

# PENGUNAAN STATISTIK SEDERHANA UNTUK PENGOLAHAN SAMPEL KECIL DALAM PENELITIAN BIOLOGI

Waminton Rajagukguk<sup>1</sup> dan Isnaini Nurwahyuni<sup>2</sup>

<sup>1</sup> Dosen Jurusan Pendidikan Matematika FMIPA, IKIP Medan

<sup>2</sup> Dosen Program Studi Biologi FMIPA, USU Medan

Kata kunci: Statistik, sampel kecil, mean, median penelitian

## ABSTRAK

Penggunaan statistik sederhana untuk pengolahan sampel kecil dalam penelitian Biologi dilelaskan dalam tulisan ini. Perhitungan statistik sederhana dapat digunakan untuk mengolah data pengamatan kecil dengan sampel  $\leq 10$ . Dijelaskan kelebihan dari median sebagai pengganti rata-rata yang berguna untuk membuat keputusan cepat. Skala dapat juga dipergunakan untuk menaksir standar deviasi dan interval konfidens dengan sedikit kehilangan dalam presisi. Metode statistik sederhana sangat cepat, tepat, dan baik dipergunakan untuk penelitian laboratorium yang menggunakan jumlah sampel sedikit, sehingga dapat menghemat waktu dan biaya penelitian.

## ABSTRACT

The use of simple statistics to process a small number of experimental data in Biology is explained. A simple statistics can be used to calculate few observation of  $\leq 10$  samples. The advantages of median to replace the averages, that is used to make a quick decision is explained. The scale is also useful to be used to estimate standard deviation and the confidence intervals at a little loss in the precision. A simple statistical method is very quick, and very good to be used for laboratory experiments that are dealing with small number of samples, and it can save time and the expenses of the research.

## PENDAHULUAN

Penelitian bidang Matematika dan Ilmu Pengetahuan Alam (MIPA), khususnya penelitian biologi, sering hanya melibatkan jumlah sampel yang sedikit ( $\leq 10$  pengamatan). Alasannya mungkin disebabkan oleh sulitnya pelaksanaan penelitian dengan jumlah sampel besar,

karena keterbatasan dan untuk menyediakan sampel banyak, mahalnya zat kimia yang digunakan, karena antisipasi resiko bahaya dalam pelaksanaan penelitian, dan alasan-alasan lainnya. Untuk penarikan kesimpulan hasil penelitian maka dalam tulisan ini dijelaskan penggunaan statistik sederhana untuk pengolahan sampel kecil dalam penelitian biologi.

Peranan matematika (statistika) dalam pelaksanaan penelitian sangat diperlukan, terutama dalam penarikan kesimpulan dan interpretasi hasil penelitian itu sendiri. Tidak jarang suatu penelitian menjadi kurang komunikatif karena pengolahan data hasil penelitian tidak menggunakan statistika yang benar. Metode statistika yang benar dapat menjadikan penelitian menjadi lebih efisien karena penggunaan sampel terbatas akan dapat menggambarkan seluruh populasi (ASTM, 1994).

Salah satu usaha pengembangan perguruan tinggi di Indonesia yang dianggap tinggi urgensinya adalah pemantapan landasan kemampuan dalam ilmu-ilmu dasar (basic sciences). Pandangan strategis ini didasarkan pada perkiraan tuntutan Pembangunan Nasional dimasa depan terhadap lulusan perguruan tinggi, serta kecenderungan yang makin mempersyaratkan penguasaan pengetahuan dan ketrampilan dalam ilmu-ilmu dasar sebagai unsur pendukung Pembangunan Nasional.

Sejalan dengan upaya pembangunan pendidikan ilmu-ilmu dasar diwaktu yang akan datang dipandang perlu membina kemampuan dalam bidang penelitian karena merupakan unsur pendorong bagi perkembangan ilmu pengetahuan dan teknologi (IPTEK). Untuk mengembangkan IPTEK diharapkan perguruan tinggi merukan institusi potensil terutama terutama dalam melaksanakan bidang penelitian dan pengembangan (research and development), sesuai dengan asas tridarma perguruan tinggi yaitu pendidikan/pengajaran, penelitian dan pengabdian dan masyarakat.

