

## Unidad 3.-Generación de VARIABLES ALEATORIAS no uniformes.

### 3.1 Introducción.

En los experimentos de simulación, por lo general existe la necesidad de generar valores de variables aleatorias (ventas, compras de inventario, llegadas de clientes, etc.) las cuales se encuentran interactuando y presentan comportamientos diferentes.

Cuando se realiza el experimento de simulación, el proceso de generar un valor de cierta variable aleatoria puede repetirse tantas veces como se requiera y con tantos comportamientos como distribuciones de probabilidad  $f(x)$  existan.

Es importante tener presente que el proceso de generación de variables aleatorias no uniformes se hace a partir de la generación de números rectangulares (pseudoaleatorios).

Las variables que interactúan en el experimento, por lo general presentan comportamientos que siguen distribuciones de probabilidad teóricas o empíricas, diferentes a la distribución uniforme. Por lo tanto, para generar estas variables, es necesario contar con un generador de números uniformes y una función que, a través de un método, proporcione los valores de la variable de la distribución deseada.

En los próximos apartados se presentarán las variables aleatorias discretas y continuas, así como las distribuciones más comunes, pero ¿Qué es una variable aleatoria?

La mayoría de los hechos de la vida real, genera resultados que se presentan como números reales, por ejemplo, el tiempo que se tarda el cajero en cobrar a un cliente, el tiempo que el vendedor atiende a un cliente, el número de accidentes automovilísticos al día en una ciudad, etc. Estos resultados numéricos que pueden cambiar en cada medición, día o experimento, se llaman *variable aleatoria*.

Cada resultado no está determinado previamente sino depende del comportamiento definido por una  $f(x)$ . El espacio muestral de un experimento son todos los resultados posibles al realizarlo, por ejemplo; al lanzar una moneda, hay dos resultados posibles: águila y sol, pero si lanzamos un dado el resultado de la cara superior puede ser 1, 2, 3, 4, 5 y 6.

Formalmente una variable aleatoria es una función de valor real (en el eje de las  $y$ ) cuyo dominio es un espacio muestral (en el eje de las  $x$ ).

### 3.2 Variables aleatorias Discretas.

**Definición.** Se denomina variable aleatoria discreta aquella que sólo puede tomar un número finito de valores dentro de un intervalo. Por ejemplo, el número de componentes de una manada de lobos, puede ser 4 ó 5 ó 6 individuos pero nunca 5.75 ó 5.87. Otros ejemplos de variable discreta serían el número de pollos de gorrión que llegan a volar del nido o el sexo de los componentes de un grupo familiar de babuinos.

Una variable aleatoria es discreta si su recorrido o rango es un conjunto discreto. Un conjunto es discreto si está formado por un número finito de elementos, o si sus elementos se pueden enumerar en secuencia de modo que haya un primer elemento, un segundo elemento, un tercer elemento, y así sucesivamente. En general podemos decir que **las variables discretas son enumerables** y se mide por conteo.

Se denomina densidad discreta a la probabilidad de que una variable aleatoria discreta  $X$  tome un valor numérico determinado ( $x$ ).

Se representa:

$$f(x) = P[X=x]$$

La suma de todos los valores (x) de la densidad será igual a 1. Esto es:  $\sum f(x) = 1$

El valor esperado de una variable aleatoria discreta es un promedio ponderado de todos los posibles resultados, donde las ponderaciones son las probabilidades asociadas con cada uno de los resultados.

$$\mu = E(x_i) = \sum x_i p(x_i)$$

Donde:  $x_i$  = i-ésimo resultado de X, la variable discreta de interés.

$P(x_i)$  = probabilidad de ocurrencia del i-ésimo resultado de X

La varianza de una variable aleatoria discreta ( $s^2$ ) se define como el promedio ponderado de los cuadrados de las diferencias entre cada resultado posible y su media (los pesos son las probabilidades de los resultados posibles). Esto es:

$$\text{Var}(x) = \sigma^2 = \sum (x_i - E(x))^2$$

### 3.3 Variables aleatorias continuas.

Una variable aleatoria es continua si su recorrido no es un conjunto numerable, intuitivamente esto significa que el conjunto de posibles valores de la variable abarca todo un intervalo de números reales. Por ejemplo, la variable que asigna la estatura a una persona extraída de una determinada población es una variable continua ya que, teóricamente, todo valor entre, supongamos por caso, 0 y 2,50 m, es posible.

### 3.4 Métodos para generar variables aleatorias.

Cuando se realiza el experimento de simulación, normalmente una o más variables tienen un comportamiento estocástico. El problema para generar estas variables aleatorias es determinar su comportamiento, es decir, se debe de probar la compatibilidad de un conjunto de frecuencias observadas con alguna frecuencia teórica, o bien, ¿pudieron originarse los datos observados de una distribución de probabilidad específica?, entonces para generar una variable aleatoria primero debemos de identificar su comportamiento, es decir que distribución de probabilidad sigue.

### Generación de variables aleatorias.

En un modelo de simulación, existen varias variables aleatorias interactuando. Generalmente siguen distribuciones de probabilidad teóricas o empíricas diferentes a la distribución uniforme, entonces se tiene que simular este tipo de variables, por lo que se hace necesario contar con un generador de números pseudoaleatorios R y una función  $f(x)$  que los transforme en los valores de la función deseada.

### Método de Monte Carlo.

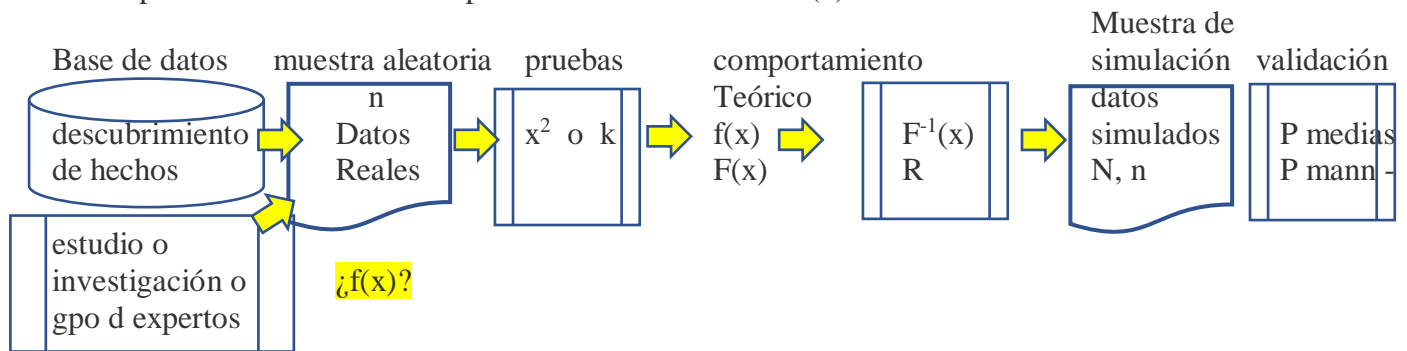
El método de Monte Carlo es una de las técnicas estadísticas más usadas en simulación. Naylor, dice que es una técnica de simulación para problemas que tienen una base estocástica o probabilística.

Su origen y nombre se remontan al trabajo de von Newman y Ulan a finales de los años cuarenta, cuando acuñaron el término y aplicaron la técnica para resolver ciertos problemas de protección nuclear. El método fue tan exitoso que se extendió su popularidad a otros campos y el término casi se ha convertido en sinónimo de simulación en la mente de muchas personas.

En el método, la experiencia o los datos artificiales se generan mediante algún generador de números aleatorios  $R$  y asociados a alguna distribución acumulada  $F(x)$  de interés. El generador de números aleatorios puede ser una tabla de dígitos aleatorios, una ruleta, una calculadora, una subrutina de computadora o cualquier otra fuente de dígitos aleatorios  $R$  distribuidos uniformemente.

