

# Fungsi Kepadatan Probabilitas



Achmad Basuki  
Politeknik Elektronika Negeri Surabaya  
2004

# Gambaran Permasalahan Fungsi Distribusi Data Dalam Statistik <sup>[1]</sup>

Perusahaan jasa penjualan telur ayam kampung yang dikelola sendiri oleh Pak Hadi, mempunyai 3 orang karyawan. Setiap bulannya pak Hadi membayar upah setiap karyawannya sebesar 1 juta rupiah, dia sendiri setiap bulannya mengambil bayaran sebesar 9 juta rupiah. Kemudian dia mengatakan bahwa rata-rata upah dalam perusahaannya adalah 3 juta rupiah

## Perhitungan statistik

<b>Hadi</b>	<b>Rp. 9 jt</b>
<b>Karyawan 1</b>	<b>Rp. 1 jt</b>
<b>Karyawan 2</b>	<b>Rp. 1 jt</b>
<b>Karyawan 3</b>	<b>Rp. 1 jt</b>
<b>Rata2</b>	<b>= Rp. 12 jt /4 orang</b>
	<b>= Rp. 3 jt/orang</b>

**Apakah ini masuk akal ?**

Kalo iya, karyawan yang mana yang mendapat upah 3 jt rupiah?

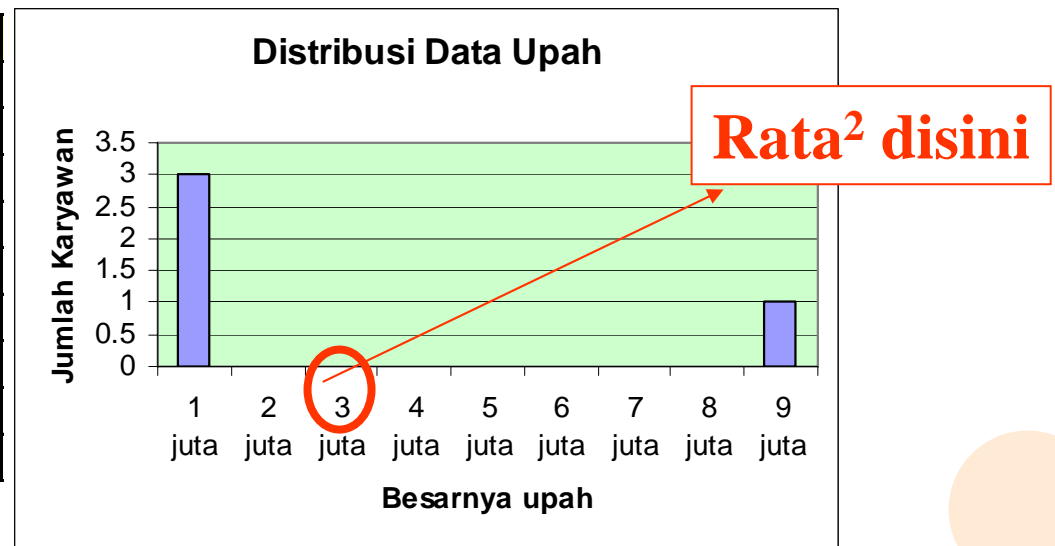
**Kenyataannya tidak ada seorang karyawanpun yang mendapat upah 3 juta**

# Gambaran Permasalahan Fungsi Distribusi Data Dalam Statistik <sup>[1]</sup>

Coba kita perhatikan masalah ini lebih jauh dengan mengamati distribusi data dari upah karyawan, dimana:

1 orang mendapat bayaran 9 juta rupiah, dan 3 orang mendapat upah 1 juta rupiah, sehingga dapat kita gambarkan distribusi data besarnya upah sebagai berikut:

Besarnya upah	Jumlah karyawan
1 juta	3
2 juta	0
3 juta	0
4 juta	0
5 juta	0
6 juta	0
7 juta	0
8 juta	0
9 juta	1



Distribusi data seperti ini adalah suatu kejadian dimana nilai rata-rata (statistik parametrik) tidak dapat menunjukkan hasil yang dapat menggambarkan kenyataan sesungguhnya.

