66.75 Procesos Estocásticos

Trabajo Práctico N^o 1

Figuroa, Gonzalo	84255
Mello, Ariel	84853
Tempone, Nicolás	84548
Borgatello, José María	67758

Índice

1. Desarrollo del trabajo Práctico - Ejercicios	2
1.1. Ejercicio 10	2
1.2. Ejercicio 13	4
1.3. Ejercicio 14	5
1.4. Ejercicio 15	6
1.5. Procesamiento de imagen	13
2. Anexo I	17
2.1. histo.m	17
2.2. pe110.m	17
2.3. pe113.m	19
2.4. pe114.m	20
2.5. pe115.m	21
2.6. histRGB.m	24

1. Desarrollo del trabajo Práctico - Ejercicios

1.1. Ejercicio 10

a) En este problema se aborda la relación que existe entre una variable aleatoria Rayleigh y una variable aleatoria Normal. Por definición, una variable aleatoria Rayleigh surge de calcular la raíz cuadrada de la suma de los cuadrados de dos variables aleatorias normales.

$$R = \sqrt{X^2 + Y^2}$$

donde X e Y son variables aleatorias normales independientes de media $\mu = 0$ y desvío $\sigma = 1$.

La interpretación geométrica que se le puede dar a esto es tomar la norma de un vector aleatorio cuyas componentes son V.A. normales. Gráficamente se puede visualizar como un triángulo rectángulo de catetos X e Y e hipotenusa R .

Aceptado esto, se puede realizar el proceso inverso también, que consiste en obtener una variable aleatoria normal a partir de una variable Rayleigh.

Simplemente con obtener la proyección de la hipotenusa sobre cualquiera de los dos ejes se podrá obtener la variable normal pedida.

Entonces $X = R \cos(\Phi)$ o, de la misma forma, $Y = R \sin(\Phi)$, donde $\Phi = \arctan\left(\frac{Y}{X}\right)$ es una variable uniforme $U(-\pi; \pi)$.

Realizando lo pedido en Matlab se obtuvo el siguiente resultado:

Figura 1: *Ejercicio 10 a*

Se ve que al realizar el histograma de $N = 20000$ muestras, la función de distribución se asemeja altamente a la función de densidad de probabilidad teórica.

b) Si se utilizara la misma realización de la variable aleatoria uniforme para generar las variables aleatorias R y Φ , la función que se obtiene deja de ser Normal pues las variables en cuestión dejan de ser independientes.

Experimentalmente se obtuvo lo siguiente:

Figura 2: *Ejercicio 10 b*

Las gráficas dejan de tener similitud.

c) El cálculo de la media y varianzas muestrales arrojó resultados cercanos a los teóricos. Tomando como parámetro α de la variable Rayleigh $\alpha = 3$, se obtuvo:

$$\bar{X} = -0,0074$$

$$S^2 = 8,9181$$

donde los valores teóricos de la distribución son:

$$\mu_x = 0$$

$$\sigma^2 = \alpha^2 = 9$$

1.2. Ejercicio 13

El problema 13 es similar al anterior, sólo que en esta oportunidad se deben generar muestras de una variable aleatoria exponencial de parámetro λ .

a) Existen dos formas de generar muestras de esta variable aleatoria. Una de ellas es utilizar el comando `EXPRND` de Matlab, mientras que otra posibilidad es utilizar el concepto de simulación para generarlas a partir de muestras de una variable random.

En esta resolución en particular se utilizó la última alternativa, para lo cual se calcula la función de probabilidad acumulada de la variable aleatoria:

$$f_X(x) = \lambda e^{-\lambda x} \implies F_X(x) = 1 - e^{-\lambda x}$$

Y a partir de aquí se utiliza el cambio de variable $Y = F_X(x)$ y se despeja x para obtener las muestras.

$$X = -\frac{1}{\lambda} \ln(1 - Y)$$

Esta ecuación permite, a partir de una muestra de una variable random(Y), obtener una variable aleatoria exponencial negativa de parámetro λ .

