

---

## TP 2.2 - GENERADORES DE NÚMEROS PSEUDOALEATORIOS DE DISTINTAS DISTRIBUCIONES DE PROBABILIDAD

---

**Leilen Avila**  
Legajo: 41610  
Mail: leilenavila@gmail.com  
UTN - FRRO  
Zeballos 1341, S2000

**Natalia Fernandez**  
Legajo: 44758  
Mail: nata.fernandez77@gmail.com  
UTN - FRRO  
Zeballos 1341, S2000

11 de Junio, 2020

### ABSTRACT

Debido a que existen diversas distribuciones de probabilidad y su extremada utilidad en los temas relacionados con la simulación, en este trabajo, sustentándonos de un generador de números pseudo-aleatorios previamente testeado, generaremos diferentes distribuciones tanto continuas como discretas. Para realizarlo modificaremos los valores obtenidos el generador y, a partir de esto generaremos diferentes distribuciones. Contaremos con la ayuda del método de la transformada inversa para los casos que sean posible y, para los que no, utilizaremos otros métodos matemáticos. A cada una de estas distribuciones las describiremos en términos formales, indicando sus parámetros de entrada y sus gráficas. Compararemos las distribuciones generadas con las esperadas y las someteremos a una evaluación matemática para comprobar si, efectivamente, los resultados esperados son semejantes a los empíricos.

### 1. Introducción

Una distribución de probabilidad es un modelo teórico que describe la forma en que varían los resultados de un experimento aleatorio, es decir, nos da todas las probabilidades de todos los posibles resultados que podrían obtenerse cuando se realiza un experimento aleatorio. Se clasifican como discretas o continuas. En la distribución de probabilidad discreta está permitido tomar sólo un número limitado de valores. En la continua, llamada función de densidad, la variable que se está considerando puede tomar cualquier valor dentro de un intervalo dado. En nuestro trabajo haremos uso de 9 diferentes distribuciones, entre las cuales se encuentran distribuciones continuas y discretas.

Utilizaremos el generador de números pseudoaleatorios de Python, el cual fue previamente testeado y verificado que su generador es óptimo y funcional para nuestros objetivos. Primordialmente haremos uso del método de la transformada inversa para generar, a partir del generador de números pseudoaleatorios de Python, diferentes distribuciones de probabilidad. En caso de que no se pueda utilizar el método de la transformada inversa, procederemos a utilizar diferentes formulas para obtener las distribuciones. Una vez generadas, testaremos mediante diferentes métodos si los resultados empíricos coinciden con los esperados, sacando nuestras propias conclusiones en cada apartado. El test que le vamos a realizar a algunas distribuciones es el test de Chi-Cuadrado junto con la prueba visual mediante la representación gráfica de las diferentes muestras obtenidas. Las distribuciones que vamos a generar en este trabajo entonces son las que se listan a continuación:

- Distribución uniforme
- Distribución exponencial
- Distribución gamma
- Distribución normal
- Distribución de pascal

- Distribución binomial
- Distribución hipergeométrica
- Distribución poisson
- Distribución empírica discreta

## 2. Generación de variables aleatorias

La variabilidad de eventos y actividades se presentan a través de funciones de densidad para fenómenos continuos, y mediante distribuciones de probabilidad para fenómenos de tipo discreto. La simulación de estos eventos o actividades se realiza con la ayuda de la generación de variables aleatorias.

### 2.1. Variables aleatorias

Una variable aleatoria es una función que asocia a cada resultado del espacio muestral un número real. Los valores posibles de una variable aleatoria pueden representar los posibles resultados de un experimento aún no realizado, o los posibles valores de una cantidad cuyo valor actualmente existente es incierto. Intuitivamente, una variable aleatoria puede tomarse como una cantidad cuyo valor no es fijo pero puede tomar diferentes valores. La distribución de probabilidad de una variable aleatoria es una función que asigna a cada valor posible de dicha variable aleatoria una probabilidad. En términos formales una variable aleatoria es una función definida sobre un espacio de probabilidad.

Existen dos tipos de variables aleatorias:

- Variable aleatoria discreta. Su recorrido es un conjunto discreto.
- Variable aleatoria continua. Su recorrido es un conjunto no numerable.

