LAPORAN PENELITIAN

VARIABEL RANDOM DAN TEORI PENGUKURAN

Oleh :

Drs. Herman . MA

Universitas Terbuka
Lembaga Penelitian
Pusat Studi Indonesia
2002

Lembar Pengesahan
Laporan Penelitian PSI-UT

1.	a. Judul Penelitian	: Variabel Random dan Teori Pengukuran
	b. Bidang Penelitian	: Statistika/ Matematika
2.	Ketua Peneliti	
	a. Nama lengkap dan gelar	: Drs. Herman, MA
	b. NIP	: 131-628-379
	c. Golongan kepangkatan	: III/d
	d. Jabatan Fungsional	: Lektor
	e. Fakultas/Unit Kerja	: FMIPA/Jurusan Statistika
3.	Lama Penelitian	: 8 bulan
5.	Biaya Penelitian	: Rp.1.995.000,- (Satu juta sembilan ratus sembilan puluh lima ribu rupiah).

Pondok Cabe, 15 Januari 2002

Mengetahui,
Dekan FMIPA

Dr. Ir. D. Djokosetiyanto
NIP 130 536 671

Ketua Peneliti

Drs. Herman, MA
NIP 131 628 379

Menyetujui,
Kepala PSI-UT

Dr. Tian Belawati
NIP 131 569 974

Menyetujui,
Ketua Lembaga
Penelitian UT

WBP Simanjuntak, MEd, PhD
NIP 130 212 017

Abstrak

Mempelajari statistika tingkat lanjut menurut beberapa ahli haruslah menggunakan integral Lebesgue. Menurut mereka, tanpa memiliki pengetahuan ini akan sulitlah membaca teori statistika tingkat lanjut. Lehmann (1986) misalnya menyatakan "kerangka matematik untuk teori keputusan statistik disediakan oleh teori peluang dimana teori ini didasarkan pada teori tentang pengukuran dan integrasi". Teori tentang integrasi ini tidak akan lepas dari konsep tentang σ-aljabar. Popoulis (1984) dan Simonnet (1996) juga menyatakan hal yang sama.

Oleh karena itu, penelitian ini mencoba menguraikan beberapa konsep statistika yang ada pada buku-buku teks pada umumnya dengan pendekatan teori pengukuran dan integrasi. Uraian yang dibuat masih sangat teoritis. Di sini diuraikan konsep peluang dan variabel random dengan memakai teori pengukuran. Teori pengukuran memang lebih rumit sebab pendekatannya abstrak sekali. Akan tetapi secara matematis teori klasik dapat diturunkan dengan pendekatan ini, sebagai hal khusus dari fungsi-fungsi yang terbentuk.

Ucapan Terima Kasih

Sehubungan dengan kegiatan penelitian ini , penulis mengucapkan banyak terima kasih kepada pihak-pihak yang sudah membantu kelancaran penyelesaian penelitian ini.

Pertama-tama penulis mengucapkan terima kasih kepada Dekan FMIPA Universitas Terbuka, Dr. Ir. D. Djokosetiyanto atas perhatian yang sudah diberikan. Juga terimakasih banyak untuk Dr. Djati Kerami yang sudah meluangkan waktunya dalam memberikan kritik dan saran untuk penelitian ini.

Selain itu juga penulis mengucapkan banyak terima kasih kepada Ketua Lembaga Penelitian, WBP Simanjuntak, Med, PhD dan Kepala Pusat Studi Indonesia, Dr. Tian Belawati yang telah memberi kesempatan kepada penulis untuk mengikuti seleksi pembiayaan penelitian di Universitas Terbuka ini.

Tidak lupa juga penulis mengucapkan terima kasih kepada rekan-rekan yang sudah membantu penelitian ini, baik berupa kritik dan saran ataupun peminjaman buku-buku yang diperlukan.

DAFTAR ISI

BAB I

Pendahuluan

Latar Belakang

Statistik sangat banyak digunakan oleh para peneliti. Penggunaannya terutama untuk pengambilan keputusan (inferensi). Selain itu statistik juga digunakan untuk menjelaskan suatu "keadaan" berdasarkan data yang ada (deskripsi). Statistik sangat erat hubungannya dengan data. Hampir setiap pengguna akan memakai statistik bila mengolah data. Cara peneliti dalam memperoleh data lebih dari satu macam; ada yang mengukur langsung pada eksperimennya, ada pula yang menggunakan kuesioner, atau mungkin menggunakan cara-cara lainnya. Biasanya, peneliti akan lebih senang bila data yang dimiliki berbentuk numerik (angka).

Data yang akan dipelajari secara keseluruhan dikenal dengan nama populasi. Istilah populasi ini dikenal juga dengan nama ruang sampel (*sample space*). Selain itu, ada pula istilah sampel random, variabel random dan peluang (*probability*). Para peneliti tentunya sudah kenal dengan istilah-istilah di atas. Akan tetapi, bagaimana sebetulnya kaitan antara ruang sampel, variabel random dan peluang ? Untuk itu, penelitian ini mencoba melihat kaitan diantara mereka dengan menggunakan teori pengukuran. Menurut Lehmann (1986), bila ingin mempelajari statistika secara lebih mendalam maka gunakanlah teori pengukuran. Teori ini dikembangkan dengan menggunakan integral Lebesgue. Simonnet (1996) mengungkapkan hal yang sama, bahwasanya teori integrasi memainkan peran yang penting dalam matematika murni ataupun terapan.

