipersilakan mempublikasikan naskahnya di jurnal lain (Shrout, 1997).

Sejumlah pakar masih mempunyai sikap menyetujui praktik uji signifikansi hipotesis nihil (Abelson, 1997; Baril & Cannon, 1995; Haggan, 1997), atau sekurang-kurangnya tidak setuju terhadap pelarangan uji signifikansi hipotesis nihil (Harris, 1997; Estes, 1997). Baril dan Cannon (1995) tidak setuju terhadap kesimpulan makalah Cohen (1994) diatas dan kedua ahli itu menegaskan bahwa mereka tidak salah menaruh keyakinan pada model hipotesis nihil. Hagen (1997) menyatakan bahwa logika uji signifikansi hipotesis nihil bersifat elegan, luar biasa kreatif, melekat secara dalam pada metode statistika inferensial, serta tidak akan pernah mampu kita cerai meskipun suatu saat kita memutuskan ingin berpisah dengannya. Sebagai penutup makalahnya yang berjudul "Pujian bagi uji signifikansi hipotesis nihil", Hagen (1997) menulis "Ketika kita menggunakan uji signifikansi hipotesis nihil, saya kira kebanyakan dari kita akan terus menggunakannya, janganlah lupa untuk merayakan kecemerlangannya sesekali waktu". Pada saat rubrik khusus mengenai kontroversi pengujian signifikansi hipotesis nihil dimuat di jurnal *Psychological Science* edisi tahun 1997, maka editor jurnal itu mengumumkan bahwa dia tidak akan melarang uji signifikansi statistik selama masa baktinya.

TANGGAPAN TERHADAP PERDEBATAN MENGENAI Uji SIGNIFIKANSI HIPOTESIS NIHIL

Sebagai tanggapan terhadap kontroversi mengenai uji signifikansi, terutama terhadap makalah Cohen di *American Psychologist* yang terbit di tahun 1994, maka *Board of Scientific Affairs (BSA)* dari *American Psychological Association (APA)* pada tahun 1996 membentuk satu komisi yang diberi nama *Task Force on Statistical Inference (TFSI)* (Wilkinson & TFSL, 1999). Salah satu tugas TFSI adalah untuk memberi penjelasan mengenai sejumlah isu kontroversial menyangkut aplikasi statistik termasuk uji signifikansi dan alternatif-alternatifnya. TFSL diketuai oleh Robert Rosenthal, Robert Abelson, dan Jacob Cohen. Perlu diketahui bahwa sebagaimana dalam uraian diatas maka Robert Abelson adalah ahli psikologi yang bersikap pro uji signifikansi, sedang Jacob Cohen adalah ahli psikologi yang telah berulang kali mengugut keberadaan uji signifikansi hipotesis nihil dalam artikel yang ditulisnya. Sayang bahwa Jacob Cohen tidak sempat melihat perkembangan kerja TFSL selanjutnya karena ia meninggal dunia pada awal tahun 1998. TFSL juga beranggotakan ahli statistik, pengajar statistik, editor jurnal, pengarang buku statistik, ahli komputer dan sepeuh yang bijaksana. Sementara itu, tokoh-tokoh seperti Lee Cronbach, Paul Meehl, Frederick Mosteller dan John Tukey bertindak sebagai penasihat senior dari TFSL.

TFSL merekomendasikan untuk merevisi *APA Publication Manual* edisi tahun 1994 terutama pada Bab Statistik. Rekomendasi tersebut telah disosialisasikan terlebih dahulu kepada publik psikologi dalam bentuk satu artikel sebelum revisi dilakukan pada pedoman publikasi APA itu (Wilkinson, & TFSL, 1999). Rekomendasi yang diajukan mencakup metode (rancangan, pemilihan subjek, dan pengukuran), hasil (komplikasi, dan analisis) dan diskusi (interpretasi, dan kesimpulan).

Rekomendasi yang berkaitan dengan analisis meliputi sejumlah aspek, diantaranya adalah (a) Memilih satu analisis yang cukup secara minimal. Prinsip parsimoni atau dikenal sebagai prinsip *Occam's razor* berlaku bagi pemilihan metode selain bagi pemilihan teori. Jika analisis yang lebih sederhana dapat menjawab permasalahan, maka tidak perlu memakai analisis yang lebih kompleks sekedar untuk memberikan kesan baik bagi pembaca. Ingat gagasan Fisher: "*Begin with an idea. Then pick a method*" (1999, h. 598). Kita dalam melakukan penelitian tidak jarang bersikap terbalik "Menentukan metode dulu, baru cari permasalahan". Misalnya, oleh karena kita baru saja belajar analisis regresi, maka sebelum membuat proposal penelitian kita berkata pada diri sendiri "permasalahan apa yang bisa saya pecahkan dengan analisis regresi", (b) Memilih program komputer. Memilih program *software*