La distribución ( $f(x)$ ) de probabilidad de interés, puede basarse en datos empíricos que se obtienen de registros anteriores (descubrimiento de hechos), ser resultado de experimentos recientes (mediciones realizadas) y puede ser una distribución teórica conocida (con las pruebas de  $\chi^2$  o kolmogorov).

Los números aleatorios servirán para producir una secuencia aleatoria de valores que representarán la experiencia esperada de acuerdo al comportamiento de la variable ( $x$ ) de interés.



Para generar la muestra aleatoria artificial (simulada) se debe de partir de alguna distribución de probabilidad. Podemos usar el siguiente procedimiento:

1. Graficar los datos de interés como una **distribución de probabilidad acumulada  $F(x)$** . En el eje  $x$  los valores de las variables y en el eje  $y$  la probabilidad ( $p(x)$ ) con rango  $[0,1]$ .
2. Seleccione **un número decimal aleatorio ( $R$ )** por medio de un generador de números aleatorios.
3. Proyecte horizontalmente el punto sobre el eje  $y$  que corresponde al  $R$ , hasta que la línea interseque la curva acumulativa.
4. Proyectar hacia abajo de este punto de intersección de la curva, al eje  $x$ .
5. Escriba el valor de  $x = F^{-1}(R)$ , el cual será el valor aleatorio para la muestra simulada.
6. Repita los pasos del 2 al 6, hasta lograr la cantidad deseada de valores de las variables aleatorias.

### Método de la transformada inversa para variables aleatorias continuas.

Este método utiliza la **distribución acumulada  $F(x)$**  de la distribución  $f(x)$  a simular.

Como  $F(x)$  se define en el intervalo  $(0,1)$ ; se genera **un número pseudoaleatorio  $R$**  y se trata de determinar el valor de la variable aleatoria ( $x$ ) para la cual su distribución acumulada  $F(x)$  es igual a  $R$ .

$F(x) = R$ , entonces la variable aleatoria debe de cumplir:

$$x = F^{-1}(R)$$

### Método de la transformada inversa para variables aleatorias.

Para distribuciones de probabilidad, primero se evalúa su distribución  $f(x)$  para cada valor de  $x$  y con ellos se determina la función acumulada  $F(x)$ .

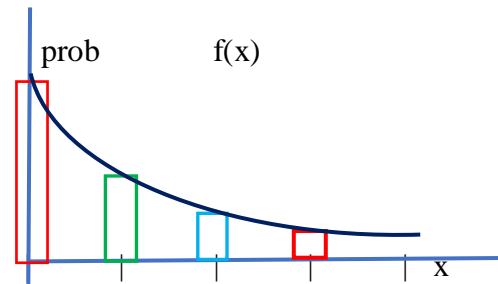
Luego se establecen los intervalos de la variable de 0 a cada valor  $F(x)$  de la variable  $x$ , y de acuerdo a la distribución acumulada se elabora una tabla que genera el valor de  $x$  de acuerdo al número pseudoaleatorio  $R$ .

obsérvese una distribución **exponencial**. Discreta.

La **distribución exponencial** suele referirse a la cantidad de tiempo que transcurre hasta que se produce algún evento específico. Por ejemplo, la cantidad de tiempo (que comienza ahora) hasta que se produzca un terremoto tiene una distribución exponencial. Otros ejemplos son la duración, en minutos, de las llamadas telefónicas de larga distancia comerciales y la cantidad de tiempo, en meses, que dura la batería de un automóvil. También se puede demostrar que el valor del cambio que se tiene en el bolsillo o en el monedero sigue una distribución exponencial aproximadamente.

La función de distribución es:

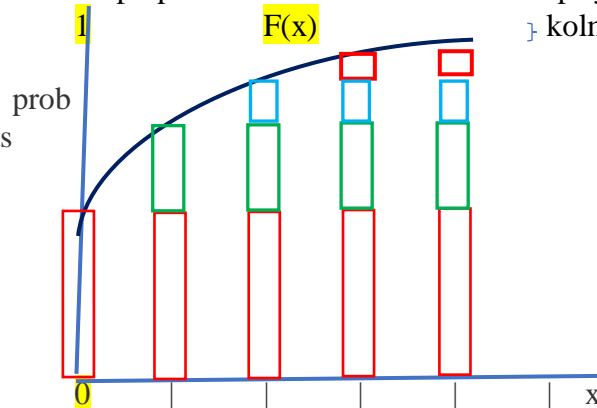
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{para } x \geq 0 \\ 0 & \text{de otro modo} \end{cases}$$



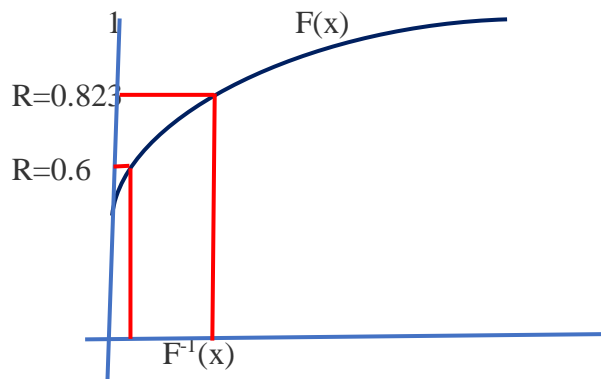
donde  $\lambda$  es el promedio de la función con las unidades de tiempo por cada unidad de la variable p.ej. 2 clientes por minuto, implica  $\lambda = (1/2)$  min/cte.

La función acumulada  $F(x)$  se obtiene acumulando las Probabilidades individuales.

En el eje “y” los valores están en el rango de  $[0,1)$   
 Los valores de la probabilidad acumulada  $F(x)$  inicia con  $(x=0, P(0))$ , luego  $(x=1, P(1))$ , etc  
 los valores de  $F(x)$  se incrementan, según el comportamiento, en este caso, exponencial



los valores de  $F(x)$  tienen rango  $[0,1)$  y un valor aleatorio uniforme  $R$  también tiene rango  $[0,1)$ .  
 Entonces si se toman valores de  $R$  se puede Conocer el valor de  $x$  correspondiente.



$$f(Y)=3x \quad x= F^{-1}(x) =R$$

## Distribución Exponencial.

Existen muchas situaciones reales que se comportan como la distribución exponencial, como por ejemplo la cantidad de tiempo (x) para que ocurra otro: nacimientos, muertes, accidentes; descompostura, llegada de un cliente el tiempo de atención de un servidor por cliente. Así x: cantidad de tiempo por cada llegada de un cliente, cantidad de tiempo para que se descomponga un equipo, cantidad de tiempo para atender a un cliente; de las obtenemos un promedio  $\mu$  entonces  $\lambda = 1/\mu$ .

La función de densidad es:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{para } x \geq 0 \\ 0 & \text{de otro modo} \end{cases}$$

En [estadística](#) la **distribución exponencial** es una [distribución de probabilidad](#) continua con un parámetro  $\lambda > 0$  cuya [función de distribución](#) es:

$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{para } x < 0 \\ 1 - e^{-\lambda x} & \text{para } x \geq 0 \end{cases}$$

donde  $e$  representa el [número e](#).

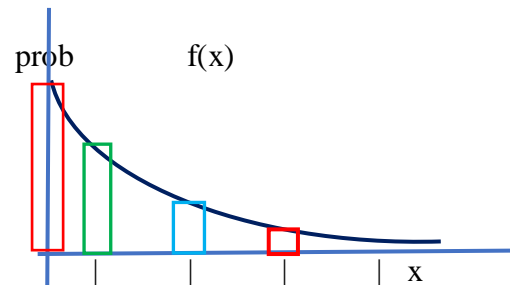
El [valor esperado](#) y la [varianza](#) de una [variable aleatoria](#) X con distribución exponencial son:

$$E[X] = \frac{1}{\lambda} \qquad V(X) = \frac{1}{\lambda^2}$$

Ejemplo. Se desean generar valores aleatorios que sigan una distribución **exponencial**.

