

574.01595
MAS
5 4

DIKTAT KULIAH

STATISTIKA

PROGRAM STUDI S-1 NUTRISI DAN MAKANAN TERNAK



Oleh

Maulana Hamonangan Nasoetion, S.Pt., M.P.

**LABORATORIUM BIOMETRIKA
JURUSAN NUTRISI DAN MAKANAN TERNAK
FAKULTAS PETERNAKAN UNIVERSITAS
DIPONEGORO**

2003

UPT-PUSTAK-UNDIP	
No. Daft:	55 / RI / FP / 10
Tgl.	27 / AS / 04

KATA PENGANTAR

Buku diktat Kuliah Statistika ini disusun untuk menunjang kegiatan proses belajar mengajar pada matakuliah Statistika yang dikelola oleh Laboratorium Biometrika Peternakan Program Studi S-1 Nutrisi dan Makanan Ternak Fakultas Peternakan Universitas Diponegoro Semarang. Buku ini diharapkan mampu meningkatkan kemampuan mahasiswa dalam memahami matakuliah Statistika sesuai dengan kurikulum yang ada.

Secara intensif matakuliah Statistika terdiri dari 15 kali tatap muka kuliah dan 1 kali mid-semester dapat dijabarkan dalam pada buku ini.

Semoga buku ini bermanfaat bagi para pembaca terutama mahasiswa yang mengambil matakuliah Statistika.

Semarang, Agustus 2003

Penyusun

DAFTAR ISI

	Halaman
KATA PENGANTAR	iii
BAB I : PENDAHULUAN	1
Tujuan Instruksional Umum	1
Tujuan Instruksional Khusus	1
Uraian dan Contoh	1
1. Pengertian Biostatistika (Biometri)	1
2. Sejarah Perkembangan Statistika	2
3. Data dalam Statistika	3
Daftar Pustaka	4
BAB II : DATA PADA BIOSTATISTIKA	5
Tujuan Instruksional Umum	5
Tujuan Instruksional Khusus	5
Uraian dan Contoh	5
1. Contoh dan Populasi	5
2. Peubah dalam Biostatistika	6
3. Ketelitian dan Ketepatan Data	7
4. Sebaran Frekuensi	9
5. Penanganan Data	10
Daftar Pustaka	12
BAB III : STATISTIKA DESKRIPTIF	13
Tujuan Instruksional Umum	13
Tujuan Instruksional Khusus	13
Uraian dan Contoh	13
1. Statistika Lokasi	14
1. Rata-rata Hitung (Mean)	14
2. Median	19
3. Modus	19
2. Statistika Penyimpangan	20
1. Kisaran	20

2. Simpangan Baku	21
3. Koefisien Varians	23
Daftar Pustaka	23
BAB IV : PENGANTAR SEBARAN PELUANG	25
Tujuan Instruksional Umum	25
Tujuan Instruksional Khusus	25
Uraian dan Contoh	25
1. Sebaran Peluang Binomial	27
2. Sebaran Peluang Poisson	28
3. Sebaran Frekuensi Peubah Kontinyu (Sebaran Normal)	31
4. Sifat-sifat Sebaran Normal	32
5. Perhitungan Peluang Sebaran Normal	34
Daftar Pustaka	37
BAB V : SAMPLING DAN SEBARAN SAMPLING	38
Tujuan Instruksional Umum	38
Tujuan Instruksional Khusus	38
Uraian dan Contoh	38
1. Sampling	38
1. Alasan Sampling	39
2. Rencana Sampling	40
3. Cara-cara Sampling	40
4. Penggunaan Sampling Acak Berstrata	43
2. Sebaran Sampling	45
1. Sebaran Sampling Rata-rata Hitung	46
2. Dalil Limit Pusat (The Central Limit Theorem)	49
Daftar Pustaka	50
BAB VI : TEORI PENDUGAAN PARAMETER	51
Tujuan Instruksional Umum	51
Tujuan Instruksional Khusus	51
Uraian dan Contoh	51
1. Ciri-ciri Penduga yang Baik	52
2. Cara-cara Pendugaan	53
1. Pendugaan Titik	53
2. Pendugaan Interval	54

3. Pendugaan Parameter dengan Sampel Besar	55
1. Pendugaan Parameter μ dengan σ Diketahui dan Populasi Tidak Terbatas	55
2. Pendugaan Parameter μ dengan σ Diketahui dan Populasi Terbatas	56
3. Pendugaan Parameter μ dengan σ Tidak Diketahui	57
4. Pendugaan Parameter $\mu_1 - \mu_2$ dimana σ_1 dan σ_2 Diketahui	58
4. Pendugaan Parameter dengan Contoh Kecil	59
1. Pendugaan Parameter μ dengan σ Tidak Diketahui dan Populasi Tak Terbatas	60
2. Pendugaan Parameter $\mu_1 - \mu_2$ dimana σ_1 dan σ_2 Tidak Diketahui	61
Daftar Pustaka	62
BAB VII : UJI HIPOTESIS	64
Tujuan Instruksional Umum	64
Tujuan Instruksional Khusus	64
Uraian dan Contoh	64
1. Kesalahan Jenis I dan Jenis II	66
2. Langkah-langkah Pengujian Hipotesis	67
3. Pengujian Hipotesis dengan Sampel Besar	67
1. Pengujian Parameter Rata - Rata, $H_0 : \mu = \mu_0$ dimana σ^2 Tidak Diketahui	67
2. Pengujian $H_0 : \mu_1 = \mu_2$ Dimana σ_p^2 Diketahui dan $\sigma_1^2 = \sigma_2^2$	69
4. Pengujian Hipotesis dengan Sampel Kecil	70
1. Pengujian $H_0 : \mu = \mu_0$ Dimana σ^2 Tidak Diketahui	70
2. Pengujian $H_0 : \mu_1 = \mu_2$ atau $\mu_1 - \mu_2 = 0$, jika σ^2 tidak diketahui dan $\sigma_1^2 \neq \sigma_2^2$	71
Daftar Pustaka	73
BAB VIII : SEBARAN CHI-SQUARE	74
Tujuan Instruksional Umum	74
Tujuan Instruksional Khusus	74
Uraian dan Contoh	74
1. Uji Kenormalan	75
2. Uji Homogenitas Varians	78
3. Uji Independensi	80
Daftar Pustaka	82

BAB IX : REGRESI	84
Tujuan Instruksional Umum	84
Tujuan Instruksional Khusus	84
Uraian dan Contoh	84
1. Persamaan Regresi Linear	84
2. Uji Signifikansi dalam Regresi Linear	86
Daftar Pustaka	87
BAB X : KORELASI	89
Tujuan Instruksional Umum	89
Tujuan Instruksional Khusus	89
Uraian dan Contoh	89
Daftar Pustaka	90

BAB I

PENDAHULUAN

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menjelaskan pengertian Biostatistika (Biometrika), sejarah perkembangan statistika dan data dalam statistika.

Tujuan Instruksional Khusus

Setelah mempelajari bab ini mahasiswa dapat menjelaskan:

1. Pengertian Biostatistika (Biometrika)
2. Sejarah perkembangan statistika
3. Data dalam statistika

Uraian dan Contoh

1.1. Pengertian Biostatistika (Biometri)

Statistika adalah suatu studi ilmiah mengenai data numerik yang berasal dari fenomena alami. Berdasarkan definisi tersebut berkembang suatu disiplin ilmu di bawah statistika yang disebut Biostatistika. Biostatistika didefinisikan sebagai penerapan metode statistika untuk pemecahan masalah biologi. Masalah biologi yang dimaksud adalah masalah yang muncul baik pada ilmu biologi dasar maupun bidang terapannya seperti ilmu-ilmu pertanian (ilmu peternakan termasuk di dalamnya). Biostatistika juga disebut statistika biologi atau biometri. Kelompok Biometrika dikembangkan dan dijadikan salah satu Kelompok Mata Kuliah yang merupakan embrio Laboratorium Biometrika dilingkungan Fakultas Peternakan UNDIP.

Dalam biologi, kebanyakan fenomena dipengaruhi oleh banyak faktor penyebab yang tidak dapat dikendalikan variasinya dan sering tidak dapat dikenali. Statistika diperlukan untuk mengukur fenomena peubah demikian itu dalam penentuan galat pengukuran dan memastikan adanya perbedaan yang meskipun kecil tapi penting.

Cara berpikir statistika hampir sama dengan cara berpikir ilmiah dalam bidang ilmu lain, semuanya berusaha agar pengamatan yang dilakukan merupakan pengamatan kuantitatif. Dalam statistika kita mengatakan seberapa jauh kita menyatakan percaya atau tidak percaya dengan menggunakan peluang, tidak berupa pernyataan samar-samar dan bersifat umum. Sebagai contoh pernyataan berikut: ayam negeri produksi telurnya lebih tinggi ($P < 0.05$ atau $P < 0.01$) dari pada ayam kampung.

1.2. Sejarah Perkembangan Statistika

Statistika adalah metode ilmiah untuk pengumpulan, pengorganisasian, penyimpulan, penyajian, dan analisis data. Statistika didefinisikan Sokal dan Rohlf sebagai studi ilmiah mengenai data numerik yang berasal dari fenomena alami. Pengertian studi ilmiah adalah statistika harus dapat diterima secara umum mengenai keabsahan bukti ilmiah. Metode statistika untuk pemecahan masalah biologi disebut Biostatistika. Masalah biologi yang dimaksud adalah masalah yang timbul pada ilmu biologi dasar atau terapannya.

Statistika modern berkembang dari dua sumber, yaitu ilmu politik dan teori matematika. Bentuk statistka yang berkembang dari ilmu politik sebagai uraian kuantitatif berbagai urusan pemerintahan dan negara. Pajak dan asuransi menyebabkan orang tertarik pada sensus, usia, dan kematian. Dua mahasiswa pertama yang mempelajari statistika di Inggris yaitu John Graunt (1620-1674) dan William Petty (1623-1687). Sumbangan penting terhadap statistika dengan sumber teori matematika adalah dari Blaise Pascal (1623-1662) dan Pierre de Fermat (1601-1665). Keduanya berkebangsaan Perancis. Jacques Bernoulli (1654-1705),

seorang warga negara Swiss meletakkan dasar modern teori peluang dalam bukunya yang berjudul *Ars Conjectandi*. Abraham de Moivre (1667-1754), warga negara Perancis yang bermukim di Inggris adalah orang pertama yang menggabungkan statistika yang ada pada masa hidupnya dengan teori peluang dalam menentukan besar pensiun yang harus dibayarkan, dan melakukan pendekatan terhadap sebaran normal melalui pemekaran binomial. Persamaan kurva normal pertama kali diumumkan pada tahun 1733 oleh Abraham de Moivre. Karyanya dipelajari oleh Karl Pearson (1857-1936), seorang ahli fisika matematik pada tahun 1924. Sumbangan abadi Karl Friedrich Gauss (1777-1855) untuk statistika adalah metode Jumlah Kuadrat Terkecil. Francis Galton (1882-1911) dianggap sebagai bapak biostatistika dan genetika. Sumbangan terbesar Galton pada biologi adalah penerapan metode statistika yang dilakukannya untuk menganalisis keragaman biologis terutama melalui analisis keragaman dan melalui regresi dan korelasi terhadap hasil pengukuran biologi.

1.3. Data dalam Statistika

Secara sempit istilah statistik digunakan untuk menyatakan data itu sendiri atau bilangan-bilangan yang diturunkan dari data, misalnya: rata-rata. Dalam pengumpulan data yang menyangkut ciri-ciri suatu kelompok, tidak praktis bahkan tidak mungkin mengamati seluruh kelompok (penyelidikan populasi). Penyelidikan sebagian kecil kelompok disebut penyelidikan sampel atau contoh. Populasi dapat terhingga atau tak terhingga.

Sampel merupakan wakil populasi, sehingga kesimpulan terhadap populasi dapat ditarik dari analisis sampel. Penarikan kesimpulan seperti itu dianggap valid, disebut statistika induktif atau inferensia. Metode ini hasilnya tidak mutlak pasti, pernyataan peluang digunakan dalam menyatakan kesimpulan. Tahap statistika yang hanya menguraikan tanpa penarikan kesimpulan disebut statistika deduktif atau deskriptif.

DAFTAR PUSTAKA

- Dajan, A. 1996. Pengantar Metode Statistika. Jilid I. Cetakan ke-18. Penerbit PT.Pustaka LP3ES, Jakarta.
- Dixon, W.J. dan F.J. Massey, Jr. 1997. Pengantar Analisis Statistik. Cetakan ke-2. Diterjemahkan oleh: Sri Kustantini S. dan Zanzawi S. Gadjah Mada University Press, Yogyakarta.
- Sokal R.R. dan F.J.Rohlf. 1991. Pengantar Biostatistika. Edisi ke-2. Diterjemahkan oleh: Nasrullah dan Setyono Setyo Sunarto. Gadjah Mada University Press, Yogyakarta.
- Spiegel, M.R., I. Y. Susila dan E. Gunawan. 1961. Statistik Edisi SI (Metrik). Schaum Publishing Company, Edinburg.
- Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.
- Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.
- Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

BAB II

DATA PADA BIOSTATISTIKA

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menjelaskan data pada Biostatistika.

Tujuan Instruksional Khusus

Setelah mempelajari bab ini mahasiswa dapat menjelaskan :

1. Pengertian contoh dan populasi
2. Peubah dalam Biostatistika
3. Ketelitian dan ketepatan data
4. Sebaran frekuensi
5. Penanganan data

Uraian dan Contoh

2.1. Contoh dan Populasi

Data pada biostatistika pada umumnya berdasarkan pengamatan individual berupa pengamatan atau pengukuran yang dilakukan pada unit contoh terkecil yakni individu itu sendiri. Jika kita menimbang 100 ekor itik Tegal, maka bobot dari masing-masing itik tersebut merupakan pengamatan individu yang disebut sebagai sampel atau contoh pengamatan.

Sampel atau contoh pengamatan didefinisikan sebagai kumpulan pengamatan individu yang dipilih melalui cara-cara tertentu. Dalam statistika, populasi adalah keseluruhan pengamatan individu yang merupakan acuan daerah pengambilan contoh tertentu dan dibatasi oleh ruang dan waktu. Kesalahan umum dalam menerapkannya pada teknik sampling adalah ketidakmampuan untuk menentukan populasi acuan dan ini menyebabkan contoh yang berasal dari populasi yang terbatas akan mengakibatkan kesimpulan hanya berlaku secara terbatas. Populasi bisa terhingga atau tak terhingga. Secara teoritis populasi biologi merupakan populasi terhingga, walaupun demikian pada umumnya ukurannya jauh lebih besar dibandingkan dengan ukuran contoh yang diambil dari populasi tersebut.

2.2. Peubah dalam Biostatistika

Peubah adalah sifat yang dapat digunakan sebagai cara untuk membedakan suatu individu dari individu lain dalam contoh. Peubah dibedakan atas peubah terukur, peubah urutan dan atribut.

Peubah terukur adalah peubah yang merupakan hasil pengukuran atau pencacahan yang dinyatakan secara numerik. Peubah terukur ada dua jenis. Jenis pertama yakni peubah kontinyu yang diantara dua nilai tertentu dapat dipunyai nilai yang tak hingga banyaknya. Contoh: diantara pengukuran nilai panjang 1,5 dan 1,6 cm terdapat banyak sekali panjang yang dapat diukur, misalnya 1,57 mm. Jenis kedua yakni peubah tidak kontinyu atau peubah diskrit adalah peubah yang hanya dapat mempunyai harga numerik tertentu tanpa ada nilai antara yang mungkin diantaranya. Misalnya: banyaknya telur yang dihasilkan (dalam butir) mungkin 4, 5 atau 10 butir, tetapi tidak pernah 5,5 atau 7,5 butir telur.

Beberapa peubah tidak dapat diukur, tapi dapat diurutkan menurut tingkatannya yang disebut sebagai peubah urutan. Contoh: suatu peubah sebagai deret urutan seperti 1,2,3,4,5. Kita tidak dapat menafsirkan bahwa perbedaan antara urutan 1 dan 2 adalah sama atau setara dengan perbedaan antara tingkat 2 dan 3.

Peubah ada juga yang tidak dapat diukur tapi harus dinyatakan secara kualitatif disebut atribut atau peubah nominal mis: hitam dan putih, hidup dan mati, dan lain-lain.

2.3. Ketelitian dan Ketepatan Data

Ketelitian dan ketepatan memiliki makna yang berbeda dalam statistika. Ketelitian menyatakan seberapa dekat suatu nilai terhitung terhadap nilai sejatinya. Ketepatan menggambarkan seberapa jauh dekatnya pengukuran-pengukuran yang diulang menghasilkan nilai yang sama. Suatu alat timbang yang bias tetapi peka menghasilkan penimbangan yang tidak teliti tetapi tepat.