Una variable aleatoria puede concebirse como un valor numérico que está afectado por el azar. Dada una variable aleatoria no es posible conocer con certeza el valor que tomará esta al ser medida o determinada, aunque sí se conoce que existe una distribución de probabilidad asociada al conjunto de valores posibles.

Para trabajar de manera sólida con variables aleatorias en general es necesario considerar un gran número de experimentos aleatorios, para su tratamiento estadístico, cuantificar los resultados de modo que se asigne un número real a cada uno de los resultados posibles del experimento. De este modo se establece una relación funcional entre elementos del espacio muestral asociado al experimento y números reales.

### 2.2. Método de la transformada inversa

Este método se utiliza cuando se desea simular variables de tipo continuo. El método utiliza la distribución acumulada  $F(x)$  de la distribución de probabilidad que se va a simular mediante integración. Ya que el rango de  $F(x)$  se encuentra en el intervalo 0 a 1, puede generarse un número aleatorio  $r_i$  para determinar el valor de la variable aleatoria cuya distribución acumulada es igual, precisamente, a  $r_i$ . La dificultad de este método radica en que algunas veces es complicado encontrar la transformada inversa. Entonces para simular variables aleatorias continuas, se puede lograr mediante la función acumulada  $F(x)$  y la generación de números pseudoaleatorios  $r_i \sim U(0, 1)$ .

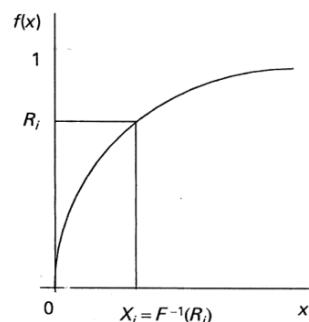


Figura 1: Método de la transformada inversa. Imagen tomada del sitio web <https://conceptosdesimulacion.wordpress.com/2016/03/10/metodo-de-transformada-inversa/> MARYDLL. (2020)

Si generamos números aleatorios uniformes correspondientes a una  $F(x)$  dada, podemos resumir, matemáticamente el método como sigue:

$$r = F(x) = \int_{-\infty}^x f(t)dt = P(X \leq x) = P(r \leq F(x)) = P(F^{-1}(r) \leq x) \quad (1)$$

y consecuentemente  $F^{-1}(r)$  es una variable aleatoria que tiene a  $f(x)$  como función de densidad de probabilidad. El método entonces consiste en:

- Definir la función de densidad  $f(x)$  en el caso continuo y la función de probabilidad puntual en el caso discreto, que representa la variable a modelar.
- Calcular la función acumulada de  $f(x)$ .
- Despejar la variable aleatoria  $x$  y obtener la función acumulada inversa  $F(x)^{-1}$ .
- Generar valores de  $x$  sustituyendo los valores generados con los números pseudoaleatorios de  $r_i \sim U(0, 1)$  en la función acumulada inversa.

### 3. Generación de variables aleatorias discretas

#### 3.1. Variables aleatorias discretas

Una variable aleatoria discreta es aquella que asociamos a experimentos aleatorios donde la variable sólo puede tomar valores enteros. Su recorrido es un conjunto finito o infinito numerable (susceptible de ser contado). Si el recorrido es un intervalo real, la variable aleatoria pasa a considerarse continua.

#### 3.2. Función de probabilidad puntual

Esta función demuestra la probabilidad de que la variable aleatoria caiga en una región específica del espacio de posibilidades estará dada por la integral de la densidad de esta variable entre uno y otro límite de dicha región. Sea  $X$  una variable aleatoria y  $\mathcal{R}_X$  su recorrido. La función que a cada valor de la variable le hace corresponder una probabilidad de que la variable  $X$  asuma ese valor, se denomina función de probabilidad puntual asociada a la variable aleatoria  $X$ . En símbolos:

$$p_x : \mathcal{R}_x \rightarrow [0, 1] \quad (2)$$

$$x \rightarrow p_x(x) = P(X = x) \quad (3)$$

El símbolo  $P(X = x)$  se lee como "probabilidad de que la variable aleatoria  $X$  asuma el valor  $x$ ". Si  $p_X$  es una función de probabilidad entonces cumple que:

1.  $p_X(x) \geq 0, \forall x \in \mathcal{R}_X$
2.  $\sum_{x \in \mathcal{R}_x} p_X(x) = 1$

A los pares ordenados  $(x, p_X(x))$  con  $x \in \mathcal{R}_X$ , como a su representación gráfica, la llamamos indistintamente distribución de probabilidad de la variable aleatoria  $X$ .