Tras utilizar este procedimiento, se obtuvieron los siguientes resultados:

Figura 3: *Ejercicio 13*

Como en el caso anterior, la similitud es notable. El parámetro λ utilizado fue $\lambda = 5$.

b) La media y varianzas muestrales arrojaron los siguientes valores:

$$\bar{X} = 0,2007$$

$$S^2 = 0,0400$$

Y los valores teóricos de dicha distribución fueron:

$$\mu_x = \frac{1}{\lambda} = \frac{1}{5}$$

$$\sigma^2 = \frac{1}{\lambda^2} = \frac{1}{25}$$

La similitud de los resultados es muy alta, en parte debido al alto número de muestras tomadas en el ejercicio.

1.3. Ejercicio 14

En este caso, las muestras a generar son de una variable aleatoria de Rayleigh de parámetro α cuya función de densidad de probabilidad es:

$$f_X(x) = \frac{x}{\alpha^2} e^{-\frac{x^2}{2\alpha^2}} \quad \text{para } x \geq 0$$

a) Como en el caso anterior, existen 2 formas de generar las muestras de la variable aleatoria, y ellas son, mediante el comando RAYLRND o mediante simulación.

La estrategia utilizada fue la segunda. Para ello se realizó el mismo procedimiento que en el caso anterior. En este caso, el cambio de variable que genera las muestras es:

$$X = \sqrt{-2\alpha^2 \ln(1 - Y)}$$

Los resultados obtenidos fueron:

Figura 4: Ejercicio 14

El parámetro utilizado para este caso fue $\alpha = 2$ y el número de muestras tomadas fue $N = 40000$.

b) La media y varianza muestrales arrojaron los siguientes valores:

$$\bar{X} = 2,5183$$

$$S^2 = 1,7270$$

Y los valores teóricos de dicha distribución fueron:

$$\mu_x = \sigma \sqrt{\frac{\pi}{2}} = 2,5066$$

$$\sigma^2 = \frac{4 - \pi}{\alpha^2} = 1,7168$$

Nuevamente el procedimiento revela que la función de distribución es altamente similar a la función de densidad de probabilidad teórica al punto en que los parámetros característicos de la misma difieran en un 1%.

1.4. Ejercicio 15

Para la resolución de este ejercicio se estudió la deducción de la expresión de la función de probabilidad gaussiana multivariable. Básicamente, el objetivo del problema es la simulación de una variable aleatoria bidimensional a partir de su matriz de covarianza. Para ello se define un vector Z , de N variables gaussianas

independientes, idénticamente distribuidas. Conforme esto, la fdp de Z se puede escribir de la siguiente manera:

$$f_Z(Z) = \frac{1}{(2\pi)^{\frac{N}{2}} \sigma^N} e^{-\frac{\|Z\|^2}{2\sigma^2}}$$

Luego, se define $A \in \mathbb{R}^{N \times N}$ tal que $X = A \cdot Z$.

Donde:

- A es la matriz de transformación.
- Z es la variable aleatoria independiente (no correlacionada).
- X es variable aleatoria correlacionada.

$$\Rightarrow C_X = E[XX^T] = E[AZ Z^T A^T] = A E[ZZ^T] A^T \underset{C_Z = I_N}{=} A I_N A^T = AA^T$$

$$\Rightarrow A = (C_X)^{\frac{1}{2}}$$

Para este ejercicio $X = [X_1 X_2]^T$, entonces:

$$C_X = \begin{bmatrix} (\sigma_1)^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & (\sigma_2)^2 \end{bmatrix}$$

Realizando el cambio de variables tenemos:

$$f_X(X) = \frac{f_Z(Z(X))}{|\det(\frac{dX}{dZ})|} = \frac{1}{2\pi(\det C_X)^{\frac{1}{2}}} e^{-\frac{1}{2}(X-m_X)^T C_X^{-1}(X-m_X)}$$

La matriz A se obtiene realizando una diagonalización de la matriz C_X .

$$\Rightarrow A = (C_X)^{\frac{1}{2}} = U(\Sigma)^{\frac{1}{2}} = \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix}$$

Debido a que C_X es definida positiva decimos que:

- Σ es la matriz que posee en la diagonal los autovalores de C_X .
- U es la matriz que posee en cada una de sus columnas los autovectores v_i correspondientes a los autovalores λ_i .