Untuk mengembangkan teknologi sangat diperlukan penelitian dasar (basic research dan penelitian terapan (applied research) yang terpadu dan berkesinambungan. Di Indonesia pada saat ini kenyataannya bahwa penelitian terapan lebih banyak dilakukan dibandingkan dengan penelitian dasar. Hal ini mudah dipahami karena kita masih berorientasi dapat teknologi yang dapat segera memberikan keuntungan dan berguna bagi masyarakat banyak, dengan investasi yang relatif kecil. Akan tetapi perlu kita sadari bahwa di masa mendatang kita juga harus menseimbangkan antara penelitian dasar dengan penelitian terapan agar

kita jangan hanya sebagai konsumen teknologi tetapi juga harus berfungsi sebagai produsen teknologi. Pembangunan penelitian dasar tersebut akan menjadikan perkembangan teknologi kita sejajar dengan teknologi maju.

Harus diakui bahwa penelitian dasar sangat sulit dilaksanakan karena membutuhkan pengamatan yang intensif, dan bahkan tidak jarang dalam penelitian harus melibatkan biaya penelitian yang cukup besar sehingga penelitian sangat jarang dilakukan dengan ulangan >10 data pengamatan yang berdiri sendiri (Dixon, 1989). Agar data penelitian dengan jumlah yang sedikit dapat dipergunakan dalam penarikan kesimpulan maka sangat diperlukan metode statistika sederhana untuk pengolahan data penelitian yang ulangan pengamatannya kecil.

Dalam tulisan ini akan dibahas tentang penggunaan statistik sederhana untuk pengolahan sampel kecil, yaitu lebih kecil atau sama dengan 10 pengamatan yang terdiri sendiri, yang dipergunakan dalam penelitian Biologi dan penelitian lain yang relevan. Metode statistik ini sangat baik terutama untuk membuat keputusan yang cepat tanpa menggunakan perhitungan matematika yang rumit. Dalam uraian, penulis membuat contoh kasus data hasil penelitian Biologi dengan sampel yang relatif kecil, yaitu hasil pengamatan terhadap pertumbuhan vegetatif tanaman tomat oleh pengaruh hara dalam tanaman hidroponik.

## **PENGOLAHAN DATA MENGGUNAKAN STATISTIK SEDERHANA**

Pada umumnya semua peneliti selalu berhati-hati dalam observasi numerik yang diperoleh dalam suatu percobaan. Sudah merupakan konsensus bahwa data penelitian bagaimanapun adanya dapat diolah dengan menggunakan statistik menjadi lebih sederhana. Ada statistik sederhana yang dapat dipergunakan untuk menyelesaikan persoalan penelitian dengan menggunakan perhitungan yang sangat sedikit (Grant, 1990).

Pengamatan berulang terhadap sampel yang memiliki kualitas data "sama" tidaklah terlalu penting dicari dalam penelitian. Secara normal kita dapat melihat arah sampel melalui sebaran data dengan menghitung harga rata-rata sebagai taksiran dari harga (nilai) yang "sebenarnya". Untuk tujuan ini kita secara mudah dapat menggunakan aritmatika mean ( $\bar{x}$ ). Telah diketahui juga, sebaran data mempunyai hubungan reabilitas pengukuran. Pengukuran sebaran ini umumnya disebut standar deviasi (SD) dan rata-rata deviasi.

Bila data penelitian umlahnya cukup besar, fluktuasinya berdiri sendiri dan menyebar dalam populasi, maka nilainya (harga) akan mengikuti kurva normal. Dalam pembahasan berikut ini kesimpulan mengikuti distribusi normal populasi. Harga-harga tabulasi yang disajikan berlaku hanya jika sampel random yang independen diambil dari populasi sebenarnya. Sebagai contoh, terhadap enam kelompok sampel tanaman tomat (*Lycopersicon esculentum* Mill.) yang diberikan perlakuan variasi konsentrasi hara dalam media tanaman hidroponik, kemudian dilakukan pengukuran pertumbuhan vegetatif dengan pengulangan sebanyak tiga kali untuk setiap sampelnya, maka hanya ada 5 kelompok data pengamatan independen sebagai mewakili hasil rata-rata seluruh sampel yang diamati tersebut (Nurwahyuni, 1996).

