Probabilidad y Estadística para IA. Práctico Clase 3

Ejercicio 1

Juan y Pedro juegan un juego de dados en el cual el que tira el dado más alto gana. Si ambos tiran el mismo número, tiran de nuevo hasta que uno gane. Juan ganó.

(a) Encontrar la probabilidad de que haya ganado con un 5. (Pista: listar todos los pares de tiradas de Juan y Pedro en las que Juan gana, y encontrar en cuáles gana con un 5)

R:

La cantidad de resultados posibles de arrojar dos dados de 6 es 36, pero de esos resultados sólo interesan aquellos en los que gana Juan. Si se restan los 6 de empate es simétrica la cantidad de resultados en los que Juan gana o pierde: $\frac{36-6}{2} = 15$. De esos 15 sólo existen cuatro casos posibles en los que gana.

$$P(X = 5|X > Y) = \frac{P(X = 5 \cap X > 5)}{P(X > Y)} = \frac{4}{15}$$

(b) Simular un dado y encontrar una estimación de la probabilidad anterior.

In [1]:

```
n_samples = 1000000;
p = unidrnd(6,[1,n_samples]); % Resultado de tirada de Pedro
j = unidrnd(6,[1,n_samples]); % Resultado de tirada de Juan
jg = sum(j>p); % Número de veces que Juan ganó
prob = sum((j==5) & (p<=4))/jg # % Número de veces que Juan ganó con 5 sobre to
```

```
prob =  0.26660
```

Ejercicio 2

Sea la longitud de una vara L. Supongamos que optamos por cortar la vara en un lugar elegido uniformemente al azar Y . Nos quedamos con la parte de vara de longitud entre [0, Y]. Luego nuevamente decidimos partir la porción restante en un lugar aleatoriamente elegido uniformemente, y llamamos a la longitud resultante X (Pista: usar la ley de esperanzas iteradas).

(a) Encontrar la expresión de E [X] en función de L.

R:

$$E[Y] = \frac{L}{2}$$

$$E[X] = E[E[X|Y]] = E[\frac{Y}{2}] = \frac{E[Y]}{2} = \frac{L}{4}$$

In [2]:

```
1/4
```

```
ans =  0.25000
```

(b) Encontrar la expresión de var[X] en función de L.

R:

Como la distribución de X es uniforme en [0,Y] su varianza está dada por:

$$var(X|Y) = \frac{(b-a)^2}{12} = \frac{Y^2}{12}$$

Y está distribuida de manera uniforme entre 0 y L, por lo tanto:

$$E[var(X|Y)] = \frac{1}{12}\int_0^L \frac{1}{L} y^2 dy = \frac{1}{12}\frac{1}{3L}y^3\Big|_0^L = \frac{L^2}{36}$$

Además $E[X|Y] = \frac{Y}{2}$, entonces:

$$var(E[X|Y]) = var(\frac{Y}{2}) = \frac{1}{4}var(Y) = \frac{1}{4}\frac{L^2}{12} = \frac{L^2}{48}$$

$$var(X) = E[var(X|Y)] + var(E[X|Y]) = \frac{L^2}{36} + \frac{L^2}{48} = \frac{7L^2}{144}$$

In [3]:

```
7/144
```

```
ans =  0.048611
```

(c) Simular el proceso con N = 1000 ensayos y encontrar la media y varianza muestral de X .

In [4]:

```
n_samples = 1000;
cuts=zeros(n_samples,1);
L = 1000;
for i = 1:n_samples
  cut_y = unidrnd(L);
  cuts(i) = unidrnd(cut_y);
endfor
mean(cuts/L),var(cuts/L)
```

```
ans =  0.25528
ans =  0.050846
```

Ejercicio 3

(a) Sean X , Y v.a. U[0, 1] independientes. Definamos Z = X + Y. Encontrar E[Z|X], E [X|Z], E [XZ|X], E[XZ|Z].

$$E[Z|X] = E[(X+Y)|X] = E[X|X] + E[Y|X] = E[X+Y|X] = E[X|X] + E[Y|X] = X + E[Y$$

$$E[X|Z] = E[X|X+Y] = X$$

Consulta: no tengo claro qué propiedad aplicar con: '|X+Y'

$$E[XZ|X] = E[X(X+Y)|X] = E[XX+XY|X] = XE[X|X] + YE[X|X] = X^2 + E[X|X] = X$$

$$E[XZ|Z] = E[XZ|(X+Y)] = E[X(X+Y)|(X+Y)] = E[XX+XY|X+Y] = E[XX|X+Y$$
$$= XE[X|X+Y] + YE[X|X+Y] = \ldots$$

(b) Simular y estimar dichas esperanzas.