En cuanto a la aplicación práctica del problema 15, se realizan los sucesivos puntos pedidos:

Se generan $N = 10000$ muestras de un vector aleatorio bidimensional de componentes normales que responden a la matriz de covarianza.

Y su vector media es $m = [m_1 m_2]^T$. Para nuestra implementación se tomó $m_1 = 1$ y $m_2 = 5$. A su vez, se tomó $\sigma_1 = 3$ $\sigma_2 = 2$.

Para los distintos valores de $\rho = 0; 0,5; 0,95; -0,5$ se grafican las muestras obtenidas en el plano $X1 - X2$ junto con las curvas de densidad de probabilidad constante. Finalmente se gráfica la función de densidad de probabilidad bidimensional.

a) $\rho = 0 \quad \sigma_1 = 3 \quad \sigma_2 = 2 \quad m_1 = 1 \quad m_2 = 5$

Figura 5: *Muestras Obtenidas*

Figura 6: *Fdp conjunta correlacionada*

b) $\rho = 0,5$ $\sigma_1 = 3$ $\sigma_2 = 2$ $m_1 = 1$ $m_2 = 5$

Figura 7: *Muestras Obtenidas*

Figura 8: *Fdp conjunta correlacionada*

c) $\rho = 0,95$ $\sigma_1 = 3$ $\sigma_2 = 2$ $m_1 = 1$ $m_2 = 5$

Figura 9: *Muestras Obtenidas*

Figura 10: *Fdp conjunta correlacionada*

d) $\rho = -0,5$ $\sigma_1 = 3$ $\sigma_2 = 2$ $m_1 = 1$ $m_2 = 5$

Figura 11: *Muestras Obtenidas*

Figura 12: *Fdp conjunta correlacionada***Estimación de parámetros:**

Para ese caso, se estimó la media con su estimador característico, el promedio muestral. El estimador del desvío es la raíz cuadrada de la varianza muestral. A su vez, para el coeficiente de correlación, se utilizó el coeficiente de correlación de Penrose, el cual se calcula mediante:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)s_x s_y}$$

Caso	Media Estimada	Desvío Estimado	ρ Estimado	C_x Estimada
$\rho = 0$	[1,0012 4,9956]	[3,0028 2,0043]	-0,0025	$\begin{bmatrix} 9,0167 & -0,0153 \\ -0,0153 & 4,0172 \end{bmatrix}$
$\rho = 0,5$	[0,9788 5,0154]	[2,9995 1,9828]	0,5027	$\begin{bmatrix} 8,9971 & -2,9894 \\ 2,9894 & 3,9313 \end{bmatrix}$
$\rho = 0,95$	[1,0717 5,0429]	[2,9671 1,9771]	0,9501	$\begin{bmatrix} 8,8039 & 5,5733 \\ 5,5733 & 3,9089 \end{bmatrix}$
$\rho = -0,5$	[1,0153 5,0031]	[2,9884 1,9803]	-0,4949	$\begin{bmatrix} 8,9307 & -2,9286 \\ -2,9286 & 3,9215 \end{bmatrix}$

Caso	C_x Teórica	C_x Estimada
$\rho = 0$	$\begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$	$\begin{bmatrix} 9,0167 & -0,0153 \\ -0,0153 & 4,0172 \end{bmatrix}$
$\rho = 0,5$	$\begin{bmatrix} 9 & 3 \\ 3 & 4 \end{bmatrix}$	$\begin{bmatrix} 8,9971 & -2,9894 \\ 2,9894 & 3,9313 \end{bmatrix}$
$\rho = 0,95$	$\begin{bmatrix} 9 & 5,7 \\ 5,7 & 4 \end{bmatrix}$	$\begin{bmatrix} 8,8039 & 5,5733 \\ 5,5733 & 3,9089 \end{bmatrix}$
$\rho = -0,5$	$\begin{bmatrix} 9 & -3 \\ -3 & 4 \end{bmatrix}$	$\begin{bmatrix} 8,9307 & -2,9286 \\ -2,9286 & 3,9215 \end{bmatrix}$

1.5. Procesamiento de imagen

Se eligió la imagen "mandrill.bmp" para hacer los histogramas para el rojo, verde y azul, a una resolución de 320 x 240 pixeles.