Dengan mengikuti pengamatan yang independen dan terdistribusi secara normal, maka statistik terbaik adalah menggunakan mean aritmatika ( $\bar{x}$ ) untuk menghitung rata-rata dan standar deviasi (SD) atau varians ( $SD^2$ ) untuk sebarannya. Istilah kata "terbaik" digunakan untuk  $\bar{x}$  dan SD adalah berdasarkan fakta bahwa perhitungan ini paling efisien untuk menafsirkan data terhadap populasi bila pengamatan mengikuti kurva normal (Gaussian). Istilah "paling efisien" menunjukkan variasi minimum atau dispersi untuk distribusi sampling statistika, misalnya jumlah terkecil dari sebaran untuk perkiraan pengulangan terhadap mean populasi atau varian populasi.

Teori statistika klasik untuk pengolahan sampel besar dalam percobaan bidang sains telah dibahas dalam beberapa buku teks seperti Fisher, (1992) dan Kendall, (1988). Akan tetapi harus disadari bahwa sebagian teori tersebut akan kurang akurat untuk sejumlah pengamatan kecil. Bila jumlah  $n$  lebih besar dari 20 sampai 30, maka  $t$ -Student hampir mendekati kesalahan kurva normal. Untuk membuat metode statistika yang benar maka peneliti Biologi menggunakan teknik untuk sejumlah pengamatan yang dipilih, karena sangat jarang penelitian Biologi, khususnya untuk perlakuan kultur jaringan tanaman, melakukan pengulangan sampai 20 kali.

• Beberapa pengamatan akan diurutkan berturut-turut sebagai  $X_1, X_2, \dots, X_{n-1}, X_n$ , dan perhitungan rata-rata harga tengah dinyatakan dengan  $\bar{X}$  (Dean, 1951):

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_{n-1} + X_n}{n} = \frac{(\sum X_i)}{n} \quad (1)$$

Sebaran data tersebut akan diukur lebih efisien dengan varian,  $SD^2$  menggunakan perhitungan:

$$SD^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n-1} = \frac{(\sum X_i - \bar{X})^2}{n-1} \quad (2)$$

dan  $SD = \sqrt{\frac{(\sum X_i - \bar{X})^2}{n-1}} \quad (3)$

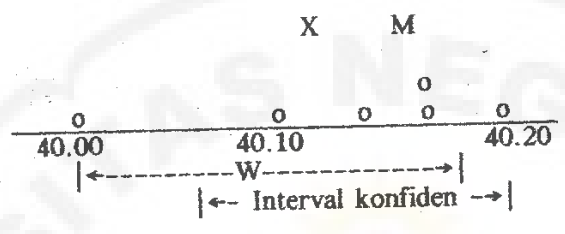
Dalam perhitungan, sangat penting menggunakan  $(n-1)$  untuk menghindari perkiraan bias dari  $\sigma^2$  varian dari populasi. Selain menggunakan rata-rata  $\bar{X}$  hitung, dapat juga dipergunakan harga median ( $M$ ) untuk menghitung nilai tengah. Nilai  $M$  adalah harga tengah atau rata-rata dari dua nilai yang saling berdekatan di bagian tengah. Bila  $n=5$  maka  $M$  adalah  $X_3$ , dan bila  $n=6$  maka  $M = (X_3 + X_4)/2$ . Besaran  $M$  adalah taksiran harga tengah dan mempunyai kelebihan yaitu tidak dipengaruhi oleh besar nilai data lain.

Median dari pengamatan yang kecil dapat ditentukan dengan inspeksi saja, walau hal itu kurang efisien dibandingkan dengan harga rata-rata, terutama terutama bila ada kesalahan mutlak. Dalam penelitian Biologi sering diperhadapkan dengan keputusan untuk menghilangkan atau mengikutsertakan data yang memiliki deviasi sangat jauh (ekstrim dari sejumlah data-data yang sudah terkumpul. Medium tidak banyak dipengaruhi oleh kesalahan mutlak dibandingkan dengan rata-rata. Untuk menghindari kehadiran kesalahan mutlak maka adalah lebih baik bila digunakan perhitungan harga median.

