

**PENGGUNAAN STATISTIK DALAM  
PENYELIDIKAN**

**Oleh  
NOOREHA HUSAIN**

Dari satu sudut, penyelidikan dapat ditakrifkan sebagai satu objektif mencari bahan dan fakta baru termasuk tafsiran yang munasabah. Penyelidikan juga adalah hasil dari pencarian ilmu pengetahuan yang berterusan, penggunaan secara praktik akan kesimpulan sains serta secara pemerhatian, bagi memastikan kesahihan maklumat yang telah diperolehi. Disini kaedah statistik dapat memberikan sumbangan yang besar dalam membantu penyelidikan, contohnya eksperimen dan penyelidikan yang menggunakan kaedah tinjauan selalunya memaksa kita menggunakan ilmu statistik.

Statistik merupakan satu kaedah penyelidikan sistematik yang membawa kepada keputusan yang lebih jitu dan mantap. Ia juga adalah satu kaedah saintifik yang digunakan untuk mengumpul, merancang, merangka dan menganalisa data serta membuat keputusan yang logik dan munasabah berdasarkan kepada analisa tersebut. Walau apapun, sebenarnya kaedah statistik mementingkan semua kesudahan yang mungkin terjadi dalam semua penyelidikan yang dijalankan.

Pada masa ini, penggunaan komputer yang meluas juga telah mempengaruhi pelbagai bidang termasuklah mempengaruhi pemikiran ahli-ahli statistik dimana mereka perlu menyelaraskan semua kaedah-kaedah statistik dengan waras supaya bersesuaian dengan penggunaan komputer. Ini akan dapat memberi pencapaian yang lebih baik dalam satu-satu penyelidikan yang dijalankan.

Skop dan konsep interpretasi data dalam statistik gunaan secara kasarnya terbahagi kepada dua jenis iaitu deskriptif dan inferensi/induktif. Statistik deskriptif lebih menitikberatkan penyusunan data yang dikumpul dan tidak ada sebarang langkah yang diambil kearah penyeluruhan data tersebut. Sebarang perlakuan data yang membenarkan ramalan atau ertian berkenaan sesuatu set data yang lebih besar dikumpulkan dibawah bidang statistik inferensi.

## STATISTIK DESKRIPTIF

Ini merupakan kaedah yang digunakan untuk persembahan data dalam bentuk yang difahami umum. Kadangkala kita ingin menerangkan tentang satu-satu pembolehubah atau mungkin juga penerangan ke atas sekutuan beberapa pembolehubah yang diperlukan. Beberapa cara yang boleh dilakukan ialah:

### (i) Penurunan Data (Data Reduction)

Penyelidikan saintifik biasanya melibatkan pengumpulan data yang besar, contohnya dalam melakukan satu kajiselidik yang mempunyai populasi seramai 2,000 orang; setiap seorang ditanya 100 soalan. Secara tak langsung kita akan dihadapkan dengan 200,000 jawapan! Tiada siapa yang mungkin dapat membaca kesemua 200,00 jawapan dan membuat sebarang kesimpulan darinya. Dengan kaedah penurunan data kesemua maklumat yang dikumpulkan dapat dirumuskan kesimpulannya.

Contoh penurunan data dapat dilihat pada matrik data mentah yang dibentuk dari satu

projek penyelidikan kuantitatif. Jadual 1 menunjukkan satu data matrik separa. Setiap baris dalam matrik menggambarkan perseorangan, setiap lajur pula mewakili satu pembolehubah dan setiap nilai dalam matrik adalah satu nilai kod yang dipunyai oleh seseorang pada sesuatu pembolehubah.

**Jadual 1 : Data matrik separa**

Perseorangan	Jantina	Umur	Pendidikan	Pendapatan	Pekerjaan	Ugama
Orang 1	1	3	2	4	1	0
2	1	4	2	4	4	1
3	2	2	5	5	2	2
4	1	5	4	4	3	2
5	2	3	7	8	6	5
6	2	1	3	3	5	1

Lajur pertama adalah jantina seseorang, katakan "1" mewakili lelaki dan "2" mewakili wanita. Dari segi umur pula, orang yang pertama "3" mungkin memberi maksud 30-39 tahun, orang kedua "4" pula mungkin memberi maksud 40-49 tahun. Secara tak langsung data yang besar telah dikecilkan kedalam bentuk matrik. Sekiranya umur diberi kod seperti yang dinyatakan, jawapan bagi umur 33 tahun akan digolongkan ke kategori "30-39". Dalam penyelidikan ini kita mungkin diberi 60 atau 70 umur-umur yang berlainan. Ini akan dapat diatasi jika ia diturunkan kepada 6 atau 7 kategori.