In [ ]:

```
n_samples = 1000;
X =rand(n_samples,1);
Y =rand(n_samples,1);
Z = X + Y;
mean(Z) % VER |
```

Ejercicio 4

Sea un proceso AWGN:

$$y(n) = 2 + w(n), \text{ donde } w(n) \sim N(0,1)$$

Estimar la media y varianza de y (n) usando los siguientes estimadores:

(c)

$$\bar{y} = \frac{1}{N}\Sigma_{i=0}^{N-1} y(i)$$

(d)

$$s_n = \frac{1}{N}\Sigma_{i=0}^{N-1}(y(i) - \bar{y})^2$$

(e)

$$s_{n-1} = \frac{1}{N-1}\Sigma_{i=0}^{N-1}(y(i) - \bar{y})^2$$

(f) Calcular la esperanza de cada estimador

In [6]:

```
%TODO
```

(g) Simular con N = 10 y N = 10000

In [7]:

```
function samples = awgn(n)
    samples = randn(1,n) + 2;
end
```

In [8]:

```
function result = ymean(s)
    n = size(s)(2);
    result = (1/n)*sum(s);
end

function result = sn(s)
    n = size(s)(2);
    mu = ymean(s);
    result = (1/n)*sum((s-mu).^2);
end

function result = sn1(s)
    n = size(s)(2);
    mu = ymean(s);
    result = (1/(n-1))*sum((s-mu).^2);
end
```

In [9]:

```
function run_sim(n_samples)
    signal = awgn(n_samples);
    plot(signal);
    ymean(signal),sn(signal),sn1(signal)
end
```

In [10]:

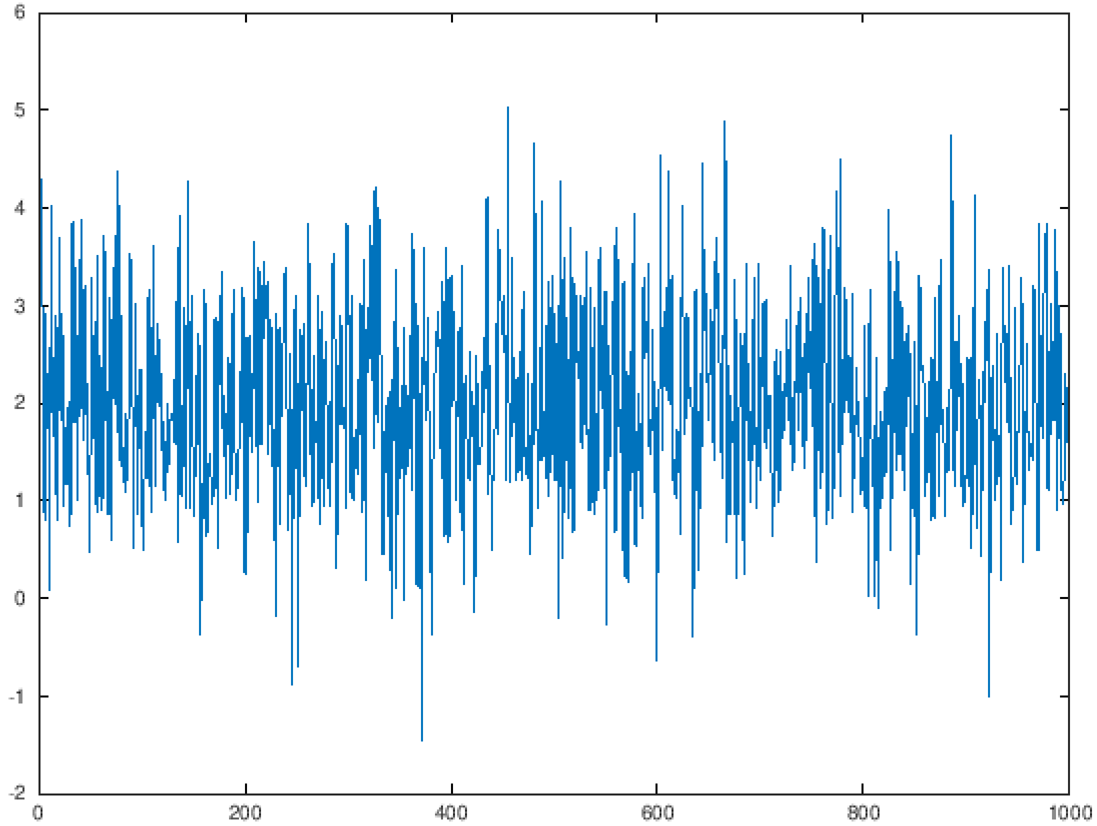
```
run_sim(10)
```

```
ans =  2.2493
ans =  1.4191
ans =  1.5768
```

In [11]:

```
run_sim(1000)
```

```
ans =  2.0060
ans =  0.94162
ans =  0.94256
```



(h) Interpretar los valores de s_n , s_{n-1} en cada caso. ¿Cuál es mejor de los dos?

R: Para números bajos de muestras S_{n-1} es mejor estimador. A medida que aumenta el número de muestras aumenta el impacto de la corrección disminuye y tienden a ser equivalentes.