PERCOBAAN 5

DISTRIBUSI PROBABILITAS KHUSUS

5.1. Tujuan :

Setelah melaksanakan praktikum ini mahasiswa diharapkan mampu :

- Membedakan beberapa jenis distribusi probabilitas variabel acak
- Menggunakan fungsi toolbox Matlab untuk penyelesaian distribusi probabilitas
- Menyelesaikan sebuah permasalahan menggunakan jenis distribusi probabilitas yang tepat

5.2. Peralatan :

- Laptop / PC Desktop yang support dengan program Matlab
- Bahasa Pemrograman Matlab versi 2009 ke atas

5.3. Teori :

Pada praktikum sebelumnya telah dijelaskan bagaimana mengkategorikan nilai-nilai dalm sebuah variabel data berdasarkan jenis data yang mendukungnya, yaitu variabel acak diskrit dan kontinyu. Selain itu juga dijelaskan tentang fungsi massa probabilitas dan fungsi densitas probabilitas yang merupakan sebaran dari nilai-nilai dalam sebuah variabel acak pada interval pengamatan tertentu. Fungsi distribusi kumulatif merupakan jumlahan dari seluruh nilai probabilitas variabel acak tersebut yang berada pada interval pengamatan tertentu.

Berdasarkan polanya, sebaran data pada variabel acak dapat dikategorikan dalam beberapa jenis distribusi. Pada variabel acak diskrit dikenal model distribusi binomial dan Poisson, sedangkan pada variabel acak kontinyu dikenal distribusi uniform, normal, eksponensial, gamma dan weibull. Dalam praktikum ini akan dijelaskan berbagai model sebaran data tersebut dan diamati karakteristik distribusinya.

5.3.1. Distribusi Binomial

Distribusi Binomial digunakan apaila sebuah proses sampling dilaksanakan sesusai dengan proses Bernoulli, yang memiliki ciri-ciri sebagai berikut:

1. Ada dua kejadian yang dapat terjadi dan saling asing pada setiap percobaan. Untuk mudahnya, dua kejadian itu disebut sukses dan gagal.
2. Urutan dari percobaan tersebut merupakan kejadian independen.

3. Probabilitas sukses dinyatakan sebagai *p*, dimana nilai *p* ini tetap dari satu percobaan ke percobaan berikutnya atau dari satu kejadian ke kejadian lainnya.

Pada distribusi Binomial, ada 3 nilai yang diperlukan yaitu : jumlah sukses (X), jumlah percobaan / observasi (n) dan probabilitas sukses dalam setiap percobaan (p). Distribusi probabilitas Binomial atau fungsi massa probabilitas (pmf) dari distribusi Binomial dinyatakan sebagai:

$$\begin{aligned} f(x \mid n, p) = P(X = x) &= \binom{n}{x} p^x (1-p)^{n-x} \\ &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \end{aligned} \qquad (1)$$

Jika sebuah variabel acak, X terdiri dari n percobaan, dimana $X = L_1 + L_2 + \ldots + L_n$ maka mean atau nilai harapan dari distribusi Binomial ini merupakan perkalian dari jumlah sampel populasinya dikalikan dengan pembobotnya yaitu nilai probabilitasnya, dinyatakan sebagai:

$$\begin{aligned} \mu = E[X] &= E(L_1) + E(L_2) + \ldots + E(L_n) \\ &= p + p + \ldots + p = n.p \end{aligned} \qquad (2)$$

Dan varians-nya dinyatakan sebagai: $\sigma^2 = \sigma_{L_1}^2 + \sigma_{L_2}^2 + \ldots + \sigma_{L_n}^2 = p.q + p.q + \ldots + p.q = n.p.q$

Fungsi distribusi kumulatif (cdf) dari distribusi Binomial merupakan jumlahan dari fungsi-fungsi probabilitasnya yang telah ditunjukkan pada persamaan (1), dan dinyatakan sebagai:

$$\begin{aligned} F(x \mid n, p) &= \sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i} \\ &= \sum_{i=0}^{x} \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \end{aligned} \qquad (3)$$

5.3.2. Distribusi Poisson

Distribusi Poisson dapat digunakan untuk menentukan probabilitas dari sejumlah sukses yang ditentukan jika kejadian-kejadian berjalan dalam kurun waktu atau ruang kontinyu tertentu. Proses Poission hampir sama dengan proses Bernoulli hanya berbeda pada sifat kontinuitasnya. Pada distribusi Poisson hanya ada 1 nilai yang diperlukan, yaitu jumlah rata-rata sukses, dinyatakan sebagai λ^x.