Variat (suatu pembacaan nilai atau pengamatan mengenai suatu peubah) yang teliti biasanya merupakan bilangan bulat. Peubah tidak kontinyu biasanya diukur sebagai angka pasti. Peubah kontinyu yang diperoleh dari peubah tidak kontinyu dalam kondisi tertentu bisa merupakan angka pasti. Contoh: sekumpulan ternak sapi terdiri dari 18 jantan dan 12 betina. Nisbah betina terhadap jantan (yang merupakan peubah kontinyu) adalah 1,5 dan merupakan suatu angka yang pasti. Pada umumnya pada peubah kontinyu kepastian nilai suatu pengukuran tidak diketahui, angka terakhir pada suatu pembacaan harus diartikan secara tidak langsung sebagai ketepatan yaitu menunjukkan batas suatu skala pengukuran yang kita percaya mengandung terdapat nilai pengukuran yang sebenarnya. Contoh lainnya adalah pengukuran panjang 12,3 mm berarti panjang yang sebenarnya terletak antara 12,25 dan 12,35 mm. Tepatnya dimana diantara kedua batas tersebut nilai sebenarnya berada kita tidak tahu. Batas tersirat selalu mempunyai satu angka di luar angka paling belakang hasil pengukuran. Kita mencatat pengukuran sebagai 12,32 berarti bahwa yang kita maksud adalah antara 12,315 dan 12,325.

Peubah tidak kontinyu, meskipun biasanya pasti, dapat ditulis sebagai pendekatan apabila melibatkan angka besar. Contoh: bila suatu pencacahan dilaporkan sampai ribuan terdekat, pencacahan 36000 serangga yang terdapat dalam satu meter kubik tanah sebagai peraga, berarti bahwa angka yang sebenarnya terletak antara 35500 sampai 36500 serangga.

Sampai berapa angka suatu pengamatan harus ditulis? Ini tergantung pada seberapa besar angka bermakna (angka yang ada artinya) berpengaruh terhadap hasil

komputasi yang akan dianalisis. Angka bermakna adalah angka cermat selain nol yang diperlukan untuk menempatkan titik desimal, mis : 65,4 mempunyai 3 angka bermakna; 4,5300 memiliki 5 angka bermakna; $0,0018 = 1,8 \times 10^{-3}$ memiliki 2 angka bermakna; $0,001800 = 1,800 \times 10^{-3}$ memiliki 4 angka bermakna. Angka nol seharusnya tidak ditulis pada akhir suatu angka yang berupa suatu pendekatan di sebelah kanan tanda desimal, kecuali angka nol tersebut dimaksudkan sebagai angka yang ada artinya, 7,80 berarti bahwa batasnya adalah 7,795 sampai 7,805. Jika 7,75 sampai 7,85 yang dimaksud, pengukuran harus ditulis sebagai 7,8.

Perhitungan yang dilakukan dalam statistika sering membutuhkan pembulatan angka. Pembulatan angka dilakukan dengan memperhatikan angka yang terdapat dibelakang angka terakhir. Bila angka tersebut lebih besar atau sama dengan dari 6, maka dibulatkan ke atas. Bila angka tersebut lebih kecil atau sama dengan dari 4, maka dibulatkan ke bawah.

Contoh :

72,8 pembulatan ke satuan terdekat adalah 73

72,8146 pembulatan ke per-ratusan terdekat adalah 72,81

72,465 pembulatan ke per-ratusan terdekat adalah 72,46

183,575 pembulatan ke per-ratusan terdekat adalah 183,58

Pembulatan angka dengan kasus angka 5 adalah angka yang terdapat dibelakang angka terakhir, maka pembulatan angka tersebut tidak boleh langsung ditetapkan ke atas seperti yang sering dilakukan selama ini. Hal itu dapat menimbulkan bias yang cukup besar, karena hasil setelah pembulatan akan jauh lebih tinggi dibandingkan dengan sebelum pembulatan. Cara menghindarkan atau memperkecil galat pembulatan kumulatif (*cumulative rounding errors*), perlu diperhatikan ketentuan sebagai berikut :

Apabila angka genap mendahului 5, maka data dibulatkan ke angka genap (tetap). Apabila angka ganjil yang mendahului 5, maka pembulatan data dilakukan dengan menambahkan 1 pada angka ganjil tersebut (ke atas).

Contoh:

23,4675 pembulatan per-ribuan 23,468

23,4685 pembulatan per-ribuan 23,468

Perhitungan yang dilakukan dalam statistika juga harus memperhatikan aspek komputasi. Dalam melakukan perhitungan yang menyangkut perkalian, pembagian dan penarikan akar bilangan, hasil terakhir tidak dapat memiliki angka bermakna lebih dari pada bilangan dengan angka bermakna tersedikit.

Contoh:

$$73,24 \times 4,52 = (73,24)(4,52) = 331$$

$$1,848/0,023 = 80$$

$$\sqrt{38,7} = 6,22$$

$$(8,416)(50) = 420,8 \text{ jika } 50 \text{ eksak}$$

2.4. Sebaran Frekuensi

Sebaran frekuensi adalah suatu daftar/tabel yang memuat frekuensi dari suatu sebaran data tertentu. Contoh dari suatu populasi mengenai produksi susu sapi PFH dalam 250 hari (kg), bila data yang digunakan terbatas ($n=10$) maka sebaran data tersebut belum terlihat bentuknya. Peningkatan jumlah data yang digunakan akan menunjukkan bentuk yang sebenarnya dari sebaran data yang disajikan. Pada contoh di atas sebarannya berbentuk genta setangkup (kurva normal). Disamping bentuk setangkup, juga terdapat sebaran yang menceng (lebih menjulur pada salah satu ujungnya dibanding dengan sisi yang lain), sebaran bentuk C, bentuk U, dan lain-lain.

Semuanya memberikan informasi yang berarti mengenai hubungan yang ditunjukkannya. Sebaran frekuensi kuantitatif berdasarkan peubah kontinyu merupakan sebaran frekuensi yang paling sering digunakan.

2.5. Penanganan Data

Dalam mekanisme pengolahan data, terdapat bagian dimana dilakukan 'pencermatan' terhadap data. Penanganan ini bertujuan untuk memilih statistik yang sesuai yang akan digunakan untuk mengolah data, sehingga efisiensi dan akurasi hasil pengolahan betul-betul tinggi. Makna penanganan data adalah mengidentifikasi data lewat pengukuran yang dilakukan dan bentuk sebarannya.

Berdasarkan peubah dalam biostatistika (terukur, urutan dan atribut) yang oleh Siegel (1985) dinyatakan sebagai data nominal, ordinal, interval dan ratio, dapat ditentukan cara penanganan data tersebut lewat test statistik yang sesuai untuk masing-masing tingkatan yang bersangkutan seperti terlihat pada tabel 2 berikut.

Berdasarkan Tabel 2. tersebut dapat diketahui bahwa test-test statistik dikelompokkan atas test statistik Parametrik dan test statistik Non Parametrik. Data skala nominal dan ordinal test statistik yang relevan adalah test statistik non parametrik, sedangkan untuk data interval dan ratio test statistik yang sesuai adalah test statistik parametrik dan non parametrik.

Beberapa metode statistik sangat mudah untuk digunakan sebab telah ada tabel-tabel khusus yang memberikan jawaban terhadap hasil pengujian hipotesis untuk diterima atau ditolak. Disamping itu dengan adanya beberapa paket program statistika untuk pengolahan data, sangat membantu mempermudah penanganan data.

Tabel 1. Skala Pengukuran dan Statistik yang Sesuai untuk Skala yang Berrangkaian.

Skala	Hubungan	Contoh-contoh	Test statistik yang sesuai
Nominal	Ekuivalensi Frekuensi Koefisien kontingensi.	Modus	Test statistik non parametrik
Ordinal	Ekuivalensi Lebih besar dari	Median Persentil Spearman rs Kendall t Kendall w	Test statistik non parametrik
Interval	Ekuivalensi Lebih besar dari Ratio semba- rang 2 interval diketahui hasil kali	Mean Deviasi standard Korelasi moment hasil kali Pearson Korelasi moment Ganda	Test statistik parametrik dan non parametrik
Ratio	Ekuivalensi Lebih besar dari Ratio semba- rang 2 interval diketahui Ratio semba- rang 2 harga skala diketahui	Mean geometrik Koefisien variasi	Test statistik parametrik dan non parametrik

Sumber: Spiegel *et al.* (1961).

DAFTAR PUSTAKA

- Dajan, A. 1996. Pengantar Metode Statistika. Jilid I. Cetakan ke-18. Penerbit PT.Pustaka LP3ES, Jakarta.
- Dixon, W.J. dan F.J. Massey, Jr. 1997. Pengantar Analisis Statistik. Cetakan ke-2. Diterjemahkan oleh: Sri Kustantini S. dan Zanzawi S. Gadjah Mada University Press, Yogyakarta.
- Sokal R.R. dan F.J.Rohlf. 1991. Pengantar Biostatistika. Edisi ke-2. Diterjemahkan oleh: Nasrullah dan Setyono Setyo Sunarto. Gadjah Mada University Press, Yogyakarta.
- Spiegel, M.R., I. Y. Susila dan E. Gunawan. 1961. Statistik Edisi SI (Metrik). Schaum Publishing Company, Edinburg.
- Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.
- Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.
- Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

BAB III

STATISTIKA DESKRIPTIF

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menjelaskan dan menghitung data dengan menggunakan statistika lokasi dan statistika penyimpangan

Tujuan Instruksional Khusus

Setelah mempelajari bab ini mahasiswa dapat menghitung:

1. Statistika lokasi yaitu : mean, median dan modus
2. Statistika penyimpangan yaitu: kisaran, varians, simpangan baku dan koefisien varians

Uraian dan Contoh

Apabila kita ingin mengetahui pengaruh suatu perlakuan mis: protein terhadap pertambahan bobot badan ayam kampung, sebelum analisis dilakukan maka diperlukan suatu informasi mengenai status keadaan ternak yang bersangkutan yang berhubungan dengan perlakuan yang akan diujicobakan. Informasi awal ini disebut stadia diskriptif. Pada stadia deskriptif ini informasi yang berasal dari sejumlah (n) pengamatan disederhanakan sedemikian rupa melalui ukuran pemusatan (statistika lokasi) dan ukuran penyimpangan (statistika penyimpangan) yang disebut sebagai statistika deskriptif.

3.1. Statistika Lokasi

Statistika lokasi (ukuran pemusatan) yakni rata-rata hitung (mean), rata-rata geometris, rata-rata harmonis, median dan modus.

3.1.1. Rata-rata Hitung (Mean)

Rata-rata hitung adalah statistik lokasi yang paling umum dipakai dalam mencari ukuran pemusatan, karena memiliki bias yang kecil terhadap harga sebenarnya. Rata-rata hitung diperoleh dengan menjumlahkan semua pengamatan dalam contoh dan membaginya dengan banyak pengamatan dalam contoh tersebut.

Bila kita menggunakan lambang statistika, maka rata-rata hitung di atas dirumuskan :

$$\bar{X} = \sum X_i / n$$

Rumus ini memberi tahu kita untuk menjumlahkan semua pengamatan (n) dan bagilah jumlah tersebut dengan n .

Dilambangkan dengan μ untuk populasi dan x untuk contoh. Didefinisikan sebagai :

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \sum_{i=1}^n x_i/n$$

Contoh :

Rata-rata hitung bilangan-bilangan 8, 3, 5, 12 dan 10 adalah :

$$\bar{x} = \frac{8 + 3 + 5 + 12 + 10}{5} = 38/5 = 7,6$$

Hasil penimbangan bobot badan 5 ekor sapi Bali

	355 kg	Jadi rata-rata hitung bobot
	367 kg	badan sapi Bali tsb.
	342 kg	
	375 kg	Rata-rata hitung = $1793/5$
	354 kg	= 358.6 kg

Jumlah	1793 kg	

Jika bilangan-bilangan x_1, x_2, \dots, x_n masing-masing terjadi sebanyak f_1, f_2, \dots, f_n kali, maka rata-rata hitungnya adalah :

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \dots + f_nx_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum_{i=1}^n f_i x_i}{\sum_{i=1}^n f_i} \quad i=1$$

Contoh :

Jika 5, 8, 6 dan 2 masing-masing terjadi dengan frekuensi 3, 2, 4 dan 1 maka rata-rata hitungnya adalah :

$$\bar{x} = \frac{(3)(5) + (2)(8) + (4)(6) + (1)(2)}{3 + 2 + 4 + 1} = \frac{15 + 16 + 24 + 2}{10} = 5,7$$

Sifat-sifat rata-rata hitung adalah sebagai berikut:

- (1). Jumlah aljabar simpangan-simpangan suatu himpunan bilangan dari rata-rata hitungnya adalah nol.

Contoh : simpangan-simpangan dari bilangan-bilangan 8, 3, 5, 12, 10 dari rata-rata hitungnya = 7,6 adalah :

$$(8 - 7,6) + (3 - 7,6) + (5 - 7,6) + (12 - 7,6) + (10 - 7,6) = (0,4) + (-4,6) + (-2,6) + (4,4) + (2,4) = 0$$

(2). Jumlah kuadrat simpangan-simpangan suatu himpunan bilangan x_j dari sebarang bilangan a adalah minimum jika dan hanya jika $a = x$

$$\sum_{j=1}^n (x_j - a)^2 \geq \sum_{j=1}^n (x_j - x)^2$$

Contoh :

$x_1 = 5, x_2 = 7, x_3 = 8$. Hitung x dan tunjukkan bahwa pernyataan tsb benar.

$$x = \frac{5 + 7 + 8}{3} = 6,67$$

$$a = 5 \rightarrow \sum (x_j - a)^2 = (5 - 5)^2 + (7 - 5)^2 + (8 - 5)^2 = 0 + 4 + 9 = 13$$

$$a = 7 \rightarrow \sum (x_j - a)^2 = (5 - 7)^2 + (7 - 7)^2 + (8 - 7)^2 = 4 + 0 + 1 = 5$$

$$a = 8 \rightarrow \sum (x_j - a)^2 = (5 - 8)^2 + (7 - 8)^2 + (8 - 8)^2 = 9 + 1 + 0 = 10$$

$$a = x \rightarrow \sum (x_j - x)^2 = (5 - 6,67)^2 + (7 - 6,67)^2 + (8 - 6,67)^2 = 4,67$$

$$\text{Jadi ternyata } \sum_{j=1}^n (x_j - a)^2 \geq \sum_{j=1}^n (x_j - x)^2$$

$$13 \geq 4,67; \quad 5 \geq 4,67; \quad 10 \geq 4,67$$

sehingga $(x_j - x)^2$ adalah minimum.

(3). Jika A sebarang nilai dugaan rata-rata hitung dan jika $d_j = x_j - A$ adalah simpangan-simpangan x_j dari A , maka persamaan $x = \sum x/n$ akan menjadi :

$$\bar{x} = A + \frac{\sum d_j}{n}$$

Contoh :

Kita buktikan rumus tersebut, dimana $\bar{x} = A + \sum d_j/n$

Bukti :

Dik. $X_1, x_2, \dots, x_i, \dots, x_n : A = \text{suatu bilangan konstan rata-rata hitung}$

→

$$d_j = x_j - A \quad \text{jadi } d_1 = x_1 - A : d_2 = x_2 - A : d_3 = x_3 - A : \dots : d_n = x_n - A$$

$$d_j = x_j - A \quad \xrightarrow{\quad} \quad \text{jadi } x_j = A + d_j$$

$$d_j = \sum x_j/n = \sum (A + d_j)/n = \frac{\sum A + \sum d_j}{n} = nA/n + \sum d_j/n = A + \sum d_j/n$$

$$\text{jadi } \bar{x} = A + \sum d_j/n$$

Bila kita harus menghitung rata-rata dari rata-rata statistik yang lain yang berbeda ukuran contohnya dan disisi lain kita ingin menghitung rata-rata dengan bobot yang berlainan, maka dalam hal ini kita menghitung rata-rata hitung tertimbang.

Rumusnya adalah sbb:

$$\bar{X}_w = \frac{\sum w_i X_i}{\sum w_i}$$

yang menunjukkan n pengamatan dirata-rata dengan pengamatan ditimbang dengan bobot w_i .

Contoh:

3 kelompok nilai rata-rata dari data jumlah anak sepelahiran babi Nias, dengan jumlah pengamatan (n) yang berbeda, seperti terlihat pada Tabel 2.

Tabel 2. Jumlah Anak Sepelahiran Babi Lokal di Nias.

Rata-rata hitung ke i	Jumlah pengamatan ke i
6,95	15
8,72	27
5,48	9

Sumber: Bulset Nasional (1985)

Dari data tersebut rata-rata hitung tertimbangya dapat dihitung sebagai berikut:

$$\begin{aligned}
 X_w &= (15)(6,95) + (27)(8,72) + (9)(5,48) / (15 + 27 + 9) \\
 &= (104,25 + 235,44 + 49,32) / 51 \\
 &= 7,63
 \end{aligned}$$

Pada data berkelompok atau dalam bentuk kelas-kelas, bilangan-bilangan x_1 , x_2 , , x_n masing-masing terjadi sebanyak f_1 , f_2 ,, f_n kali, maka rata-rata hitungnya adalah :

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \dots + f_nx_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum_{i=1}^n f_iy_i}{\sum_{i=1}^n f_i}$$

Contoh :

Jika 5, 8, 6 dan 2 masing-masing terjadi dengan frekuensi 3, 2, 4 dan 1 maka rata-rata hitungnya adalah :

$$\bar{x} = \frac{(3)(5) + (2)(8) + (4)(6) + (1)(2)}{3 + 2 + 4 + 1} = \frac{15 + 16 + 24 + 2}{10} = 5,7$$

Modus jarang digunakan dalam bidang biologi. Sebaran yang mempunyai dua puncak disebut dwimodus, dan yang lebih dari dua disebut multimodus.

