

BAB II
METODE STATISTIK, KONSEP PROBABILITAS
UNTUK KARAKTERISTIK DATA

Statistik adalah cara pengambilan keputusan tentang suatu proses atau populasi berdasarkan pada suatu analisis informasi yang terkandung didalam suatu sampel dari populasi itu. Metode statistik memainkan peranan penting dalam jaminan kualitas. Metode statistik memberikan cara-cara pokok dalam pengambilan sampel produk, pengujian serta evaluasinya dan informasi didalam data itu digunakan untuk mengendalikan dan meningkatkan proses pembuatan.

2.1 Tujuan dan Dasar-Dasar Sampling

Populasi merupakan keseluruhan unsur-unsur yang memiliki satu atau beberapa ciri/karakteristik yang sama. Misalnya kita dapat saja berbicara tentang suatu populasi yang terdiri dari besi dengan ukuran yang tertentu dan diproduksi dengan proses yang tertentu pula dalam suatu periode yang tertentu. Pada hakekatnya perumusan tentang populasi yang terdiri dari benda atau manusia yang tertentu, kita dapat juga berbicara tentang populasi yang terdiri dari nilai-nilai kuantitatif yang diperoleh dari pengukuran atau observasi dari satu atau beberapa unsur-unsur populasi yang terdiri dari benda-benda atau manusia sendiri. Misalnya populasi yang terdiri dari

angka-angka yang menyatakan tinggi badan mahasiswa atau yang menyatakan diameter pipa besi. Kita dapat melakukan suatu pendugaan atau menarik suatu kesimpulan tentang suatu populasi yang memiliki karakteristik yang berbeda dari populasi yang datanya kita pergunakan sebagai dasar pendugaan atau penarikan kesimpulan. Ciri-ciri suatu populasi akan lebih tetap diketahui jika observasi atau pengukuran dilakukan terhadap tiap-tiap unsur populasi tanpa pengecualian. Cara semacam ini bukan merupakan suatu kelaziman karena pertimbangan-pertimbangan energi maupun waktu. Pada umumnya penelitian terhadap populasi dilakukan dengan jalan melakukan observasi atau pengukuran terhadap sebagian dari keseluruhan populasi. Bagian yang diobservasi digunakan dengan tujuan penelitian populasi atau karakteristiknya dinamakan sampel. Sampel itu dapat dikumpulkan atau dipilih dalam pelbagai cara seperti misalnya dipilih secara random atau secara sistematis.

Sampel sistematis dipilih berdasarkan suatu sistem tertentu, katakanlah dengan cara memilih sebagai sampel tiap unit pipa besi yang dihasilkan pada tiap sesudah produksi sebanyak 100 unit. Sampel demikian itu sering digunakan dalam proses pengawasan kualitas. Jika kita melakukan pemilihan sampel dari populasi yang tidak terhingga atau melakukan pemilihan sampel sedemikian rupa sehingga unsur dalam populasi akan memiliki kesempatan yang sama dan secara independen terpilih. Jika jumlah unsur dalam populasi terbatas sedangkan pemilihan

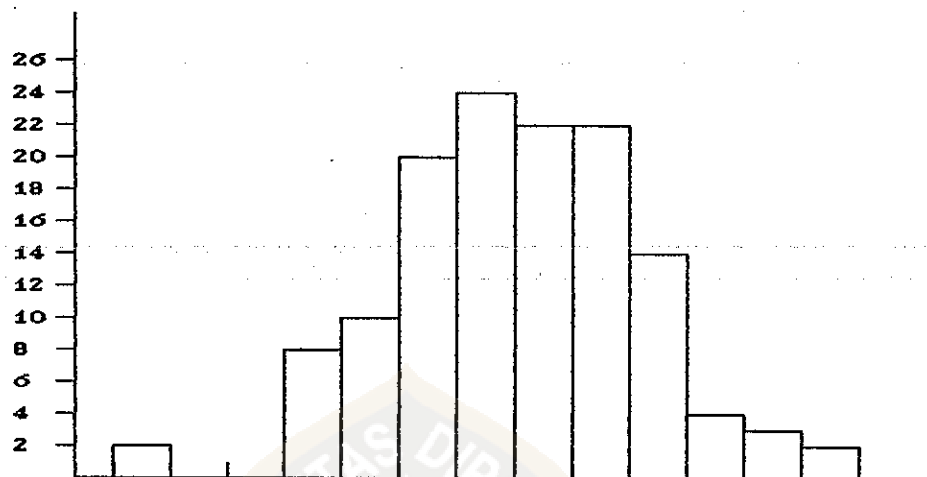
sampelnya dilakukan tanpa pemilihan, maka sampelnya masih dikatakan random jika dipilih sedemikian rupa agar tiap sampel yang besarnya sama dan yang mungkin dipilih memiliki kesempatan yang sama sebagai sampel yang dipilih.