Distribusi Poisson efektif digunakan untuk jumlah pengamatan, *n* yang sangat besar, sementara probabilitas, *p* untuk satu kejadian sangat kecil (biasanya kurang dari 0,5). Contoh penggunaan distribusi ini antara lain: pendudukan traffic telepon dalam satu jam di sentral telepon, banyaknya kesalahan ketik dalam 1 halaman laporan, jumlah cacat dalam 1 lembar kain.

Distribusi massa probabilitas (pmf) dari Poisson dinyatakan sebagai berikut:

$$f\left(x \mid \lambda\right) = P\left(X \mid \lambda\right) = \left(\frac{\lambda^x}{x!}\right).e^{-\lambda} = \frac{e^{-\lambda}.\lambda^x}{x!} \qquad (4)$$

Nilai harapan dan varians dari distribusi Poisson adalah: $E\left[X\right] = \lambda$ dan $V\left[X\right] = \lambda$

Fungsi distribusi kumulatif (cdf) dari Poisson dinyatakan sebagai:

$$F\left(x \mid \lambda\right) = e^{-\lambda} \sum_{i=0}^{floor(x)} \frac{\lambda^i}{i!} \qquad (5)$$

5.3.3. Distribusi Uniform

Distribusi Uniform termasuk dalam distribusi variabel acak kontinyu, dimana pada jenis ini setiap nilai pada variabel acak tersebut memiliki probabilitas yang sama, dimana: $f\left(x \mid k\right) = \frac{1}{k}$; untuk $x_1, x_2, ..., x_k$.

Jika nilai dari variabel acak berada pada rentang interval (a,b) maka fungsi densitas probabilitas (pdf) dari distribusi uniform dinyatakan sebagai:

$$f\left(x \mid a,b\right) = \frac{1}{b-a}; \quad a < x < b \qquad (6)$$

Mean dan varians dari distribusi uniform ini dinyatakan sebagai:

$$\mu = E\left[X\right] = \frac{a+b}{2} \quad \text{dan} \quad \sigma^2 = V\left[X\right] = \frac{\left(b-a\right)^2}{12}$$

Sedangkan fungsi distributif kumulatif (cdf) dari distribusi uniform ini dinyatakan:

$$F\left(x\right) = \begin{cases} 0; & x \le a \\ \frac{x-a}{b-a}; & a < x < b \\ 1; & x \ge b \end{cases} \qquad (7)$$

5.3.4. Distribusi Eksponensial

Jika kejadian sukses berjalan secara kontinyu dan distribusi probabilitasnyapun bersifat kontinyu dalam kurun waktu tertentu, maka distribusi probabilitas tersebut dinamakan distribusi eksponensial. Distribusi eksponensial ini digunakan untuk memodelkan jumlahan waktu hingga kemunculan sebuah kejadian tertentu. Beberapa aplikasi distribusi eksponensial di antaranya adalah: pemodelan waktu hingga komputer log off, pemodelan waktu antar waktu kedatangan panggilan telepon, pemodelan waktu pintu gerbang otomatis terbuka jika ada obyek di depannya dll.

Fungsi kerapatan probabilitas (pdf) dari distribusi eksponensial dinyatakan sebagai:

$$f\left(x \mid \lambda\right) = \begin{cases} \lambda e^{-\lambda x}; & x \geq 0; \quad \lambda > 0 \\ 0; & x \text{ lainnya} \end{cases} \tag{8}$$

Mean dan varians dari distribusi eksponensial dinyatakan sebagai: $E\left[X\right] = \frac{1}{\lambda}$ dan $V\left[X\right] = \frac{1}{\lambda^2}$

Fungsi distribusi kumulatif (cdf) dari distribusi eksponensial dinyatakan sebagai:

$$F\left(x \mid \lambda\right) = \begin{cases} 1 - e^{-\lambda x}; & x \geq 0; \quad \lambda > 0 \\ 0; & x \text{ lainnya} \end{cases} \tag{9}$$