#### 3.3. Función de probabilidad acumulada

Esta función sirve para calcular la probabilidad de que una variable aleatoria asuma un valor menor o igual a un cierto valor dado. Esto motiva a la siguiente definición:

Sea  $X$  una variable aleatoria discreta con recorrido  $\mathcal{R}_x$  y su función de probabilidad  $p_X$ . La función  $F_X$  tal que

$$F_x : \mathbb{R} \rightarrow [0, 1] \quad (4)$$

$$x \rightarrow F_x(x) = P(X \leq x) \quad (5)$$

la denominamos función de distribución acumulada de la variable aleatoria  $X$ .

### 3.4. Estadísticos: Esperanza matemática y varianza

- **Esperanza matemática:** Sea  $X$  una variable aleatoria discreta con recorrido  $\mathcal{R}_X$  y  $p_X$  su función de probabilidad asociada. Definimos la esperanza matemática de la variable aleatoria  $X$  o media poblacional de la variable aleatoria  $X$  y la notamos indistintamente  $E(X)$  o  $\mu_X$  al numero:

$$E(X) = \mu_X = \sum_{x \in \mathcal{R}_x} x * p_X(x) \quad (6)$$

Si  $\mathcal{R}_X$  es infinito numerable, la suma de la definición es una serie numérica. En tal caso, la esperanza de la variable existe si la serie es absolutamente convergente.

- **Varianza:** Sea  $X$  una variable aleatoria discreta con recorrido  $\mathcal{R}_X$  y  $p_X$  su función de probabilidad asociada. Se define varianza de una V.A.  $X$  y la notamos indistintamente  $V(X)$  o  $\sigma_X^2$  al numero:

$$V(X) = \sigma_X^2 = \sum_{x \in \mathcal{R}_x} (x - \mu_X)^2 p_X(x) \quad (7)$$

### 3.5. Distribuciones discretas

Una distribución discreta describe la probabilidad de ocurrencia de cada valor de una variable aleatoria discreta. Una variable aleatoria discreta, como bien se dijo, es una variable aleatoria que tiene valores contables, tales como una lista de enteros no negativos. Con una distribución de probabilidad discreta, cada valor posible de la variable aleatoria discreta puede estar asociado con una probabilidad distinta de cero. Por lo tanto, una distribución de probabilidad discreta suele representarse en forma tabular.

A continuación pasaremos a generar diferentes distribuciones discretas, para esto vamos a proveernos del generador de números pseudoaleatorio que tiene una distribución uniforme (0,1). Con este generador y con distintos métodos matemáticos intentaremos aproximarnos a la obtención de diferentes distribuciones discretas. Compararemos estas distribuciones empíricas con la esperada (real) para cada una. Mediante representaciones gráficas y mediante la implementación de un test evaluaremos si los resultados generados son óptimos o no, es decir si se asemejan entre si.

Comenzaremos haciendo una introducción sobre cada una de las diferentes distribuciones, definiendo sus respectivas funciones de probabilidad puntuales y acumuladas asociadas y marcando los métodos utilizados para convertir el generador de números pseudoaleatorios en un generador de números con determinada distribución.

### 3.6. La distribución de Pascal

Esta distribución, también conocida como distribución Binomial negativa, es una generalización de la distribución geométrica. La variable con distribución geométrica se define como "cantidad de ensayos independientes que se realizan hasta que el suceso  $A$  se presenta por primera vez"; en el caso de una distribución binomial negativa, la variable se define como "cantidad de ensayos que se realizan hasta que el suceso  $A$  se presenta por  $r$ -ésima vez".

Una distribución de Pascal de parámetros " $r$ " y " $p$ " surge como una secuencia infinita de intentos de tipo Bernoulli en los que:

- Cada secuencia es independiente de las otras.
- En cada intento solamente son posibles dos resultados (éxito o fracaso).
- La probabilidad de éxito es constante en cada secuencia.
- Los intentos continúan hasta que se consigan  $r$  éxitos.