2.2. Distribusi Frekuensi dan Histogram

Tidak ada 2 unit produk pun yang dihasilkan oleh suatu produksi itu identik. Tidak dapat dihindarkan adanya variasi. Misalnya, isi satu kaleng minuman ringan sedikit berbeda dari satu kaleng ke kaleng yang lain dan voltase penyediaan tenaga listrik tidak tepat sama dari satu unit ke unit berikutnya.

Suatu grafik observasi pengamatan terhadap diameter cincin ditunjukkan dalam gambar 2.1 Grafik ini dinamakan Histogram. Tinggi tiap persegi panjang dalam gambar 2.2 sama dengan frekuensi kejadian diameter cincin. Histogram meyakinkan peragaan visual data, yang dengan itu seseorang dapat lebih mudah melihat 3 sifat yaitu :

1. Bentuk
2. Lokasi, atau kecenderungan tengah
3. Pemencaran atau penyebaran



Gambar 2-1 Histogram untuk data diameter cincin piston

2.3 Ringkasan Numerik Data

Histogram menyajikan peragaan visual tiga sifat data sampel : bentuk distribusi data, kecenderungan tengah dalam data, dan pemencaran atau variabilitas dalam data. Adalah juga bermanfaat untuk menggunakan ukuran numerik bagian nilai tengah dan pemencaran.

Pemencaran atau penyebaran dalam data sampel diukur dengan variansi sampel.

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Kita perhatikan bahwa variasi sampel adalah jumlah kuadrat deviasi tiap observasi dari rata-rata sampel \bar{x} , dibagi dengan ukuran sampel kurang satu. Jika tidak ada variabilitas didalam sampel, maka tiap observasi sampel

$x_i = \bar{x}$ dan variasi sampel $s^2 = 0$. Umumnya makin besar

deviasi s^2 makin besar pula pemencaran didalam data sampel. Standar deviasi tidak mencerminkan besarnya data sampel, melainkan hanya pemencaran terhadap rata-rata.

2.4-Distribusi Probabilitas

Histogram digunakan untuk menggambarkan data sampel. Dengan menggunakan metode statistik. Kita mampu menganalisa data sampel diameter cincin Piston dan menarik kesimpulan-kesimpulan tertentu tentang proses yang menghasilkan cincin itu. Distribusi probabilitas adalah model matematika yang menghubungkan nilai variabel random. Karena diameter itu menjalani nilai-nilai yang berbeda dalam populasi itu menurut mekanisme random, maka distribusi probabilitas diameter cincin menggambarkan probabilitas terjadinya setiap nilai diameter didalam populasi itu. Ada 2 macam distribusi probabilitas :

- 1) Distribusi kontinyu. Apabila variabel yang diukur dinyatakan dalam skala kontinyu.

Distribusi probabilitasnya dinamakan distribusi kontinyu.

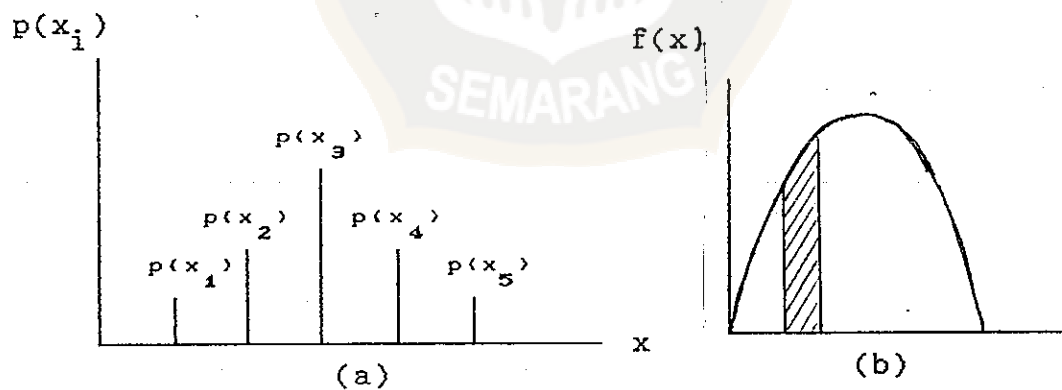
- 2) Distribusi Diskrit. Apabila variabel yang diukur hanya dapat menjalani nilai-nilai tertentu, seperti bilangan bulat 0,1,2,... Distribusi probabilitasnya dinamakan distribusi diskrit. Misalnya banyak yang tak sesuai atau cacat dalam papan untaian tercetak adalah distribusi diskrit.