La función de distribución es:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{para } x \geq 0 \\ 0 & \text{de otro modo} \end{cases}$$

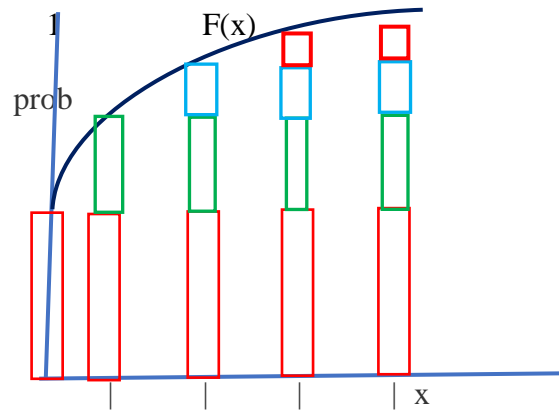


donde  $\lambda$  es el promedio de la función con las unidades de tiempo por cada unidad de la variables. Como  $\mu=2$  min por cliente Implica que  $\lambda = (1/2)$  **min / cliente**

$f(x) = \lambda e^{-\lambda x}$ ; si  $\lambda = (1/2)$  min/cliente. Entonces,  $f(0) = (1/2)e^{-(1/2)*0} = 0.5$

$f(1) = (1/2)e^{-(1/2)*1} = 0.3032$

$f(2) = (1/2)e^{-(1/2)*2} = 0.1838$



La función acumulada  $F(x)$  se calcula con:

$$F(x) = \int f(x) dx = \int \lambda e^{-\lambda x} dx,$$

con límite inferior 0 y superior x

$$\text{Como } \int e^u du = e^u ;$$

Si  $u = -\lambda x$  entonces  $du = -\lambda dx$ . y cambiando x por t :

$$F(t) = - \int_0^x e^{-\lambda t} (-\lambda) dt = - e^{-\lambda t} \Big|_0^x \text{ sustituyendo los limites obtenemos } x) = - (e^{-\lambda x} - 1) = -e^{-\lambda x} + 1 \text{ o bien}$$

$$F(x) = 1 - e^{-\lambda x}$$

Igualando  $F(x)$  con  $R$ ; por que tienen el mismo rango de  $[0, 1)$   $R = 1 - e^{-\lambda x}$  .

$$R - 1 = -e^{-\lambda x}$$

$$1 - R = e^{-\lambda x}$$

Como  $R$  tiene un rango de  $[0, 1)$ , y  $1-R$  tiene el mismo rango, entonces:

$$e^{-\lambda x} = R, \text{ para despejar } x :$$

$$\ln(e^{-\lambda x}) = \ln R$$

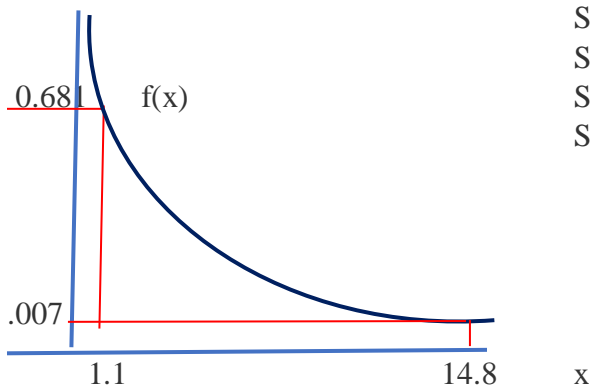
$$-\lambda x = \ln(R)$$

$$x = - (1/\lambda) \ln(R)$$

es la transformada inversa de la función exponencial.

3 ctes/2min

Suponga que la llegada de clientes a una tienda ocurre en promedio de  $\mu=3$  clientes por minuto de un comportamiento exponencial. Esto es que  $\lambda = (1/3)$  (min / cte) . Entonces para simular la llegada de clientes se requiere de valores de  $R$ .



$$\text{Si } R = 0.681 \text{ entonces } x = - (1/(1/3)) \ln(0.681) = 1.152 \cong 1 \text{ cte}$$

$$\text{Si } R = 0.337 \text{ entonces } x = - (1/(1/3)) \ln(0.337) = 3.263 \cong 3 \text{ cte}$$

$$\text{Si } R = 0.007 \text{ entonces } x = - (1/(1/3)) \ln(0.007) = 14.8855 \cong 15 \text{ cte}$$

$$\text{Si } R = 0.991 \text{ entonces } x = - (1/(1/3)) \ln(0.991) = 0.0271 \cong 0 \text{ cte}$$

El tiempo que los cónyuges dedican a la compra de tarjetas de aniversario se puede modelar mediante una distribución exponencial con un promedio de tiempo igual a ocho minutos. Escriba la distribución, indique la función de densidad de probabilidad y haga un gráfico de la distribución.



## Distribución Poisson.

La distribución de Poisson es una distribución de probabilidad discreta. Expresa la probabilidad de que un número  $k$  de eventos está ocurriendo en un tiempo fijo, si estos eventos ocurren con una frecuencia media conocida y son independientes del tiempo discurrido desde el último evento. Fue descubierta por Siméon-Denis Poisson.

La distribución se puede utilizar para el modelar **las llegadas** a una caja registradora, el número de defectos en un metro cuadrado de tela, el número de elementos de una colonia de bacterias en un centímetro cúbico de agua, el número de máquinas que fallan en el transcurso de una semana, el número de accidentes que ocurren al día en cierto cruce de la ciudad, etc.

La función de densidad de probabilidad es:

$$f(k; \lambda) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad \text{donde } \lambda \text{ es un parámetro positivo que representa la frecuencia esperada del fenómeno modelado por la distribución.}$$

Tanto el **valor esperado** como la **varianza** de una variable aleatoria con distribución de Poisson **son iguales a**  $\lambda$ . Esto es que:  $E(x) = \lambda$ , y  $V(x) = \lambda$ .

Ejemplo. Elaborar el generador de la variable aleatoria, cantidad de ventas semanales, que sigue una **distribución Poisson** con  $\lambda = 5$  **pzas/sem.**

$$f(k; \lambda) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad \text{con } k = (x) = 0, 1, 2, 3, 4, \dots$$

Primero evaluamos la función para diversos valores de  $k$ , por ejemplo:

$$f(0) = 5^0 e^{-5} / 0! = 0.0067$$

$$f(1) = 5^1 e^{-5} / 1! = 0.0337$$

$$f(2) = 5^2 e^{-5} / 2! = 0.0842$$

...

A continuación se elabora la tabla **equivalente** a la transformada inversa de la función.

$k = (x)$	$f(x)$ Probabilidad del evento	$F(x)$ Funcion acumulada
0	0.0067	0.0067
1	0.0337	0.0404
2	0.0842	0.1246
3	0.1409	0.2650
4	0.1755	0.4405
5	0.1755	0.6160
6	0.1462	0.7622
7	0.1044	0.8666
8	0.0653	0.9319
9	0.0363	0.9682
10	0.0181	0.9863
11	0.0082	0.9945
12	0.0034	0.9979
13	0.0013	0.9992
14	0.0005	0.9997
15	0.0002	0.9999



Valor del número pseudoaleatorio	Valor generado de $k=x$
$0.0000 \leq R < 0.0067$	0
$0.0067 \leq R < 0.0404$	1
$0.0404 \leq R < 0.1246$	2
$0.1246 \leq R < 0.2650$ 0.129	3
$0.2650 \leq R < 0.4405$ 0.3854	4
$0.4405 \leq R < 0.6160$ .568	5
$0.6160 \leq R < 0.7622$ 0.653      0.7321	6
$0.7622 \leq R < 0.8666$	7
$0.8666 \leq R < 0.9319$	8
$0.9319 \leq R < 0.9682$	9
$0.9682 \leq R < 0.9882$	10
$0.9882 \leq R < 0.9945$	11
$0.9945 \leq R < 0.9979$	12
$0.9979 \leq R < 0.9992$	13
$0.9992 \leq R < 0.9997$	14
$0.9997 \leq R < 0.9999$	15

Obteniendo  $R_1=0.3854$ , implica un valor de  $x_1=4$ , es decir la venta para la semana 1 será 4 pzas.  
Obteniendo  $R_2=0.7321$ , implica un valor de  $x=6$ , es decir la venta para la semana 2 será 6 pzas.  
Obteniendo  $R_3=0.129$ , implica un valor de  $x=3$ , es decir la venta para la semana 3 será 3 pzas.