Contoh:

12, 13, 13, 12, 14, 16, 14, 17, 14 Modus = 14
 12, 13, 13, 13, 17, 14, 12, 14, 14 Modus = 13 dan 14 (bimodus)

Diantara ketiga statistik lokasi, rata-rata hitung memiliki keuntungan nisbi yang lebih baik dibanding median dan modus, karena mempunyai simpangan baku yang lebih kecil, lebih mudah ditangani secara matematis, cenderung mengikuti sebaran normal meskipun data aslinya tidak menyebar normal dan lebih peka terhadap perubahan bentuk sebaran frekuensi. Pada sebaran setangkup (normal) dengan satu puncak, rata-rata hitung, median dan modus memberikan hasil yang sama.

3.2. Statistika Penyimpangan

3.2.1. Kisaran

Bila kita memiliki beberapa kelompok data yang rata-rata hitungnya sama, sebarannya normal tapi tingkat keruncingannya berbeda, maka dibutuhkan statistik penyimpangan untuk menjelaskan ciri dari sebaran data tersebut. Salah satu statistik penyimpangan yang paling sederhana adalah kisaran yang didefinisikan sebagai selisih antara pengamatan terbesar dan terkecil dalam contoh.

Contoh:

12, 15, 16, 18, 21, 26, 29, 30 Kisaran = $30 - 12 = 18$

Karena kisaran merupakan besaran yang mengukur bentangan data pada skala pengukuran, maka satuannya sama dengan satuan pengukuran asli.

3.2.2. Simpangan Baku

Simpangan baku adalah besaran yang menimbang semua pengamatan dari suatu sebaran, yang bobotnya merupakan jarak dari pusat sebaran tersebut. Jarak tiap variat dari rata-rata hitungny adalah penyimpangan yang dirumuskan sbb:

$$x = X_i - \bar{X}$$

Dari rumus di atas diketahui bahwa simpangan adalah selisih tiap pengamatan terhadap rata-rata hitungny, dan bukan sebaliknya. Simpangan dilambangkan dengan huruf kecil yang sesuai dengan huruf besar peubah. Tabel 3. berikut memperlihatkan simpangan baku yang dihitung seperti rumus di atas.

Tabel 3. Simpangan Baku dari Data Neutrofil (dibagi 1000) per-Mi kroliter untuk 15 Pasien dengan Tumor Nonhematological.

(1)	(2)	(3)
X_i	$x = X_i - \bar{X}$	x^2
4,9	-2,81	7,9148
4,6	3,11	9,6928
5,5	-2,21	4,8988
9,1	1,39	1,9228
16,3	8,59	73,7308
12,7	4,99	24,8668
6,4	-1,31	1,7248
7,1	-0,61	0,3762
2,3	-5,41	29,3042
3,6	-4,11	16,9195
18,0	10,29	105,8155
3,7	-4,01	16,1068

	7,3	-0,41	0,1708
	4,4	-3,31	10,9782
	9,8	2,09	4,3542
	<hr/>	<hr/>	<hr/>
Total	115,7	0,05	308,7770

$$\text{Rata-rata hitung } \bar{X} = \sum X_i / n = 115,7 / 15 = 7,713$$

Sumber: Liu *et al.* (1983)

Kuadrat simpangan ditunjukkan pada kolom (3) dari Tabel 4. -dan memungkinkan kita untuk mendapatkan hasil yang bukan nol. Jumlah kuadrat simpangan dalam statistika disebut "Jumlah Kuadrat" dan diberi lambang $\sum x^2$ atau JK. Langkah berikutnya adalah menghitung rata-rata hitung dari n kuadrat simpangan. Besaran yang diperoleh dikenal sebagai varians atau Kuadrat Tengah (KT)

$$\text{Varians} = \sum x^2 / n = 308,7770 / 15 = 20,5851$$

Varians merupakan besaran yang penting sekali dalam statistika dan sering digunakan. Satuan dari varians adalah satuan kuadrat, untuk menghilangkan pengaruh pengkuadratan, diambil akar positif varians dan diperoleh simpangan baku.

$$\text{Simpangan baku} = \pm \sqrt{\sum x^2 / n} = 4,5371$$

Simpangan baku kembali dinyatakan dalam satuan asli pengukuran, karena merupakan akar satuan kuadrat varians.

Penjelasan di atas merupakan tiga langkah untuk menghitung simpangan baku, cara tersebut menjelaskan secara gamblang arti dari jumlah kuadrat. Rumus hitung yang lebih cepat untuk besaran ini adalah :

$$\sum x^2 = X^2 - [(X)^2 / n]$$

Kita perhatikan rumus diatas, suku pertama ruas kanan persamaan, $\sum X^2$, adalah jumlah semua X yang masing-masing telah dikuadratkan:

$$\sum X^2 = X_1^2 + X_2^2 + X_3^2 + \dots + X_n^2$$

$\sum X^2$ disebut 'Jumlah X Kuadrat' dan dibedakan dari $\sum x^2$. Jumlah Kuadrat X. Besaran lain dari persamaan tadi adalah $(\sum X)^2 / n$. Nilai ini sering disebut sebagai Faktor Koreksi disingkat FK.

3.2.3. Koefisien Varians

Sesudah besaran simpangan baku kita peroleh, maka kita dapat menentukan dugaan besarnya ragam dari populasi. Sehingga bila diinginkan untuk membandingkan besar simpangan baku satu populasi terhadap populasi yang lain dapat kita lakukan. Apabila populasi-populasi tersebut berbeda rata-rata hitungnya, digunakan koefisien keragaman (CV) dalam membedakan besaran simpangan bakunya. Rumusnya adalah:

$$CV = \left(\frac{S_b}{\bar{X}} \right) \times 100\%$$

Jadi koefisien keragaman digunakan bila diinginkan untuk membandingkan keragaman dua populasi tanpa memperhatikan besar rata-rata hitungnya.

DAFTAR PUSTAKA

- Dajan, A. 1996. Pengantar Metode Statistika. Jilid I. Cetakan ke-18. Penerbit PT.Pustaka LP3ES, Jakarta.
- Dixon, W.J. dan F.J. Massey, Jr. 1997. Pengantar Analisis Statistik. Cetakan ke-2. Diterjemahkan oleh: Sri Kustantini S. dan Zanzawi S. Gadjah Mada University Press, Yogyakarta.

- Sokal R.R. dan F.J.Rohlf. 1991. Pengantar Biostatistika. Edisi ke-2. Diterjemahkan oleh: Nasrullah dan Setyono Setyo Sunarto. Gadjah Mada University Press, Yogyakarta.
- Spiegel, M.R., I. Y. Susila dan E. Gunawan. 1961. Statistik Edisi SI (Metrik). Schaum Publishing Company, Edinburg.
- Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.
- Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.
- Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

BAB IV

PENGANTAR SEBARAN PELUANG

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menjelaskan dan menghitung sebaran peluang.

Tujuan Instruksional Khusus

Setelah mempelajari bab ini mahasiswa dapat:

1. Menghitung sebaran binomial
2. Menghitung sebaran poisson
3. Menghitung sebaran frekuensi peubah kontinyu (sebaran normal)
4. Menjelaskan sifat-sifat sebaran normal
5. Menghitung peluang sebaran normal

Uraian dan Contoh

Generalisasi yang berkaitan dengan inferensia statistik mempunyai unsur ketidakpastian, karena kita hanya mendasarkan pada informasi parsial yang diperoleh. Penanggulangan ketidakpastian itu, pemahaman teori peluang sangatlah mendasar, agar kita dapat menyusun model matematik yang secara teori dapat menjelaskan perilaku populasi yang dibangkitkan oleh percobaannya. Model-model teoritik itu, yang sangat mirip dengan sebaran frekuensi relatif yang disebut sebaran peluang.

Pada sebaran frekuensi telah dibahas mengenai jumlah pengamatan (frekuensi) dari suatu variat tertentu.

Contoh:

Produksi susu sapi PFH 250 hari (Kg). Tabel 5. berikut memberikan informasi mengenai frekuensi mutlak dan relatif dari produksi susu tersebut yang berasal dari 58 pengamatan.

Tabel 4. Produksi Susu Sapi PFH selama 250 Hari (Kg).

Produksi susu (kg)	Freq.Mutlak	Freq.Relatif
2000	10	0,1724
3000	12	0,2069
4000	14	0,2414
5000	12	0,2069
6000	10	0,1724
Total	58	1,0000

Sumber: Warwick *et.al.* (1984)

Dari Tabel 4. tersebut terlihat bahwa dari total 58 pengamatan untuk masing-masing variat (2000 s/d 6000) frekuensi mutlaknya berturut-turut adalah 10, 12, 14, 12 dan 10. Sedangkan frekuensi relatifnya 0,1724 , 0,2069 , 0,2414 , 0,2069 dan 0,1724.

Satu kasus saja dari sebaran frekuensi di atas misalnya variat 4000 dimana jumlah pengamatannya adalah 14 ekor, kemudian membaginya dengan total pengamatan $n = 58$ diperoleh 0,2414 yang merupakan frekuensi relatif dari variat produksi susu kelompok 4000 kg. Nilai 0,2414 adalah pernyataan peluang dari kemungkinan terambilnya variat 4000 apabila dilakukan pengambilan secara acak terhadap contoh $n = 58$ tersebut. Secara sederhana dapat dikatakan bahwa variat 4000 peluangnya terambil sebagai anggota contoh sebesar 24,14 persen.

Jadi nilai-nilai frekuensi relatif tersebut adalah kumpulan nilai peluang yang memiliki distribusi atau sebaran tertentu. Makin banyak (n) pengamatan maka makin jelas dan pasti bentuk dari sebaran peluangnya.

4.1. Sebaran Peluang Binomial

Nilai $X_1, X_2, X_3, \dots, X_n$ masing-masing merupakan peubah acak yang bebas stokastik sesamanya, serta menyebar secara Bernoulli (ada dua kejadian di dalam ruang contoh), maka peubah tersebut merupakan peubah acak yang menyebar menurut kaidah Binomial, dimana timbulnya kejadian $X = x$, adalah:

$$p(x) = C(n,x) q^x (1-q)^{n-x}, x = 0, 1, 2, \dots, n$$

Contoh:

Pengalaman 51% dari pada telur yang menetas akan menjadi ayam pejantan. Berapakah peluang agar dari 10 butir telur yang menetas terdapat 3 ekor ayam jantan dan 7 ekor ayam betina.

Jawab:

Ditetaskannya seekor ayam jantan dapat dimisalkan sebagai timbulnya kejadian $X=x$, $x = 0, 1, 2, \dots, 10$. Maka $x=3$ sedang $n = 10$ dan $q = 0,51$ sehingga peluang yang ditanyakan ialah:

$$P(x=\text{jantan}; n-x=\text{betina}) = C(10,3) (0,51)^3 (0,49)^7 = 0,1083$$

Sebaran peluang Binomial adalah sebaran peluang data diskret. Berikut ini adalah nilai-nilai koefisien binomium Newton.

Tabel 5. Nilai-nilai Koefisien Binomium Newton

		C(n,x) untuk x =									
n	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1
11	1	11	55	165	330	462	462	330	165	55	11
12	1	12	66	220	495	792	924	792	495	220	66
13	1	13	78	286	715	1287	1716	1716	1287	715	286
14	1	14	91	364	1001	2002	3003	3432	3003	2002	1001
15	1	15	105	455	1365	3003	5005	5005	6435	6435	3003

4.2. Sebaran Peluang Poisson

Apabila dari fungsi peluang binomial, dengan suatu peluang timbulnya kejadian sangat kecil, maka fungsi peluang binomial berubah menjadi fungsi peluang Poisson. Sebaran ini berparameter tunggal. Kaidah peluang Poisson berguna untuk menentukan suatu pendekatan bagi timbulnya suatu kejadian $X=x$ yang menyebar

secara binomial, kalau n besar dan q kecil nilainya. Misalnya saja untuk menghitung kerapatan bakteri di dalam sel-sel kisi suatu kaca obyek di bawah mikroskop. Jumlah sel-sel kisi biasanya sebanyak 10×10 buah. Jika X = jumlah bakteri yang kelihatan dalam satu sel kisi tersebut, maka X akan menyebar mendekati sebaran Poisson.

Pada umumnya rumus untuk sebaran peluang binomium dapat digunakan secara memuaskan untuk menghitung nilai-nilai $b(x; N, p)$ dimana $x = 0, 1, 2, \dots, N$. Jika parameter N ternyata besar sekali ($N > 50$), sedangkan p kecil ($p < 0,10$) sehingga hasil perkalian Np - nya menjadi moderat, maka perhitungan $p(X=x)$ tidak mudah dilakukan. Maka pemecahan $b(x; N, p)$ akan lebih mudah dilakukan dengan menggunakan pendekatan poisson. Bila kita persamakan $Np = \mu$, maka rumus poisson dapat diberikan sebagai berikut :

$$P(x; N, p) = p(x) = p(X=x) = \frac{\mu^x e^{-\mu}}{x!}$$

Dimana x dapat merupakan nilai-nilai $0, 1, 2, \dots, N$ dan $e = 2,71828$. Nilai $e^{-\mu}$ dapat dicari pada tabel nilai $e^{-\mu}$. Sebaran peluang poisson mempunyai parameter rata-rata μ dan simpangan baku σ , sebagai berikut :

$$\begin{aligned}\mu &= Np = \sigma^2 \\ \sigma &= \sqrt{Np}\end{aligned}$$

Contoh :

Berdasarkan pengalaman, sebuah mesin tetas yang berkapasitas 2000 butir telur, pada saat dioperasikan hanya sebutir telur yang tidak menetas. Mahasiswa Fak. Peternakan UNDIP ingin mengetahui berapa peluang memperoleh 0, 1, 2, 3, 4, dan

5 butir telur yang tidak menetas dari proses penetasan yang hanya diisi 1000 butir telur.

Jawab :

Disini $p = 1/2000$ (lebih kecil dari 0,10) dan $N = 1000$ (lebih besar dari 50), sehingga pendekatan poisson dapat dilakukan. Dari nilai p dan N tersebut maka $\mu = 0,50$. Sebaran poisson secara berturut-turut dapat diberikan sebagai berikut :

$$(a). P(X=x=0) = \frac{(0,50)^0 e^{-0,50}}{0!} = 0,6066$$

$$(b). P(X=x=1) = \frac{(0,50)^1 e^{-0,50}}{1!} = 0,3033$$

$$(c). P(X=x=2) = \frac{(0,50)^2 e^{-0,50}}{2!} = 0,07582$$

$$(d). P(X=x=3) = \frac{(0,50)^3 e^{-0,50}}{3!} = 0,01264$$

$$(e). P(X=x=4) = \frac{(0,50)^4 e^{-0,50}}{4!} = 0,00158$$

$$(f). P(X=x=5) = \frac{(0,50)^5 e^{-0,50}}{5!} = 0,000184$$

Contoh:

Suatu kebun bibit menjamin, bahwa bibit yang berasal dari kebun itu memiliki daya tunas sebesar 98%. Berapa peluang bahwa seorang pembeli mendapatkan didalam 100 batang tanaman yang dibelinya 0, 1, 2, 3,....., 100 tanaman yang tidak mampu tumbuh ?

Jawab:

Pada persoalan ini, $n=100$ $q=1,00 - 0,98 = 0,02$ $=0,02(100)=2$

Kalau X ialah jumlah bibit yang gagal bertunas, maka:

$$P(x : N,p) = p(x) = p(X=x) = \frac{\mu^x e^{-\mu}}{x!}$$

Nilai-nilainya untuk $x = 0, 1, 2, \dots, 100$ tersusun sebagai berikut:

x	P (X=x)
0	0,1353
1	0,2707
2	0,2707
3	0,1804
4	0,0902
5	0,0361
6	0,0120
7	0,0034
8	0,0009
9	0,0002
10	0,0000
100	0,0000

4.3. Sebaran Frekuensi Peubah Kontinyu (Sebaran Normal)

Peubah acak kontinyu dan fungsi kepekatannya muncul bila data percobaan adalah data kontinyu. Hasil pengukuran selang waktu, bobot badan, tinggi pundak,

volume dan lain sebagainya dapat dinyatakan dengan suatu sebaran kontinyu. Seperti juga terdapat beberapa sebaran peluang diskret yang khusus, kita juga mengenal banyak sekali sebaran kontinyu, yang grafiknya mungkin menunjukkan adanya kemenjuluran atau dalam beberapa kasus setangkup sempurna. Namun yang paling penting adalah suatu sebaran kontinyu yang grafiknya berbentuk genta dan menjulur tak terbatas dalam kedua arah. Sebaran inilah yang merupakan landasan bagi sebagian besar teori inferensia statistik.

Banyak fenomena biologi yang sebaran datanya mendekati normal, sehingga sebaran ini menjadi dasar bagi teori statistik yang digunakan oleh para ahli biologi. Grafik sebaran normal juga disebut kurva Gauss. Lokasi pusat kurva ini terletak pada μ , gemuk-kurusnya kurva tergantung pada besarnya varians. Nilai varians yang kecil akan menyebabkan kurvanya tinggi dan ramping, sedangkan varians yang besar menyebabkan kurvanya pendek dan gemuk.