Contoh distribusi Probabilitas diskrit dan kontinue ditunjukkan pada gambar 2-2a dan 2-2b. Bentuk distribusi diskrit adalah sederetan garis tegak lurus sumbu x yang tingginya proporsional dengan probabilitas. Kita tulis bahwa probabilitas variabel random x mengambil nilai tertentu x_i sebagai

$$P(x = x_i) = P(x_i)$$

Bentuk distribusi kontinue adalah kurva yang halus, dengan luasan dibawah kurva sama dengan probabilitas, sehingga probabilitas x terletak didalam interval dari a sampai b ditulis

$$P(a \leq x \leq b) = \int_a^b f(x) dx$$



gambar 2-2 distribusi probabilitas
(a) kasus diskrit
(b) kasus kontinu

Contoh 2.1

Suatu proses memproduksi ribuan diode tiap hari rata-rata 1 % dari diode ini tidak memenuhi spesifikasi. Setiap jam, seorang pemeriksa memilih sampel random 50 diode dan

mengklasifikasikan tiap diode dalam sampel itu sebagian memenuhi atau tidak memenuhi spesifikasi. Apabila kita misalkan x sebagai variabel random yang menunjukkan banyak diode yang tidak memenuhi spesifikasi dalam sampel itu, maka distribusi probabilitas x adalah :

$$P(x) = \binom{50}{x} (0,01)^x (0,99)^{50-x}$$

dimana $x = 1,2,3,4,\dots,50$

dengan $\binom{50}{x} = (50!/x!(50-x)!)$ adalah banyaknya kombinasi 50 diode yang diambil x diode setiap kali. Ini adalah distribusi diskrit, karena banyaknya diode tidak memenuhi spesifikasi yang diamati adalah $x = 0,1,2,\dots,50$. dan dinamakan distribusi binomial. Kita dapat menghitung probabilitas akan mendapatkan satu atau kurang diode yang tidak memenuhi spesifikasi didalam sampel itu sebagai

$$\begin{aligned} P(x \leq 1) &= P(x = 0) + P(x = 1) \\ &= P(0) + P(1) \\ &= \sum_{x=0}^1 \binom{50}{x} (0,01)^x (0,99)^{50-x} \\ &= \frac{50!}{0! 50!} (0,99)^{50} (0,01)^0 + \frac{50!}{1! 49!} (0,99)^{49} (0,01)^1 \\ &= 0,6050 + 0,3056 = 0,9106 \end{aligned}$$

Contoh 2.2

Distribusi Kontinyu

Misalkan suatu variabel random yang menunjukkan isi yang sebesarnya dalam ons kaleng kopi 1 pon. Distribusi x

dianggap sebagai

$$f(x) = \frac{1}{1,5} ; \quad 15,5 \leq x \leq 17,0$$

Ini adalah distribusi kontinu, karena ruang rentang x adalah interval $(15,5 ; 17,0)$. Distribusi ini dinamakan distribusi uniform, dan ditunjukkan dengan grafik dalam gambar 2-3. Perhatikan bahwa luasan dibawah fungsi $f(x)$ berkaitan dengan probabilitas sehingga probabilitas suatu kaleng akan berisi kurang dari 16,0 ons adalah

$$\begin{aligned} P(x \leq 16,0) &= \int_{15,5}^{16} f(x) dx = \int_{15,5}^{16} \frac{1}{1,5} dx \\ &= \frac{x}{1,5} \Big|_{15,5}^{16,0} = \frac{16,0 - 15,5}{1,5} \\ &= 0,3333 \end{aligned}$$