5.3.5. Distribusi Normal

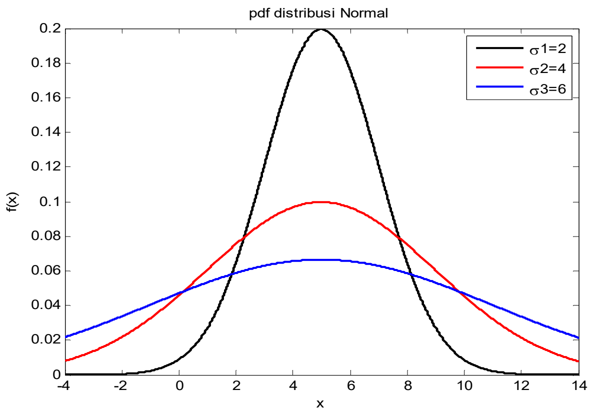
Distribusi probabilitas normal adalah distribusi probabilitas dari variabel acak kontinyu yang simetris dan mesokurtik. Kurva yang dihasilkan benrbentuk lonceng yang simetris kiri dan kanan. Distribusi normal ini disebut juga sebagai distribusi Gauss, dimana model noise Gaussian tersebar dalam distribusi normal ini. Dua parameter yang menentukan kurva distribusi normal adalah rata-rata (mean) dan simpangan baku (standard deviasi). Pengujian terhadap normalitas dapat dilakukan dengan menghitung prosentase data observasi berada dalam plus minus satu standard deviasi atau plus minus dua standard deviasi dari rata-rata. Dengan cara ini, suatu distribusi data observasi dikatakan normal jika kurang 68% data-data tersebut berada dalam satu standard deviasi atau kurang lebih 95% berada dalam dua standard deviasi.

Fungsi kerapatan probabilitas (pdf) dari distribusi normal yang memiliki mean μ dan varians σ^2 dinyatakan sebagai:

$$f\left(x\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{10}$$

Dimana $e = 2,7182$ dan $\pi = 3,14159$

Nilai simpangan baku pada distribusi normal menghasilkan kurva yang mengerucut ke arah nilai harapan atau menyebar menjauhinya. Beberapa bentuk kurva distribusi normal dengan mean yang konstan dan simpangan baku yang variatif ditunjukkan pada gambar 5.1.



Gambar 5.1. Kurva distribusi normal dengan simpangan baku variatif

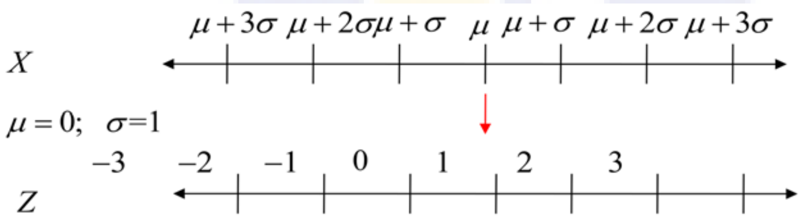
Untuk mendapatkan probabilitas dari data-data yang berada di dalam suatu interval tertentu dapat diselesaikan dengan meng-integralkan persamaan (10) dengan batas interval tersebut menjadi:

$$P\left(a < x < b\right) = \int_a^b f\left(x\right) dx = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \qquad (11)$$

Namun penyelesaian persamaan (11) di atas sangat rumit. Untuk mempermudahnya dilakukan pen-translasi-an persamaan dengan parameter x, μ dan σ menjadi persamaan dengan parameter Z, dimana hubungan antara ke empat parameter tersebut adalah:

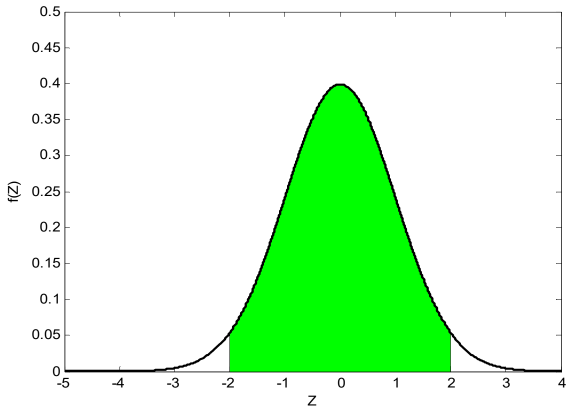
$$Z = \frac{x-\mu}{\sigma} \qquad (12)$$

Grafik yang menyatakan hubungan antara parameter Z dengan x, μ dan σ ditunjukkan pada gambar 5.2.



Gambar 5.2. Grafik hubungan Z dengan x, μ dan σ

Setelah melakukan konversi masing-masing data observasi ke dalam bentuk Z, gunakan tabel distribusi normal untuk mendapatkan *P(Z)* nya. Gambar 5.3 menyatakan gambar kurva terdistribusi normal dengan sumbu Z dan P(Z) dimana luasan berwarna hijau menyatakan daerah probabilitas untuk $P\left(-2 < Z < 2\right)$.