4.4. Sifat-sifat Sebaran Normal

Fungsi kepekatan normal digambarkan oleh rumus berikut:

$$Z = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}$$

Di sini Z menunjukkan tinggi ordinat kurva yang menggambarkan kerapatan variat. Z merupakan peubah tidak bebas karena merupakan fungsi peubah X. Terdapat dua konstanta dalam rumus di atas, yakni

$$\frac{1}{\sigma\sqrt{2\pi}} = 3,14159 \quad e = 2,71828$$

Pada fungsi kepekatan normal, terdapat 2 parameter yaitu parameter rata-rata hitung (μ) dan parameter simpangan baku (σ), yang menentukan lokasi dan bentuk

penyimpangan. Kurva normal adalah setangkup terhadap rata-rata hitungnya. Oleh karena itu rata-rata hitung, median dan modus sebaran normalnya semuanya berada pada titik yang sama. Persentase variat yang mengikuti sebaran normal, akan terletak pada batas-batas yang ditunjukkan sebagai berikut:

$\mu \pm$ berisi 68,27% variat

$\mu \pm 2$ berisi 95,45% variat

$\mu \pm 3$ berisi 99,73% variat

sebaliknya:

50% variat berada dalam kisaran $\mu \pm 0,674$

95% variat berada dalam kisaran $\mu \pm 1,960$

99% variat berada dalam kisaran $\mu \pm 2,576$

Sebaran normal merupakan sebaran yang paling banyak digunakan dalam statistika, dimana penerapannya dapat dibagi sebagai berikut:

1. Sebelum kita menerapkan suatu uji tertentu terhadap suatu contoh, kita harus mengetahui terlebih dahulu apakah contoh tersebut sebarannya normal atau tidak. Untuk menguji hal ini, maka kita harus menghitung frekuensi harapan untuk kurva normal dimana rata-rata hitung dan simpangan bakunya sama, dengan menggunakan tabel luas daerah kurva normal.
2. Suatu contoh telah mengikuti sebaran normal, maka kita dapat menerima atau menolak hipotesis mengenai sifat-sifat faktor yang mempengaruhi fenomena yang dipelajari. Beberapa keadaan yang cenderung menghasilkan sebaran normal, yaitu: (a) adanya banyak faktor (b) faktor-faktor ini kemunculannya saling tidak gayut (c) faktor-faktor ini pengaruhnya aditif dan (d)

sumbangannya ke varians adalah sama besar. Jadi jika kita mendapatkan bahwa suatu peubah mengikuti sebaran normal, kita tidak mempunyai alasan untuk menolak hipotesis bahwa faktor penyebab yang mempengaruhi peubah adalah yang bersifat aditif, saling tidak gayut dan mempunyai varians yang sama. Sebaliknya bila kita mendapatkan penyimpangan terhadap kaidah normal, hal ini memberi petunjuk akan adanya kekuatan tertentu, seperti seleksi, yang mempengaruhi peubah yang telah dipelajari. Sebagai contoh, sebaran dengan dwimodus menunjukkan tercampurnya pengamatan-pengamatan dari dua populasi. Kemencengan pada hasil susu menunjukkan bahwa catatan hasil susu tersebut merupakan catatan sapi-sapi terpilih, dan sapi-sapi perah yang tidak unggul tidak disertakan dalam catatan tersebut.

3. Jika kita berasumsi bahwa suatu sebaran adalah normal, kita dapat membuat suatu peramalan dan menguji suatu hipotesis berdasarkan atas asumsi ini.

4.5. Perhitungan Peluang Sebaran Normal

Luas daerah dibawah kurva diantara 2 ordinat $X = a$ dan $X = b$, dimana $a < b$, adalah menyatakan peluang bahwa X terletak antara a dan b , dan dinyatakan : Sebaran normal adalah salah satu sebaran. Peluang variabel acak kontinyu, mempunyai fungsi densitas $f(x)$ darimana probabilitasnya dapat dihitung.

$$\int_{-\infty}^{\infty} f(x) dx = 1 \longrightarrow \text{untuk } -\infty < x < \infty$$

$$\text{jika } f(x) = Y = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2}$$

maka :

$$\int_{-\infty}^{\infty} (\sigma \sqrt{2\pi})^{-1} e^{-1/2 \left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

untuk menentukan probabilitas harga x antara a dan b yakni $p(a < x < b)$, maka :

$$p(a < x < b) = \int_b^a (\sigma \sqrt{2\pi})^{-1} e^{-1/2 \left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Penggunaan rumus diatas tak perlu digunakan, karena sudah ada daftar distribusi normal standar, yaitu suatu distribusi normal dengan rata-rata $\mu = 0$ dan simpangan baku $\sigma = 1$. Agar daftar distribusi normal standar dapat digunakan maka distribusi normal umum yang fungsi densitasnya :

jika $f(x) = Y = \frac{1}{\sigma \sqrt{2\pi}} e^{-1/2 \left(\frac{x-\mu}{\sigma}\right)^2}$

dirubah menjadi distribusi normal standar yang fungsi densitasnya :

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-1/2 z^2}$$

Sebaran normal didefinisikan oleh persamaan

$$Y = \frac{1}{\sigma \sqrt{2\pi}} e^{-1/2 \left(\frac{x-\mu}{\sigma}\right)^2}$$

dimana :

π = Nilai konstan yang bila ditulis hingga 4 desimal = 3,1416

e = Bilangan konstan yang bila ditulis dalam 4 desimal = 2,7183

μ = Parameter, ternyata merupakan rata-rata untuk distribusi

σ = Parameter, merupakan simpangan baku untuk distribusi

Rumus diatas tidak perlu kita gunakan, karena telah ada suatu daftar normal standard. Untuk menggunakan daftar (Tabel Z) ini, data yang sebarannya normal umum diubah ke normal standard dengan menggunakan transformasi:

$$Z = \frac{\bar{X} - \mu}{\sigma}$$

Setelah sebaran normal baku diperoleh, maka daftar sebaran normal baku dapat kita gunakan.

Caranya :

1. Hitung Z hingga 2 desimal
2. Gambar kurvanya
3. Letakkan harga Z pada sumbu datar lalu tarik garis vertikal memotong kurva
4. Luas yang tertera dalam daftar adalah luas daerah antara garis vertikal yang memotong kurva tsb. dengan garis tegak dititik 0 (nol)
5. Dalam daftar cari tempat harga Z pada kolom paling kiri, hanya hingga satu desimal dan desimal kedua dicari pada baris paling atas
6. Dari Z dikolom kiri maju kekanan dan dari Z dibaris atas turun ke bawah, maka diperoleh bilangan yang merupakan luas yang dicari. Bilangan yang didapat ditulis dalam bentuk 4 desimal (0,xxxx).

Contoh :

Antara $Z = 0$ dan $Z = 1,2$

Di bawah Z pada kolom kiri cari 1,2 dan di atas cari angka nol. Dari 1,2 maju ke kanan dan dari nol turun, didapat angka 3849. Maka luas daerah yang dicari (daerah diarsir) = 0,3849

Penelitian yang menggunakan contoh perlu dilakukan uji normalitas terhadap data yang diperoleh dari contoh tersebut. Pengujian dilakukan melalui test kecocokan dengan membandingkan antara frekuensi hasil pengamatan (observasi) dari penelitian dengan frekuensi yang diharapkan berdasarkan model sebaran normal (Baca Bab IX).

DAFTAR PUSTAKA

- Dajan, A. 1996. Pengantar Metode Statistika. Jilid I. Cetakan ke-18. Penerbit PT.Pustaka LP3ES, Jakarta.
- Dixon, W.J. dan F.J. Massey, Jr. 1997. Pengantar Analisis Statistik. Cetakan ke-2. Diterjemahkan oleh: Sri Kustantini S. dan Zanzawi S. Gadjah Mada University Press, Yogyakarta.
- Sokal R.R. dan F.J.Rohlf. 1991. Pengantar Biostatistika. Edisi ke-2. Diterjemahkan oleh: Nasrullah dan Setyono Setyo Sunarto. Gadjah Mada University Press, Yogyakarta.
- Spiegel, M.R., I. Y. Susila dan E. Gunawan. 1961. Statistik Edisi SI (Metrik). Schaum Publishing Company, Edinburg.
- Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.
- Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.
- Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

BAB V

SAMPLING DAN SEBARAN SAMPLING

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menjelaskan alasan dan cara sampling, serta menghitung sebaran sampling.

Tujuan Instruksional Khusus

Setelah mempelajari bab ini mahasiswa dapat:

1. Menjelaskan pengertian sampling
2. Menyebutkan dan menjelaskan alasan sampling
3. Menyebutkan dan menjelaskan rencana sampling
4. Menyebutkan dan menjelaskan cara-cara sampling
5. Menjelaskan penggunaan sampling acak berstrata
6. Menjelaskan dan menghitung sebaran sampling rata-rata hitung
7. Menjelaskan dalil limit pusat

Uraian dan Contoh

5.1. SAMPLING

Statistika terbagi atas dua tahapan yakni stadia deskriptif dan inferensia. Fase pertama dikerjakan untuk memberikan informasi mengenai latar belakang suatu

Statistic tertentu untuk mengantarkan ke fase kedua, yakni stadia inferensia. Pada stadia kedua ini karakteristik populasi disimpulkan berdasarkan data contoh yang diambil dari populasi yang bersangkutan. Kesimpulan yang dapat dipertanggungjawabkan harus ditempuh sesuai metode-metode yang benar dalam setiap langkah termasuk cara-cara pengambilan contoh yang disebut metode sampling.

5.1.1. Alasan Sampling

Data yang akan dianalisis diperoleh berdasarkan fakta yang ada dengan sensus dan cara sampling (sering dilakukan). Contoh diambil dan dikumpulkan dari populasi berupa data yang akan dianalisis lebih lanjut.

Beberapa alasan mengapa sampling dilakukan, yaitu:

1. Biaya dan faktor ekonomis

Dengan menggunakan contoh biaya akan lebih ringan dibanding melakukan sensus, jadi faktor ekonomis menentukan pilihan terhadap penggunaan contoh.

2. Ketelitian dalam penelitian

Pemilihan contoh sebagai sumber data memungkinkan kita untuk meningkatkan ketelitian dan memperbanyak variabel yang akan diamati.

3. Penghematan waktu

Karena ukuran contoh lebih kecil dari populasi, maka waktu yang dibutuhkan untuk mendapatkan data tentunya lebih singkat.

4. Percobaan yang sifatnya merusak

Percobaan yang sifatnya merusak maka pilihan contoh sebagai sumber data tidak bisa dihindari.

5. Populasi yang tak hingga

Populasi yang tidak hingga (tidak diketahui seberapa besar jumlahnya) maka pilihan contoh sebagai sumber data, tidak dapat dihindari.

5.1.2. Rencana Sampling

Jika penelitian yang dilakukan ternyata menggunakan sampling, maka sampling perlu direncanakan dengan baik.

Beberapa hal yang perlu diperhatikan sehubungan dengan sampling, yaitu:

1. Rumuskan persoalan yang ingin diketahui
2. Tentukan dengan jelas batas populasi mengenai persoalan yang ingin diketahui itu. Sering kesimpulan tidak benar karena dibuat berdasarkan contoh yang diambil dari populasi yang salah.
3. Definisikan dengan jelas dan tepat segala unit dan istilah yang diperlukan
4. Tentukan unit sampling yang diperlukan. Unit sampling adalah satuan terkecil yang menjadi anggota populasi. Untuk menyelidiki macam beras yang digunakan misalnya, apakah unit samplingnya keluarga atau perorangan anggota keluarga.
5. Tentukan dan rumuskan cara-cara pengukuran dan penilaian yang akan dilakukan, ukuran apa yang dipakai.
6. Tentukan ukuran contoh, yakni berapa obyek yang harus diambil dari populasi sehingga kesimpulan lebih tepat.
7. Tentukan cara sampling mana yang akan ditempuh agar contoh yang diperoleh cukup representatif.
8. Tentukan cara pengumpulan data yang mana akan dilakukan apakah wawancara langsung atau mengumpulkan dari sumber-sumber yang telah ada
9. Tentukan metode analisa mana yang akan digunakan.

5.1.3. Cara-cara Sampling

Sebuah populasi terhingga berukuran N diambil contoh berukuran n . Berapa buah contoh yang dapat diambil dari populasi tersebut? Hal ini tergantung, pada bagaimana anggota populasi diperlakukan ketika contoh diambil.

Dua perlakuan anggota populasi yang digunakan sebagai sampel:

1. Anggota yang telah diambil untuk dijadikan anggota contoh dikembalikan atau disatukan dengan anggota lainnya. Anggota contoh tersebut mempunyai kesempatan untuk diambil kembali pada pengambilan berikutnya. Cara pengambilan contoh ini disebut sampling dengan pengembalian.

Conoh:

Populasi berukuran $N = 4$ dengan anggota A, B, C, D dan contoh yang diambil berukuran $n = 2$ (termasuk contoh beranggotakan sama), maka diperoleh:

contoh 1 : AA	contoh 9 : CA
contoh 2 : AB	contoh 10 : CB
contoh 3 : AC	contoh 11 : CC
contoh 4 : AD	contoh 12 : CD
contoh 5 : BA	contoh 13 : DA
contoh 6 : BB	contoh 14 : DB
contoh 7 : BC	contoh 15 : DC
contoh 8 : BD	contoh 16 : DD

Semuanya ada $4^2 = 16$ buah contoh

Secara umum dikatakan bahwa bila dari populasi berukuran N diambil contoh berukuran n dengan pengembalian, maka semuanya ada N^n buah contoh. Oleh karena anggota yang sudah diambil dikembalikan lagi maka populasi dianggap populasi tak hingga. Cara sampling seperti ini dalam praktek tidak dilakukan, kecuali untuk simulasi untuk mencari hasil-hasil yang mungkin diperoleh dari sampling terhadap populasi tak hingga.

2. Anggota yang telah diambil untuk dijadikan anggota contoh tidak dikembalikan ke populasi. Dengan demikian stiap anggota hanya bisa diambil satu kali. Cara pengambilan demikian dinamakan sampling tanpa pengembalian.

Contoh:

Populasi beranggotakan $N = 5$ terdiri atas A, B, C, D, E. Contoh berukuran $n = 2$ diambil dari populasi itu tanpa pengembalian, maka diperoleh:

contoh 1 : AB	contoh 6 : BD
contoh 2 : AC	contoh 7 : BE
contoh 3 : AD	contoh 8 : CD
contoh 4 : AE	contoh 9 : CE
contoh 5 : BC	contoh 10: DE

Semuanya ada 10 buah contoh yang berlainan.

Jika $N = 4$ dengan anggota A, B, C, D dan $n = 3$, maka semuanya ada 4 buah contoh yang berlainan yakni ABC, ABD, ACD, BCD.

Secara umum: banyaknya contoh berukuran n yang dapat diambil dari sebuah populasi berukuran N adalah:

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

Meskipun dari sebuah populasi bisa diambil lebih dari sebuah contoh, dalam prakteknya, biasanya kesimpulan dibuat berdasarkan hanya sebuah contoh saja dan cara yang digunakan biasanya sampling tanpa pengembalian.

Secara garis besar ada 3 cara sampling yang dapat digunakan agar diperoleh contoh yang cukup representatif, yaitu:

1. Sampling seadanya biasa juga disebut sebagai *convenience sampling*, *accidental sampling* atau *haphazard sampling*. Contoh yang diperoleh dari cara atau teknik ini sangat tersamar dengan populasi dari contoh diambil. Hal ini berakibat kesimpulan yang ditarik bersifat kasar dan sementara.

Contoh:

Pengumpulan data pemilikan ternak di suatu desa,, dengan menanyai orang-orang yang kebetulan ditemui dijalan. Pengacakan tidak berlaku pada sampling ini.

2. Sampling Purposive, dikenal sebagai sampling pertimbangan. Hal ini terjadi bila pengambilan contoh dilakukan berdasarkan pertimbangan tertentu dari si peneliti. Hasilnya dianggap layak bila yang memberi pertimbangan adalah seorang yang ahli dan mengenal populasi dari mana contoh akan diambil. Teknik sampling seperti ini tidak memerlukan pengacakan dalam menentukan pertimbangan dan sangat cocok untuk penelitian studi kasus.
3. Sampling Probabilitas adalah teknik pengambilan contoh dengan cara acak dengan kesalahan pengambilan contoh ditentukan pada taraf tertentu. Setiap anggota populasi mempunyai kesempatan yang sama untuk menjadi anggota contoh. Cara pengacakan bisa dilakukan dengan pemberian nomor (untuk populasi terbatas) atau menggunakan tabel acak yang biasanya tersedia di lampiran pada pustaka-pustaka statistika.

Seberapa sampling acak, diantaranya sampling acak berstrata, sampling sistematis, sampling kluster, sampling ganda, dan lain-lain.. Penjelasan berikut adalah contoh mengoperasionalkan teknik sampling berdasarkan sampling acak berstrata.

5.1.4. Penggunaan Sampling Acak Berstrata

Stratifikasi dilakukan diawali dengan melaksanakan *pre-iventory*, kemudian berdasarkan informasi yang diperoleh pada stadia ini dilakukan strata.