Figura 13: *Imagen: Mandrill; Resolución: 320 × 240*

Los histogramas correspondientes, producidos con los scripts de Matlab que se incluyen en el anexo, fueron los siguientes:

Figura 14: *Histograma de las intensidades de rojo de la imagen*

Figura 15: *Histograma de las intensidades de verde de la imagen*

Figura 16: *Histograma de las intensidades de azul de la imagen*

Por otra parte, para el histograma de la escala de grises, se utilizó la imagen 'lena.bmp' con una resolución de 640 x 480 píxeles.

Figura 17: *Imagen: Lena; Resolución: 640 × 480*

El histograma obtenido de manera análoga a los anteriores fue el siguiente:

Figura 18: *Histograma de la escala de grises de la imagen*

2. Anexo I

A continuación se presentan los programas en Matlab mencionados en este informe.

2.1. histo.m

```
function histo(input,bins)

n = length(input);

[frecuencia,intervalo] = hist(input,bins);
title('Histograma');
delta      = intervalo(bins)-intervalo(bins-1);
probabilidad= frecuencia/n;
valor_fdp  = probabilidad/delta;
bar(intervalo,valor_fdp,'b');
```

2.2. pe110.m

```
%-----%
%El presente es un script que apunta a la resolucio del Problema 10 de %
%la guia 1 de la materia. %
%Para su realizacion se utilizara la propiedad que una variable aleatoria %
%Rayleigh se obtiene como la raiz cuadrada de la suma de los cuadrados de %
%dos variables aleatorias normales independientes. %
%De la misma manera si se piensa a la variable Rayleigh como la hipotenusa %
%de un triangulo rectangulo cuyos catetos son variables normales(X, Y), para%
%obtener la longitud de uno de los catetos a partir de la hipotenusa, se %
%utilizan las propiedades trigonometricas, tal que R = X.cos(theta) donde %
%theta sera una variable uniforme. Esto ultimo se debe a que el angulo %
%theta surge de realizar el arco tangente de Y/X, donde Y/X es una %
%variable aleatoria Cauchy que tras realizar el cambio de variable arroja %
%el resultado de ser uniforme. %
%-----%
hold off;

N = 20000;          %Numero de muestras a generar de cada variable aleatoria
M= 4;              %Limite de la escala horizontal para los graficos
s= 1/50;           %Step del plot
x = -M:s:M;        %Escala horizontal
```

```
u = rand(1,N);           %Muestras de la funcion uniforme
bins = ceil(sqrt(N));   %Numero de bins a utilizar en el histograma
alfa = 3                 %Parametro de la variable Rayleigh

disp('Se generan N muestras de una variable aleatoria Rayleigh de parametro alfa')
N
alfa

%La funcion de densidad de probabilidad de la variable aleatoria normal x es:

fx = normpdf(x,0,1);

%Se grafica la funcion de densidad de probabilidad

hold on;
subplot(1,2,1);
title('PDF Teorica');
plot(x,fx,'r');

%Se obtienen las muestras de la variable Rayleigh

r = raylrnd(alfa,1,N);

%Al cambiar el comando de arriba por este, se utiliza la misma realizacion
%de la variable aleatoria uniforme para generar las variables R y U.
%r = sqrt(-2*(alfa^2)*log(1-u));

%Se extiende el rango de variabilidad de la variable uniforme

u = u*2*pi;

%Se realiza la transformacion

w= r.*cos(u);

subplot(1,2,2)
histo(w/alfa,bins);
title('Histograma a partir de las muestras');
disp('La media teorica de la variable aleatoria es:')
0
disp('La media de la poblacion generada es:')
mean(w)
disp('La varianza teorica de la variable aleatoria es:')
alfa^2
disp('La varianza de la poblacion generada es:')
var(w)
```

```
clear;
```