Entonces, formalizando, sea  $X$  una variable aleatoria definida como "cantidad de ensayos independientes de Bernoulli que se realizan hasta que el suceso  $A$  se presenta por  $r$ -ésima vez" y sea  $P(A) = p$ . Entonces  $X$  tiene una distribución de Pascal con parámetros  $p$  y  $k$  ( $r = k + x$ ). Donde  $p$  representa la probabilidad de que el suceso  $A$  ocurra. Entonces sea  $X \sim NB(k, p)$ , su función de probabilidad puntual queda determinada por:

$$P(X = k) = \binom{k+x-1}{x} p^k q^x \quad x = 0, 1, 2, \dots \quad (8)$$

donde  $k$  es el número total de éxitos en una sucesión de  $r = k + x$  ensayos, con  $x$  el número de fallas que ocurren antes de obtener  $k$  éxitos. Además, el valor esperado y la varianza de  $X$  se encuentran determinados por:

$$E(X) = \frac{kq}{p} \quad V(X) = \frac{kq}{p^2} \quad (9)$$

La función de probabilidad puntual de esta distribución se encuentra graficada debajo. En la siguiente sección vamos a generar, a partir de una fórmula matemática, números pseudoaleatorios que deberían corresponder y generar una función de probabilidad similar a la mostrada recientemente.

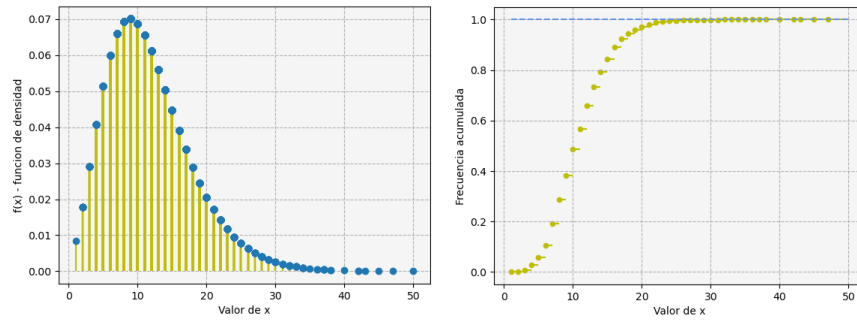


Figura 2: Distribución de Pascal.

### 3.6.1. Generación de la distribución de Pascal

Para generar esta distribución a partir de el generador de números pseudoaleatorios no vamos a utilizar el método de la transformada inversa de forma directa. La razón de esto es que, cuando  $k$  es un entero, los valores de la variable aleatoria con distribución de Pascal se pueden generar con solo considerar la suma de  $k$  valores con distribución geométrica. Sabiendo entonces que la transformada inversa de la distribución geométrica presenta la siguiente fórmula:  $x = \frac{\log r}{\log q}$ . Entonces el valor de  $x$  con distribución de Pascal se puede determinar por:

$$x = \frac{\sum_{i=1}^k \log r_i}{\log q} = \frac{\log (\prod_{i=1}^k r_i)}{\log q} \quad (10)$$

Este valor generado tiene que ser redondeado con respecto al menor entero mas próximo al valor calculado. Utilizaremos esta fórmula para generar una muestra de tamaño  $n$ , cuya gráfica debería tender a ser similar a la presentada anteriormente. Para esto ejecutaremos el fragmento de código que se encuentra en la figura siguiente un conjunto de  $n$  veces:

```

1 from math import log, exp, factorial
2 from generator import nro_random
3
4 def pascal(k, p):
5     x = 0
6     producto = 1
7     for i in range(k):
8         r = nro_random()           # Genera nro
9         producto *= r              # pseudoaleatorio
10        x = log(producto) / log(1 - p)  # T.Inv
11        indirecta
12    return int(x)                   # Redondeo al
13    entero mas proximo

```

Listing 1: Distribución Binomial



Figura 3: Diag. de flujo.

Procederemos a graficar el código mostrado anteriormente para  $k = 5$  y  $p = 0,3$  con un  $n = 5,000$ . La primer figura corresponde a un diagrama de barras en donde se representan la cantidad de veces que ocurrió cada valor en la muestra. La segunda figura nos indica el valor de la función de probabilidad puntual de cada valor que ocurrió en la muestra. Y la tercer figura muestra la frecuencia acumulada de todos los valores generados.