Perumusan Masalah

- Variabel random adalah istilah yang mungkin tidak asing lagi bagi para peneliti. Akan tetapi apakah sebetulnya variabel random itu ? Bagaimana kaitan antara kasus diskrit dan kontinu ?
- Bagaimanakah konsep variabel random dapat diperoleh dengan menggunakan integral Lebesgue ?
- Bagaimana konsep peluang masuk ke konsep variabel random ?

Tujuan Penelitian

- Mempelajari konsep teori pengukuran dan kaitannya dengan konsep variabel random serta konsep peluang.

Metodologi Penelitian

Konsep tentang σ-aljabar (*σ-field*) akan diberikan pada tinjauan pustaka. Setelah itu akan didiskusikan tentang konsep peluang. Konsep peluang ini akan didekati dari dua arah, yaitu pendekatan klasik dan pendekatan dengan menggunakan konsep σ-aljabar (*σ-field*). Langkah selanjutnya adalah menurunkan pengertian ruang sampel dari konsep σ-aljabar atau *σ-field*. Pengertian tentang σ-aljabar ini juga dilengkapi dengan konsep-konsep lainnya seperti konsep borel. Konsep Borel inilah yang akan membedakan antara data diskrit dan data kontinu. Penjelasan konsep variabel random juga akan diturunkan berdasarkan konsep σ-aljabar. Setelah itu dipelajari kaitan antara ruang sampel, variabel random dan peluang.

Manfaat Penelitian

Penelitian ini diharapkan dapat membuat pengertian tentang ruang sampel, variabel random dan peluang secara lebih mendalam dan lebih teoritis. Selain itu, diharapkan penelitian ini dapat menjembatani konsep tentang variabel random yang dipelajari sebelumnya dengan konsep variabel random yang diperoleh dari teori pengukuran.

Tinjauan Pustaka

Menurut Lehmann (1986) kerangka matematis untuk teori keputusan yang berdasarkan statistik disediakan oleh teori peluang. Teori peluang ini dibangun berdasarkan teori pengukuran dan integrasi. Kecenderungan teori peluang adalah pada keadaan-keadaan yang mungkin diperoleh dari hasil-hasil percobaan (*outcome*) yang berbeda. Keseluruhan dari hasil-hasil ini ditunjukkan secara abstrak oleh keseluruhan titik-titik yang ada di dalam suatu ruang S. Karena kejadian-kejadian (*events*) merupakan kesatuan dari hasil-hasil (*outcome*) yang diperoleh, maka mereka direpresentasikan oleh subset-subset S.

Notasi yang digunakan di sini adalah notasi yang biasanya dipakai pada himpunan. Gabungan dari dua himpunan C_1 dan C_2 ditunjukkan oleh $C_1 \cup C_2$, sedangkan irisannya ditunjukkan oleh $C_1 \cap C_2$, dan komplemen C ditunjukkan oleh $C^C = S - C$. Selain itu, himpunan kosong diberi simbol ∅. Perlu diperhatikan bahwa peluang P(C) dari kejadian C adalah suatu bilangan real antara 0 dan 1. Secara khusus dinyatakan bahwa

$$P(\emptyset) = 0 \text{ dan } P(S) = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

Peluang memiliki sifat "dapat dijumlahkan secara terbilang (*countable additivity*)",

$$P(\cup C_i) = \Sigma\, P(C_i) \text{ jika } C_i \cap C_j = \emptyset \text{ untuk semua } i \neq j \quad \ldots\ldots\ldots \quad (2)$$

Kalau seandainya harus memenuhi (2), maka himpunan fungsi-fungsi yang akan dibahas biasanya tak dapat didefinisikan untuk semua himpunan-bagian S. Sebagai contoh , bagaimana mungkin mendefinisikan daerah (*area*) untuk semua himpunan-bagian dari suatu unit luas pada suatu bidang. Disinilah letaknya perbedaan antara data diskrit dan data kontinu.

Contoh berikut ini adalah tentang variabel random untuk eksperimen berbentuk diskrit. Misalkan dibuat suatu eksperimen dengan melemparkan mata uang logam (koin), dimana ruang sample yang berkaitan dengan eksperimen itu adalah S. Ruang sampel S adalah suatu himpunan $\{c|$ dengan c adalah H atau T$\}$. H menunjukkan KEPALA (Head), sedangkan T menunjukkan EKOR (Tail). Misal X adalah suatu fungsi sedemikian hingga $X(c) = 0$ bila $c =$ T dan $X(c) = 1$ bilamana $c =$ H. Tampak bahwa X adalah fungsi bernilai real yang didefinisikan pada ruang sample S ke ruang real A, dengan $A = \{x|\ x = 0,\ 1\}$. Pada matakuliah pengantar Statistika, X inilah yang dikenal sebagai *variabel random*. Ruang yang berkaitan dengan X adalah himpunan $A = \{x|\ x = 0,\ 1\}$. Hogg & Craig (1978) mendefinisikan variabel random sebagai :

Definisi Diberikan suatu eksperimen random dengan ruang sample S. Suatu fungsi X, yang memetakan setiap elemen $c \in S$ ke satu-dan-hanya satu bilangan real $X(c) = x$ disebut sebagai variabel random. Ruang dari X adalah himpunan bilangan real $A = \{x|\ x = X(c),\ c \in S\}$.