2.3. pe113.m

```
%-----%
%El presente es un script que apunta a la resolucio del Problema 13 de      %
%la guia 1 de la materia.                                                %
%Para su realizacion se utilizara el concepto de Simulacion en el contexto %
%de la Probabilidad y Estadistica                                         %
%-----%

N = 20000;           %Numero de muestras a generar
M= 2;               %Limite de la escala horizontal para los graficos
s= 1/50;           %Step del plot
x = 0:s:M-s;       %Escala horizontal

y = rand(1,N);      %Muestras de la funcion uniforme
lambda = 5;        %Parametro de la variable aleatoria
bins = ceil(sqrt(N)); %Numero de bins a utilizar en el histograma

disp('Se generan N muestras de una variable aleatoria exponencial de parametro lambda')
N
lambda

%La funcion de densidad de probabilidad de la variable aleatoria x es:

fx= lambda*exp(-lambda*x);

%para x >= 0

%Se grafica la funcion de densidad de probabilidad

hold on;
subplot(1,2,1);
plot(x,fx,'r');

%Se define la funcion de acumulacion de probabilidad correspondiente a la
%variable aleatoria exponencial de para poder realizar la simulacion.

%Fy(y) = 1 - exp(-lambda*y);

%Y luego se le realiza el cambio de variable z = Fy(y) y se despeja z
% para obtener las muestras
```

```

z = -1/lambda*log(1-y);
subplot(1,2,2)
histo(z,bins);
disp('La media teorica de la variable aleatoria es:')
1/lambda
disp('La media de la poblacion generada es:')
mean(z)
disp('La varianza teorica de la variable aleatoria es:')
1/(lambda^2)
disp('La varianza de la poblacion generada es:')
var(z)

```

2.4. pe114.m

```

%-----%
%El presente es un script que apunta a la resolucio del Problema 14 de      %
%la guia 1 de la materia.                                                %
%Para su realizacion se utilizara el concepto de Simulacion en el contexto %
%de la Probabilidad y Estadistica                                         %
%-----%

N = 40000;           %Numero de muestras a generar
M= 10;              %Limite de la escala horizontal para los graficos
s= 1/50;            %Step del plot
x = 0:s:M-s;        %Escala horizontal

y = rand(1,N);      %Muestras de la funcion uniforme
alfa = 2;           %Parametro de la variable aleatoria
bins = ceil(sqrt(N)); %Numero de bins a utilizar en el histograma

disp('Se generan N muestras de una variable aleatoria Rayleigh de parametro alfa')
N
alfa

%La funcion de densidad de probabilidad de la variable aleatoria x es:

fx= x/(alfa^2).*exp(-x.^2/(2*alfa^2));

%para x >= 0

%Se grafica la funcion de densidad de probabilidad

```

```

hold on;
subplot(1,2,1);
plot(x,fx,'r');

%Se define la funcion de acumulacion de probabilidad correspondiente a la
%variable aleatoria Rayleigh de para poder realizar la simulacion.

%Fy(y) = 1 - exp(-(y.^2)/(2*alfa^2));

%Y luego se le realiza el cambio de variable z = Fy(y) y se despeja z
% para obtener las muestras

z = sqrt(-2*(alfa^2)*log(1-y));
subplot(1,2,2)
histo(z,bins);
disp('La media teorica de la variable aleatoria es:')
alfa*sqrt(pi/2)
disp('La media de la poblacion generada es:')
mean(z)
disp('La varianza teorica de la variable aleatoria es:')
(4-pi)/2*alfa^2
disp('La varianza de la poblacion generada es:')
var(z)

```

2.5. pe115.m

```

%-----
% El presente es el script correspondiente al ejercicio 15 de la guia 1 de
% la materia. Aborda los temas de correlacion en una funcion de densidad
% bidimensional gaussiana.
%-----

close all;
clear all;
hold off;

N = input('ingrese el valor del numero de muestras a generar= ');
rho = input('ingrese el valor de rho= ');
m1 = input('ingrese el valor de m1= ');
m2 = input('ingrese el valor de m2= ');
sigma1 = input('ingrese el valor de sigma1= ');
sigma2 = input('ingrese el valor de sigma2= ');