Contoh:

Luas pemilikan tanah distratifikasi karena alasan pemikiran kemampuan tanah untuk memberikan penghidupan kepada petani.

Stratum I : 0 - 200 m² (Landless)

Stratum II : 201 - 1500 m² (Subsistance)

Stratum III: 1501 - 3000 m² (Petani Kecil)

Stratum IV : 3001 - 10.000 m² (Petani Menengah)

Stratum V : > 10.000 m² (Petani Besar)

Selanjutnya tentukan petani contoh untuk setiap strata , caranya sebagai berikut:

1. Strata ditentukan seperti diatas dan jumlah petani pra-iventory ditempatkan pada setiap strata.
2. Ratio imbangan dihitung berdasarkan jumlah petani per stratum dengan total petani (W_i).
3. Varians setiap stratum dihitung (dalam penelitian peternakan pakai animal unit)
4. $W_i S_i$ dan $W_i S_i^2$ dihitung.
5. Jumlah responden dihitung.
6. Kemudian jumlah responden ditebarkan per-strata dengan timbangan $W_i S_i / \sum W_i S_i$
7. $S_i^2 = \text{varians per stratum} = (Y_i - \bar{Y}) / N_i - 1$
8. Hitung total responden (contoh) yang diambil $n = \text{total responden yang diambil}$
 $= (W_i S_i)^2 / v + 1/N W_i S_i^2$
 $v = \text{kesalahan yang diinginkan (1% atau 5%)}$
9. Kemudian jumlah total contoh yang diambil didistribusikan kesetiap stratum dengan timbangan $W_i S_i / \sum W_i S_i$
10. Prosedur sampling selesai. Tabel 9.- Prosedur dalam Sampling.

Strata	$N_i/N=W_i$	S_i	$W_i S_i$	$W_i S_i^2$	n_i
N1	W1	S1	W1S1	W1S1 ²	W1S1/ $W_i S_i \times n$
N2	W2	S2	W2S2	W2S2 ²	W2S2/ $W_i S_i \times n$
N3	W3	S3	W3S3	W3S3 ²	W3S3/ $W_i S_i \times n$
N4	W4	S4	W4S4	W4S4 ²	W4S4/ $W_i S_i \times n$
N5	W5	S5	W5S5	W5S5 ²	W5S5/ $W_i S_i \times n$
N			$W_i S_i$	$W_i S_i^2$	n

Sumber: M.Sabrani *et.al.*1982. Laporan Survey Baseline Ternak Kambing dan Domba.

5.2. Sebaran Sampling

Telaah karakteristik populasi dilakukan berdasarkan contoh yang diambil dari populasi tersebut. Biasanya kesimpulan diambil hanya berdasarkan sebuah contoh. Agar kesimpulan yang diperoleh tidak bias, perlu diketahui pengetahuan sebaran sampling

Pada pengambilan contoh tanpa pengembalian, untuk setiap contoh yang diambil, dihitung nilai statistiknya (rata-rata hitung, simpangan baku atau proporsi). Maka nilai-nilai statistik tersebut akan berbeda untuk setiap contoh. Jika nilai statistik yang sejenis dikumpulkan, lalu disusun dalam suatu daftar sehingga terdapat hubungan antara nilai statistik dan frekuensinya, maka diperoleh kumpulan statistik yang biasa disebut sebaran sampling. Jika statistik yang dikumpulkan itu rata-rata hitung, maka diperoleh sebaran sampling rata-rata hitung.

5.2.1. Sebaran Sampling Rata-rata Hitung

Sebaran sampling rata-rata hitung adalah kumpulan bilangan yang masing-masing merupakan nilai \bar{X} (rata-rata hitung). Statistik rata-rata hitung dari sebaran sampling rata-rata hitung dinotasikan sebagai μ_x , sedangkan simpangan baku dari sebaran sampling rata-rata hitung dinotasi sebagai σ_x , dimana:

$$\mu_x = \mu$$

$$\sigma_x = \sigma / \sqrt{n} \times \sqrt{(N-n)/(N-1)}$$

Kita buktikan rumus di atas, bahwa rata-rata hitung dari sebaran sampling rata-rata hitung adalah rata-rata hitung populasi dan simpangan baku dari sebaran sampling rata-rata hitung adalah simpangan baku populasi dibagi akar jumlah contoh (dikalikan faktor koreksi untuk populasi terbatas).

Contoh:

Suatu populasi dengan $N = 10$ dengan data : 98, 99, 97, 98, 99, 98, 97, 97, 98 dan 99. Setelah dihitung $\mu = 98$ dan $\sigma = 0,78$. Kemudian diambil contoh berukuran $n = 2$, sehingga semuanya ada 45 buah contoh. Setiap contoh kita hitung rata-rata hitungnya. Data dalam tiap contoh dan rata-rata hitungnya, bisa dilihat pada Tabel 7.

Tabel 7. Jumlah Contoh berukuran $n = 2$ dan Rata-rata Hitungnya Diambil dari Populasi berukuran $N = 10$.

Contoh	Rerata	Contoh	Rerata	Contoh	Rerata
(98,99)	98,5	(99,98)	98,5	(99,98)	98,5
(98,97)	97,5	(99,99)	99,0	(99,97)	98,0
(98,98)	98,0	(97,98)	97,5	(99,97)	98,0

(98,99)	98,5	(97,99)	98,0	(99,98)	98,5
(98,98)	98,0	(97,98)	97,5	(99,99)	99,0
(98,97)	97,5	(97,97)	97,0	(98,97)	97,5
(98,97)	97,5	(97,97)	97,0	(98,97)	97,5
(98,98)	98,0	(97,98)	97,5	(98,98)	98,0
(98,99)	98,5	(97,99)	98,0	(98,99)	98,5
(99,97)	98,0	(98,99)	98,5	(97,97)	97,0
(99,98)	98,5	(98,98)	98,0	(97,98)	87,5
(99,99)	99,0	(98,97)	97,5	(97,99)	98,0
(99,98)	98,5	(98,97)	97,5	(97,98)	97,5
(99,97)	98,0	(98,98)	98,0	(97,99)	98,0
(99,97)	98,0	(98,99)	98,5	(98,99)	98,5

Jumlah semua rata-rata hitung = 4410

Sumber: Sudjana (1975).

$$\begin{aligned} \mu_x &= 4410/45 = 98,0 && \text{maka terbukti } \mu_x = \mu \\ x &= \overline{X_i} - \overline{X}/n-1 = 0,53 && \text{maka hasil perhitungan populasi } \sigma_x = 0,78 \\ \text{Untuk } n &= 2 \text{ dan } N = 10, \text{ maka } \sigma_x &= \sigma / n \times (N-n / N-1) \\ &&& = 0,78 / 2 \times 10-2/10-1 \\ &&& = 0,53 \end{aligned}$$

Hasil ini sama dengan σ_x yang dihitung dari kombinasi contoh yang diambil.

$$\begin{aligned} \sigma_x &= \sigma / n \times (N-n / N-1) && \text{untuk } n \text{ cukup besar dibandingkan dengan ukuran} \\ &&& \text{populasi } N \text{ atau } (n/N > 5\%) \\ \sigma_x &= \sigma / n && \text{untuk ukuran } n \text{ relatif kecil dibandingkan ukuran} \\ &&& \text{populasi } N \text{ atau } (n/N < 5\%) \end{aligned}$$

Catatan:

Contoh atau sampel yang berukuran cukup besar, ternyata distribusi sampling rata-rata hitung memenuhi sebuah dalil:

“Jika ukuran contoh n cukup besar, maka sebaran sampling rata-rata hitung, ternyata mendekati sekali ke sebaran normal.”

Transformasi ke normal standard digunakan rumus:

$$\text{Rumus Umum } Z = \frac{X - \mu_x}{x} \quad \text{diperinci: } Z = \frac{X - \mu_x}{\sigma / n} \quad \text{pada } (n/N < 5\%)$$

$$Z = \frac{X - \mu_x}{\sigma / n \times (N - n / N - 1)} \quad \text{pada } (n/N > 5\%)$$

Contoh:

Jika dari populasi 10.000 ekor ayam, $\mu = 158$ butir (rata-rata jumlah telur yang dihasilkan selama 6 bulan) dan $\sigma = 6$ butir, diambil sebuah contoh terdiri atas 50 ekor. Berapa peluang nilai rata-rata hitung produksi telur dari ke 50 ekor tersebut. (periode 6 bulan):

- a). Antara 154 butir dan 160 butir
- b). Paling sedikit 162 butir

Jawab:

$$N = 10.000 \quad \mu = 158$$

$$n = 50 \quad \sigma = 6,0$$

maka: $\mu_x = \mu = 158$

$$\sigma_x = \sigma / n = 6 / 50 = 0,85$$

a). $X_1 = 154$ dan $X_2 = 160$

$$Z_1 = 154 - (158 / 0,85) = -4,70 \quad Z_2 = 160 - (158 / 0,85) = 2,35$$

Luas daerah dari grafik normal standard = $0,5 + 0,4906 = 0,9906$

Luas tersebut merupakan peluang bahwa rata-rata produksi telur selama 6 bulan antara 154 dan 160 butir untuk $n = 50$

b). $X = 162$

$$Z = 162 - (158 / 0,85) = 4,70$$

Luas daerah dari grafik normal standard adalah dari $Z = 4,70$ ke sebelah kanan, luasnya praktis = 0

Catatan: σ_x : kekeliruan standard = standard error (SE), mengukur besarnya perbedaan rata-rata yang diharapkan dari contoh ke contoh.

5.2.2. Dalil Limit Pusat (*The Central Limit Theorem*)

Apabila kita melakukan sampling dari populasi yang tidak memiliki sebaran normal, bagaimana bentuk sebaran sampling dari rata-rata contohnya ? Secara matematis, sebaran rata-rata contoh akan menyerupai sebaran normal jika besarnya contoh n bertambah tanpa batas. Dalil ini dinyatakan secara matematis dan berguna dalam analisis statistik inferensia sebagai berikut:

“Dalil Limit Pusat (*Central Limit Theorema*)” :

Jika contoh acak dipilih dari populasi dengan rata-rata (μ) dan varians (σ)², dan jika besarnya contoh n bertambah secara tidak terbatas, maka rata-rata contohnya akan memiliki sebaran yang mendekati sebaran normal dengan rata-rata $\mu_x = \mu$ dan simpangan baku $\sigma_x = \sigma / n$.

Catatan: Bahwa sebaran rata-rata contoh akan menyebar normal jika diambil dari populasi yang sebarannya normal atau apabila contoh n besar sekali walaupun diambil dari populasi yang sebarannya tidak diketahui.

DAFTAR PUSTAKA

- Dixon, W.J. dan F.J. Massey, Jr. 1997. Pengantar Analisis Statistik. Cetakan ke-2. Diterjemahkan oleh: Sri Kustantini S. dan Zanzawi S. Gadjah Mada University Press, Yogyakarta.
- Sokal R.R. dan F.J.Rohlf. 1991. Pengantar Biostatistika. Edisi ke-2. Diterjemahkan oleh: Nasrullah dan Setyono Setyo Sunarto. Gadjah Mada University Press, Yogyakarta.
- Spiegel, M.R., I. Y. Susila dan E. Gunawan. 1961. Statistik Edisi SI (Metrik). Schaum Publishing Company, Edinburg.
- Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.
- Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.
- Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

BAB VI

TEORI PENDUGAAN PARAMETER

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menjelaskan pengertian pendugaan parameter dan mampu menghitung pendugaan parameter berdasarkan statistik dari sampel

Tujuan Instruksional Khusus

Setelah mempelajari bab ini mahasiswa dapat:

1. Menjelaskan ciri-ciri penduga yang baik
2. Menjelaskan cara-cara pendugaan (pendugaan titik dan interval)
3. Menghitung pendugaan parameter dengan sampel besar
4. Menghitung pendugaan parameter dengan sampel kecil

Uraian dan Contoh

Penelitian dengan menggunakan sampling, selalu menimbulkan pertanyaan:

1. Seberapa jauh hasil yang diperoleh dapat dipercaya
2. Seberapa jauh kemungkinannya bahwa perbedaan hasil yang teramati dan yang diharapkan secara hipotesis hanya disebabkan oleh faktor kebetulan.

Jawaban untuk pertanyaan 1. mengenai tingkat keyakinan diantisipasi dengan pembentukan batas kepercayaan untuk statistik contoh, yakni pada pendugaan secara statistik. Sedangkan pertanyaan 2. dijawab melalui test hipotesis.

Penelitian (percobaan atau survei) dengan sampling, kesimpulan tentang fenomena yang diselidiki didasarkan atas data yang dikumpulkan. Penarikan kesimpulan tentang populasi berdasarkan sampel yang dipilih dapat berbentuk pendugaan parameter θ (dimana θ bisa μ atau σ) oleh statistik $\hat{\theta}$ (dimana $\hat{\theta}$ dalam hal ini adalah \bar{x} atau S_d). Karena penduga merupakan fungsi dari nilai-nilai sampel, maka penduga juga merupakan variabel acak dan memiliki sebaran sampling. Untuk sampel yang cukup besar ($n \geq 30$) sebarannya adalah normal.

6.1. Ciri-ciri Penduga yang Baik

Ciri-ciri penduga yang baik terdiri dari 3, yaitu:

1. Tidak bias (*Unbiased*)

Apabila rata-rata dari distribusi sampling suatu statistik adalah sama dengan parameter populasi, maka statistik itu disebut penduga yang tidak bias bagi parameter. Namun bila tidak demikian maka nilai statistik yang bersangkutan disebut sebagai penduga yang bias terhadap parameter. Rata-rata sample (\bar{X}) pada hakekatnya penduga tidak bias terhadap μ . s^2 penduga tidak bias bagi σ^2 asal digunakan pembagi $(n-1)$.

2. Efisien

Apabila distribusi sampling dari dua statistik mempunyai rata-rata (atau harapan) yang sama, maka statistik dengan varians yang lebih kecil disebut penduga yang efisien dari rata-rata, sedangkan statistik yang lain disebut penduga yang tidak efisien. Jika kita membandingkan Mean dan Median contoh sebagai penduga parameter μ , maka mean merupakan penduga yang lebih efisien dibandingkan

dengan median sebesar 157%. Artinya mean sebagai penduga μ , 57% lebih efisien dibandingkan dengan median.

3. Konsisten

Penduga yang konsisten merupakan penduga yang berkonsentrasi secara sempurna pada parameter, jika besarnya contoh bertambah secara tak terhingga. Rata-rata hitung contoh \bar{X} merupakan penduga μ yang konsisten karena bias rata-ratanya = 0 bagi sebarang contoh n , dan $\sigma^2(X) = \sigma^2 / \sqrt{n} \rightarrow 0$ jika contoh n

6.2. Cara-cara Pendugaan

Dua jenis pendugaan secara statistik, yaitu:

1. Pendugaan Titik (Point estimation)
2. Pendugaan Interval (Interval Estimation)

6.2.1. Pendugaan Titik

Jika parameter nilainya diduga oleh sebuah nilai yang tertentu, maka dinamakan titik dugaan. Pada pendugaan titik, tingkat keyakinan (*level of confidence*) tidak bisa diukur karena tidak adanya ulangan pengamatan (pendugaan dilakukan terhadap pengamatan tunggal).

Contoh:

10 ekor sapi diberi makan tertentu (kondisi fisik, lingkungan dan cara pemberian pakan diatur sama). Sebulan kemudian pertambahan bobot badannya dicatat (kg), datanya adalah sbb:

45 109 61 80 79

93 48 35 57 65

$$\sum X = 670 \quad n = 10 \quad \bar{X} = 670 / 10 = 67$$

$$s^2 = 530,4 \quad sd = \sqrt{530,4} = 23$$

Hasil perhitungan statistik lokasi dan sebaran, 67 kg adalah nilai titik dugaan untuk μ dan 530,4 adalah nilai titik dugaan untuk σ^2 . Sehingga rata-rata dari contoh ke contoh menyimpang sebesar $530,4/10$ (atau $23/10$) = 7,28 kg.

6.2.2. Pendugaan Interval

Pendugaan interval memberikan nilai-nilai statistik dalam suatu interval, sehingga kita dapat menyatakan berapa besar tingkat kepercayaan kita, bahwa interval dugaan diatas betul-betul mencakup parameter yang kita duga.

Rumus umum:

$$St - Z \frac{1}{2} \sigma_{st} < \text{Parameter} < St + Z \frac{1}{2} \sigma_{st}$$

dimana:

St = Statistik contoh (penduga)

σ_{st} = Standard deviasi statistik contoh

$Z \frac{1}{2}$ = Koefisien yang sesuai dengan interval keyakinan yang digunakan dalam pendugaan interval dan nilainya diberikan dalam tabel Z. Misalkan dalam pendugaan, kita gunakan interval keyakinan sebesar 95%, maka dalam jangka panjang kita akan mentolerir kesalahan duga (*error of estimate*) sebesar = 5%

Pendugaan dengan interval dibatasi oleh batas keyakinan atas dan batas keyakinan bawah. Statistik contoh merupakan variabel acak yang tersebar secara normal, maka luas interval keyakinan dapat dicari dengan mengubah variabel acak normal umum ini ke normal standard Z. Misalkan bila kita gunakan interval keyakinan sebesar 95% dalam pendugaan parameter, maka kesalahan duga yang merupakan nilai-nilai statistik contoh yang lebih besar atau yang lebih kecil dari pada yang dispesifikasikan dalam kedua batas, masing-masing akan sebesar $\frac{1}{2}$, dimana = kesalahan duga sebesar 5%. Dengan kata lain kesalahan duga yang terdapat di luar

batas atas atau batas bawah masing-masing menjadi sebesar 0.025 atau 2.5%. Koefisien standard $Z_{\frac{1}{2}\alpha}$ yang sesuai dengan kedua nilai batas keyakinan menjadi $0.5000 - 0.0250 = 0.4750$. Per tabel luas kurva normal terbaca sebesar 1.96, sehingga nilai koefisien standard $Z_{\frac{1}{2}\alpha}$ pada kedua batas keyakinan ialah ± 1.96 . Jika pendugaan dilakukan terhadap parameter μ , maka luas intervalnya menjadi $\mu \pm 1.96 / \sqrt{n}$.