Seterusnya pengukuran-pengukuran lain seperti ukuran kecenderungan memusat iaitu mod, median, min ; ukuran serakan seperti julat, sisihan piawai dan lain-lain ukuran boleh diperolehi dari data yang telah diringkaskan ini.

**(ii) Ukuran-Ukuran Sekutuan**

Sekutuan diantara mana-mana dua pembolehubah boleh dipersembahkan dengan satu data matrik yang merupakan satu taburan kekerapan tercantum bagi dua pembolehubah. Jadual 2 menunjukkan contoh bagi kes di atas. Ia dapat memberi semua maklumat yang diperlukan untuk menentukan hubungan asal antara kepercayaan tahyul dan tahap akademik seseorang.

**Jadual 2 : Data mentah bagi tahap akademik dan kepercayaan tahyul**

Kepercayaan Tahyul	Tahap Akademik				
	Tiada	Sek. Rendah	Sek. Menengah	Kolej	Universiti
KUAT	23	34	156	67	16
SEDERHANA	11	21	123	102	23
SEDIKIT	6	12	95	164	77

Matrik di atas dapat memberi banyak maklumat. Jika dikaji jadual di atas, didapati apabila tahap akademik meningkat dari tiada pendidikan ke peringkat universiti, kecenderungan bagi kepercayaan tahyul adalah menurun.

## STATISTIK INFERENSI

Kebanyakan penyelidikan kemasyarakatan merangkumi ujian data yang dikumpulkan dari satu sampel yang diambil dari satu populasi yang besar, contohnya satu sampel mengandungi orang yang akan ditemuduga dan satu sampel laporan perceraian yang dikodkan.

Sebenarnya tujuan penyelidikan yang paling utama ialah untuk mendapatkan rumusan/kesimpulan tentang populasi besar yang dikaji melalui kajian yang dilakukan ke atas sampel. Oleh itu sampel univariat dan multivariat sebagai asas kepada pentakbiran yang hendak dibuat berkenaan populasi perlu dijelaskan.

## INFERENSI UNIVARIAT

Analisa univariat ialah satu kaedah di mana kita ingin mengkaji satu pembolehubah pada satu masa tertentu. Dua ukuran univariat yang biasa diambil kira ialah peratusan dan min. Contoh bagi peratusan; katakan 50% dari satu sampel orang yang menghidap selsema ditahun lepas diambil kira; 50% juga adalah merupakan anggaran terbaik bagi kadar mereka yang selsema dalam jumlah populasi bagi mewakili sampel yang diambil. (Andaian yang digunakan ialah dengan sampel rawak mudah). Walau bagaimanapun, mungkin agak mustahil jika 50% dari populasi mengalami selsema; tetapi sekiranya pensampelan dilakukan dengan terperinci; maka kita mungkin boleh menganggarkan ralat yang berlaku.

Ralat piawai yang dikira sebagai kuantiti adalah:  $\sqrt{\frac{p \times q}{n}}$  dimana p ialah peratusan : q = 1-p dan n ialah saiz sampel. Apabila membicarakan tentang ralat pensampelan, 2 perkara perlu diambil kira iaitu Aras Keyakinan (mungkin 90%, 95%, 99%, dll...) dan Selang Keyakinan (contoh  $\pm 1 \sigma$ ,  $\pm 2 \sigma$ ,  $\pm 3 \sigma$ ).

Contohnya: Jika 50% bagi satu sampel yang mengandungi 1,600 orang yang mengatakan mereka mengalami selsema ditahun itu, kita mungkin kata yang kita 95% yakin bahawa peratus mereka yang mengalami selsema itu adalah diantara 46.5% dan 52.5%.