## Distribución Binomial.

La **distribución binomial** es una distribución de probabilidad discreta que mide el número de éxitos en una secuencia de  $n$  ensayos independientes de Bernoulli con una probabilidad fija  $p$  de ocurrencia del éxito entre los ensayos.

Un experimento de Bernoulli se caracteriza por ser dicotómico, esto es, sólo son posibles dos resultados. A uno de estos se denomina éxito y tiene una probabilidad de ocurrencia  $p$  y al otro, fracaso, con una probabilidad  $q = 1 - p$ . En la distribución binomial el anterior experimento se repite  $n$  veces, de forma independiente, y se trata de calcular la probabilidad de un determinado número de éxitos. Para  $n = 1$ , la binomial se convierte, de hecho, en una distribución de Bernoulli.

Para representar que una variable aleatoria  $X$  sigue una distribución binomial de parámetros  $n$  y  $p$ , se escribe:

$$X \sim B(n, p) \qquad f(x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

Existen muchas situaciones en las que se presenta una experiencia binomial. Este tipo de experiencias se caracteriza por estar formada por un número predeterminado  $n$  de experimentos iguales. Cada uno de los experimentos es independiente de los restantes (la probabilidad del resultado de un experimento no depende del resultado del resto). El resultado de cada experimento ha de admitir sólo dos categorías (a las que se denomina éxito y fracaso). Las probabilidades de ambas posibilidades han de ser constantes en todos los experimentos (se denotan como  $p$  y  $q$  o  $p$  y  $1-p$ ).

Se designa por  $X$  a la variable que mide el número de éxitos que se han producido en los  $n$  experimentos.

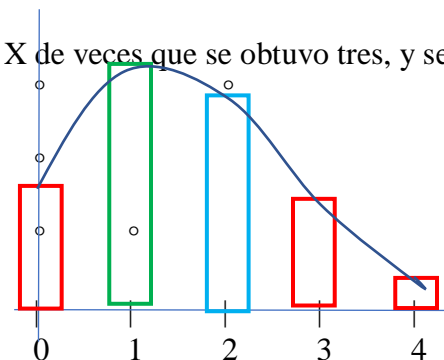
Cuando se dan estas circunstancias, se dice que la variable  $X$  sigue una distribución de probabilidad binomial, y se nota  $B(n, p)$ .

Una variable aleatoria tiene una distribución Binomial si existen las siguientes condiciones:

1. El experimento consiste en un número fijo de  $n$  intentos idénticos.
2. Cada intento solo puede tener un resultado de dos posibles, que se llama *éxito* y *fracaso*.
3. La probabilidad  $p$  de éxito es constante de intento a intento.
4. Los intentos son independientes.
5. Se define a  $X$  como el número de éxitos en el intento.

Por ejemplo si se lanza un dado diez veces  $n=10$  y se cuenta el número  $X$  de veces que se obtuvo tres, y se sabe que  $p=1/6$ ; entonces:  $X \sim B(10, 1/6)$ ;

$$\begin{aligned} f(0) &= {}_{10}C_0 * (1/6)^0 * (5/6)^{10} = 0.1615 \\ f(1) &= {}_{10}C_1 * (1/6)^1 * (5/6)^9 = 10 * (1/6) * 0.1938 = 0.323 \\ f(2) &= {}_{10}C_2 * (1/6)^2 * (5/6)^8 = 45 * 0.0277 * 0.2325 = 0.2907 \\ f(3) &= {}_{10}C_3 * (1/6)^3 * (5/6)^7 = 120 * 0.0043 * 0.279 = 0.155 \\ f(4) &= {}_{10}C_4 * (1/6)^4 * (5/6)^6 = 210 * 0.0007 * 0.1938 = 0.0542 \end{aligned}$$



otro ejemplo, se lanza una moneda dos veces  $n=2$  y se cuenta el número  $X$  de águilas obtenidas, y se sabe que  $p=1/2$ ; entonces:  $X \sim B(2, 1/2)$ .

La función de probabilidad Binomial está dada por:

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

donde  $x = \{0, 1, 2, \dots, n\}$ , siendo

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \quad \text{las combinaciones de } n \text{ en } x (n \text{ elementos tomados de } x \text{ en } x)$$

$$\mathbb{E}[X] = np \quad ; \quad \text{Var}[X] = np(1-p)$$

## Generación de valores de variable binomial

Supongamos que un producto tiene 5% de elementos defectuosos. Si se toma una muestra de 10 de esos artículos, simule la cantidad de defectuosos, observe la  $f(x)$

datos:  $n=10$ ,  $p=.05$ ,  $q=.95$ .

$$p(x=0) = \binom{10}{0} (.05)^0 (.95)^{10} = (1) * (1) * (.5987) = 0.5987$$

$$p(x=1) = \binom{10}{1} (.05)^1 (.95)^9 = 10 * (.05) * (.6302) = 0.3151$$

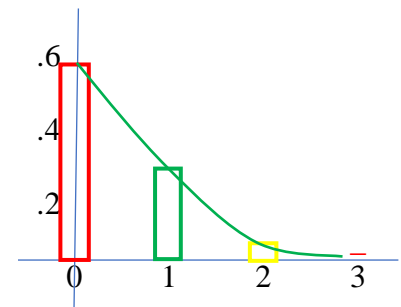
$$p(x=2) = \binom{10}{2} (.05)^2 (.95)^8 = 45 * (.0025) * (.6634) = 0.0746$$

$$p(x=3) = \binom{10}{3} (.05)^3 (.95)^7 = 120 * (.0001) * (.6983) = 0.0083$$

$$p(x=4) = \binom{10}{4} (.05)^4 (.95)^6 = 210 * (.0) * (.735) =$$

$$p(x=5) = \binom{10}{5} (.05)^5 (.95)^5 = 252 * (0) * (0) =$$

$$p(x=6) = \binom{10}{6} (.05)^6 (.95)^4 = 210 * (0) * (0) =$$



R1=0.325

R2=0.005

R3=0.854

R4=0.321

R5=0.03

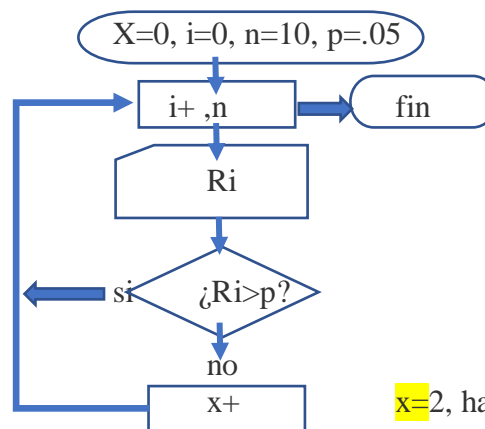
R6=0.186

R7=0.543

R8=0.379

R9=0.734

R10=0.811



$x=2$ , hay 2 piezas defectuosas en la muestra.

De nuevo:

R1, R2, R3, R4, ... R10. Y se obtiene otro valor de  $x$  que tiene comportamiento binomial, si se cuenta cuantos valores son menores o iguales que  $p=0.05$ . esa cantidad es el valor de  $x$ .