6.3. Pendugaan Parameter dengan Sampel Besar

6.3.1. Pendugaan Parameter μ dengan σ Diketahui dan Populasi Tidak Terbatas

Parameter μ diduga oleh \bar{x} sebagai penduga tidak bias dengan interval keyakinan sebesar 95%, sedangkan σ diketahui. Interval keyakinannya dapat diberikan sebagai:

$$\bar{X} - Z_{0,025} \sigma_x < \mu < \bar{X} + Z_{0,025} \sigma_x$$

dimana $\sigma_x = \sigma / \sqrt{n}$ dan $Z_{0,025} = 1,96$ sehingga

$$\bar{X} - 1,96 \sigma / \sqrt{n} < \mu < \bar{X} + 1,96 \sigma / \sqrt{n}$$

pertidaksamaan di atas dinyatakan dalam koefisien probabilitas atau koefisien keyakinan $1 - \alpha$ sebagai berikut :

$$P(\bar{X} - Z_{\frac{1}{2}\alpha} \sigma / \sqrt{n} < \mu < \bar{X} + Z_{\frac{1}{2}\alpha} \sigma / \sqrt{n}) = 1 - \alpha$$

Untuk $\alpha = 0,05$ maka :

$$P(\bar{X} - Z_{\frac{1}{2}\alpha} \sigma / \sqrt{n} < \mu < \bar{X} + Z_{\frac{1}{2}\alpha} \sigma / \sqrt{n}) = 0,95$$

Contoh :

Penelitian survei kepariwisataan telah dilakukan untuk memperkirakan pengeluaran rata-rata para wisman di Jateng. Sampel acak yang terdiri atas 100

wisman dipilih dari populasi yang dianggap tidak terkira. Dari hasil interview diperoleh data bahwa rata-rata pengeluaran per kunjungan adalah \$ 800 per orang. Jika kita anggap $\sigma = \$120$, maka buatlah interval keyakinan 95% untuk menduga rata-rata pengeluaran wisman per kunjungannya di Jateng.

Jawab :

Dari soal diatas diketahui: $n = 100$ $\bar{X} = \$800$ $\sigma = \$120$.
 $\alpha = 0,05$ $1 - \alpha = 0,95$ $Z_{0,025} = 1,96$

sehingga $\sigma_x = 120 / \sqrt{100} = 12$

maka : $P (800 - 1,96 (12) < \mu < 800 + 1,96 (12)) = 0,95$
 $P (776,48 < \mu < 823,53) = 0,95$

Disimpulkan bahwa rata-rata pengeluaran wisman/orang/kunjungan berkisar antara \$776,48 s/d \$823,53

6.3.2. Pendugaan Parameter μ dengan σ Diketahui dan Populasi Terbatas

Jika contoh acak dipilih dari populasi yang terbatas tanpa pemulihan, berkecendrungan untuk kurang dari σ / n . Untuk itu penggunaan penduga yang tidak bias \bar{X} untuk menduga parameter μ dengan interval keyakinan 95% dan σ diketahui, berubah menjadi:

$$P(\bar{X} - 1.96 \cdot (\sigma / \sqrt{n}) \times (N-n / N-1) < \bar{\mu} < \bar{X} + 1.96 \cdot (\sigma / \sqrt{n}) \times (N-n / N-1)) = 95\%$$

Ketentuan populasi terbatas:

Bila $n / N > 5\%$ pakai faktor koreksi populasi terbatas $(N-n / N-1)$

Bila $n / N < 5\%$ tidak memakai faktor koreksi

Contoh:

Jika diketahui contoh acak sebesar $n = 64$ dan $\bar{X} = 0.1165$ diambil dari populasi terbatas sebesar $N = 300$ dan yang diketahui memiliki $\sigma = 0.0120$, maka pendugaan parameter μ dengan interval keyakinan sebesar 95.45%, dirumuskan sbb:

$$\begin{aligned}
 P(0.1165 - (2 \times 0.012 / \sqrt{64}) \times (300-64 / 300-1) < \mu < 0.1165 + (2 \times 0.012 / \sqrt{64}) \times (300-64 / 300-1) &= 0.9545 \\
 P(0.1165 - 2(0.00134) < \mu < 0.1165 + 2(0.00134)) &= 0.9545 \\
 P(0.11382 < \mu < 0.11918) &= 0.9545
 \end{aligned}$$

6.3.3. Pendugaan Parameter μ dengan σ Tidak Diketahui

Pada hakekatnya tergantung pada simpangan kuadrat dari μ , sehingga bila μ tidak diketahui maka σ juga tidak diketahui. Jika μ dan σ tidak diketahui, pendugaan terhadap μ dapat dilakukan dengan jalan melakukan pendugaan terhadap σ . Pendugaan σ adalah simpangan baku contoh (sd), sehingga pendugaan μ dengan interval keyakinan 95%, adalah sbb:

$$P(\bar{X} - Z_{1/2} * sd / \sqrt{n} < \mu < \bar{X} + Z_{1/2} * sd / \sqrt{n}) = 0.95$$

Jika kita menggunakan jumlah contoh yang besar ($n = 25$ s/d 50), pendugaan ini akan memberikan hasil yang memuaskan. Tetapi bila contoh yang digunakan $n < 25$, maka pendugaan nilai σ oleh sd akan menghasilkan kesalahan (selisih dugaan) yang berarti. Kondisi seperti ini kita harus menggunakan sebaran t-student. Jika contoh acak dipilih dari populasi yang terbatas tanpa pemulihan, maka pendugaan μ juga harus dilakukan dengan menggunakan sebaran t-student.

6.3.4. Pendugaan Parameter $\mu_1 - \mu_2$ dimana σ_1 dan σ_2 Diketahui

Dua rata-rata populasi yang tidak diketahui dapat dibandingkan dengan cara menduga beda $\mu_1 - \mu_2$. Penduga yang tidak bias dari perbedaan tersebut adalah $X_1 - X_2$ yang merupakan beda antara dua rata-rata contoh. Pendugaan $\mu_1 - \mu_2$ dengan interval keyakinan sebesar 0.95 dapat diberikan sbb:

$$P((X_1 - X_2) - Z \frac{1}{2} \times \sqrt{\sigma^2_1/n_1 + \sigma^2_2/n_2} < \mu_1 - \mu_2 < (X_1 - X_2) + Z \frac{1}{2} \times \sqrt{\sigma^2_1/n_1 + \sigma^2_2/n_2}) = 0.95$$

Jika $\sigma^2_1 = \sigma^2_2 = \sigma^2$, maka σ^2 merupakan penduga dari varians populasi umum yang dipersatukan, dengan $1 - \alpha = 0.95$ dapat dinyatakan sebagai berikut:

$$P(\overline{X_1 - X_2} - Z \frac{1}{2} \times \sigma \sqrt{(1/n_1 + 1/n_2)} < \mu_1 - \mu_2 < \overline{X_1 - X_2} + Z \frac{1}{2} \times \sigma \sqrt{(1/n_1 + 1/n_2)}) = 0.95$$

Catatan:

1. Jika contoh acak dipilih dari populasi terbatas atau tanpa pemulihan, maka varians statistik harus dihitung dengan faktor koreksi populasi terbatas (* $N-n/N-1$),
2. Jika σ^2_1 dan σ^2_2 tidak diketahui, maka kedua varians populasi tersebut diduga dengan varians contoh s^2_1 dan s^2_2 . Hasil dugaan akan memuaskan jika contoh n yang digunakan cukup besar.
3. Jika n kecil, maka pendugaan $\mu_1 - \mu_2$ harus menggunakan sebaran t-student, dimana teori sebarannya akan dibicarakan pada sub pokok bahasan berikut.

Contoh:

Seorang importir menerima kiriman 2 strain DOC dalam jumlah besar, masing-masing dari Thailand dan USA. Untuk mengetahui berapa lama (hari) bobot 2 kg. dicapai, diambil contoh acak masing-masing sebesar 50 ekor. Hasil penelitian menunjukkan, waktu yang dibutuhkan untuk mencapai bobot 2 kg., untuk DOC Thailand rata-rata 50 hari sedangkan DOC USA rata-rata 60 hari. Berdasarkan pengalaman sebagai importir diketahui bahwa simpangan baku ke-2 DOC tersebut

konstan masing-masing 12 dan 15 hari. Hitung pendugaan parameter dari beda rata-rata lama waktu untuk mencapai bobot 2 kg diatas dengan menggunakan interval keyakinan 95%.

Jawab:

$$\frac{n_1 = 50}{X_1 = 50} \quad \frac{n_2 = 50}{X_2 = 60}$$

$$\sigma_1 = 12 \quad \sigma_2 = 15$$

$$\alpha = 0.05 \quad 1 - \alpha = 0.95$$

$$\sigma_1 - \sigma_2 = (12)^2 / 50 + (15)^2 / 50 = 2.72$$

$$P((60-50) - 1.96(2.72) < \mu_1 - \mu_2 < (60-50) + 1.96(2.72)) = 0.95$$

$$P(4.67 < \mu_1 - \mu_2 < 15.33) = 0.95$$

Jadi dugaan tentang beda rata-rata lama waktu untuk mencapai bobot 2 kg.adalah antara 4.67 sampai dengan 15.33 hari.

6.4. Pendugaan Parameter dengan Contoh Kecil

Pada pendugaan parameter dengan contoh besar, penduga sd bagi σ akan diperoleh hasil pendugaan yang memuaskan bila contoh acak yang digunakan besar. Bila contoh acak yang digunakan kecil ($n < 30$) maka hasil dugaan kekeliruannya cukup besar.

Pada umumnya jika contoh acak kecil, pendugaan parameter populasi sebaiknya dilakukan dengan sebaran t. Sebaran t sama seperti sebaran normal dalam hal sifat setangkup dan bernilai dari negatif tak terhingga sampai positif tak terhingga. Perbedaannya dengan sebaran normal adalah pada bentuknya yang berbeda-beda yang tergantung pada derajat kebebasannya yang kita artikan sebagai besaran n-1 dengan n sebagai ukuran contoh yang digunakan untuk mendapatkan varians. Perlu diingat bahwa n-1 adalah pembagi yang digunakan untuk mendapatkan penduga yang tidak

bias bagi varians dari jumlah kuadratnya. Besar derajat bebas suatu sebaran t , sama dengan besar derajat bebas simpangan baku pada perbandingan $(\bar{X} - \mu) / \frac{s}{\sqrt{n}}$. Derajat bebas (db) dapat berkisar dari 1 sampai tidak terhingga. Sebaran t dengan db=1 menyimpang jauh dari kaidah normal. Bertambahnya nilai db, sebaran t mendekati sebaran normal baku (0,1). Sebaran t dengan db =30 tidak dapat dibedakan dari sebaran normal. Untuk db = 30 , sebaran t merupakan sebaran normal.

Rumus:

$$t = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

6.4.1. Pendugaan Parameter μ dengan σ Tidak Diketahui dan Populasi Tak Terebatas

Interval keyakinannya dalam pernyataan probabilitas koefisien keyakinan $1 - \alpha$, dinyatakan sebagai :

$$P (\bar{x} - t_{(0,025; n-1)} \frac{sd}{\sqrt{n}} < \mu < \bar{x} + t_{(0,025; n-1)} \frac{sd}{\sqrt{n}}) = 0,95$$

Untuk $\alpha = 0,05$

Contoh:

Diketahui : $n = 10$ $\bar{x} = 112$ $sd = 11$ $\alpha = 0,05$ dan $1 - \alpha = 0,95$

Maka :

$$P (112 - t_{(0,025; 9)} \frac{11}{\sqrt{10}} < \mu < 112 + t_{(0,025; 9)} \frac{11}{\sqrt{10}}) = 0,95$$

$$P (112 - 2,262 (11 / \sqrt{10}) < \mu < 112 + 2,262 (11 / \sqrt{10})) = 0,95$$

$$P (104 < \mu < 120) = 0,95$$

Catatan :

Untuk populasi terbatas $n / N > 5\%$ maka digunakan faktor koreksi $\frac{N - n}{N - 1}$

6.4.2. Pendugaan Parameter $\mu_1 - \mu_2$ dimana σ_1 dan σ_2 Tidak Diketahui

Interval keyakinannya dalam pernyataan probabilitas koefisien keyakinan $1 - \alpha$, dinyatakan sebagai :

$$P((\bar{X}_1 - \bar{X}_2) - t_{(0,025; db)} Sp \sqrt{1/n_1 + 1/n_2} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{(0,025; db)} Sp \sqrt{1/n_1 + 1/n_2}) = 0,95$$

Dimana Sp = nilai duga sd gabungan

$$Sp = \frac{\frac{\sum X_1^2 - (\sum X_1)^2}{n_1} + \frac{\sum X_2^2 - (\sum X_2)^2}{n_2}}{(n_1 - 1) + (n_2 - 1)}$$

distribusi t ini mempunyai derajat bebas (db) = $n - 2$

Contoh :

Sebuah sampel acak sebesar $n_1 = 7$ dipilih dari populasi normal dengan μ_1 , sedangkan sampel acak sebesar $n_2 = 6$ dipilih dari populasi normal dengan μ_2 . Hasil observasi sampel diatas ditunjukkan pada tabel berikut :

X_1	57,8	56,2	61,9	54,4	53,6	56,4	53,2
X_2	64,2	58,7	63,1	62,5	59,8	59,2	####

Buatlah interval keyakinan sebesar 95% untuk menduga $\mu_1 - \mu_2$

Jawab :

Dari persoalan tersebut diperoleh informasi bahwa :

$$N_1 = 7 \quad n_2 = 6 \quad \bar{X}_1 = 56,21 \quad \bar{X}_2 = 61,25$$

- Spiegel, M.R., I. Y. Susila dan E. Gunawan. 1961. Statistik Edisi SI (Metrik). Schaum Publishing Company, Edinburg.
- Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.
- Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.
- Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

BAB VII

UJI HIPOTESIS

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menjelaskan pengertian uji hipotesis dan menghitung pengujian hipotesis dengan sampel besar dan sampel kecil

Tujuan Instruksional Khusus

Setelah mempelajari bab ini mahasiswa dapat:

1. Menjelaskan kesalahan jenis I dan II
2. Menjelaskan langkah-langkah pengujian hipotesis
3. Menghitung pendugaan parameter dengan sampel besar
4. Menghitung pendugaan parameter dengan sampel kecil

Uraian dan Contoh

Penerapan statistika yang paling sering pada penelitian biologi adalah pengujian hipotesis. Pada penelitian biologi, hasil yang diperoleh (variat) dari satu pengamatan ke pengamatan berikutnya biasanya tidak begitu jelas perbedaannya. Untuk ini diperlukan uji statistik dalam pengambilan keputusan terhadap hipotesis yang disusun.

Hipotesis yang bersifat statistik adalah suatu asumsi mengenai parameter fungsi frekuensi variabel acak. Misalnya sebaran peluang yang merupakan model bagi sebaran X , katakanlah tentang sex ratio dari suatu hasil penetasan telur ayam.

Berapa peluang 3 sex jantan yang dilahirkan pada penetasan 10 butir telur, jika diketahui $p = 0,51$. Fungsi frekuensinya dinyatakan sebagai:

$$f(X=\text{jantan}) = \frac{10!}{3!(10-3)!} (0,51)^3 (0,49)^7 = 0,1083$$

Parameter p diatas merupakan peluang lahirnya jenis kelamin jantan dan merupakan asumsi yang memiliki karakteristik hipotesis karena $p = 0,51$ merupakan parameter fungsi frekuensi variabel acak X . Jika kita meragukan kebenaran hipotesis tersebut, dapat dilakukan pengujian secara statistik jika ada data yang mendukung dan dianalisis menurut kriteria statistik yang berlaku. Pengujian statistik tersebut dianggap suatu prosedur untuk menentukan apakah hipotesis diatas diterima atau ditolak.

Hipotesis yang kita uji dapat dinyatakan sbb:

$$H_0 : p = 0,51$$

$$H_1 : p \neq 0,51$$

H_0 adalah hipotesis nol (null hypothesis) dan merupakan hipotesis yang akan diuji dan yang nantinya akan diterima atau ditolak tergantung pada hasil penelitian atau samplangnya. H_1 merupakan hipotesis alternatif atau hipotesis tandingan (*alternative hypothesis*).