Gambar 5.3. Grafik untuk $P(-2 < Z < 2)$

Fungsi distribusi kumulatif dari distribusi normal dinyatakan sebagai:

$$F(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \tag{13}$$

5.4. Prosedur

5.4.1. Distribusi Binomial

Sebuah perusahaan sepatu mengelompokkan hasil produksinya menjadi dua bagian, yaitu kualitas ekspor, biasanya 40%, dan sisanya merupakan kualitas non ekspor, 60%. Jika diambil secara acak 10 pasang sepatu, hitung

a. sekurang-kurangnya enam puluh persennya adalah kualitas ekspor, $P(X > 0,6 \mid)$
b. berapa mean dan varians nya.
c. Gambarkan grafik pmf dan pdf.

Gunakan:

A. Pemrograman biasa dengan faktorial
B. Pemrograman menggunakan fungsi **binopdf()** dan **binocdf().**

```
% Program: DBino.m
% Distribusi Binomial

clear all;clc;
% Dapatkan nilai untuk variabel acak X
x=1:1:10;
% Nilai probabilitas
p=0.4;q=0.6;
% Jumlah pengamatan
n=10;

% Fungsi binomial dengan kombinasional
% nCx p^(x+i)q^(n-i)

% fungsi massa probabilitas (pmf) Binomial
for i=1:n
    b=i;
    depan(i)=(factorial(n)/(factorial(b)*factorial(n-b)));
    blk(i)=(p^b)*(1-p)^(n-b);
    fx(i)=depan(i)*blk(i);
end

% probabilitas sekurang-kurangnya 60%
sum60=sum(fx(6:10))
% mean dan varians
mi=n*p
vi=n*p*(1-p)

% fungsi distribusi kumulatif
sumx=sum(fx);
sumpy(1)=fx(1);
for i=2:n
    sumpy(i)=sumpy(i-1)+fx(i);
end

% fungsi massa probabilitas dengan binopdf
pdf1=binopdf(x,n,0.4);

% fungsi distribusi kumulatif dengan binocdf
pdf2=binocdf(x,n,0.4);

figure(1)
%plot grafik pmf pakai faktorial dan binopdf
Y=[fx',pdf1'];
bar(Y,'group');
xlabel('x');
ylabel('f(x)');
title('pmf distribusi Binomial dengan n=10, p=0,4');
legend('faktorial','binopdf');

figure(2)
%plot grafik cdf pakai faktorial dan binocdf
S=[sumpy',pdf2'];
bar(S,'group');
xlabel('x');
ylabel('F(x)');
title('cdf distribusi Binomial');
legend('faktorial','binocdf');
```

5.4.2. Distribusi Poisson

Sensus penduduk pedalaman Watampone pada tahun 2012 menunjukkan keberadaan 2,5 orang albino per 175 orang. Jika diambil sampel 525 orang pada sensus tersebut dengan menggunakan pendekatan Poisson, tentukan:

a. Probabilitas tidak terdapat orang albino
b. Terdapat orang albino
c. Terdapat paling sedikit 3 orang albino

d. Gambarkan grafik cdf dan pmf nya, dengan program faktorial dan fungsi **poisspdf()** dan **poisscdf().**

```
% Program: DPois.m
% Distribusi Poisson

clear all;clc;
% Dapatkan nilai untuk X
x=1:1:30;
N=30;
% Nilai lambda
L=2.5;  %untuk 175 orang
L1=7.5; % untuk 3x175 = 525 orang sampel

% 1. Probabilitas tidak ada albino
% a. Dengan pemrograman faktorial
y=0;    % Untuk sampel tidak ada albino
Pr0=(exp(-L1)*L1^y)/factorial(y);

% Untuk semua sampel jumlah albino
for i=1:N
    f=i;
    PR(i)=(exp(-L1)*L1^f)/factorial(f);
end

% b. Fungsi Massa Probabilitas (pmf)dengan poisspdf
pdf1=poisspdf(x,L1);

% 2. Probabilitas ada albino
pdf2=1-pdf1;

% 3. Probabilitas paling sedikit ada 3 orang albino:
for i=1:3
    f=i;
    Pr(i)=(exp(-L1)*L1^f)/factorial(f);
end

Prt=1-sum(Pr);

% % Fungsi Kumulatif Distribusi (cdf)
% a. Menggunakan pemrograman faktorial
for i=1:N
    f=i;
    PC(i)=(exp(-L1)*L1^f)/factorial(f)
end

sumPC=sum(PC); %nilai cdf

sumP(1)=PC(1);
for i=2:N
    sumP(i)=sumP(i-1)+PC(i);
end

% CDF menggunakan poisscdf
pdf3=poisscdf(x,L1);

figure(1)
%plot grafik pmf pakai faktorial dan poisspdf
Y=[PR',pdf1'];
bar(Y,'group');
xlabel('x');
ylabel('f(x)');
title('pmf distribusi Poisson dengan lamda=7,5');
legend('faktorial','poisspdf');

figure(2)
%plot grafik cdf pakai faktorial dan binocdf
S=[sumP',pdf3'];
bar(S,'group');
xlabel('x');
ylabel('F(x)');
title('cdf distribusi Poisson');
legend('faktorial','poisscdf');
```