Dapat saja terjadi himpunan S sudah memiliki elemen-elemen bilangan real. Dalam hal itu dituliskan $X(c) = c$. Dengan demikian $A = S$.

Pandang ruang sample S sebagai himpunan yang unsurnya adalah hasil yang mungkin terjadi dari suatu eksperimen random. Akan didefinisikan suatu himpunan fungsi $P(C)$ sedemikian hingga jika C adalah himpunan bagian dari S maka $P(C)$ adalah peluang bahwa hasil dari eksperimen random tersebut adalah C.

Definisi Bila $P(C)$ didefinisikan untuk jenis sub-himpunan dari ruang S, dan jika

(a) $P(C) \geq 0$,

(b) $P(C_1 \cup C_2 \cup C_3 \cup \ldots) = P(C_1) + P(C_2) + P(C_3) + \ldots$ dengan $C_i \cap C_j = \varnothing$, $i \neq j$.

(c) $P(S) = 1$,

maka $P(C)$ disebut *himpunan fungsi peluang* dari hasil eksperimen random. Untuk setiap $C \subset S$, bilangan $P(C)$ disebut peluang bahwa hasil dari eksperimen random adalah C, atau peluang dari kejadian C, atau ukuran peluang dari himpunan C.

Misalkan X adalah variabel random yang didefinisikan pada ruang sample S, dan misalkan A adalah ruang dari X. Misalkan A adalah himpunan bagian A. Mirip seperti dalam penggunaan "kejadian C", dengan $C \subset S$, juga akan digunakan "kejadian A". Peluang $P(C)$ untuk kejadian C sudah didefinisikan. Sekarang akan didefinisikan peluang untuk kejadian A yang ditunjukkan dengan Pr (X$\in$ A). Dengan A $\subset A$, misalkan $C \subset S$ sedemikian hingga $C = \{c|\ c\in S$ dan $X(c)\in$ A$\}$. Pr (X$\in$ A) adalah sama dengan $P(C)$, dengan $C = \{c|\ c\in S$ dan $X(c)\in$ A$\}$. Jadi Pr (X$\in$ A) adalah suatu peluang yang dimiliki oleh himpunan $A \subset A$ yang terkait dengan variabel random X. Peluang ini ditentukan oleh himpunan fungsi peluang P dan variabel random X, yang kadangkala ditunjukkan sebagai P_x(A). Dengan demikian, Pr (X$\in$ A) = P_x(A) = $P(C)$, dengan $C = \{c|\ c\in S$ dan $X(c)\in$ A$\}$.
Jadi, suatu variabel random X adalah suatu fungsi yang membawa peluang dari ruang sampel S ke ruang $A \subset$ R. Fungsi P_x(A) memenuhi kondisi a), b) dan c) pada definisi di atas.

Berikut ini akan diuraikan konsep tentang *σ-field* atau *σ-aljabar* yang berkaitan dengan pengertian terukur. Himpunan-himpunan dimana fungsi peluang P akan didefinisikan adalah terukur (*measurable*). Domain dari definisi P adalah termasuk himpunan C, komplemen C, sejumlah terbilang banyaknya kejadian-kejadian dari gabungan mereka. Dari (1) tampak bahwa himpunan terukur tersebut harus mengandung S.

Definisi Suatu kelas himpunan-himpunan yang mengandung S; tertutup dibawah operasi komplemen; serta tertutup dibawah gabungan terbilang disebut sebagai ***σ-field*** atau ***σ-aljabar***.

Kelas ini secara otomatis juga akan tertutup dibawah operasi irisan yang terbilang.

Definisi yang lebih rinci tentang *σ-field* akan diberikan berikut ini. Pandang suatu himpunan X, dan misalkan $\mathbf{X}$ adalah keluarga dari sub-himpunan sub-himpunan X. Keluarga ini sudah tentu mengandung himpunan kosong $\varnothing$ dan himpunan semesta X itu sendiri. Selain tertutup untuk operasi komplemen, $\mathbf{X}$ juga tertutup untuk operasi gabungan yang terbilang. Dari informasi ini, Bartle (1966) mendefinisikan σ-aljabar.

Definisi Keluarga $\mathbf{X}$ dari sub-himpunan sub-himpunan suatu himpunan X disebut suatu **σ-aljabar**, bilamana :

(i) $\varnothing, X \in \mathbf{X}$.

(ii) Jika $A \in \mathbf{X}$, maka komplemen $A^c = X - A$ adalah unsur dari $\mathbf{X}$.

(iii) Jika (A_n) adalah suatu barisan himpunan-himpunan dalam $\mathbf{X}$, maka gabungan $\cup_{n=1}^{\infty} A_n$ termasuk di dalam $\mathbf{X}$.

Pasangan terurut $(X, \mathbf{X})$ yang terdiri dari himpunan X dan σ-aljabar $\mathbf{X}$ dari sub-himpunan sub-himpunan X disebut **ruang terukur** (*measureable space*).