Salah satu tahapan terpenting dari prosedur pengujian hipotesis adalah menentukan nilai statistik contoh yang merupakan dasar untuk menerima atau menolak H_0 . Nilai statistik ini menentukan daerah kritis (critical region) pengujian tersebut.

Interval keyakinan, yang diformulasikan sebagai $1 - \alpha$ akan menentukan batas keyakinan dari sebaran $f(X)$. Prosedur pengujian hipotesis, kita menerima atau menolak pernyataan katakanlah $\mu > \mu_0$, $\mu = \mu_0$ atau $\mu < \mu_0$ tergantung pada apakah μ_0 terletak atau tidak dalam interval keyakinan yang kita tentukan.

7.1. Kesalahan Jenis I dan Jenis II

Setiap proses pengambilan keputusan tentang apakah kita menerima atau menolak suatu hipotesis, kita selalu dihadapkan pada 2 macam kesalahan pengambilan keputusan yang berbeda. Kesalahan menolak H_0 benar dinamakan kesalahan jenis I (Type I error) atau kesalahan (α error) dengan peluang sebesar α yakni taraf nyata pengujiannya. Kesalahan menerima H_0 palsu dinamakan kesalahan jenis II (Type II error) atau kesalahan β (β error) dengan peluang sebesar β

Secara skematis, kedua jenis kesalahan (kesalahan jenis I dan jenis II) serta hubungannya dapat dilihat pada Tabel 8.

Tabel 8. Beberapa Kemungkinan Hasil Pengujian Hipotesis

KEPUTUSAN	HIPOTESIS	
	Jika H_0 Benar	Jika H_1 benar
Terima H_0	Keputusan betul = peluang = $1 - \alpha =$ "Tingkat Kepercayaan"	Kesalahan Jenis II, peluang = β
Terima H_1	Kesalahan jenis I, peluang = $\alpha =$ taraf nyata	Keputusan betul = peluang = $1 - \beta =$ Kuasa Pengujian.

Dalam merencanakan suatu penelitian untuk menguji suatu hipotesis, jelas bahwa kedua jenis kesalahan harus sekecil mungkin dan dinyatakan dalam peluang. Kesalahan jenis I dinyatakan sebagai α dan biasanya ditentukan terlebih dahulu dengan harga yang biasa digunakan yakni $\alpha = 0,01$ dan $\alpha = 0,05$ dan disebut sebagai taraf nyata (*level of significance*). Bila $\alpha = 0,05$ berarti 5 dari setiap 100 kesimpulan bahwa kita akan menolak H_0 yang seharusnya diterima atau dengan kata lain kira-kira 95% "percaya" bahwa kita telah membuat kesimpulan yang benar. Kesalahan jenis II dinyatakan sebagai β . Setiap pengujian dengan α yang ditentukan,

maka besar β dapat dihitung. Kuantitas α , β dan n berhubungan sedemikian rupa sehingga jika 2 dari ketiga kuantitas tersebut dispesifikasikan, maka yang ketiga akan dapat ditentukan secara otomatis. Nilai $1-\beta$ disebut 'Kuasa Pengujian'

7.2. Langkah-langkah Pengujian Hipotesis

Prosedur umum pengujian hipotesis yang harus diikuti terdiri dari tahap:

1. Nyatakan hipotesis nol (H_0) dan alternatifnya (H_1)
2. Pilih taraf nyata dan tentukan besarnya contoh (n)
3. Pilih statistik uji yang sesuai sebagai dasar bagi prosedur pengujian. Hal ini tergantung pada asumsi tentang bentuk sebaran dan bentuk hipotesisnya.
4. Tentukan daerah kritis. Hal ini sebagian juga tergantung pada H_1 yang kita terima/ tentukan.
5. Tentukan data dari contoh yang diambil dan hitung statistiknya dan transformasi ke Z .
6. Jika statistik yang dihitung terletak didaerah penolakan, maka tolak H_0 dan sebaliknya.

7.3. Pengujian Hipotesis dengan Sampel Besar

7.3.1. *Pengujian Parameter Rata - Rata, $H_0 : \mu = \mu_0$ dimana σ^2 Tidak Diketahui*

Rumus :

$$Z = \frac{\bar{X} - \mu_0}{Sd / \sqrt{n}}$$

Daerah kritis pengujian untuk populasi tak terbatas :

$$\frac{\bar{X} - \mu_0}{Sd/\sqrt{n}} + Z \frac{1}{2} \alpha \quad \text{dan} \quad \frac{\bar{X} - \mu_0}{Sd/\sqrt{n}} - Z \frac{1}{2} \alpha$$

Contoh :

Suatu populasi yang terdiri atas: garam beriodium (dalam pak) yang dihasilkan oleh suatu perusahaan, mengandung 50% iod. Team evaluasi Departemen Perindustrian melakukan pengecekan dengan mengambil 100 pak garam beriodium dari populasi. Hasil analisis menunjukkan bahwa rata-rata kandungan iod dari sampel tersebut adalah 54% dan simpangan baku 2,5%. Apakah garam iod yang dihasilkan perusahaan tersebut kandungan ioidnya benar 50% ?

Jawab :

1. Rumuskan hipotesia nol-nya

$$H_0 : \mu = 50\% ; H_1 : \mu \neq 50\%$$

2. $\alpha = 0,05$

$$\frac{\bar{X} - \mu_0}{Sd/\sqrt{n}}$$

3. $Z = \frac{\bar{X} - \mu_0}{Sd/\sqrt{n}}$

4. Statistik uji diatas merupakan distribusi normal dengan $\mu = 0$ dan $\sigma = 1$

5. Daerah kritis dan taraf nyata $\alpha = 0,05$ dalam pengujian 2 arah ialah ;

$$Z > 1,96 \text{ dan } Z < -1,96$$

6. $Z = \frac{54 - 50}{2,5 / \sqrt{100}} = 16$

$$2,5 / \sqrt{100}$$

7. Karena $Z > 1,96$ (Z hitung lebih besar dari Z tabel), maka kita nyatakan bahwa beda antara hasil sampel sebesar 54% dengan rata-rata hipotesis $\mu = 50\%$ adalah nyata. Ini berarti $\mu = 50\%$ ditolak dan penolakan ini membawa resiko kesalahan H_0 sebesar 5% dari keseluruhan waktu.

7.3.2. Pengujian $H_0 : \mu_1 = \mu_2$ Dimana σ_p^2 Diketahui dan $\sigma_1^2 = \sigma_2^2$

$$\text{Statistik uji } Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{(\bar{x}_1 - \bar{x}_2)}}$$

dimana $\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sigma_p \sqrt{1/n_1 + 1/n_2}$

Contoh ;

Dua orang teknisi melakukan observasi secara sendiri-sendiri mengenai hasil rata-rata perjam dari penggunaan suatu mesin gergaji kayu teknisi (A) : 12 observasi dan memperoleh hasil rata-rata 120 lembar kayu. Sedangkan teknisi (B) : 8 observasi rata-rata 115 lembar kayu. Pengalaman menunjukkan bahwa $\sigma^2 = 40$ lembar kayu. Apakah ke-2 teknisi yakin bahwa beda antara ke-2 hasil rata-rata diatas betul-betul nyata, bukan karena faktor kebetulan.

Jawab :

1. $H_0 : \mu_1 = \mu_2$ dan $\mu_1 \neq \mu^2$

2. $\alpha = 0,05$

3.

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_p \sqrt{1/n_1 + 1/n_2}}$$

4. Daerah kritis dengan $\alpha = 0,05$ secara 2 arah :

$$Z > Z_{\frac{1}{2}\alpha} \text{ dan } Z < -Z_{\frac{1}{2}\alpha}$$

$$Z > 1,96 \text{ dan } Z < -1,96$$

$$5. \quad Z = \frac{(120) - (115)}{6,325 \sqrt{1/12 + 1/8}} = 1,73358$$

6. Karena $1,733 < 1,96$ maka H_0 diterima, beda rata-rata hanya disebabkan faktor kebetulan dan tidak nyata serta : $\mu_1 = \mu_2$

7.4. Pengujian Hipotesis dengan Sampel Kecil

7.4.1 Pengujian $H_0 : \mu = \mu_0$ Dimana σ^2 Tidak Diketahui

Rumus:

$$\text{uji - } t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

Statistik uji - t memiliki distribusi t dengan derajat bebas $(n - 1)$. Daerah kritis pengujian untuk populasi tak terbatas :

$$\frac{\bar{X} - \mu_0}{S / \sqrt{n}} > t(1/2 \alpha; n-1) \quad \text{dan} \quad \frac{\bar{X} - \mu_0}{S / \sqrt{n}} < -t(1/2 \alpha; n-1)$$

Contoh :

Secara hipotesis dikatakan mesin stensil x dapat menstensil 6500 helai kertas per-jam. Suatu perusahaan ingin membuktikan kebenaran hipotesis tersebut dengan mengambil sampel secara acak sebanyak 12 buah mesin stensil 'X' tersebut.

Hasilnya adalah sebagai berikut :

6000	5900	6200	6200	6100	5500
5800	6400	6500	5400	6700	6200

Apakah ada alasan bagi perusahaan untuk mempercayai hipotesis tersebut ?

Jawab : $n = 12$ $\bar{x} = 6075$ $sd = 348,06$

1. $H_0 : \mu = 6500$ $H_1 : \mu \neq 6500$

2. $\alpha = 0,05$

3.
$$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

4. Daerah kritis dengan taraf nyata $\alpha = 0,05$ secara 2 arah :

$$t > t(0,025; 11) \text{ dan } t < -t(0,025; 11)$$

$$t > 2,201 \text{ dan } t < -2,201$$

5. $t = -3,8178$

6. $-3,8178 < -2,201$ sehingga tolak $H_0 : \mu = 6500$

7.4.2. Pengujian $H_0 : \mu_1 = \mu_2$ atau $\mu_1 - \mu_2 = 0$, jika σ^2 tidak diketahui dan $\sigma_1^2 \neq \sigma_2^2$

Stratistik t dirumuskan sebagai berikut :

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$$

$$db = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 + 1} + \frac{(S_2^2/n_2)^2}{n_2 + 2}} - 2$$

atau $t_{1/2 \alpha}$ dihitung secara langsung dengan rumus :

$$t = \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$$

$$\text{dimana : } w_1 = S_1^2 / n_1 \quad t_1 = t(1/2 \alpha; n_1 - 1)$$

$$w_2 = S_2^2 / n_2 \quad t_2 = t(1/2 \alpha; n_2 - 1)$$

sehingga kriteria test untuk uji 2 arah :

$$- \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} < t < \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$$

Contoh :

Semalam barang diproduksi dengan menggunakan 2 proses, ingin diketahui apakah ke-2 proses tersebut mempunyai pengaruh yang sama atau tidak terhadap kualitas barang tersebut, ditinjau dari rata-rata daya tekannya. Untuk itu dilakukan percobaan sebanyak 20 dari hasil proses I dan 20 dari proses II. Rata-rata dan standar deviasinya berturut-turut:

$$x_1 = 9,25 \text{ kg} \quad x_2 = 10,40 \text{ kg} \quad S_1 = 2,24 \text{ kg} \quad S_2 = 3,12 \text{ kg}$$

Jika variansi ke-2 populasi tidak sama, dengan taraf nyata $\alpha = 0,05$. Bagaimana hasilnya ?

Jawab :

1. $H_0 : \mu_1 = \mu_2$ dan $H_1 : \mu_1 \neq \mu_2$

2. $\alpha = 0,05$

3. $t = \frac{(x_1 - x_2) - 0}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} = 1,339$

4. daerah kritisnya ;

$$\frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} < t < \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$$

$$-2,09 < t < 2,09$$

5. Karena t hitung berada pada daerah H_0 : $-2,09 < 1,339 < 2,09$ maka terima

$$H_0: \mu_1 = \mu_2$$

DAFTAR PUSTAKA

- Dajan, A. 1996. Pengantar Metode Statistika. Jilid II. Cetakan ke-18. Penerbit PT.Pustaka LP3ES, Jakarta.
- Dixon, W.J. dan F.J. Massey, Jr. 1997. Pengantar Analisis Statistik. Cetakan ke-2. Diterjemahkan oleh: Sri Kustantini S. dan Zanzawi S. Gadjah Mada University Press, Yogyakarta.
- Sokal R.R. dan F.J.Rohlf. 1991. Pengantar Biostatistika. Edisi ke-2. Diterjemahkan oleh: Nasrullah dan Setyono Setyo Sunarto. Gadjah Mada University Press, Yogyakarta.
- Spiegel, M.R., I. Y. Susila dan E. Gunawan. 1961. Statistik Edisi SI (Metrik). Schaum Publishing Company, Edinburg.
- Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.
- Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.
- Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

BAB VIII

SEBARAN CHI-SQUARE

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menghitung uji kenormalan data, uji homogenitas varians dan uji independensi.

Tujuan Instruksional Khusus

Setelah mempelajari bab ini mahasiswa dapat menghitung:

1. Uji kenormalan
2. Uji homogenitas varians
3. Uji independensi

Uraian dan Contoh

Hasil penelitian sering atau bahkan selalu menunjukkan data sample tidak selalu tepat sama dengan hasil-hasil yang secara teoritis diharapkan sesuai dengan aturan atau probabilitas. Misalnya suatu sample tertentu suatu himpunan kemungkinan peristiwa E_1, E_2, \dots, E_k memiliki frekuensi pengamatan atau frekuensi yang sebenarnya terjadi sebesar o_1, o_2, \dots, o_k . Sebenarnya menurut aturan teori probabilitas menyatakan bahwa frekuensi peristiwa tersebut sebesar e_1, e_2, \dots, e_k yang disebut juga frekuensi yang diharapkan atau frekuensi teoritis. Lihat tabel di bawah ini:

Peristiwa	E1	E2	E3	...	Ek
Frekuensi Observasi	o1	o2	o3	...	ok
Frekuensi Diharapkan	e1	e2	e3	...	ek

Hasil penelitian yang demikian mengakibatkan kita ingin tahu apakah frekuensi yang diobservasi berbeda secara nyata dengan frekuensi yang diharapkan. Pengujian ini dilakukan secara statistika dengan uji chi square (χ^2) (kai kuadrat).

Suatu ukuran mengenai perbedaan terdapat antara frekuensi yang diobservasi dan yang diharapkan diuji dengan chi square dengan ketentuan:

$$\chi^2 = \frac{(o1-e1)^2}{e1} + \frac{(o2-e2)^2}{e2} + \dots + \frac{(ok-ek)^2}{ek} = \sum_{j=1}^k \frac{(o1-e1)^2}{e_j}$$

9.1. Uji Kenormalan

Uji statistik yang digunakan adalah Test Chi-square

$$\chi^2 = \frac{(O_i - E_i)^2}{E_i}$$

dimana :

O_i = data hasil pengamatan ke i

E_i = nilai harapan dari data hasil pengamatan ke i

derajat bebas-nya = $k - p - 1$

k = jumlah kelas interval

p = banyak parameter yang ditaksir

Contoh:

Uji apakah tabel sebaran frekuensi bobot badan sapi Bali umur 6 sampai 8 bulan berikut, menyebar normal?

Tabel 6. Sebaran frekuensi Bobot Badan Sapi Bali Umur 6 sampai 8 Bulan.

Bobot Badan (kg)	Jumlah (ekor)
60 - 62	6
63 - 65	18
66 - 68	42
69 - 71	27
72 - 74	8
Total	100
	$\bar{X} = 67,45 \quad sd = 2,92$

Penyelesaian:

Transformasi data ke $Z = (X_i - \bar{X}) / sd$ untuk menghitung luas di bawah kurva normal.

Langkah-langkah penyelesaian:

1. Hitung Batas Kelas (BK)
2. Hitung nilai Z dari BK
3. Kemudian tentukan nilai peluangnya
4. Cari luas untuk tiap kelas, yang merupakan frekuensi yang diharapkan
5. Hitung Chi-kuadrat
6. Bandingkan dengan tabel Chi-kuadrat, dan simpulkan hasilnya. Bila Chi-kuadrat Hitung lebih kecil dari Chi-kuadrat Tabel, terima Hipotesis Nol (H_0) dan lainnya ditolak (H_1 ditolak). Artinya data menyebar normal.

Bobot Badan	Batas Kelas	Z	Luas di bawah kurva normal Dari 0 -Z	Luas tiap kelas	Frekuensi diharapkan	Frekuensi pengamatan
60-62	59,5	-2,72	-0,4967	0,0413	4,13(4)	5
63-65	62,5	-1,70	-0,4554	0,2068	20,68(21)	18
66-68	65,5	-0,67	-0,2486	0,3892	38,92(39)	42
69-71	68,5	0,36	0,1406	0,2771	27,71(28)	27
72-74	71,5	1,39	0,4177	0,0743	7,43(7)	8
	74,5	2,41	0,4920			

$$\text{Chi-kuadrat Hitung} = (5-4)^2/4 + (18-20)^2/20 + \dots + (8-7)^2/7$$

$$= 0,959$$

$$\text{Derajat Bebas (db)} = k - p - 1$$

$$= 5 - 2 - 1 = 2$$

Untuk db = 2, χ^2 tabel (0,95) = 5,99

Kesimpulan: Chi-kuadrat Hitung lebih kecil dari Chi-kuadrat Tabel
Terima Ho --- data menyebar normal.