5.4.3. Distribusi Eksponensial

Waktu kedatangan seorang pelanggan di sebuah restoran kota kecil diasumsikan terdistribusi eksponensial dengan rata-rata 3,2 pelanggan per 30 menit.

a. Berapa menit waktu rata-rata antar pelanggan di restoran tersebut ?

b. Berapa probabilitas kedatangan pelanggan ada selang 1 jam atau kurang, $P = (X \leq 2)$?

c. Berapa probabilitas kedatangan pelanggan ada selang 15 menit atau lebih, $P = (X \geq 0{,}5)$

Gambar kan kurva pdf dan cdf totalnya untuk $\lambda = 3/2,\ \lambda = 1,\ \lambda = 1/2$

```
% Program: DEksp.m
% Distribusi Eksponensial
clear all; clc;

% Interval pengamatan
x=0:.1:10;

lamda=3.2;
% a. Waktu rata-rata antar pelanggan:
% Nilai mu = 1/lamda untuk 30 menit, x=1
mu1=(1/3.2)*30;

% b. Untuk selisih waktu kedatangan 1 jam atau kurang
% x=2 (karena 2x30 menit=1 jam)
%     P(x<=2)
xx=2;
y=1-exp(-lamda*xx);

% c. Untuk selisih waktu kedatangan 15 menit atau lebih
% x=0,5 (karena 0,5x30 menit=15 menit)
%     P(x>=0,5)
x1=0.5;
y1=exp(-lamda*x1);

lamda1=1; lamda2=1/2;
m=exppdf(x,lamda);
m1=exppdf(x,lamda1);
m2=exppdf(x,lamda2);

figure(1)
plot(x,m,'-k','Linewidth',2);
hold on;
plot(x,m1,'-b','Linewidth',2);
hold on;
plot(x,m2,'-r','Linewidth',2);
hold on;
xlabel('x');
ylabel('f(x)');
legend('\lambda=3.2','\lambda=1','\lambda=0,5');
title('pdf distribusi eksponensial P(X<10)');

v=expcdf(x,lamda);
v1=expcdf(x,lamda1);
v2=expcdf(x,lamda2);

figure(2)
plot(x,v,'-k','Linewidth',2);
hold on;
plot(x,v1,'-b','Linewidth',2);
hold on;
plot(x,v2,'-r','Linewidth',2);
hold on;
xlabel('x');
ylabel('F(x)');
title('cdf distribusi eksponensial P(X<10)');
legend('\lambda=3.2','\lambda=1','\lambda=0,5');
grid on;
```

5.4.4. Distribusi Normal

Nilai rata-rata hasil ujian suatu mata kuliah tertentu adalah 60, dengan varians 36. Untuk mendapatkan nilai A, nilai angka peserta ujian tersebut minimal adalah 81 dan mendapatkan nilai B jika nilai angkanya antara 65 sampai 70. Dengan sebaran nilai peserta yang terdistribusi normal maka jika diambil seorang peserta ujian secara acak, tentukan:

a. Probabilitas peserta tersebut mendapat nilai A
b. Probabilitas peserta tersebut mendapat nilai B
c. Probabilitas peserta tersebut mendapat nilai di bawah B