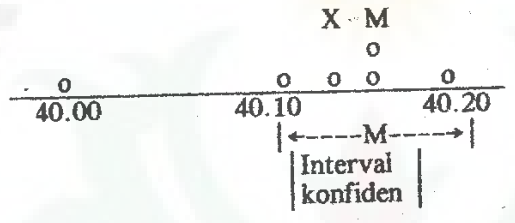
Data yang disajikan dalam Tabel 1 berikut adalah merupakan hasil rata-rata pengamatan terhadap pertumbuhan vegetatif tanaman tomat yang diberikan perlakuan variasi hara ke dalam media dalam tanaman hidroponik dalam penelitian Biologi. Data hasil pengamatan untuk enam kelompok perlakuan (variasi hara,  $H_1 - H_6$ ) yang hanya dilakukan dengan tiga kali ulangan, dan rata-rata kelompoknya diilustrasikan secara berturut-turut pada Gambar 1 dan 2.

Tabel 1. Hasil pengamatan pertumbuhan vegetatif tanaman tomat yang diberi variasi hara dalam hidroponik. Kelompok data yang disajikan merupakan rata-rata tiga ulangan pengamatan.

Jenis hara	Pertumbuhan Vegetatif (cm)	
H <sub>1</sub>	40.02	X = 40.14
H <sub>2</sub>	40.12	M = 40.17
H <sub>3</sub>	40.16	
H <sub>4</sub>	40.18	
H <sub>5</sub>	40.18	
H <sub>6</sub>	40.20	



Gambar 1 Grafik dari presentasi data Tabel 1



Gambar 2. Grafik dari penolakan data ekstrim Tabel 1

Nilai pengamatan rata-rata adalah 40.14 dan harga median adalah 40.14. Sebaiknya kita lebih cenderung meletakkan nilai tengah pada median 40.17 dibandingkan dengan rata-rata 40.14. Skala observasi (W) adalah batas paling panjang dengan paling pendek, yaitu  $W = X_n - X_1$ . Ini akan memberikan efisiensi tinggi bila pada pengamatan 10 atau lebih kecil.

Sejalan dengan bertambahnya  $n$  maka efisiensi dari skala akan menjadi semakin berkurang. Bila data disajikan secara random, misalnya sebagai urutan produksi dan bukan sebagai urutan hasil,



maka rata-rata dari skala untuk sebanyak 6 atau 8 sub kelompok akan lebih efisien bila dibanding terhadap skala tunggal.

Walaupun SD dan SD<sub>w</sub> sangat berguna dalam mengukur dispersi data yang sebenarnya, kita biasanya lebih tertarik pada konfidens interval atau konfidens batas. Dengan menggunakan Interval konfidens yaitu rata-rata dua sisi  $x$  sebagaimana yang dicari secara peluang pada titik tengah. Misalnya kita mengharapkan rata-rata pada batas konfidens 95%. Dengan menggunakan konfidens yang lebih lebar, misalnya 99%, maka kita dapat meningkatkan kesempatan untuk mendapatkan rata-rata yang sebenarnya akan tetapi intervalnya akan menjadi lebih panjang. Interval terpendek untuk peluang yang diberikan adalah sesuai dengan harga  $t$  test. Interval konfidensnya adalah  $x = ts/\sqrt{n}$ . Harga  $t$  bervariasi berdasarkan jumlah pengamatan dan derajat konfidens yang diinginkan.

Konfidens batas dapat dihitung dengan cara yang sama, dengan menggunakan SD<sub>w</sub> yang diperoleh dari skala yang menyerupai tetapi tidak sama dengan tabel untuk  $t$ . Akan tetapi lebih meyakinkan bila menghitung batas langsung dari skal sebagai  $x = Wt_w$  Faktor untuk mengubah  $W$  ke  $s_w$  telah dimasukkan dalam kuantitas  $t_w$  yang ditabulasi dalam kolom untuk konfidens 95% dan konfidens 99% dan faktor ini dapat ditemui dalam tabel Fisher, (1992).

Di dalam tabel (Fisher, 1992) untuk dat dengan 6 kelompok pengamatan diperoleh besar harga  $t_w = 0.40$  dan skala adalah 0.18, sehingga harga  $Wt_w$  adalah 0.072. Dengan demikian untuk hasil pengamatan yang disajikan pada Tabel 1 kita dapat katakan bahwa pada interval konfidensnya 95%, rata-rata hitung adalah  $40.14 \pm 0.072$ . Bila kita menghitung interval konfidens dengan menggunakan tabel  $t$ , dan berdasarkan perhitungan SD pada konfidens 95% diperoleh  $t = 2.6$ , sehingga  $ts/6 = 0.078$ . Dengan demikian rata-rata hitung pengamatan dapat juga dinyatakan sebesar  $40.14 \pm 0.078$ .