Contoh-contoh σ-aljabar

1. Misalkan X adalah suatu himpunan sebarang dan misalkan $\mathbf{X}$ adalah keluarga dari semua sub himpunan X.
2. Misalkan X himpunan sebarang dan $\mathbf{X}$ adalah terdiri dari $\varnothing$ dan X.
3. Misalkan $X = \{1, 2, 3, \ldots\}$ adalah himpunan bilangan asli dan misalkan $\mathbf{X}$ terdiri dari $\varnothing$, $\{1, 3, 5, \ldots\}$, $\{2, 4, 6, \ldots\}$, X.
4. Misalkan X adalah suatu himpunan tak terbilang dan $\mathbf{X}$ adalah kumpulan dari sub-himpunan sub-himpunan dari X yang terbilang atau memiliki komplemen yang terbilang.
5. Jika $\mathbf{X}_1$ dan $\mathbf{X}_2$ adalah dua σ-aljabar dari sub-himpuan sub-himpuan X, misalkan $\mathbf{X}_3$ adalah irisan dari $\mathbf{X}_1$ dan $\mathbf{X}_2$, maka $\mathbf{X}_3$ adalah juga σ-aljabar.
6. Misalkan $\boldsymbol{A}$ adalah kumpulan sub himpunan tak kosong dari X. Ada σ-aljabar terkecil dari sub himpunan sub-himpunan X yang mengandung $\boldsymbol{A}$. Untuk melihat ini, perhatikan keluarga dari semua sub himpunan sub himpunan X yang merupakan σ-aljabar yang mengandung $\boldsymbol{A}$. Irisan dari semua σ-aljabar σ-aljabar yang mengandung $\boldsymbol{A}$ ini adalah juga suatu σ-aljabar yang mengandung $\boldsymbol{A}$. Bentuk σ-aljabar terkecil ini disebut juga sebagai **σ-aljabar yang dibangun oleh $\boldsymbol{A}$**.
7. Misalkan X adalah himpunan $\boldsymbol{R}$ dari bilangan bilangan real. **Aljabar Borel** adalah σ-aljabar $\boldsymbol{B}$ yang dibangun oleh semua interval buka (a,b) dalam $\boldsymbol{R}$. Definisi tentang ini akan dituliskan juga nanti.
8. Misalkan X adalah himpunan $\overline{\boldsymbol{R}}$ dari bilangan bilangan real yang diperluas (*extended real numbers*). Jika E adalah sub himpunan Borel dari $\boldsymbol{R}$, misalkan $E_1 = E \cup \{-\infty\}$, $E_2 = E \cup \{\infty\}$, $E_3 = E \cup \{-\infty, \infty\}$, dan misalkan $\overline{\boldsymbol{B}}$ adalah kumpulan dari semua himpunan-himpunan E, E_1, E_2, E_3 sebagai E bervariasi

terhadap B. Sudah diselidiki bahwa $\overline{B}$ adalah suatu σ-aljabar dan disebut sebagai **Aljabar Borel diperluas**.

Definisi Misalkan X adalah himpunan bilangan real **R**, maka **aljabar Borel** adalah σ-aljabar B yang dibentuk oleh semua interval buka (a,b) di dalam **R**.

Definisi Suatu fungsi f dari X ke **R** disebut terukur jika untuk setiap bilangan real α, himpunan $\{x \in X | f(x) > \alpha\}$ adalah unsur dari X.

Kerangka matematis untuk teori keputusan statistik didasari oleh teori peluang, yang memiliki fondasi pada teori pengukuran dan integrasi (Lehmann, 1986). Teori peluang berkaitan dengan situasi dari hasil pengamatan. Keseluruhan dari hasil yang mungkin diperoleh ditunjukkan secara abstrak oleh keseluruhan titik-titik dalam ruang X. Hal ini karena, kejadian-kejadian yang akan diteliti merupakan kesatuan dari hasil-hasil pengamatan. Menurut Lehmann (1986) bilamana suatu fungsi μ dikenakan pada X sehingga nilai $\mu(\varnothing) = 0$ dan $\mu(X) = 1$, maka μ disebut ukuran peluang. Secara lebih umum, μ terhingga jika $\mu(X) < \infty$ dan σ-terhingga jika ada $C_1, C_2, \ldots$ di dalam X (yang saling bebas) sehingga $\cup C_i = X$ dan $\mu(C_i) < \infty$ untuk $i = 1, 2, \ldots$.
$P(\cup C_i) = \Sigma P(C_i)$, jika $C_i \cap C_j = \varnothing$ untuk semua $i \neq j$.

Bab II

Pembahasan

Teori peluang berkaitan dengan rata-rata dari suatu phenomena besar yang terjadi secara sekuensial, ataupun secara simultan (Papoulis, 1984), misalnya emisi electron, panggilan telepon, deteksi radar, kontrol kualitas, kegagalan sistem, turbulensi, kebisingan, tingkat kelahiran dan tingkat kematian, teori antrian dan lain sebagainya.