Gambarkan grafik pdf dan cdf untuk $\sigma = 6,\ \sigma = 10,\ \sigma = 3$

Gunakan persamaan biasa dan persamaan **normpdf()** dan **normcdf()**.

```
% Program: DNorm.m
% Distribusi Normal (Gaussian)

clear all; clc;

% Interval pengamatan
x = 0:1:100;

% Dapatkan nilai mean dan standard deviasi
mu=60;
sigma=6;   %varians=36

% a. Probabilitas dapat nilai A (nilai >81)
% P(X>81) --> Z=(X-mu)/sigma
X1=81;
Z1=(X1-mu)/sigma;  % Selanjutnya cari dengan tabel distribusi normal

% dengan normpdf
y1=normpdf(X1,mu,sigma);

% b. Probabilitas dapat nilai B (nilai >66, dan nilai <70)
% P(65<X<70) --> Z=(X-mu)/sigma
X2=65; X3=70;
Z2=(X2-mu)/sigma;  % Selanjutnya cari dengan tabel distribusi normal
Z3=(X3-mu)/sigma;

% dengan normpdf
y2=normpdf(X2,mu,sigma);
y3=normpdf(X3,mu,sigma);

% c. Probabilitas dapat nilai A (nilai <65)
% P(X<65) --> Z=(X-mu)/sigma
X4=65;
Z4=(X4-mu)/sigma;  % Selanjutnya cari dengan tabel distribusi normal

% dengan normpdf
y4=normpdf(X4,mu,sigma);

% Plot pdf
sigma1=10; sigma2=3;
w1=normpdf(x,mu,sigma);
w2=normpdf(x,mu,sigma1);
w3=normpdf(x,mu,sigma2);

figure(1)
plot(x,w1,'k',x,w2,'r',x,w3,'b','Linewidth',2);
axis([0 100 0 0.15]);
xlabel('x');
ylabel('f(x)');
legend('\sigma1=6', '\sigma2=10','\sigma3=3');
title('pdf distribusi Normal ');
```

```
% Plot cdf
m1=normcdf(x,mu,sigma);
m2=normcdf(x,mu,sigma1);
m3=normcdf(x,mu,sigma2);

figure(2)
plot(x,m1,'k',x,m2,'r',x,m3,'b','Linewidth',2);
axis([0 100 0 1]);
xlabel('x');
ylabel('F(x)');
legend('\sigma1=6', '\sigma2=10','\sigma3=3');
title('cdf distribusi Normal ');
grid on;
```

5.5. Analisa dan Kesimpulan

Analisalah dari masing-masing kasus yang telah diselesaikan dengan Matlab tersebut, kapan distribusi binomial akan mendekati distribusi normal ? dan kapan distribusi normal mendekati distribusi Eksponensial ?

Apa yang terjadi dengan perubahan nilai simpangan baku pada distribusi normal ? dan bagaimana pula dengan perubahan lamda pada distribusi eksponensial ? Apa pengaruhnya kedua parameter tersebut pada data-data observasinya ?

5.6. Tugas

Bangkitkan bilangan acak berukuran (2,*n*) menggunakan beberapa jenis distribusi yang telah anda kenal (binomial, poisson, normal dan eksponensial), dimana *n* adalah panjang data yang variatif, yaitu 100,400 dan 1000. Gunakan fungsi Matlab: **binornd(), poissrnd(), normrnd()** dan **eksprnd()**. Perhatikan batasan-batasan berikut ini:

a. Pada data terdistribusi binomial, tentukan sebuah nilai *p* tertentu dimana $\left(0 < p < 1\right)$

b. Gunakan $\mu = \frac{1}{3}n$ dan $\sigma = \frac{1}{10}n$ untuk mean dan simpangan baku dari data acak normal dan $\lambda = \mu$ untuk data acak poisson dan eksponensial.

<u>Contoh:</u> *A=binornd(50,1/3,2,10),* artinya membangkitkan bilangan acak terdistribusi binomial pada matriks dengan ukuran 2x10, dimana p=1/3 dan paramater 50.

Plot-lah data acak dengan matriks 2 dimensi yang anda dapatkan tadi ke dalam luasan observasi dengan ukuran *nxn*.

Contoh:

```
% Plot data dua dimensi dengan distribusi eksponensial
% Pada luasan 2 dimensi nxn

clear all; clc;
n=100;
p=1/3; lamda=30;

A=exprnd(lamda,2,n);
x=A(1,:);  % Baris pertama A sebagai koordinat x
y=A(2,:);  % Baris kedua A sebagai koordinat y

plot(x,y,'ob');
grid on;
axis([0 n 0 n]);
xlabel('X');
ylabel('Y');
```