Entre las compras de materia prima de una empresa se espera que el 3% de los artículos esté defectuoso. Un inspector selecciona 6 piezas de un artículo al azar, simule 3 selecciones y determine cuantos artículos defectuosos hay en cada selección.

Si  $n=6$  y  $p=0.03$  1ª muestra, COND INICIO:  $x=0$ ,  $i=0$ ,  $n=10$ ,  $p=.03$

0.301	0.830	0.690	0.807	0.480	0.600
-------	-------	-------	-------	-------	-------

La cantidad de valores menores que 0.03, entonces  $x_1 = 0$

2ª muestra, COND INICIO:  $x=0$ ,  $i=0$ ,  $n=10$ ,  $p=.03$

0.190	0.654	0.500	0.096	0.013	0.054
-------	-------	-------	-------	-------	-------

La cantidad de valores menores que 0.03, entonces  $x_2 = 1$

3ª muestra, COND INICIO:  $x=0$ ,  $i=0$ ,  $n=10$ ,  $p=.03$

.745	.256	.012	.345	.015	.009
------	------	------	------	------	------

La cantidad de valores menores que 0.03, entonces  $x_3 = 3$

El departamento de control de calidad de una empresa que fabrica tornillos, sabe que el 15% de los tornillos que fabrica tienen defectos. Si en un proceso cualquiera de fabricación se toma una muestra de 5 tornillos para verificar si tienen o no defectos, simule tres muestras y determine cuantos tornillos defectuosos hay en cada muestra.

Si  $n=5$  y  $p=0.15$ , 1ª muestra:

0.254	0.810	0.567	0.046	0.091	
-------	-------	-------	-------	-------	--

La cantidad de valores menores que 0.15, entonces  $x_1 = 2$

## Distribución Normal.

Se llama **distribución normal**, **distribución de Gauss** o **distribución gaussiana**, a una de las distribuciones de probabilidad de variable continua que con más frecuencia aparece en fenómenos reales.

La gráfica de su Función de densidad tiene una forma acampanada y es simétrica respecto de un determinado parámetro, la media. Esta curva se conoce como campana de Gaus.

La importancia de esta distribución radica en que permite modelar numerosos fenómenos naturales, sociales y psicológicos. Mientras que los mecanismos que subyacen a gran parte de este tipo de fenómenos son desconocidos, por la ingente cantidad de variables incontrolables que en ellos intervienen, el uso del modelo normal puede justificarse asumiendo que cada observación se obtiene como la suma de unas pocas causas independientes. Algunos **ejemplos** de variables asociadas a fenómenos naturales que siguen el modelo de la normal son:

- caracteres morfológicos de individuos como la estatura;
- caracteres fisiológicos como el efecto de un fármaco;
- caracteres sociológicos como el consumo de cierto producto por un mismo grupo de individuos;
- caracteres psicológicos como el cociente intelectual;
- nivel de ruido medido en un ambiente.
- Errores experimentales cometidos al medir ciertas magnitudes, etc.
- Se dice que una variable aleatoria continua  $X$  sigue una distribución normal de parámetros  $\mu$  y  $\sigma$  y se denota  $X \sim N(\mu, \sigma)$  y su función de densidad está dada por:

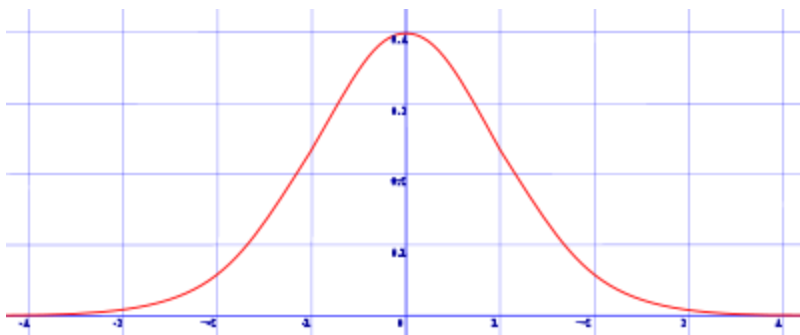
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad x \in \mathbb{R},$$

donde  $\mu$  (mu) es la media y  $\sigma$  (sigma) es la desviación típica ( $\sigma^2$  es la varianza).

Se llama **distribución normal "estándar" o tipificada** a aquella en la que sus parámetros toman los valores  $\mu = 0$  y  $\sigma = 1$ . En este caso la función de densidad tiene la siguiente expresión:

$$f(x) = f_{0,1}(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}, \quad x \in \mathbb{R},$$

Su gráfica se muestra abajo y con frecuencia se usan tablas para el cálculo de los valores de su distribución.



$$E(x) = \mu \quad \text{y} \quad \text{Var}(x) = \sigma^2$$

NORMAL tipificada.  $z = (x - \mu) / \sigma$ . Donde: x son los valores de la variable,  $\mu$  la media de los valores y  $\sigma$  la desviación estándar.

La media de los pesos de 400 estudiantes de un cierto colegio es de 63 kg., con una desviación estándar de 4.8kg. Suponiendo que los pesos se distribuyen normalmente, hallar cuántos Q estudiantes pesan:

$n=400$  alumnos,  $\mu = 63$  kg,  $\sigma = 4.8$  kg

a) cuantos pesan entre 58 y 67 kg.

$$Z_{58} = (58 - 63) / 4.8 = -1.0416$$

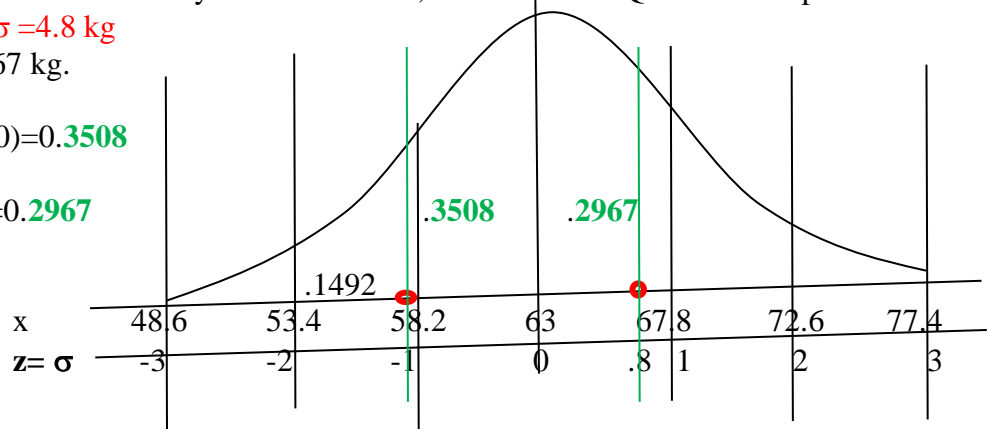
$$P(58 < x < 63) = P(-1.0416 < z < 0) = 0.3508$$

$$Z_{67} = (67 - 63) / 4.8 = 0.8333$$

$$P(63 < x < 67) = P(0 < z < .8333) = 0.2967$$

$$0.5 - .1492 = 0.3508$$

$$0.5 - 0.2033 = 0.2967$$



por último  $Q = P(58 < x < 67) * n = (.3508 + .2967) * 400 = 259$  alumnos

b) hallar cuántos Q estudiantes pesan más de 68 kg.