Sudah diteliti bahwa pada lapangan-lapangan kerja yang disebutkan di atas ataupun lapangan lainnya nilai rata-rata akan mendekati suatu nilai konstan bilamana jumlah observasi semakin bertambah. Nilai ini juga tidak akan berubah bila nilai rata-rata diteliti pada sebarang sub-barisan data yang sudah ditentukan sebelum eksperimen. Contohnya adalah persentase munculnya *muka* (head) pada penelitian satu koin adalah 0,5 atau suatu nilai konstan lainnya. Nilai yang sama akan akan muncul bila perhitungan dilakukan pada setiap lemparan koin yang ke-4. Dengan jalur berfikir seperti di atas akan dijelaskan tentang nilai rata-rata yang berkaitan dengan peluang dari kejadian. Misalkan suatu kejadian A dengan peluang terjadinya adalah bilangan $P(A)$. Bilangan $P(A)$ ini oleh Papoulis (1984) diartikan sebagai berikut :

> Jika suatu eksperimen dilakukan n kali dan kejadian A terjadi n_A kali, maka, dengan "derajat hampir pasti" , frekuensi relatif n_A/n dari kejadian A akan "mendekati" nilai $P(A)$. Sehingga $P(A) \cong n_A/n$ asalkan nilai n "cukup besar".

Perhatikan bahwa istilah "derajat hampir pasti", "mendekati" dan "cukup besar" tidak memiliki arti yang jelas (Papoulis,1984). Tetapi kelemahan istilah ini agak sulit dihindarkan. Kalau dipaksa untuk mendefinisikan istilah "derajat hampir pasti" dalam terminologi peluang, maka yang akan diperoleh hanyalah penundaan kesimpulan (yang tak dapat dihindarkan) bahwa peluang hanya berkaitan dengan fenomena fisik dalam bentuk-bentuk tidak eksak saja. Padahal teori adalah disiplin yang eksak yang dikembangkan secara logis dari aksioma-aksioma yang jernih dan jelas dan bila teori itu diterapkan ke problem yang nyata maka ia akan berfungsi. Untuk memperjelas arti dari peluang perhatikan contoh berikut ini yang akan membedakan observasi, deduksi dan prediksi (Papoulis,1984). Ada tiga langkah yang akan dilakukan.

Sebagai langkah pertama, akan ditentukan secara fisik suatu peluang proses tak-pasti $P(A_t)$ dari kejadian pasti A_t. Proses ini dapat dilakukan berdasarkan hubungan 1-1 antara peluang dengan observasi. Peluang data $P(A_t)$ akan sama dengan ratio observasi yaitu n_{A_i}/n. Hal yang sama dapat juga dilakukan dengan "alasan" kesimetrian yaitu bila dari N percobaan terdapat hasil N_A (banyaknya kejadian yang mungkin untuk A), maka peluang $P(A) \cong N_A/N$. Contohnya adalah bila suatu dadu dilemparkan (toss) 1000 kali dan angka "5" muncul sebanyak 203 kali maka peluang munculnya angka "5" adalah 0,2. Tetapi bila menggunakan alasan kesimetrian bahwa kalau dadu tersebut *fair* maka peluang munculnya angka "5" adalah sebesar 1/6.

Sebagai langkah kedua adalah mengasumsikan secara konseptual bahwa peluang memenuhi axioma-axioma tertentu. Secara deduktif, dari informasi peluang kejadian pasti A_t adalah $P(A_t)$ maka dapat diturunkan bahwa kejadian pasti yang lain B_j memiliki peluang $P(B_j)$. Sebagai contoh adalah : bilamana dadu tersebut *fair* (artinya peluang masing-masing angka untuk muncul adalah 1/6) maka peluang munculnya bilangan genap adalah 3/6.

Langkah yang ketiga adalah membuat prediksi secara fisik berdasarkan bilangan P(B_j) yang terbentuk. Hal ini dapat dilakukan dengan dasar hubungan 1-1, tetapi kerjanya dibalik. Jadi bila dilakukan eksperimen n kali dan kejadian B muncul n_B kali maka $n_B \cong n$ P(B). Contohnya, dadu yang *fair* dilemparkan 1000 kali, prediksi untuk muncul angka genap adalah 500 kali.

Dalam membuat solusi suatu masalah, ke tiga langkah di atas tidaklah dapat secara ketat dibedakan. Yang harus dapat dibedakan adalah antara data yang diperoleh secara empiris dan hasil yang diperoleh secara deduktif.

Langkah pertama dan ke tiga didasarkan pada alasan induktif. Misal akan ditentukan peluang munculnya kepala dari suatu koin. Berapa banyakkah koin tersebut akan dilemparkan : 100 kali atau 1000 kali ? Katakanlah koin akan dilempar 1000 kali dan rata-rata munculnya kepala adalah 0,48. Prediksi apa yang dapat dibuat berdasarkan informasi ini ? Dapatkah diduga bahwa untuk pelemparan koin sebanyak 1000 kali, maka kepala akan muncul sebanyak 480 kali ? Pertanyaan ini hanya dapat dijawab secara induktif.

Pendekatan aksiomatik untuk peluang menurut Papoulis (1984) didasarkan pada tiga postulat, tidak ada yang lain, yaitu :

- Peluang P(A) dari kejadian A adalah suatu bilangan tak negatif yang diberikan pada kejadian tersebut. Dengan perkataan lain, P(A) ≥ 0.
- Peluang dari suatu kejadian yang pasti terjadi adalah 1, P(S) = 1, S = semesta.
- Jika kejadian A dan B saling bebas (*mutually exclusive*) maka P($A \cup B$) = P(A) + P(B).

Menurut Papoulis (1984), pendekatan ini tampaknya adalah cara yang paling baik untuk memperkenalkan konsep peluang di tingkat elementer. Cara ini menekankan sifat deduktif suatu teori. Ia juga dapat menghindari kebingungan konsepsi. Pendekatan ini dapat diaplikasikan. Selain itu, ia juga merupakan pengantar awal untuk mempelajari peluang secara lebih mendalam.