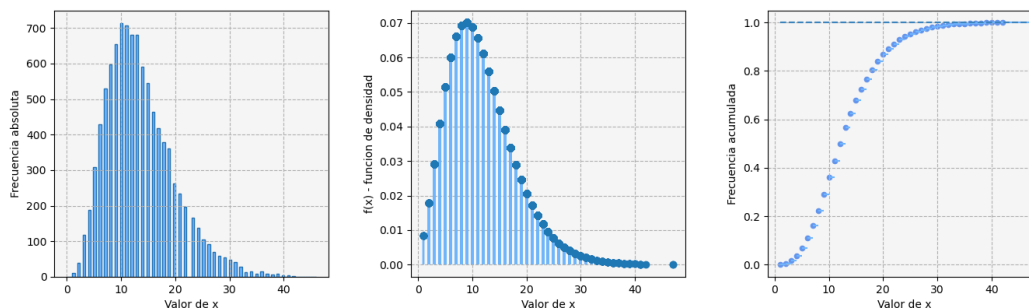


Figura 4: Muestra de la Dist. de pascal generada.

La muestra generada aparenta asemejarse a lo esperado. Para tener un mayor nivel de precisión graficaremos varias muestras pero cada una con diferentes parámetros de la distribución de probabilidad. De esta forma podremos ver las semejanzas y la certeza con respecto a diferentes variaciones.

También se presenta la Función de probabilidad acumulada de cada una de las variables aleatorias. Aclaración: La primera, segunda y tercer curva corresponden al valor de  $k = 10, 20$  y  $70$  respectivamente, con  $p = 0.5$ . Puede verse como efectivamente las muestras aparentan asemejarse de una forma correcta a lo esperado. A continuación se presenta la función de probabilidad puntual y la función acumulada de las 3 variables aleatorias con diferentes parámetros.

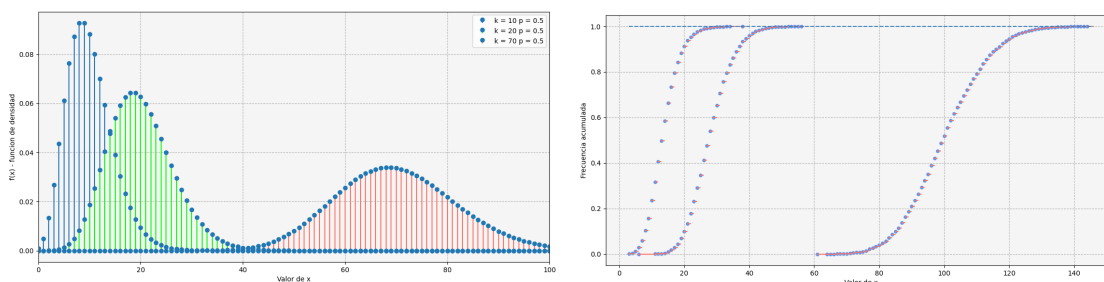


Figura 5: Distribución de Pascal.

### 3.7. La distribución binomial

Para definir esta distribución, previamente deberemos llamar ensayo de Bernoulli a toda experiencia aleatoria en relación a la cual observamos si ocurre el suceso  $A$ , o si ocurre  $\bar{A}$ . La probabilidad del suceso  $A$  es  $P(A)$ . Realizamos  $n$  repeticiones independientes del ensayo bajo el supuesto de que  $P(A)$  permanece constante en cada ensayo. En tales condiciones consideremos la variable aleatoria  $Y$ : "cantidad de veces que ocurre el suceso  $A$  en  $n$  repeticiones independientes". Para la variable aleatoria así definida resulta  $\mathcal{R}_Y = 0, 1, 2, \dots, n$ .

Si al realizar  $n$  repeticiones de un ensayo de Bernoulli el suceso  $A$  ocurre en los primeros  $k$  ensayos y  $\bar{A}$  en los  $n - k$  ensayos restantes, es decir cuando ocurre la secuencia  $AAA\dots A$   $k$  veces y  $\bar{A}\bar{A}\bar{A}\dots\bar{A}$   $n - k$  veces y estas en conjunto suman  $n$  veces. Entonces en este caso la variable aleatoria  $Y$  asume el valor  $k$ . La reciproca no es cierta. La variable aleatoria puede tomar el valor  $k$ , pero no darse necesariamente esta secuencia.