# Gambaran Permasalahan Fungsi Distribusi Data Dalam Statistik [2]

Data hasil ujian pemrograman dari 20 mahasiswa Jurusan TI adalah seperti tabel di sebelah kanan. Dapat dinyatakan bahwa nilai rata-rata programming adalah 69.

## **Kesimpulan:**

Pemrograman mahasiswa TI rata-rata lemah, karena tidak mencapai nilai 75 sebagai standard yang sudah ditentukan sebelumnya.

## **Kenyataannya adalah:**

Hanya beberapa orang yang lemah, sedang sebagian besar (13 mhs) yang nilainya di atas 75.

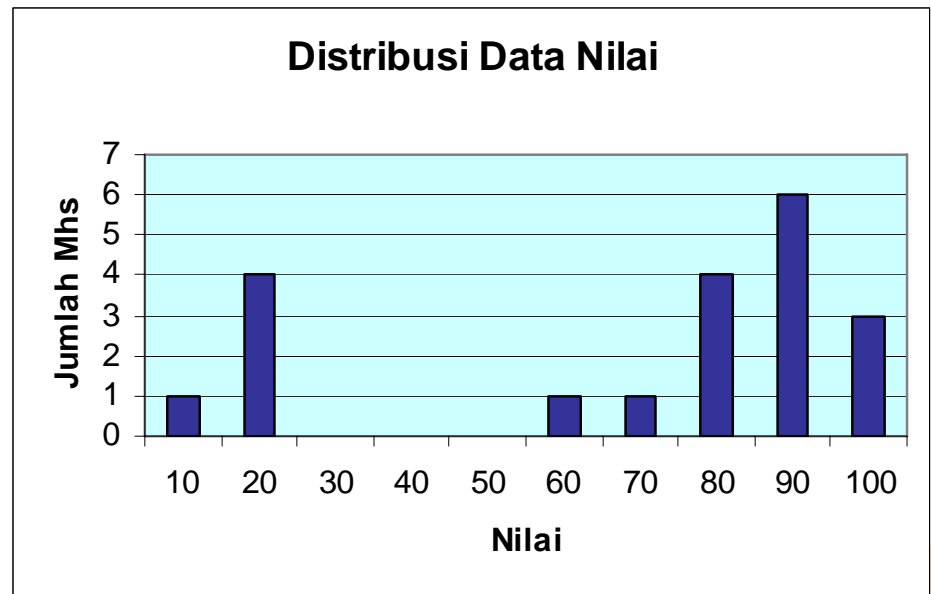
no.mhs	Nilai
1	90
2	70
3	20
4	80
5	10
6	100
7	60
8	80
9	90
10	80
11	20
12	80
13	100
14	20
15	90
16	20
17	100
18	90
19	90
20	90

# Gambaran Permasalahan Fungsi Distribusi Data Dalam Statistik [2]

no.mhs	Nilai
1	90
2	70
3	20
4	80
5	10
6	100
7	60
8	80
9	90
10	80
11	20
12	80
13	100
14	20
15	90
16	20
17	100
18	90
19	90
20	90

Coba kita perhatikan distribusi nilai pemrograman dari mahasiswa TI ini.

Nilai	Jumlah Mhs
10	1
20	4
30	0
40	0
50	0
60	1
70	1
80	4
90	6
100	3



**Dengan distribusi data ini, nilai rata-rata tidak dapat menunjukkan keadaan sebenarnya.**

**Bagaimana cara menunjukkan distribusi data agar kesimpulan yang diambil dapat menunjukkan keadaan sesungguhnya ?**



Fungsi Kepadatan  
Probabilitas



# Definisi

- Fungsi kepadatan probabilitas atau *probability density function* (pdf) menyatakan nilai probabilitas dari setiap kejadian  $X$  dan dituliskan dengan  $p(X)$
- Karena  $p(X)$  menyatakan nilai probabilitas maka  $0 \leq p(X) \leq 1$
- Untuk semua kejadian maka jumlah nilai probabilitasnya adalah satu atau dituliskan dengan:

$$\sum_n p(X = x_n) = 1$$

# Ciri-ciri Fungsi Kepadatan Probabilitas

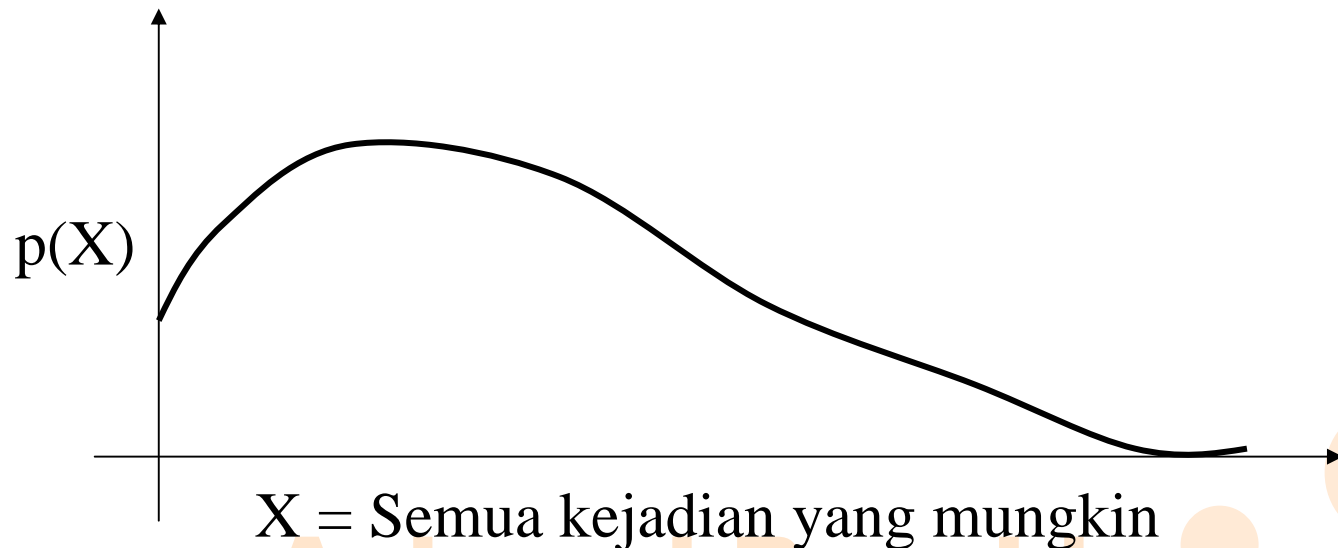
- $X = \{x_1, x_2, x_3, \dots, x_n\}$  menyatakan semua kejadian yang mungkin
- $0 \leq p(X) \leq 1$
- Nilai probabilitas untuk semua kejadian:

$$\sum_n p(X = x_n) = 1$$



# Grafik dari Fungsi Kepadatan Probabilitas

- Grafik yang menyatakan nilai kemungkinan dari setiap kejadian.
- Absis menyatakan kejadian yang mungkin.
- Ordinat menyatakan nilai kemungkinan  $p(x_i)$



# Kontinu vs Diskrit

Pada dasarnya fungsi-fungsi di dalam statistik berdasarkan sifat kejadiannya dibedakan menjadi dua macam yaitu kontinu dan diskrit.

- Kontinu: kejadian yang mungkin jumlahnya tak berhingga dan operasionalnya dilakukan dalam bentuk kalkulus, misalkan untuk menghitung jumlah peluang semua kejadian dituliskan dengan:

$$\int_{\forall x} f(x) dx = 1$$

- Diskrit: kejadian yang mungkin jumlahnya berhingga dan dapat berarti dilakukan secara berkala, operasionalnya menggunakan operasional fungsi diskrit, misalkan untuk menghitung jumlah peluang semua kejadian dituliskan dengan:

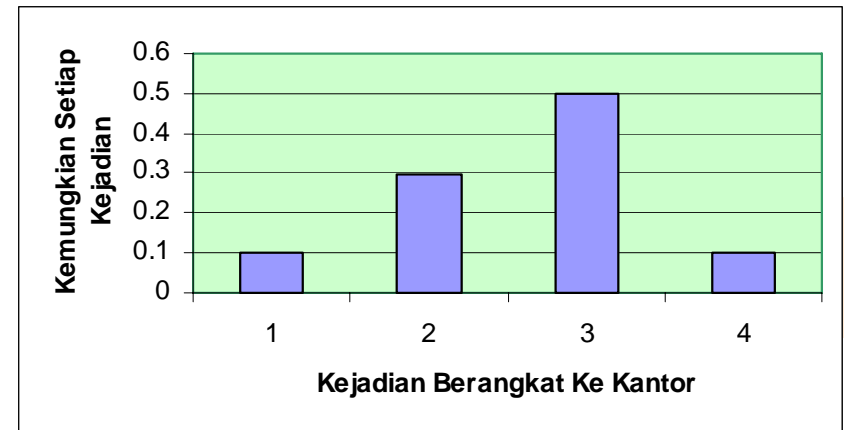
$$\sum_n p(X = x_n) = 1$$



*Pembahasan banyak dilakukan pada model diskrit*

# Contoh 1

- X adalah suatu kejadian seseorang akan berangkat ke kantor: kemungkinan dia berangkat naik mobil adalah 0.1, kemungkinan naik kendaraan umum 0.3, kemungkinan naik sepeda motor 0.5 dan kemungkinan tidak berangkat 0.1
- Fungsi kepadatan probabilitas dinyatakan dengan:  $p(x_1)=0.1$ ,  $p(x_2)=0.3$ ,  $p(x_3)=0.5$  dan  $p(x_4)=0.1$  dimana  $X=\{x_1, x_2, x_3, x_4\}$  menyatakan kejadian-kejadian yang mungkin.
- Nilai probabilitas dari semua kemungkinan adalah  $0.1+0.3+0.5+0.1 = 1$



## Contoh 2

- Dari hasil pencatatan jumlah mobil bemo yang lewat setiap setengah jam di depan ITS diperoleh: tidak ada yang lewat: 4 kali, 1 bemo lewat: 5 kali, 2 bemo lewat: 8 kali, 3 bemo lewat: 9 kali, 4 bemo lewat: 6 kali, 5 bemo lewat: 3 kali, 6 bemo lewat: 1 kali, 7 bemo lewat: 1 kali.
- Absis (X) menyatakan jumlah bemo lewat dalam setengah jam
- Ordinat (Y) menyatakan kemunculan atau frekwensi kejadian dibagi dengan jumlah seluruh kejadian (37)



# HISTOGRAM

- Histogram adalah suatu teknik untuk menyatakan jumlah munculnya setiap kejadian dari semua kejadian yang muncul.
- $H(x_n)$  menyatakan jumlah munculnya kejadian  $x_n$ .
- Fungsi kepadatan probabilitas  $p(x_n)$  dapat dinyatakan