Berikut ini akan diperkenalkan sedikit tentang frekuensi relatif. Pendekatan ini berdasarkan definisi berikut.

Peluang P(A) dari kejadian A adalah :

$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$; n_A adalah besarnya kejadian A dan n adalah besarnya percobaan.

Karena peluang digunakan untuk menjelaskan frekuensi relatif maka adalah lumrah bila mendefinisikannya sebagai limit dari frekuensi. Dengan demikian persoalan-persoalan yang muncul sebelum ini dapat dieliminasi. Walaupun konsep frekuensi relatif menjadi dasar dalam penggunaan peluang (langkah 1 dan 3), penggunaannya sebagai dasar penurunan teori pada langkah 2 dapat dipertanyakan. Secara fisik penggunaan n_A dan n dalam percobaan walaupun sangat besar tetapi tetaplah terbatas.

Sehingga perbandingan nilai mereka tidak dapat dijadikan limit. Kalau definisi frekuensi relatif digunakan untuk mendefiniskan P(A), maka bentuk "limit" haruslah diterima sebagai hipotesa, bukan sebagai bilangan yang dapat ditentukan dalam eksperimen.

Pendekatan Klasik untuk Peluang

Menurut Papoulis (1984), teori peluang berdasarkan definisi klasik sudah digunakan beberapa abad lamanya. Konsep ini pula yang digunakan saat ini untuk menentukan data probabilistik dan sebagai hipotesa. Pendekatan klasik itu akan dijelaskan berikut ini.

Berdasarkan definisi klasik, peluang P(A) dari kejadian A ditentukan sebelumnya tanpa percobaan yang sebenarnya. Bentuknya adalah :

$$\mathrm{P}(A) = \frac{N_A}{N}$$

N_A adalah besarnya hasil yang mungkin untuk kejadian A dan
N adalah besarnya hasil yang mungkin terjadi

Pada percobaan pelemparan dadu maka banyaknya hasil yang mungkin terjadi adalah enam (6). Sedangkan banyak kejadian munculnya bilangan genap adalah tiga (3). Dengan demikian peluang munculya bilangan genap P(*genap*) = 3/6. Seperti yang sudah dijelaskan sebelumnya, sigifikansi angka N dan N_A tidaklah terlalu jelas. Untuk itu perhatikan contoh berikut ini.

> Suatu eksperimen dilakukan dengan melempar dua buah dadu. Akan dicari besarnya peluang p bahwa jumlah angka yang muncul adalah 7.

Untuk menggunakan pendekatan klasik ada beberapa cara yang dapat dilakukan :

- Menentukan dulu hasil yang mungkin yaitu : 2, 3, ..., 12. Dari sini terlihat ada 11 macam angka yang mungkin muncul. Karena angka 7 hanya muncul satu kali, maka peluang munculnya dua dadu dengan jumlah 7 adalah 1/11. Hasil ini tentu saja salah.
- Tentukan pasangan yang mungkin muncul tanpa membedakan urutan kejadiannya. Dengan demikian terdapat $N = 21$ kejadian yang dapat muncul. Jumlah 7 yang mungkin muncul adalah (3,4), (5,2) dan (6,1), $N_A = 3$. Sehingga $p = 3/21 = 1/7$. Hasil inipun salah.
- Ke dua cara di atas salah karena hasil mereka bukanlah berasal dari kejadian yang memiliki peluang sama (*equally likely*). Untuk itu, haruslah diperhitungkan semua kejadian (pasangan) dengan membedakan urutan kejadian Dengan demikian $N = 36$. Banyaknya kejadian yang berjumlah 7 ada 6 yaitu, (3,4), (4,3), (2,5), (5,2) dan (1,6), (6,1). Karena itu, $p = 6/36 = 1/6$.

Contoh di atas memperlihatkan kebingungan yang dapat muncul pada penggunaan definisi klasik tentang peluang. Karena itu, Papoulis (1984) menyarankan sedikit modifikasi pada definisi itu sebagai berikut:

> Peluang suatu kejadian A sama dengan ratio dari banyaknya kejadian A yang mungkin muncul terhadap total kejadian asalkan setiap kejadian tersebut memiliki peluang yang sama besarnya.

Modifikasi definisi ini, menurut Papoulis (1984), tidaklah menghilangkan problem yang berkaitan dengan definisi klasik. Menurut Bernauli (1713), definisi klasik diperkenalkan sebagai akibat dari suatu "*prinsip alasan yang tidak cukup*". Bila tidak memiliki pengetahuan awal, haruslah diasumsikan bahwa kejadian-kejadian A_i memiliki peluang yang sama. Kesimpulan ini didasarkan pada pemahaman peluang secara subjektif sebagai ukuran dari pengetahuan seseorang tentang kejadian-kejadian A_i. Sehingga kalau memang kejadian-kejadian A_i tidak memiliki peluang yang sama, maka bila mengubah indeks A_i, maka akan dihasilkan peluang yang berbeda, tanpa mengubah pengetahuan yang dimiliki orang itu (Papoulis, 1984).