Pengiraannya adalah seperti berikut:-

Selang keyakinan 95% bagi kadar di atas ialah

$$P \pm Z \frac{\alpha}{2} \sqrt{\frac{p \cdot q}{n}}$$

Aras keyakinan 95%,

$$\alpha = 5\%, Z_{\alpha/2} = \pm$$

$$1.96, p = 0.5, q = 0.5$$

$$0.5 \pm 1.96$$

$$\sqrt{\frac{0.5(0.5)}{1600}}$$

$$0.5 \pm 0.0245$$

$$(0.4755, 0.5245)$$

Selain daripada menyatakan ukuran univariat dalam peratusan, kita boleh mengkaji data secara purata. Pemilihan kita pada kali ini adalah berkisar diantara mod (nilai yang paling kerap berlaku), min aritmetik ataupun median (nilai penengah dalam satu taburan cerapan yang telah disusuntertibkan).

Contohnya, dijalankan satu kaji selidik keatas sekumpulan remaja yang berumur diantara 13 hingga 19 tahun yang dianggap sebagai pemboleh ubah seperti dinyatakan dibawah:

Umur	Bilangan Remaja
13	3
14	4
15	6
16	8
17	4
18	3
19	3

Dari sini, kita dapat melihat umur sebenar bagi kesemua 31 cerapan. Soalannya, apakah yang dapat kita katakan tentang umur remaja secara umum atau pada hitung panjangnya? Ini boleh dilihat dari tiga cara yang berlainan dalam memberikan jawapan di atas.

Purata yang paling mudah untuk dikira ialah mod. Seperti yang dilihat, kebanyakan daripada mereka berumur 16 tahun (8 orang remaja), jadi umur modnya adalah 16 tahun (rujuk pada gambarajah 1). Gambarajah 1 juga menerangkan bagaimana pengiraan min dapat dilakukan dan diperolehi min umur remaja ialah 15.87 tahun. Daripada 31 cerapan tersebut juga, didapati nilai ditengah-tengah setelah disusuntertib ialah 16 tahun. Ini bermakna umur salah seorang daripada 8 orang yang berumur 16 tahun itu mewakili nilai median.

**GAMBARAJAH 1:**

**Tiga "Purata"**

Umur Bilangan

13	○○○
14	○○○○
15	○○○○○○
16	○○○○○○○○
17	○○○○
18	○○○
19	○○○

Nilai mod = 16  
Paling kerap

Umur Bilangan

13	○○○
14	○○○○
15	○○○○○○
16	○○○○○○○○
17	○○○○
18	○○○
19	○○○

$13 \times 3 = 39$

$14 \times 4 = 56$

$15 \times 6 = 90$

$16 \times 8 = 128$

$17 \times 4 = 68$

$18 \times 3 = 54$

$19 \times 3 = \frac{57}{492} + 31 = 15.87$

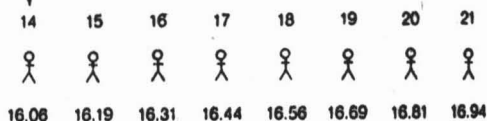
Jumlah (Kes)

Min = 15.87  
min aritmetik

Umur Bilangan

13	○○○	1-3
14	○○○○	4-7
15	○○○○○○	8-13
16	○○○○○○○○	
17	○○○○	22-25
18	○○○	26-28
19	○○○	29-31

Median = 16.31  
Titik tengah



## INFERENSI MULTIVARIAT

Analisa multivariat ialah satu kaedah untuk mengetahui hubungan serentak bagi beberapa pembolehubah dan boleh juga digunakan untuk memahami dengan lebih mendalam hubungan di antara dua pembolehubah tak bersandar. Jadual multivariat boleh dibentuk sekiranya terdapat beberapa sub-kumpulan yang sukar dijelaskan iaitu apabila kita ingin menerangkan lebih daripada satu pembolehubah tak bersandar. Contohnya, tinjauan berkenaan dengan pandangan/pendapat orang ramai terhadap persamaan hak kaum lelaki dan wanita. Dalam tinjauan ini, dipercayai umur juga merupakan salah satu faktor yang mempengaruhi pandangan/pendapat terhadap persamaan hak kaum lelaki dan wanita.