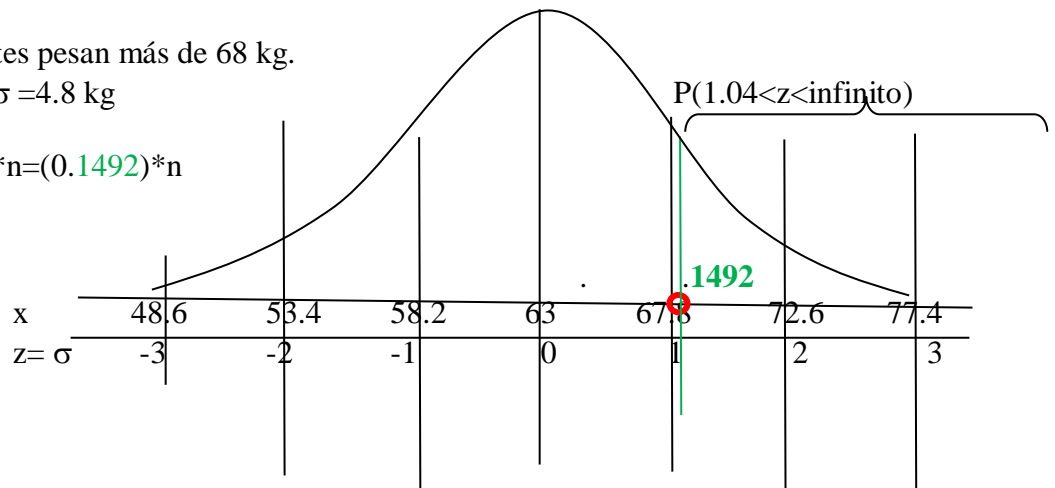
$n=400$  alumnos,  $\mu = 63$  kg,  $\sigma = 4.8$  kg

$$Z_{68} = (68 - 63) / 4.8 = +1.0416$$

$$Q = P(x > 68) * n = P(z > 1.0416) * n = (0.1492) * n$$

$$Q = (.1492) * 400 = 59.68$$

aprox  $Q = 60$  alumnos



el tiempo promedio para terminar un examen es

$\mu = 70$  min, con  $\sigma = 13$  min

a) cual es la prob de q termine entre 65 y 68 min.,

$$Z_{65} = (65 - 70) / 13 = -.3846; P(-.3846 < z < 0) = .148$$

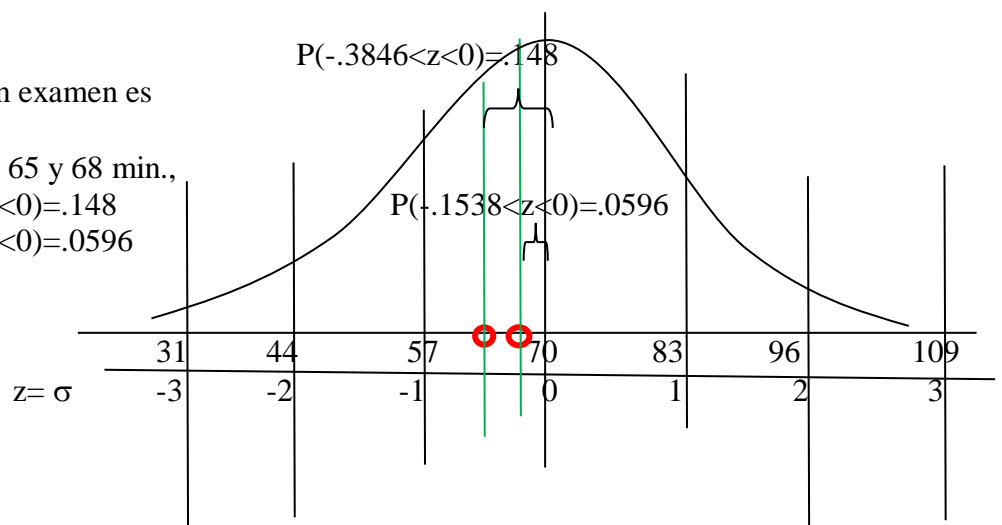
$$Z_{68} = (68 - 70) / 13 = -.1538; P(-.1538 < z < 0) = .0596$$

$$P(65 < x < 68) = .148 - .0596 = .0884$$

la prob q un alumno termine

entre 65 y 68 min

es de 8.84%



$$0.5 - .352 = 0.148$$

$$0.5 - .4404 = .0596$$

en una fábrica de bombillas, cuya duración en uso, se distribuye normalmente con una media de 1000 horas y una desviación estándar de 20 hrs. ¿Cual es la probabilidad de que se quemé entre 960 y 1020 horas?

$$Z_{1020} = (1020 - 1000) / 20 = 1$$

$$Z_{960} = (960 - 1000) / 20 = -2$$

$$P(1000 < x < 1020) = P(0 < z < 1) = .3413 \quad .5 - .1587$$

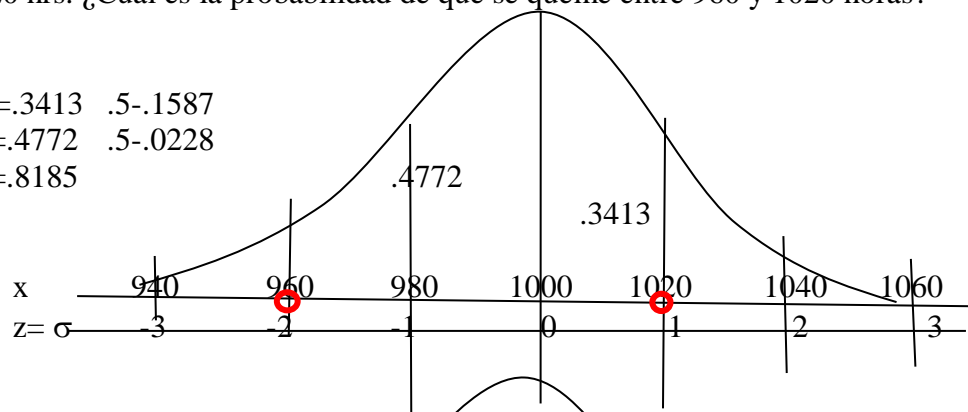
$$P(1000 > x > 960) = P(0 > z > -2) = .4772 \quad .5 - .0228$$

$$P(960 < x < 1020) = P(-2 < z < 1) = .8185$$

la prob de q se quemé

la bombilla entre

960 y 1020 hrs es 81.85%.



Una computadora cuya duración en uso, se distribuye normalmente con una media de 48 meses y una desviación estándar de 4 meses. ¿Cual es la probabilidad de que se descomponga entre 38 y 42 meses?

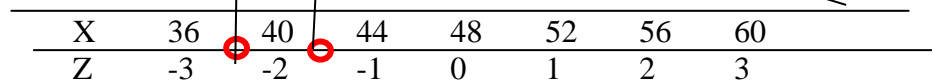
$$Z_{38} = (38 - 48) / 4 = -2.5$$

$$p(-2.5) = 0.0062$$

$$Z_{42} = (42 - 48) / 4 = -1.5$$

$$p(-1.5) = 0.0668$$

$$p(38 < x < 42) = 0.0606$$



Supongamos que se tomaron 120 lecturas de la densidad relativa del cobre, obteniéndose una media de 8.6 y una desviación estándar de 0.3.

a) ¿Cuántas lecturas tendrán una densidad mayor de 9.3? b) ¿Cuántas tendrán una densidad entre 8.3 y 8.9?

c) Si se selecciona al azar una de estas lecturas, ¿cuál es la probabilidad de que su densidad sea inferior a 8.2?