```

```

m1g=ones(51,1)*m1;
m2g=ones(51,1)*m2;
t=-25:1:25;

%Se genera el vector bidimensional cuyas componentes son variables
%aleatorias gaussianas indeptes entre si
X = randn(2,N);

%Se ingresa la matriz de covarianza con los parametros correspondientes
C = [sigma1^2 rho*sigma1*sigma2 ; rho*sigma1*sigma2 sigma2^2];

%Se realiza una descomposicion de la matriz de covarianza en autovectores y
%autovalores.
[V,D] = eig(C);

%Se define la matriz A de transformacion que permite obtener una matriz de
%covarianza C tras ser aplicada al vector X
A = V * sqrt(D);
Y = A*X;

%Se suma la media correspondiente a cada variable aleatoria.
Y(1,:) = Y(1,)+m1;
Y(2,:) = Y(2,)+m2;

hold on;

%Graficos
%-----
%Muestras obtenidas en la simulacion
plot(Y(1,:),Y(2,:), 'o');

%Mesh de plotting
Xa=[min(Y(1,:)): ((max(Y(1,:))-min(Y(1,:)))/50) : max(Y(1,:))];
Xb=[min(Y(2,:)): ((max(Y(2,:))-min(Y(2,:)))/50) : max(Y(2,:))];
[Xd,Yd]= meshgrid(Xa,Xb);

%Definicion de la FDP Conjunta correlacionada
CI=inv(C);
J=det(C);
v=(1/(1-rho^2))*(((Xd-m1)./sigma1).^2+((Yd-m2)./sigma2).^2-((2*rho.*(Xd-m1).*(Yd-m2))./(sigma1*sigma2)));
Z=1/(2*pi*sqrt(J))*exp(-v/2);

%Curvas de densidad de probabilidad constante
contour(Xd,Yd,Z, 'r');
%hold on;

```

```

plot(m1g,t,'r--')
plot(t,m2g,'r--')
axis([-sigma1*5+m1 sigma1*5+m1 -sigma2*5+m2 sigma2*5+m2])
%hold off;
title('Muestras obtenidas');

%Grafico de la FDP Conjunta
figure (2)
mesh(Xd,Yd,Z)
tope =max(max(Z));
axis([-sigma1*5+m1 sigma1*5+m1 -sigma2*5+m2 sigma2*5+m2 0 tope*1.05])
title('FDP conjunta correlacionada');

%Estimacion de parametros
%-----
medial= mean(Y(1,:));
media2= mean(Y(2,:));
disp('Medias teoricas')
[m1;m2]
disp('Medias estimadas')
[medial;media2]

%Estimacion de la matriz de covarianza
cov_estimada = cov(Y');
disp('Matriz de covarianza teorica: ')
C
disp('Matriz de covarianza estimada: ')
cov_estimada
disp('Vector de desvios teorica: ')
[sigma1;sigma2]
disp('Vector de desvios estimados: ')
[std(Y(1,:));std(Y(2,:))]
disp('Coeficiente de correlacion teorico: ')
rho
disp('Coeficiente de correlacion estimado: ')
covarianza_estimada = 0;
for(i=1:N)
    % covarianza_estimada = covarianza_estimada+(Y(1,i)-medial)*(Y(2,i)-media2);
end
%rho_estimado = covarianza_estimada/((N-1)*std(Y(1,:))*std(Y(2,:)));
rho_estimado
clear;

```

2.6. histRGB.m

```
function histRGB(archivo,num)

%Como variable se pasan la imagen y el color con el que se realizará el histograma. num puede ser 1, 2, 3 o 4.

X=imread(archivo); %Lee el archivo imagen.

X=double(X); % Pasa los valores de X de uint8 a double.

if (num==4)
% Si num=4 se guarda en M la suma de los colores RGB promediados
    M=(X(:,:,1)+X(:,:,2)+X(:,:,3))/3;
else
    M=X(:,:,num); % Guarda la información de la matriz del color num.
end

%Una vez obtenida la matriz de un color determinado se pasa dicha matriz
%a un vector para poder usar la función histo.m y realizar el histograma.

d=size(M); % Guarda las dimensiones de M

L=d(1)*d(2); %Dimensiones que va a tener el vector

V=zeros(L,1); % Inicializa un vector de L espacios

%A continuación se utilizará un ciclo para guardar la matriz M en el vector V.

t1=length(X(1,:,1));
t2=length(X(:,1,1));

for i=1:t1,
    for j=1:t2,
        V(j+t2*(i-1))=M(j,i);
    end
end

histo(V,ceil(sqrt(length(V)))) % Histograma de V
```