Jadual 3 dapat memberikan maklumat-maklumat yang telah dicatit dalam peratusan dan interpretasi jadual dapat dirumuskan seperti berikut:

- i) Diantara lelaki dan wanita, golongan muda didapati lebih menyokong kepada persamaan hak. Diantara golongan wanita 90% daripada wanita yang berusia kurang daripada 30 tahun dan 60% daripada wanita yang berusia lebih daripada 30 tahun didapati menyokong usul tersebut.
- ii) Dari jadual, umur merupakan faktor yang kuat mempengaruhi pandangan persamaan hak berbanding dengan jantina.
- iii) Umur dan jantina adalah dua pembolehubah tak bersandar yang mempunyai pengaruh yang berbeza keatas pandangan terhadap persamaan hak.

Walau bagaimanapun, jadual 3 boleh diringkaskan kepada jadual 4.

**Jadual 3: Hubungan Multivariat : (Pandangan/Pendapat, Jantina dan Umur)**

"Adakah anda menyokong atau sebaliknya pada cadangan supaya lelaki dan wanita diberi persamaan hak".				
	Umur < 30 tahun		Umur ≥ 30 tahun	
	Lelaki	Wanita	Lelaki	Wanita
Sokong	78%	90%	48%	60%
Membanggang	22	10	52	40
	100%	100%	100%	100%
	(200)	(200)	(200)	(200)
Tiada Jawapan	(2)	(10)	(2)	(3)

**Jadual 4 : Ringkasan daripada Jadual 3.**

	Peratus Yang Menyokong	
	Lelaki	Wanita
Umur < 30	78 (200)	90 (200)
Umur ≥ 30	48 (200)	60 (200)

Daripada 200 responden, 180 daripada wanita yang berumur kurang daripada 30 tahun (atau 90%) menyokong persamaan hak. Secara tidak langsung, ini bermakna 10% daripada responden wanita dalam umur kurang daripada 30 tahun tidak menyokong akan usul tersebut.

Teknik-teknik statistik yang berlainan kadangkala dapat membantu atau mungkin juga menyusahkan sesetengah penyelidikan. Masalah ini boleh diatasi jika dibuat beberapa garis kasar yang mudah untuk menentukan teknik-teknik yang benar-benar bersesuaian dalam penyelidikan. Justeru itu ini dapat mengurangkan beban dalam membuat satu-satu kataputus.

Dalam pemilihan teknik yang paling sesuai beberapa perkara perlu diambilkira:

- i. Keperluan alat ukuran yang minima : samada nominal, ordinal, selangan, dan lain-lain.
- ii. Bilangan kumpulan atau sampel-sampel yang diuji (dan dibandingkan).
- iii. Bilangan pembolehubah yang perlu dianalisa secara serentak, Adalah lebih baik jika sesuatu penyelidikan itu dapat dilakukan dengan sempurna dimana setiap penyelidikan *empirik* dapat memberi gambaran yang logik selain dari hanya merupakan satu operasi matematik sahaja.

Akhirnya adalah mustahak untuk menyedari bahawa *semua* teknik-teknik yang dipilih sehinggakan yang paling mudah pun mempunyai ANDAIAN-ANDAIAN disebaliknya. Perhatian hendaklah diberi kepada semua keadaan alam sekitar yang berhubung kait dengan penyelidikan supaya bersesuaian dengan keperluan teknik-teknik statistik.



## RUJUKAN:

1. *H.T. Hayslett, MS, "Statistics : Made Simple" . Heinemann, 1983.*
2. *Wong Sin Mong, "Statistik Asas STPM" , Federal Publications, 1985.*
3. *Lam Swee Kin, "Statistik STPM" , Penerbit Fajar Bakti, 1987.*
4. *Zainal Mat Saat, "Pengantar Statistik" , Penerbit Fajar Bakti, 1985.*
5. *Earl Babbie, "The Practice of Social Research" , Wadsworth, 1983.*
6. *Richard L. Mills, "Statistic for Applied Economics and Business, Magraw-Hill, 1977.*
7. *Abdul Wahab Jusoh, "Ilmu Statistik Vital" , USM, 1980.*
8. *K.S. Ramaswari, "Quantitative Techniques" , S. Chand & Co., 1984.*