Pendekatan dengan Teori Pengukuran

Menurut Lehmann (1986), kalau berbicara mengenai peluang maka ada dua hal yang harus menjadi perhatian, yaitu ruang S dan σ-field $\boldsymbol{S}$. Ruang S menunjukkan hasil percobaan (*outcomes*) yang mungkin terjadi; dan σ-field $\boldsymbol{S}$ dari himpunan-bagian himpunan-bagian S menunjukkan kejadian-kejadian (*events*) yang peluangnya akan didefinisikan. Pasangan $(S,\boldsymbol{S})$ disebut **ruang terukur** (*measureable space*), dan elemen-elemen dari $\boldsymbol{S}$ secara bersama-sama membentuk **himpuan-himpunan terukur** (*measureable sets*). Suatu himpunan fungsi tak negatif μ yang didefinisikan pada $\boldsymbol{S}$ yang bersifat dapat dijumlahkan, dimana nilai $\mu(\varnothing) = 0$ disebut suatu **ukuran** (*measure*). Bila diberikan nilai μ (S) = 1, maka himpunan fungsi tak negatif tersebut adalah **ukuran peluang** (*probability measure*). Secara umum, μ terbatas bila $\mu(S) < \infty$ dan σ-terbatas bila ada $C_1, C_2, \ldots \in \boldsymbol{S}$ (yang saling lepas) $\ni \cup C_i = S$ dengan $\mu(C_i) < \infty$, untuk $i = 1, 2, \ldots$.

Dalam aplikasinya, peluang-peluang di $(S,\boldsymbol{S})$ mengacu pada eksperimen random (*random experiment*), dimana hasil-hasil yang mungkin diperoleh dari eksperimen adalah titik-titik $z \in S$. Menurut Lehmann (1986), dalam mencatat hasil-hasil eksperimen, orang tertarik pada aspek-aspek tertentu saja, biasanya berbentuk angka. Dalam bentuk fungsi, hal ini dapat diwakilkan oleh suatu fungsi T yang mengambil nilai-nilai dalam ruang τ.

Di dalam ruang τ fungsi seperti ini membentuk σ-field β' dari himpunan-himpunan B yang invers fungsinya, yaitu $C = T^{-1}(B) = \{z \mid z \in S, T(z) \in B\}$ berada di dalam $\boldsymbol{S}$. Untuk sebarang ukuran peluang P pada $(S, \boldsymbol{S})$ didefinisikan suatu ukuran peluang Q pada (τ, β') sebagai $Q(B) = P(T^{-1}(B))$.

Seringkali, terdapat σ-field β dari himpunan-himpunan di dalam τ sedemikian hingga peluang B akan terdefinisi jika dan hanya jika $B \in \beta$. Ini artinya $T^{-1}(B) \in \boldsymbol{S}$ untuk semua $B \in \beta$. Fungsi T dari $(S, \boldsymbol{S})$ ke dalam (τ, β) disebut sebagai $\boldsymbol{S}$-terukur.

Sebagai ilustrasi adalah kasus pengukuran tunggal dimana fungsi T bernilai real. Misalkan hasil pengukuran tunggal adalah X, dan misalkan $\boldsymbol{A}$ adalah klas dari himpunan-himpunan Borel pada garis real R. Ukuran bernilai real seperti X di atas disebut sebagai ***variabel random***, dimana ukuran peluang yang dikenakan pada $(R, \boldsymbol{A})$ akan ditunjukkan dengan P^x dan disebut distribusi peluang X. Nilai ukuran ini yang diberikan pada himpunan $A \in \boldsymbol{A}$ akan ditunjukkan secara bergantian oleh $P^X(A)$ dan $P(X \in A)$.

Karena interval-interval $\{x|x \leq a\}$ berada di $\mathcal{A}$, maka peluang $F(a) = P(X \leq a)$ terdefinisi untuk semua a. Fungsi F yaitu *fungsi distribusi komulatif* dari X adalah fungsi yang tidak naik ; kontinu di bagian kanan ; $F(-\infty) = 0$, $F(+\infty) = 1$. Sebaliknya, bila F adalah fungsi dengan sifat-sifat di atas, suatu ukuran dapat didefinisikan pada interval dengan $P\{a < X \leq b\} = F(b) - F(a)$. Dengan demikian, distribusi peluang p^X dan fungsi distribusi komulatif F akan saling menentukan. Hal ini dapat diperluas ke ruang Euclides berdimensi-n. Bentuk distribusi komulatif untuk ruang Euclides yang berdimensi-n adalah :

$$F(a_1,\ldots,a_n) = P\{X_1 \leq a_1,\ldots,X_n \leq a_n\}.$$

Ruang $(S, \mathcal{S})$ yang berkaitan dengan keseluruhan hasil yang mungkin muncul, biasanya tidak dinyatakan. Hal yang muncul adalah himpunan X dari pencatatan observasi. Himpunan X ini adalah hasil pencatatan observasi yang dikenal dengan istilah data dan berkaitan dengan ruang terukur $(A, \mathcal{A})$ dan juga berkaitan dengan ruang sampel. Variabel-variabel random atau vektor-vektor yang merupakan transformasi terukur T dari $(A, \mathcal{A})$ ke dalam (τ,β) disebut *statistik*.