$$p(x_i) = \frac{H(x_i)}{\sum_{j=1}^n H(x_j)}$$

# Contoh 3

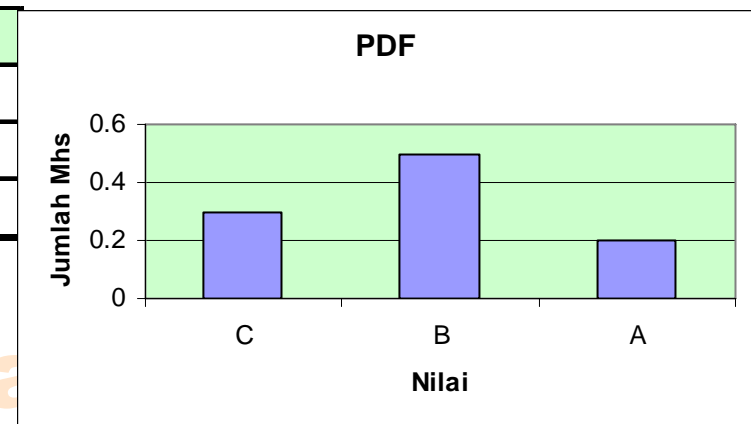
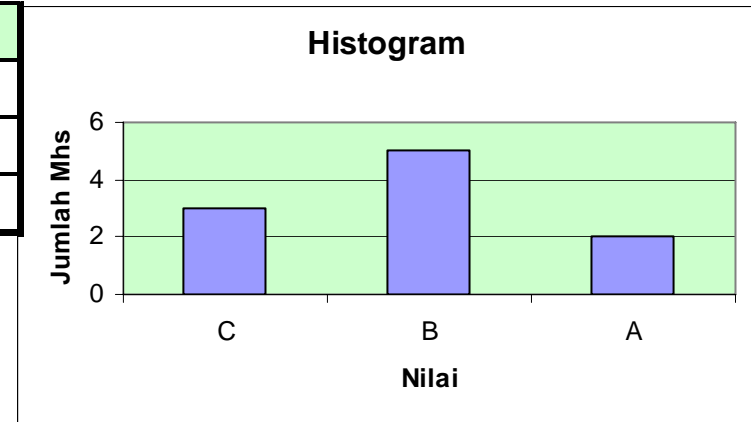
Data nilai test dasar pemrograman yang diikuti oleh 10 orang mahasiswa adalah sebagai berikut:

no. mhs	nilai
1	A
2	B
3	B
4	B
5	A
6	B
7	C
8	C
9	B
10	C

Nilai	Jumlah
C	3
B	5
A	2

Nilai	pdf
C	0.3
B	0.5
A	0.2

Berdasarkan nilai yang diperoleh dapat dinyatakan bahwa yang mendapat nilai C sebanyak 3 orang, yang mendapat nilai B sebanyak 5 orang dan yang mendapat nilai A sebanyak 2 orang. Nilai-nilai kemunculan ini disebut dengan histogram dari nilai ujian.



# Diagram Pareto

- Suatu diagram yang digunakan untuk mencatat kemunculan setiap kejadian.
- Model diagram Pareto ini banyak digunakan untuk pencatatan kerusakan produksi.
- Diagram Pareto ini dapat juga disebut dengan diagram counting.
- Diagram Pareto ini banyak digunakan untuk menghasilkan nilai histogram dari suatu kejadian secara manual.



# Contoh 4

## Diagram Pareto

Pencatatan Kerusakan Produksi Tempe Setiap 1000 Bungkus Tempe Standard

Jenis Kerusakan	Counting	Jumlah
Kurang Masak		24
Kedelai Terlalu Busuk		16
Kedelai Hancur		10
Tempe Pecah		8

Histogram dan pdf dari kejadian di atas adalah:

Jenis Kerusakan	Histogram	PDF
Kurang Masak	24	0.4138
Kedelai Terlalu Busuk	16	0.2759
Kedelai Hancur	10	0.1724
Tempe Pecah	8	0.1379



# Distribusi Frekwensi

- Distribusi frekwensi adalah suatu model perhitungan histogram dengan menggunakan pengelompokan data.
- Satu kelompok dapat dinyatakan sebagai satu range nilai dengan nilai tengah dianggap sebagai nilai yang mewakili kelompok tersebut.
- Kemunculan suatu kelompok dinamakan dengan frekwensi.