Catatan: Uji normalitas suatu sebaran data juga dapat dilakukan dengan cara grafis yang didasarkan atas distribusi frekuensi kumulatif. Sebaran data normal bila digrafikkan dalam bentuk kumulatifnya, akan menghasilkan kurva berbentuk S yang disebut kurva Sigmoid. Sebaran data teramati dibandingkan dengan sebaran harapan dalam bentuk sebaran frekuensi kumulatif (mengepaskan sebaran normal sebagai kurva yang ditumpang-tindihkan dengan sebaran frekuensi teramati) dapat diperkirakan suatu data menyebar normal atau tidak.

9.2. Uji Homogenitas Varians

Data hasil percobaan yang akan diuji lebih lanjut perlu dilakukan uji homogenitas ragam, agar terpenuhi asumsi ragam yang homogen. Uji homogenitas varians dikembangkan oleh Barlett (1937) yang merupakan suatu modifikasi dari uji nisbah kemungkinan Neyman-Pearson.

Perhitungan pertama yang dilakukan adalah rumus varians sample dari populasi ke-i :

$$s_i^2 = \frac{r_i \sum Y_{ij}^2 - (\sum Y_i)^2}{r_i (r_i - 1)}$$

Hipotesis yang akan diuji adalah :

Ho : $\sigma^2_1 = \sigma^2_2 = \dots = \sigma^2_n$ (Varians semua perlakuan sama)

Hi : Minimal ada satu perlakuan yang variansnya tidak sama dengan yang lainnya.

Perhitungan selanjutnya dengan rumus-rumus sebagai berikut:

$$\chi^2 = 2,3026 \{ [\sum (r_i - 1) \log s^2 - \sum (r_i - 1) \log s_i^2] \}$$

$$C = 1 + \left[\frac{1}{3(t-1)} \right] \left\{ \sum \frac{1}{(r_i - 1)} - \left[\frac{1}{\sum (r_i - 1)} \right] \right\} \quad \text{C adalah Faktor Koreksi}$$

$$\chi^2 \text{ terkoreksi} = \chi^2 \times (1/C)$$

$$JK = \sum Y^2 - (\sum Y)^2 / r$$

Contoh :

Data:

Ket.	Perlakuan						Total
	A	B	C	D	E	F	
	19.4	17.7	17.0	20.7	14.3	17.3	
	32.6	24.8	19.4	21.0	14.4	19.4	
	27.0	27.9	9.1	20.5	11.8	19.1	
	32.1	25.2	11.9	18.8	11.6	16.9	
	33.0	24.3	15.8	18.6	14.2	20.8	
Total	144.1	119.9	73.2	99.6	66.3	93.5	596.6
Rerata	28.8	24.0	14.6	19.9	13.3	18.7	

Perhitungan uji Barlett :

Perl.	Db ($r_i - 1$)	$1/(r_i - 1)$	JK	s_i^2	$\text{Log } s_i^2$	$(r_i - 1) \times s_i^2$
A	4	0.25	134.568	33.642	1.527	6.108
B	4	0.25	57.068	14.267	1.154	4.616
C	4	0.25	67.772	16.943	1.229	4.916
D	4	0.25	5.108	1.277	0.106	1.424
E	4	0.25	8.152	2.038	0.309	1.236
F	4	0.25	10.260	2.565	0.409	1.636
Total	24	1.50	282.928	-	-	18.936

$$JK(A) = (19.4)^2 + (32.6)^2 + \dots + (33.0)^2 - (144.1)^2 / 5 = 134.568$$

$$s^2 \text{ (varians gabungan)} = \text{total JK} / \text{total db} = 282.928 / 24 = 11.789$$

$$\log s^2 = 1.071$$

Masukkan dalam rumus-rumus di atas :

$$\chi^2 = 2.3026 \{(24)(1.071) - 18.936\} = 15.584$$

$$db = t - 1 = 6 - 1 = 5$$

$$C \text{ (faktor koreksi)} = 1 + \frac{1}{3(5)} \left[(1.50 - \frac{1}{24}) \right] = 1.09722$$

$$\chi^2 \text{ terkoreksi} = \chi^2 \times (1/C) = 15.584 \times (1/1.09722) = 14.203$$

χ^2 (5%) tabel = 11.1, Maka : χ^2 terkoreksi > χ^2 (5%) tabel , Tolak Ho

Data perlakuan tersebut tidak memenuhi asumsi kehomogenan varians atau data tidak homogen.

9.2. Uji Independensi

Dalam penelitian kadang-kadang data yang diperlukan dicatat dalam bentuk frekuensi dan dikelompokkan ke dalam dua atau lebih kriteria klasifikasi. Tujuannya adalah ingin diketahui apakah kriteria untuk satu klasifikasi yang dipergunakan independent satu terhadap yang lain.

Data penelitian disusun dalam tabel yang dimiliki lajur dan kolom. Persilangan lajur dan kolom membentuk sel. Apabila terdapat 1 lajur dan 3 kolom maka akan terbentuk sel-sel ij.

Pemakaian notasi:

Y_{ij} = banyak observasi dalam sel ij

$Y_{.i}$ = total lajur i.

$Y_{.j}$ = total kolom j.

$j_{..}$ = n = total seluruh observasi.

Pengujian chi square dilakukan untuk menguji hipotesis:

HO : Independen

HA : Tidak Independen

Rumus :

$$q = \sum_{i=1}^r \sum_{j=1}^c [(o_{ij} - e_{ij})^2 / e_{ij}]$$

Nilai kritis tabel chi square :

$$X^2 \alpha, (r-1)(c-1)$$

Contoh:

Pengujian efektifitas suatu vaksin maka dilakukan vaksinasi 120 hewan dan tidak divaksinasi 180 hewan. Kemudian seluruh hewan tersebut (300) diinfeksi dengan penyakit bersangkutan. Diantara hewan yang divaksin 6 ekor mati, hewan tidak divaksin 18 ekor mati. Apakah vaksin tersebut efektif dengan tingkat kesalahan 5% ?

Jawab:

Keterangan	Mati	Hidup	Total
Kontrol	18	162	180
Vaksinasi	6	144	120
Total	24	276	300

Jumlah yang diharapkan

$$e_{11} = (180 \times 24) / 300 = 14,4$$

$$e_{12} = (180 \times 276) / 300 = 164,6$$

$$e_{21} = (120 \times 24) / 300 = 9,6$$

$$e_{22} = (120 \times 276) / 300 = 110,4$$

Jumlah yang diamati

$$o_{11} = 18$$

$$o_{12} = 162$$

$$o_{21} = 6$$

$$o_{22} = 114$$

Pengujian statistika:

$$q = (18-14,4)^2 / 14,4 + (162-165,6)^2 / 165,6 + (6-9,6)^2 / 9,6 + (144-110,4)^2 / 110,4$$

$$= 2,45$$

$$X^2_{5\%, 1} = 3,841$$

Maka hipotesis independent diterima. Kesimpulannya bahwa vaksin tersebut tidak efektif dan berpengaruh terhadap tingkat kematian (mortalitas).

DAFTAR PUSTAKA

- Dajan, A. 1996. Pengantar Metode Statistika. Jilid II. Cetakan ke-18. Penerbit PT.Pustaka LP3ES, Jakarta.
- Dixon, W.J. dan F.J. Massey, Jr. 1997. Pengantar Analisis Statistik. Cetakan ke-2. Diterjemahkan oleh: Sri Kustamtini S. dan Zanzawi S. Gadjah Mada University Press, Yogyakarta.
- Gaspersz, V. 1991. Teknik Analisis Dalam Penelitian Percobaan. Penerbit Tarsito, Bandung.
- Sokal R.R. dan F.J.Rohlf. 1991. Pengantar Biostatistika. Edisi ke-2. Diterjemahkan oleh: Nasrullah dan Setyono Setyo Sunarto. Gadjah Mada University Press, Yogyakarta.
- Spiegel, M.R., I. Y. Susila dan E. Gunawan. 1961. Statistik Edisi SI (Metrik). Schaum Publishing Company, Edinburg.
- Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.

Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh:
B.Sumantri. Penerbit PT.Gramedia, Jakarta.

BAB IX

REGRESI

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menghitung persamaan regresi linear dan uji signifikansinya

Tujuan Instruksional Khusus

Setelah mempelajari bab ini mahasiswa dapat menghitung:

1. Persamaan regresi linier
2. Uji signifikansi dalam regresi linear

Uraian dan Contoh

Pemikiran ilmiah mengenai hubungan pasangan peubah dihipotesiskan sebagai hubungan sebab akibat. Penentuan hubungan sebab akibat berupa signifikansi hubungan fungsional antara dua peubah dengan prosedur ilmiah diperlukan suatu *fungsi*. *Fungsi* adalah hubungan matematika yang memungkinkan penentuan nilai peubah Y pada nilai tertentu peubah X.

10.1. Persamaan Regresi Linear

Persamaan regresi linear merupakan persamaan yang menggambarkan hubungan dua peubah menjadi garis lurus dalam grafiknya. Hubungan kedua peubah tersebut dinyatakan dengan rumus : $Y = a + bX$.

Pada persamaan tersebut Y adalah fungsi X [$Y = f(X)$]. Peubah Y disebut sebagai peubah tidak bebas, karena besarnya nilai Y tergantung dari besarnya nilai X . Peubah X disebut sebagai peubah bebas, karena nilai yang akan menggantikan X bebas untuk berubah sesuai keinginan peneliti. Nilai koefisien a merupakan perpotongan garis fungsi $f(X)$ dengan sumbu Y yang disebut intersep Y . Jika peubah bebas (X) bernilai 0, maka peubah tidak bebas (Y) akan bernilai sebesar a . Nilai koefisien b menunjukkan kemiringan (lereng) garis linear yang disebut koefisien regresi. Peningkatan 1 unit peubah bebas (X) akan meningkatkan sebesar b unit peubah tidak bebas (Y).

Penelitian sesungguhnya menunjukkan bahwa hasil pengamatan yang dilakukan tidak akan terletak tepat pada garis regresi, tetapi menyebar pada kedua sisi garis. Sebaran ini biasanya disebabkan keragaman alami yang dimiliki data (karena factor genetik dan lingkungan) atau karena sesatan ukur. Jadi regresi hubungan fungsional tidak berarti bahwa untuk suatu nilai X tertentu, nilai Y harus sama dengan $a + bX$, melainkan rerata (nilai harapan) Y adalah $a + bX$.

Model yang paling umum dan cocok untuk percobaan adalah regresi model I, berdasarkan 4 asumsi:

1. Peubah bebas X diukur tanpa kesalahan (tertentu). Peubah tidak bebas (Y) adalah peubah acak akibat pengaruh peubah bebas (X) yang tidak berubah secara acak tetapi di bawah kendali peneliti.
2. Nilai harapan peubah untuk sembarang nilai X digambarkan sebagai fungsi linear: $\mu_Y = \alpha + \beta X$. Fungsi ini sama dengan di atas, namun menggunakan huruf Yunani (sebagai a dan b), karena kita menggambarkan parameter hubungan.
3. Untuk sembarang nilai X_i yang merupakan nilai dari X , Y menyebar normal dan saling bebas. Hal ini dapat dituliskan sebagai persamaan: $Y_i = \alpha + \beta X_i + \epsilon_i$ dengan ϵ_i diasumsikan sebagai sesatan yang distribusinya adalah normal dengan rerata nol (0).

4. Asumsi yang terakhir bahwa contoh-contoh di sepanjang garis regresi adalah *homoscedastic*; yaitu bahwa contoh mempunyai varians σ^2 yang sama. Kita menganggap bahwa varians di sekitar garis regresi adalah tetap dan tidak tergantung pada besar X atau Y.

10.2. Uji Signifikansi dalam Regresi Linear

Penafsiran regresi yang dijelaskan di atas adalah suatu metode untuk memperoleh penduga Y untuk tiap nilai X tertentu. Penafsiran lainnya adalah metode untuk menjelaskan sebagian keragaman peubah tidak bebas (Y) dipandang dari keragaman peubah bebas (X). Penentuan ketepatan pendekatan persamaan regresi linear pada 2 peubah (X dan Y) harus diuji dahulu. Uji hipotesis dari persamaan $Y_i = \alpha + \beta X_i + \epsilon_i$ yang dimaksud adalah: $H_0: \beta = 0$; $H_1: \beta \neq 0$. Apabila terima H_0 tolak H_1 , maka persamaan tersebut tidak tepat sebagai persamaan regresi linear. Sebaliknya apabila tolak H_0 terima H_1 , persamaan tersebut tepat sebagai persamaan regresi linear. Uji yang dilakukan dengan uji F.

Contoh:

X	8,98	8,14	6,67	6,08	5,90	5,83	4,68	4,20	3,72
Y	0	12,0	29,5	43,0	53,0	62,5	75,5	85,0	93,0

Hasil perhitungan

$$n = 9; \quad \sum X = 453,5; \quad \sum Y = 54,20; \quad \bar{X} = 50,389$$

$$\sum XY = 2289,260; \quad \sum X^2 = 31152,75; \quad \sum Y^2 = 350,5350; \quad \bar{Y} = 6,022$$

$$\sum x^2 = \sum X^2 - \frac{(\sum X)^2}{n}; \quad \sum y^2 = \sum Y^2 - \frac{(\sum Y)^2}{n}; \quad \sum xy = \sum XY - \frac{(\sum X)(\sum Y)}{n}$$

$$= 8301,3889 \quad = 24,1306 \quad = -441,81$$

Koefisien regresi (b):

$$b = \frac{\sum xy}{\sum x^2} = -441,81 / 8301,3889 = -0,05322$$

Intesep Y (a):

$$a = \bar{Y} - (b \times \bar{X}) = 6,022 - (-0,05322 \times 50,389) = 8,7037$$

Persaman regresi linear:

$$Y = 8,7037 + (-0,05322 \times X) = 8,7037 - 0,05322 \cdot X$$

Uji Hipotesis:

$$JK \text{ regresi} = (\sum xy)^2 / \sum x^2 = (-441,81)^2 / 8301,3889 = 23,5154$$

$$JK \text{ total} = \sum y^2 = 24,1306$$

$$JK \text{ galat} = JK \text{ total} - JK \text{ regresi} = 24,1306 - 23,5154 = 0,6161$$

$$KT \text{ regresi} = JK \text{ regresi} / db \text{ regresi} = 23,5154 / 1 = 23,5154$$

$$KT \text{ galat} = JK \text{ galat} / db \text{ galat} = 0,6161 / 7 = 0,08801$$

$$F \text{ hitung} = KT \text{ regresi} / KT \text{ galat} = 23,5154 / 0,08801 = 267,18$$

F hitung > F table 1% , maka **

ANOVA

SK	db	JK	KT	F hitung.
Regresi	1	23,5154	23,5154	267,18**
Galat	7	0,6161	0,08801	
Total	8	24,1306		

DAFTAR PUSTAKA

Dajan, A. 1996. Pengantar Metode Statistika. Jilid II. Cetakan ke-18. Penerbit PT.Pustaka LP3ES, Jakarta.

- Dixon, W.J. dan F.J. Massey, Jr. 1997. Pengantar Analisis Statistik. Cetakan ke-2. Diterjemahkan oleh: Sri Kustantini S. dan Zanzawi S. Gadjah Mada University Press, Yogyakarta.
- Sokal R.R. dan F.J.Rohlf. 1991. Pengantar Biostatistika. Edisi ke-2. Diterjemahkan oleh: Nasrullah dan Setyono Setyo Sunarto. Gadjah Mada University Press, Yogyakarta.
- Spiegel, M.R., I. Y. Susila dan E. Gunawan. 1961. Statistik Edisi SI (Metrik). Schaum Publishing Company, Edinburg.
- Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.
- Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.
- Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

BAB X

KORELASI

Tujuan Instruksional Umum

Setelah mengikuti kuliah ini mahasiswa dapat menghitung koefisien korelasi

Tujuan Instruksional Khusus

Setelah mengikuti kuliah ini mahasiswa dapat menghitung koefisien korelasi

Uraian dan Contoh

Pengertian mengenai regresi dan korelasi sering rancu di kalangan masyarakat awam. Masalah koelasi sering diperlakukan sebagai masalah regresi, demikian pula sebaliknya. Alasan yang menjadi penyebab kerancuan tersebut adalah:

1. Hubungan matematika antara kedua metode analisis tersebut sangat dekat,
2. Secara matematis seseorang dapat dengan mudah berpindah dari metode satu ke metode yang lain.

Sebagian besar langkah-langkah perhitungan analisis regresi dan korelasi sama. Besaran dasar yang diperlukan untuk analisis regresi adalah jumlah hasil kali. Jumlah hasil kali ini merupakan kuantitas yang sama yang digunakan untuk perhitungan koefisien korelasi.

Analisis regresi menjelaskan ketergantungan suatu peubah Y pada peubah bebas X. Persamaan regresi bertujuan untuk mendukung hipotesis mengenai

Steel R.G.D. dan J.H. Torrie. 1991. Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik. Cetakan ke-2. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.

Sudjana. 1975. Metode Statistika. Cetakan ke-1. Penerbit Tarsito. Bandung.

Walpole, R.E. 1988. Pengantar Statistika. Cetakan ke-3. Diterjemahkan oleh: B.Sumantri. Penerbit PT.Gramedia, Jakarta.