$$\mu = 8.6, \sigma = 0.3$$

a) ¿Cuántas tendrán una densidad mayor de 9.3?.,

$$Z_{9.3} = (9.3 - 8.6) / 0.3 = 2.33$$

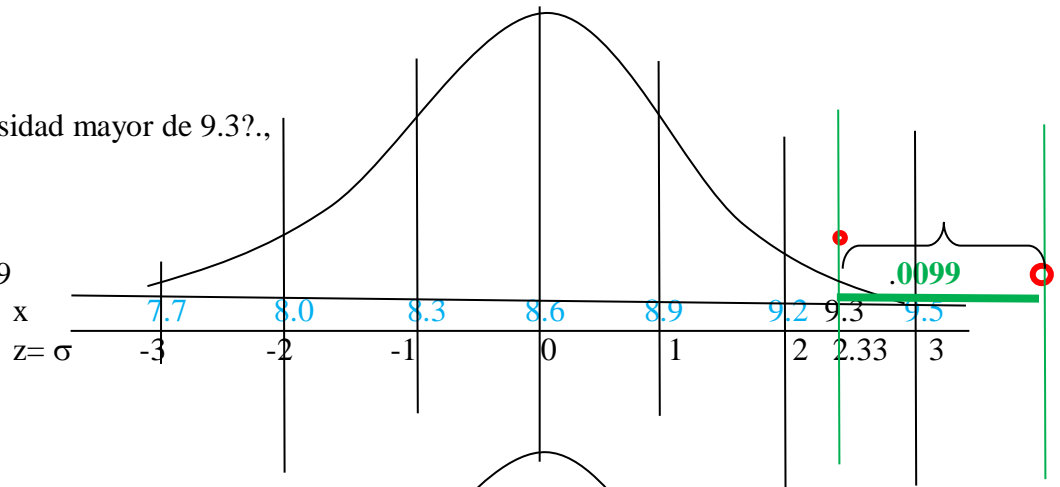
$$P(0 < z < 2.33) = 0.4901$$

$$P(x > 9.3) = p(z > 2.33) =$$

$$P(x > 9.3) = 0.5 - 0.4901 = 0.0099$$

$$Q = P(x > 9.3) * n = .0099 * 120$$

$$Q = 1.188 \text{ aprox } Q = 1 \text{ lectura}$$



$$\mu = 8.6, \sigma = 0.3$$

b) ¿Cuántas tendrán una densidad ente 8.3 y 8.9?.,

$$Z_{8.3} = (8.3 - 8.6) / 0.3 = -1; P(0 < z < -1) = .3413$$

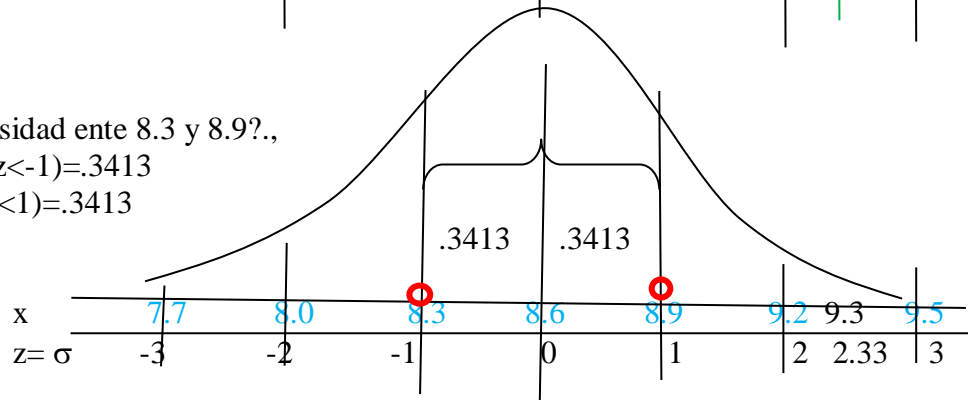
$$Z_{8.9} = (8.9 - 8.6) / 0.3 = 1; P(0 < z < 1) = .3413$$

$$P(8.3 < x < 8.9) = .6826$$

$$.5 - .1587 = .3413$$

$$Q = .6826 * 120 = 81.912$$

$$\text{aprox } Q = 82 \text{ lecturas}$$



$$\mu = 8.6, \sigma = 0.3 \text{ c)}$$

Si se selecciona al azar una de estas lecturas, ¿cuál es la probabilidad de que su densidad sea inferior a 8.2?

$$Z_{8.2} = (8.2 - 8.6) / 0.3 = -1.3333$$

$$P(Z < -1.3333) = 0.0918$$

R: es de 9.18%

Se toma una muestra aleatoria de tamaño  $n = 16$ , para analizar un proceso y se obtiene que la media de la muestra es  $\bar{x} = 2.5$ . Si la desviación estándar del proceso es conocido y se sabe que es igual a 0.4; x

a) hallar el intervalo de confianza para la media del proceso, con un nivel de confianza del 90%.

$$\mu = 2.5, \sigma = 0.4$$

$$-Z = (x - \mu) / (\sigma / \sqrt{n}) = Z$$

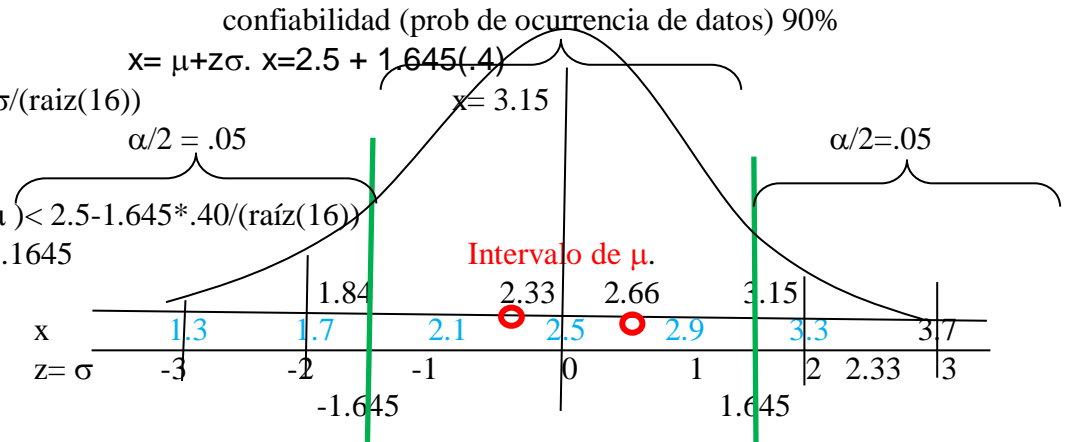
$$X - Z(\sigma / \sqrt{n}) < \mu < X + Z(\sigma / \sqrt{n})$$

$$\alpha/2 = .05$$

$$2.5 - 1.645 \cdot 0.4 / (\sqrt{16}) < \mu < 2.5 + 1.645 \cdot 0.4 / (\sqrt{16})$$

$$2.5 - 0.1645 < \mu < 2.5 + 0.1645$$

$$2.3355 < \mu < 2.6645$$



Para aplicar el método de la transformada inversa se requiere de la determinación de la función acumulada  $F(x)$ , como no es posible expresarla explícitamente, es decir calcular la integral de  $f(x)$ , no es posible utilizar este método.

Para lograr el generador se usa el teorema del límite central, que establece que la suma de  $n$  variables aleatorias independientes se aproxima a una **distribución normal** a medida que  $n$  se aproxima al infinito.

En forma de teorema: Si  $x_1, x_2, x_3, \dots, x_n$  es una secuencia de  $n$  variables aleatorias independientes con  $E(x_i) = \mu_i$  y  $\text{var}(x_i) = \sigma_i^2$  (ambas finitas) y  $Y = a_1x_1 + a_2x_2 + \dots + a_nx_n$

Entonces, bajo ciertas condiciones:

$Z = [Y - \sum a_i \mu_i] / \sqrt{(\sum a_i^2 \sigma_i^2)}$  tiene una distribución normal a medida que  $n$  se aproxima al infinito.