## Contoh 5

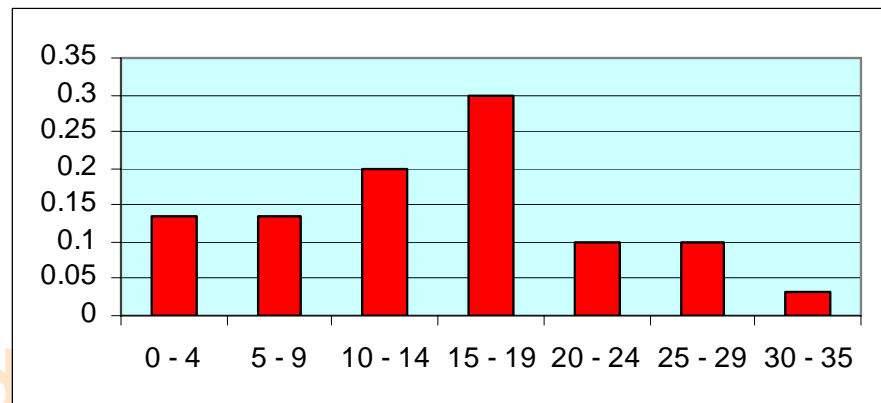
Data penjualan telur kampung setiap harinya pada toko MAJU MAKMUR dicatat selama 30 hari adalah sebagai berikut:

30	25	18	15	21	12	0	15	6	12
0	10	15	24	6	18	27	12	0	15
12	15	20	3	9	25	12	15	6	15

Distribusi frekwensi dengan range 5 adalah sebagai berikut:

Range	Median	Frekwensi
0 - 4	2	4
5 - 9	7	4
10 - 14	12	6
15 - 19	17	9
20 - 24	22	3
25 - 29	27	3
30 - 35	32	1

PDF dapat dihitung dengan frekwensi dibagi dengan jumlah seluruh kejadian (30)



# Fungsi Kepadatan Kumulatif

- Fungsi Kepadatan Kumulatif atau Cumulative Density Function (CDF) adalah fungsi yang menjumlahkan nilai kemungkinan sampai suatu kejadian tertentu. Atau dituliskan dengan  $p(X \leq x_i)$
- Bila  $X = \{x_1, x_2, x_3, \dots, x_n\}$ , maka fungsi kepadatan kumulatif untuk  $X = x_k$  dituliskan dengan:

$$p(X \leq x_k) = p(x_1) + p(x_2) + \dots + p(x_k)$$

*atau*

$$p(X \leq x_k) = \sum_{i=1}^k p(x_i)$$

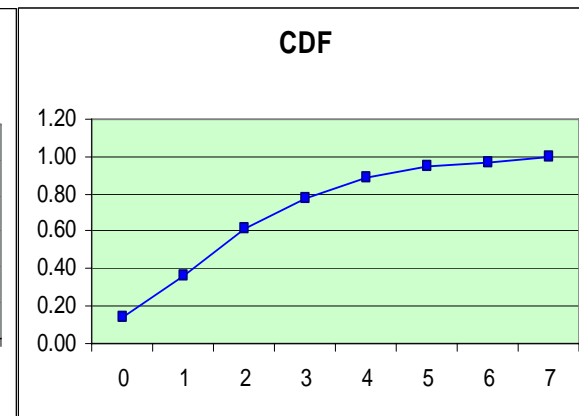
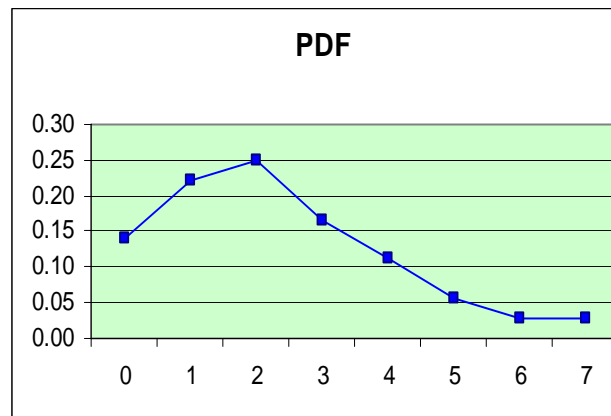
# Contoh 6

Diketahui frekwensi jumlah pelanggan yang melalui pintu kasir untuk setiap 5 menit sebuah supermarket adalah sebagai berikut:

Perhitungan PDF dan CDF adalah sebagai berikut:

Jumlah Plg	Frekwensi	PDF	CDF
0	5	$5/36 = 0.14$	0.14
1	8	$8/36 = 0.22$	$0.14+0.22 = 0.36$
2	9	$9/36 = 0.25$	$0.36+0.25 = 0.61$
3	6	$6/36 = 0.17$	$0.61+0.17 = 0.78$
4	4	$4/36 = 0.11$	$0.78+0.11 = 0.89$
5	2	$2/36 = 0.06$	$0.89+0.06 = 0.94$
6	1	$1/36 = 0.03$	$0.94+0.03 = 0.97$
7	1	$1/36 = 0.03$	$0.97+0.03 = 1.00$

Jumlah Plg	Frekwensi
0	5
1	8
2	9
3	6
4	4
5	2
6	1
7	1



# Contoh Aplikasi [1]

Pengamatan terhadap nilai matematika mahasiswa Jurusan TI

Nilai matematika 2 dari 30 mahasiswa Jurusan TI (kelas 2 TI-a) adalah sebagai berikut:

no.mhs	nilai		no.mhs	nilai		no.mhs	nilai
1	B		11	C		21	C
2	C		12	C		22	B
3	C		13	A		23	A
4	B		14	B		24	D
5	A		15	C		25	C
6	C		16	B		26	B
7	B		17	B		27	B
8	C		18	C		28	B
9	D		19	B		29	A
10	B		20	B		30	C

Nyatakan Histogram, PDF dan CDF dari data nilai mahasiswa di atas

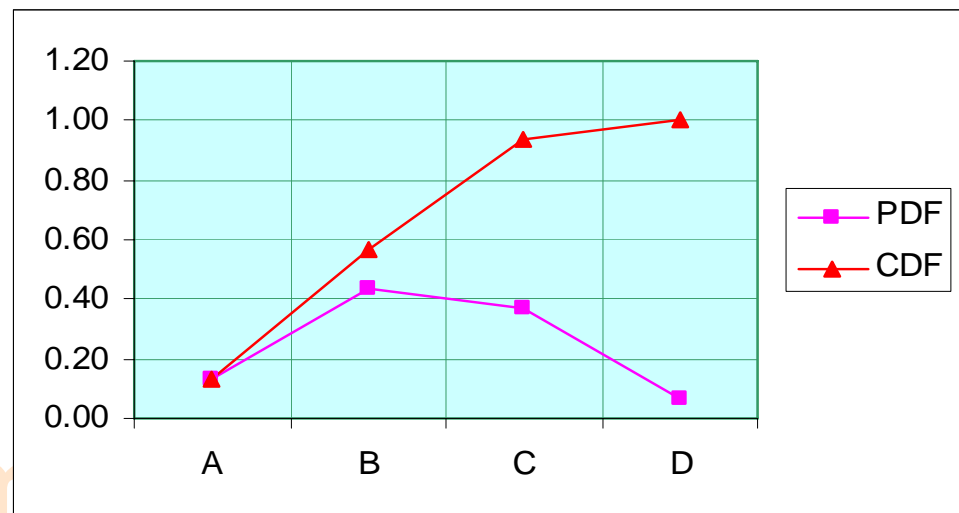
# Contoh Aplikasi [2]

Diagram distribusi frekwensi dari data nilai matematika tersebut adalah

Nilai	Jumlah mhs yang mendapat nilai
A	4
B	13
C	11
D	2

Histogram, PDF dan CDF diperoleh sebagai berikut:

Nilai	Histogram	PDF	CDF
A	4	0.13	0.13
B	13	0.43	0.57
C	11	0.37	0.93
D	2	0.07	1.00



# Tugas 1

**Anda lakukan survey terhadap 20 orang teman anda yang dipilih secara acak. Tanyakan jenis acara TV yang sering ditonton oleh mereka dari acara-acara TV berikut ini:**

- (1) Olahraga**
- (2) Info Selebriti**
- (3) Berita**
- (4) Horor dan Misteri**
- (5) Film**
- (6) Film Kartun**
- (7) Komedi**
- (8) Sinetron**

**Buatlah Histogram, PDF dan CDF dari hasil survey tersebut, dan jangan lupa sebutkan segmen mahasiswa yang anda pilih berdasarkan jenis kelamin (berapa laki2 dan berapa wanita).**