Secara intuitif terlihat dari pengamatan gab. 2-3

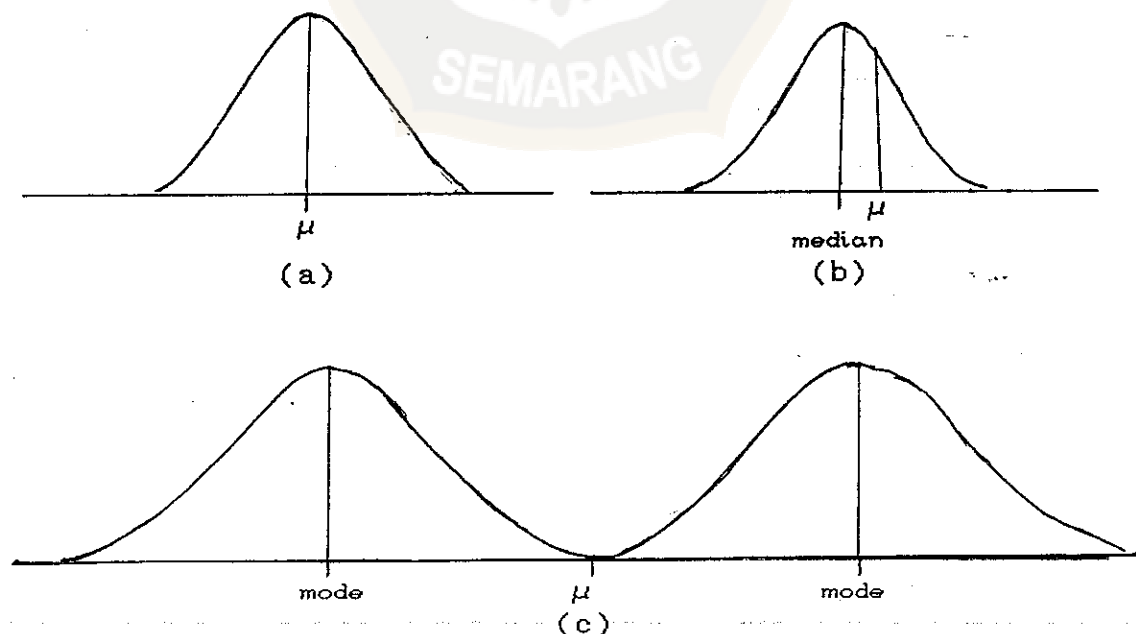


gambar 2-3. Distribusi uniform untuk contoh 2-2

Mean μ suatu distribusi probabilitas adalah suatu ukuran kecenderungan tengah dalam distribusi itu atau lokasinya. Apabila populasi itu terdiri dari N maka mean didefinisikan sebagai

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

Perhatikan kesamaan dengan rata-rata sampel \bar{x} yang didefinisikan dalam persamaan $\bar{x} = \sum X_i/n$. Mean adalah titik tempat distribusi itu benar-benar "seimbang" (lihat gb. 2-4). Jadi mean adalah pusat massa distribusi probabilitas. Perhatikan gb. 2-4b. Bahwa mean tidak harus merupakan persentil ke-50 distribusi itu (yang dinamakan median) Dan gambar 2-4c tidak harus merupakan nilai yang paling mungkin variabel itu (yang dinamakan modus). Mean hanya menentukan lokasi distribusi itu, seperti ditunjukkan dalam gb. 2-5.



gambar 2-4

mean suatu distribusi

Pemencaran penyebaran atau variabilitas dalam suatu distribusi dinyatakan dengan variansi σ^2 . Jika ada N dalam populasi, variansinya adalah :

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

Yakni, variasi adalah rata-rata jarak kuadrat tiap elemen populasi terhadap meannya. Perhatikan kesamaannya dengan variasi sampel s^2 , yang didefinisikan dalam persamaan

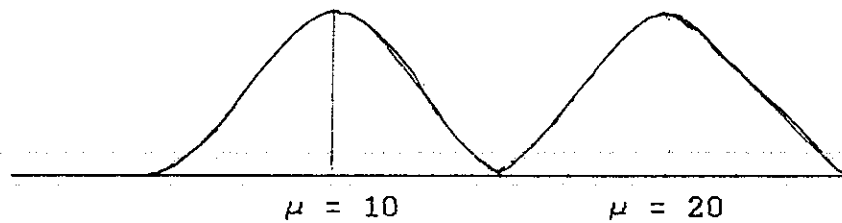
$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

Jika $\sigma^2 = 0$, maka tidak ada

variabilitas didalam populasinya. Apabila variabilitas bertambah maka varansi σ^2 bertambah pula. Variansi dinyatakan dalam kuadrat satuan variabel aslinya. Misalnya, apabila kita mengukur voltase, satuan variansinya adalah (volt)². Maka biasanya orang bekerja dengan akar variansi yang dinamakan deviasi standart σ .

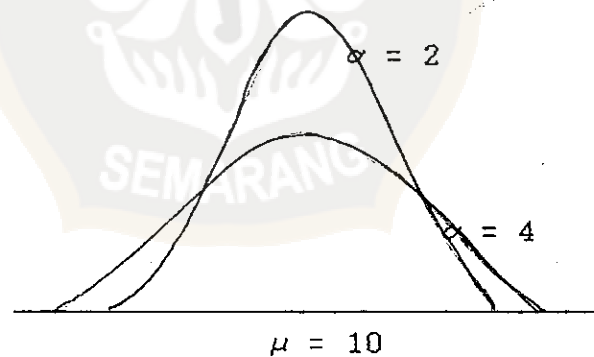
$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

Jelas bahwa diviasi standar mengukur penyebaran atau pemencaran dalam populasinya yang dinyatakan dengan satuan asli.



gambar 2-5 Dua Distribusi probabilitas dengan mean berbeda

Dua distribusi dengan mean sama tetapi deviasi standart berbeda ditunjukkan dalam gambar 2-6



gambar 2-6 dua distribusi probabilitas dengan dengan standar deviasi berbeda.