Si las variables que se están sumando son uniformes en el intervalo (0,1), entonces

$$Z = \left( \sum R_i - n/2 \right) / \sqrt{n/12} \text{ tiene una distribución normal estandar.}$$

Como la variable aleatoria  $x$  distribuida normalmente se obtiene como:

$$Z = (x - \mu) / \sigma, \text{ entonces la simulación de la variable aleatoria } x \text{ se haría con la siguiente expresión.}$$

$x = \mu + \sigma \left[ \left( \sum R_i - n/2 \right) / \sqrt{n/12} \right]$ , se ha comprobado que con  $n=12$  los valores simulados son bastante aceptables, simplificando la expresión a:



$$x = \mu + \sigma \left( \sum_{i=1}^{i=12} R_i - 6 \right), \text{ con } i=1 \text{ hasta } 12$$

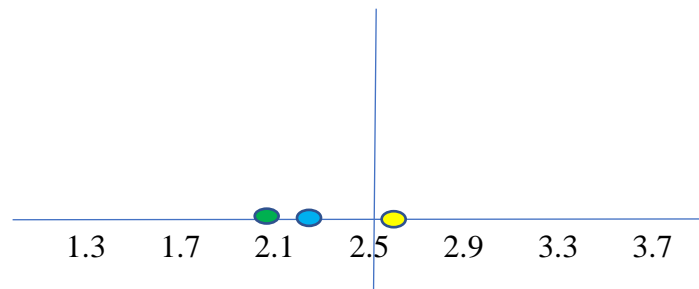
Se toma una muestra aleatoria de tamaño  $n = 16$ , para analizar un proceso que tiene comportamiento Normal y se obtiene que la media de la muestra es  $= 2.5$ . Si la desviación estándar del proceso es conocido y se sabe que es igual a  $0.4$ ; genere 3 valores que tengan este comportamiento.

R1=.44	R2=.476	R3=.309	.417	.083	.157	.82	.457	.542	.873	.12	.576
.381	.964	.47	.913	.437	.352	.188	.546	.574	.703	.31	.189
.018	.431	.634	.776	.782	.36	.189	.191	.166	.295	.43	.611

$$X_1 = 2.5 + .4(5.187 - 6) = 2.1748$$

$$X_2 = 2.5 + .4(6.027 - 6) = 2.608$$

$$X_3 = 2.5 + .4(4.883 - 6) = 2.0532$$



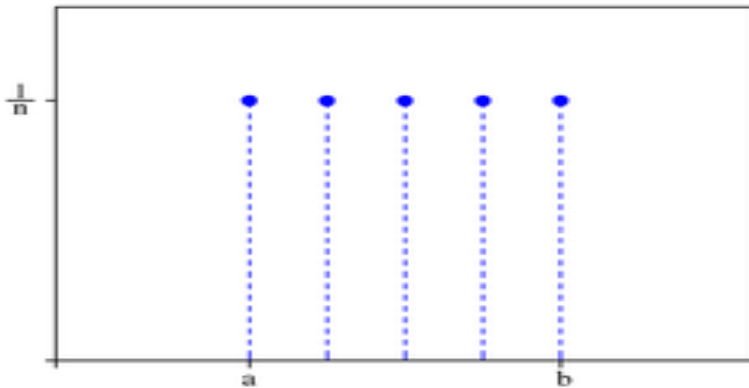
## Distribución Uniforme discreta.

Es la más simple, y se usa en fenómenos que tienen la misma probabilidad de ocurrencia. Por ejemplo, la ocurrencia de la cara de un dado, la ocurrencia de un número en una ruleta legal, etc.

La función de densidad es:

$$F(x) = \frac{1}{n} \sum_i 1_{(-\infty, x]}(x_i).$$

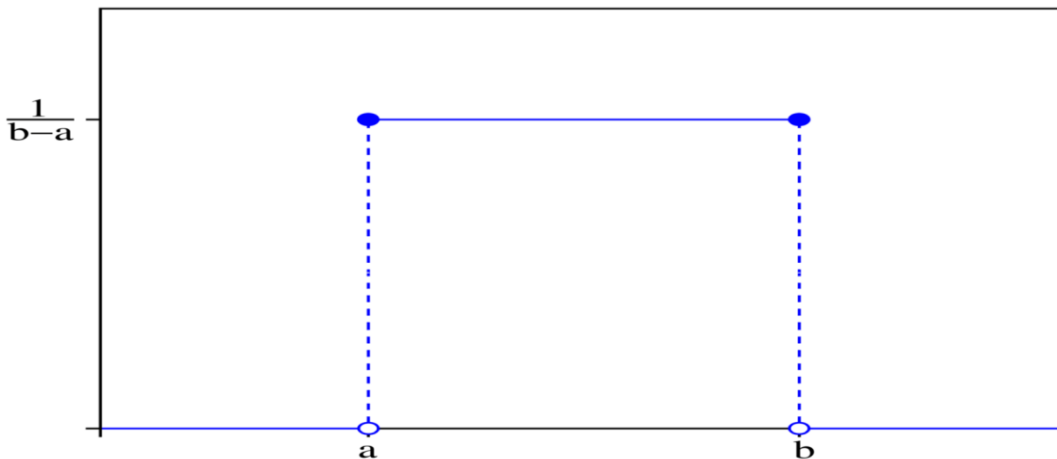
$$P(x_i) = 1/n; \quad E(x_i) = \sum (x_i) / n \quad y \quad \text{Var}(x_i) = \sigma^2 = \sum (x_i - \mu)^2 / n$$



## Distribución Uniforme o rectangular. Variable continua

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{para } a \leq x \leq b, \\ 0 & \text{para } x < a \text{ o } x > b, \end{cases}$$

$$E(x) = (a + b) / 2 \quad y \quad \text{Var}(x) = \sigma^2 = (b - a)^2 / 12$$



Ejemplo. Elabore el generador de números aleatorios que sigan una distribución uniforme.

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{para } a \leq x \leq b, \\ 0 & \text{para } x < a \text{ o } x > b, \end{cases}$$

La distribución acumulada es:

$$F(x) = \int f(x) dx = \int [1 / (b - a)] dx, \text{ con limite inferior } 0 \text{ y superior } x.$$

Integrando,  $F(x) = (x - a) / (b - a)$ , igualando con  $R$  y despejando se obtiene:

$$x = a + (b - a) R$$

### Método de Rechazo.

Este método consiste en: primero generar un valor de la variable aleatoria y enseguida probar que dicho valor simulado proviene de la distribución de probabilidad que se está analizando. La distribución debe estar acotada y con un rango finito., es decir:  $a \leq x \leq b$ .

La aplicación del método de rechazo implica el desarrollo de los siguientes pasos:

1. Generar dos números pseudoaleatorios  $R_1$  y  $R_2$ .
2. Determinar el valor de la variable aleatoria  $x$  de acuerdo a la relación lineal de  $R_1$ :

$$x = a + (b - a) R_1$$

3. Evaluar la función de probabilidad en  $x = a + (b - a) R_1$
4. Determinar si se cumple la siguiente desigualdad:

$$R_2 \leq f(a + (b - a) R_1) / M, \text{ donde } M \text{ es la moda.}$$

5. Si la respuesta es afirmativa se utiliza el valor de  $x$  como un valor simulado de la variable aleatoria, de lo contrario se rechaza. Es necesario repetir el procedimiento cuantas veces sea necesario.

Algunos autores como Tocher, que este método podría ser ineficiente para ciertas distribuciones de probabilidad en las cuales la moda sea grande.

### Método de Composición.

Mediante este método la distribución de probabilidad  $f(x)$ , se expresa como una mezcla de varias distribuciones de probabilidad  $f(x)$  seleccionadas adecuadamente.

Los pasos requeridos son:

Dividir la distribución de probabilidad original en sub-areas.

Definir la distribución de probabilidad para cada area.

Expresar la distribución de probabilidad original de la forma siguiente:

$$f(x) = A_1 f_1(x) + A_2 f_2(x) + A_3 f_3(x) + \dots + A_n f_n(x) \quad \text{y} \quad \sum A_i = 1$$

Obtener la distribución acumulada por areas.

Generar dos números aleatorios  $R_1$  y  $R_2$ .

Seleccionar la distribución de probabilidad  $f(x)$  con la cual se va a simular el valor de  $x$  ( usando la transformada inversa) usando  $R_1$ .

Usar  $R_2$  para simular el valor de  $x$  de acuerdo a la distribución y procedimiento utilizado.

Si se desea estudiar detenidamente estos dos últimos métodos consultar el libro: Simulación: Un enfoque práctico de Raul Coos Bú.