Bila dalam perhitungan standardeviasi telah diketahui, atau misalnya dari data sebelumnya, kita dapat menggunakan kurva normal untuk menghitung konfidens batas. Situasi ini bisa muncul bila analisis digunakan untuk sejumlah 50 atau lebih sampel yang sama. Standar deviasi dari populasi dapat ditaksir dari rata-rata varians  $SD^2$  dari kelompok pengamatan. Interval dari  $x = 1.96S/\sqrt{m}$ , dimana  $x$  dihitung dari set pengamatan, dan dapat diharapkan telah mewakili populasi rata-rata 95%.

Sebagaimana dijelaskan bahwa statistik sederhana dapat menghitung nilai tengah dan mendapatkan konfidens batas dari hasil pengamatan. Penggunaan median akan mengurangi pengaruh sampel yang ekstrim pada penghitungan nilai tengah. Nilai ekstrim dapat dihilangkan dengan menggunakan tes skrining, disebut *test-Q* sebagai berikut: Hitung jarak pengamatan dari nilai yang berdekatan, kemudian bagi jarak tersebut dengan skala sehingga rasio untuk *Q* adalah :

$$Q = (x_2 - x_1) / w \quad \text{atau} \quad Q = (x_n - x_{n-1}) / w$$

Bila *Q* berada diluar nilai dalam tabulasi maka pengamatan dapat dibuang pada konfidens 90% atau 95%. Dalam contoh pengamatan data yang disajikan Tabel 1, maka data pengamatan 40.02 dapat diterima pada konfidens 90% karena konfidens interval 90% memberikan harga tabel 0.56, akan tetapi untuk konfidens 95% data tersebut harus ditolak.

$$Q = (40.12 - 40.02) / (40.20 - 40.02) = 0.56$$

Dengan menolak data pertama maka kita meningkatkan median dari 40.17 menjadi 40.18, dan rata-rata dari 40.14 menjadi 40.17 (Gambar 2). Standar deviasi (SD) berubah dari 0.067 menjadi 0.030 (bila data 40.02 telah ditolak), dan  $SD_w$  berubah dari 0.072 menjadi 0.034. Maka konfidens interval yang sesuai dengan *test-t* sekarang menjadi  $40.17 \pm 0.038$ , dan dari median dan skala menjadi  $40.18 \pm 0.040$ , berkurang sekitar setengah dalam penjaga skala interval.

#### KESIMPULAN

Statistik sederhana dapat digunakan untuk menganalisis data sampel kecil dalam penelitian Biologi. Statistik ini juga dapat dikembangkan untuk analisis data untuk percobaan Kimia yang mahal, sehingga dapat meningkatkan efisiensi dana dan waktu dalam penelitian dan dengan hasil yang bisa mewakili populasi.



## DAFTAR PUSTAKA

- American Soc. Testing Materials, (1994) *Manual on Presentation of Data ASTM*
- Dean, R.B., dan Dixon, W.J., (1951), Simplified statistical for small numbers of observations, *Analytical Chem.* 23: 636638.
- Dixon, W.J., dan Massey, F.J., (1989) *Introduction to Statistical Analysis*, New Yourk McGraw Hill Book Co.
- Fisher, R.A., dan Yates, F., (1992) *Statistical Tables for Biological, Agricultural, and Medical Research*, Edinburgh, Oliver and Boyd
- Grant, E.L., (1990), *Statistical Quality Control*, New Yourk McGraw Hill Book Co.
- Kendall, M.G., (1988) *Advanced Theory of Statistics Vol XII*, London , Charles Griffin Co.
- Nurwahyuni, I., Sudrajadhat, R., Miffahuddin, dan Wibowo, N.L., (1996) Potensi limbah tanaman tomat sebagai larutan hara hidroponik. *Majalah Universitas Sumatera Utara*, Vol 8(2) 1996, on Press

oooooooooooo00000oooooooooooo