statistik tertentu yang baik itu perlu dilakukan oleh peneliti. Namun ada hal yang lebih penting lagi, yakni peneliti harus memverifikasi hasil analisis statistik, memahami apa maksudnya, dan mengerti bagaimana menghitungnya. *TFSI* juga menambahkan: "Do not report statistics to a greater precision than is supported by your data simply because they are printed that way by the program" (1999, h. 598). Tidak jarang juga disekeliling kita diemukkan sejumlah laporan penelitian yang diampiri berpuluh-puluh halaman hasil perhitungan statistik, meskipun ketika ditanya arti dari angka-angka tadi si pembuat laporan kurang mengerti, (c) Mempertimbangkan asumsi-asumsi. *TFSI* menulis: "You should take efforts to assure that the underlying assumptions required for the analysis are reasonable given the data", (d) Pengujian hipotesis. Melaporkan nilai p aktual akan lebih bagus daripada membuat keputusan dikotomis niolak-menerima hipotesis, dan akan lebih bagus lagi melaporkan interval kepercayaan. Senantiasa memberikan *effect size estimate* jika melaporkan nilai p. *TFSI* menulis: "Never use the unfortunate expression 'accept the null hypothesis'". *Always provide some effect size estimate when reporting a p value*" (1999, h. 399), (e) Menyajikan *Effect Size*. Senantiasa menyajikan *effect size* untuk hasil-hasil pokok/utama, (f) Menyajikan *Interval Estimates*. Perkiraan interval sebaiknya disajikan untuk setiap *effect size* untuk hasil-hasil utama.

Sebenarnya masih ada beberapa aspek lagi yang direkomendasikan untuk bab analisis pada khususnya atau untuk bab metode pada umumnya, namun untuk kepentingan makalah ini hanya disajikan materi yang relevan saja. Memang rekomendasi *TFSI* tidak sampai pada keputusan melarang praktek uji signifikansi hipotesis nihil, sebab *TFSI* memandang ada sejumlah contoh yang melawan kritikan terhadap uji signifikansi (1999, h. 603; lihat juga Abelson, 1997). Akan tetapi rekomendasi *TFSI* menegaskan bahwa laporan penelitian sebaiknya menyatakan juga interval kepercayaan dan *effect size*. Disamping itu, *TFSI* juga merekomendasikan untuk menyatakan signifikansi teoretis (*theoretical/practical significance*) (1999, h. 603; lihat juga Kirk, 1996 dan makalah Swediat & Bastari dalam edisi ini). Dampak dari rekomendasi ini sudah terlihat, yakni sekitar 13 jurnal telah mengharuskan para pengirim naskah untuk melaporkan *effect size* (Thompson, 2000). Jurnal itu sebagian diantaranya adalah *Contemporary - Educational Psychology*, *Educational and Psychological Measurement*, *Journal of Applied Psychology*, *Journal of Consulting and Clinical Psychology*, *Journal of Experimental Education*, *Journal of Learning and Disabilities*, *Language Learning*, *Measurement and Evaluation in Counseling and Development*. Satu pertanyaan yang timbul adalah apakah rekomendasi itu atau nantinya revisi pedoman publikasi APA akan berdampak juga bagi praktek-praktek publikasi penelitian ilmiah di jurnal psikologi di Indonesia atau secara umum pada

Bagian Psikologi Umum dan Eksperimen, Fakultas Psikologi, UGM, Yogyakarta

Thompson, B. 2000. Where the Social Sciences Are Headed: Away from NHST, toward Effect Size & Replicability Evidence. *Hip://acc.lamu.edu/bbr/61471*.

Tyron, W.W. 1998. The inscrutable null hypothesis. *American Psychologist*, vol.53, 796.

Wilkinson, L, & Task Force on Statistical Inference. 1999. Statistical Methods in *Psychology Journals: Guidelines and Explanations. American Psychologist*, vol. 54, 8, 594-604.

PENGUJIAN HIPOTESIS NIHIL: UJI SIGNIFIKANSI DAN INTERVAL KEPERCAYAAN

Djemari Mardapi

PENDAHULUAN

1. Sikap orang terhadap statistik bermacam-macam, ada yang senang, kagum, ada yang sinis, curiga, dan tidak senang. Bahkan ada yang mengatakan bahwa orang mudah berbohong dengan statistik. Seseorang yang sebagian anggota badannya di almari pendingin dan sebagian lainnya di alat pemanas, disimpulkan keadaan orang tersebut normal. Tentunya kesimpulan ini hanya melihat harga rerata saja tanpa memperhatikan besarnya *range* dan simpangan baku.
2. Sebagian orang menghindari statistik karena bias frisafat, kekakuan statistik, atau karena salah konsep terhadap disiplin statistik (Glass & Hopkins, 1984). Mereka cenderung bekerja berdasarkan tradisi, intuisi, kekuasaan atau rasa kewajaran. Namun lambat laun, orang memerlukan adanya cara yang sistematis dan objektif yang dapat digunakan pada penelitian empiris, yaitu statistik.
3. Statistik adalah ilmu menjawab pertanyaan berdasarkan data empiris (Olson, 1987). Definisi ini menunjukkan bahwa statistik adalah alat atau teknik untuk menyajikan dan mengolah data yang diperoleh melalui pengukuran sehingga menjadi informasi yang berguna bagi pembuat kebijakan. Sesuai fungsinya sebagai alat, statistik tidak dapat mengubah data yang tidak menjadi baik.
4. Data yang diolah dengan statistik harus mengandung kesalahan yang kecil. Untuk memperoleh data yang memiliki kesalahan yang kecil, alat ukur yang digunakan harus sah dan handal, dan cara penggunaan alat ukur juga harus baku. Data dengan banyak kesalahan tidak akan dapat diubah menjadi baik dengan teknik statistik apapun. Sebagai alat, kualitas keluarannya tergantung