Bab III

Kesimpulan

Mempelajari statistika dimulai dengan mempelajari himpunan. Lebih spesifik lagi himpunan yang nantinya akan dikaitkan dengan peluang haruslah memenuhi kriteria sebagai suatu σ-field .Definisi σ-field adalah sebagai berikut :

Diberikan suatu himpunan X dan $\mathbf{X}$ adalah keluarga dari sub-himpunan X. $\mathbf{X}$ disebut sebagai suatu σ-field atau σ-aljabar bila :

1. $\varnothing$ dan X adalah unsur $\mathbf{X}$.
2. Bila A adalah unsur $\mathbf{X}$ maka komplemen A atau $X - A$ adalah juga unsur $\mathbf{X}$.
3. Bila $<A_n>$ adalah barisan himpunan di dalam $\mathbf{X}$, maka gabungan $\cup_{n=1}^{\infty} A_n$ adalah juga berada di dalam $\mathbf{X}$.

Pasangan terurut $(X, \mathbf{X})$ yang terdiri dari himpunan X dan σ-field $\mathbf{X}$ dari sub-himpunan sub-himpunan X disebut suatu *ruang terukur*. Suatu fungsi tak-negatif μ yang dikenakan pada $\mathbf{X}$ sedemikian hingga $\mu(\varnothing) = 0$ disebut suatu *ukuran*. Jika juga didefinisikan bahwa $\mu(X) = 1$, maka μ disebut *ukuran probabiliti* atau *ukuran peluang*. Definisi inilah yang dipakai untuk mendefinisikan peluang seperti yang terdapat pada buku-buku teks.

Pada aplikasinya, peluang yang dikenakan pada $(X, \mathbf{X})$ berkaitan dengan percobaan random atau observasi. Unsur-unsur atau datanya adalah $x \in X$. Biasanya orang akan lebih senang kalau mengukur atau mencatat data yang berbentuk angka. Untuk itu pandang suatu fungsi T yang mengambil nilainya dari ruang τ. Suatu fungsi membuat σ-field β' dari himpunan-himpunan B yang invers imagenya

$$C = T^{-1}(B) = \{z \mid z \in X, T(z) \in B\}$$

berada di dalam $\mathbf{X}$, dan untuk sebarang ukuran peluang P yang dikenakan pada $(X, \mathbf{X})$ maka suatu ukuran peluang Q yang dikenakan pada (τ, β') didefinisikan sebagai

$$Q(B) = P(T^{-1}(B)).$$

Jadi untuk mempelajari konsep statistik dengan dasar matematika yang lebih dalam, maka teori pengukuran dengan mengunakan ukuran Lebesgue tak dapat dihindarkan. Kalau pada matakuliah dasar definisi tenang variabel random diberikan sebagai fungsi dari suatu ruang sampel ke bilangan real, maka dengan menggunakan konsep terukur dan dengan mengunakan ukuran Lebesgue definisi tentang variabel random tidak langsung dilompatkan sebagai fungsi dari ruang sampel ke bilangan real. Untuk konsep peluang, ide dasarnya tidaklah berbeda karena nilai peluang tetap diberikan kepada anggota dari suatu himpunan yang berasal dari ruang sampel.

Sebagai ilustrasi adalah pelemparan sebuah mata uang dengan H sebagai Head (kepala) dan T menunjukkan Tail (ekor). Untuk itu misalkan X = {H, T}. Dari sini diperoleh $\mathbf{X} = \{\varnothing, \{H\}, \{T\}, \{H, T\}\}$. Ruang sampelnya adalah X = {H, T} dan sampel randomnya adalah {H} atau {T}. Peluang yang diberikan pada unsur-unsur $\mathbf{X}$ dengan asumsi mata uangnya "*fair*" adalah P(∅)= 0; P(H) = 0,5; P(T) = 0,5; P(H, T) = 1. Misalkan fungsi f : X → R, dengan f(H) = 1, dan f(T)= 0. Fungsi f inilah yang disebut sebagai variabel random.

Daftar Pustaka

Bartle, R.G. (1966). The Element of Integration. John Wiley & Sons, Inc: NewYork.

Bernoulli, H. (1713). Arts Conjectandi.

Hogg, R.V & Craig, A.T. (1978). Introduction to Mathematical Statistics. 4 th. Ed. Collier Macmilan International Edition: New York.

Lehmann, E.L. (1986). Testing Statistical Hypotheses. John Wiley & Sons, Inc: New York.

Papoulis, A (1984). Probability, Random Variables, and Stochastic Processes. 2^{nd}. Ed. McGraw-Hill Inc: Singapore.

Simonnet, M (1996). Measures and Probabilities. Springer-Verlag New York, Inc: New York.

Riwayat Hidup Peneliti

Nama : Drs. Herman, MA

Unit : Jurusan Statistika FMIPA - Universitas Terbuka

Tempat / Tgl. Lahir : Palembang / 25 Mei 1956

Pendidikan : - S1, Matematika, Institut Teknologi Bandung, 1984

- S2, Educational Psychology, University of Victoria, 1993

Pengalaman Penelitian : - A Study of Relationship between Achievement in Prerequisite Course in Applied Statistics and Economics Study Programs at Universitas Terbuka, 1993.

- Item Analisis, Statistika dan Penurunan Rumus Item Analisis, 1997.

- Uji Umur dan Penaksiran Keandalan Alat-alat yang Berdistribusi Eksponensial dengan Pendekatan Bayes, 1997.

- Metode Bayes untuk Penaksiran Parameter μ dari $N(\mu,\sigma^2)$, 2000.