Podrían haberse dado cualquier otra combinación de dichos sucesos. Cualquiera de estas secuencias tiene una probabilidad de ocurrencia igual a  $p^k(1 - p)^{n-k}$ . Para determinar la probabilidad  $P(Y = k)$  necesitamos conocer de cuentas formas diferentes se puede presentar  $k$  veces el suceso  $A$  en  $n$  ensayos independientes de Bernoulli.

Usando un análisis combinatorio podemos deducir que la respuesta es  $\binom{n}{k}$ . Además las diferentes secuencias son mutuamente excluyentes (no pueden presentarse simultáneamente dos secuencias diferentes). Entonces formalmente:

Sea A un suceso con probabilidad p. Entonces, la variable aleatoria X: "cantidad de veces que se presenta el suceso A en n repeticiones independientes" tiene comportamiento o distribución binomial con parámetros n y p. En símbolos  $X \sim Bi(n, p)$  es decir:

$$p_X(x) = P(X = x) = \binom{n}{x} p^x q^{n-x} \quad (11)$$

Donde x se toma como un entero definido en el intervalo finito 0,1,2,...,n y al que se le asocia el valor  $q = (1 - p)$ . El valor esperado y la varianza de la variable binomial X queda determinado por:

$$E(X) = np \quad V(X) = npq \quad (12)$$

La función de probabilidad puntual y la respectiva función acumulada de esta distribución se encuentra graficada debajo. En la siguiente sección vamos a generar, a partir de una formula matemática, números pseudoaleatorios que deberían corresponder y generar una función de probabilidad similar a la mostrada recientemente.

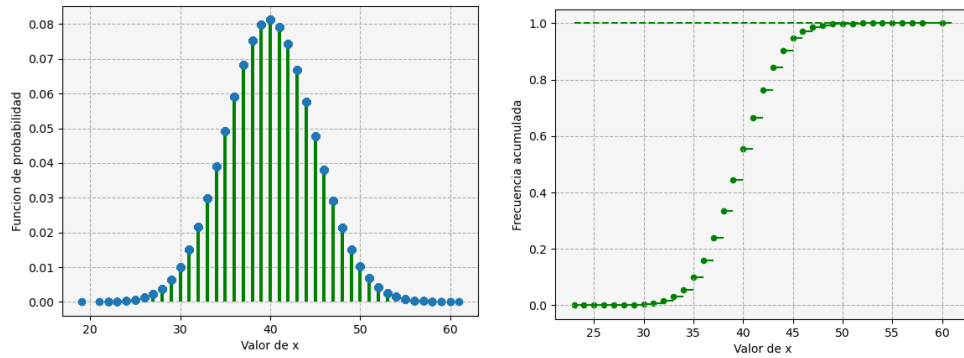


Figura 6: Función acumulada Binomial.

Ahora graficamos diferentes variables aleatorias de la misma distribución, pero variando los parámetros n y p. Estas gráficas serán las que en la sección siguiente desearemos generar mediante métodos matemáticos. Entonces, sus funciones de probabilidad puntual:

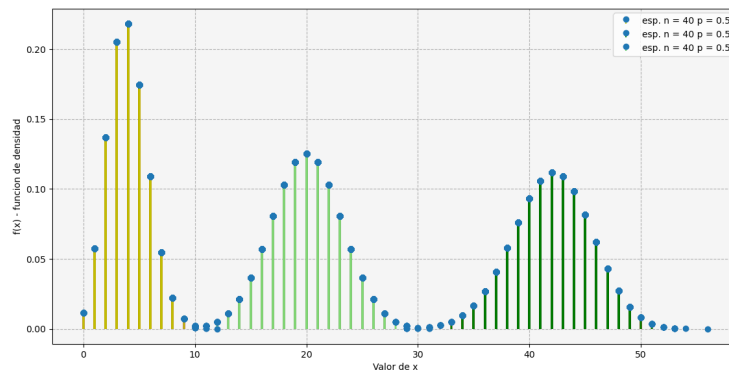


Figura 7: Múltiples muestras de dist. Binomial.

### 3.7.1. Generación de la distribución binomial

Puesto que una variable  $X \sim Bi(n, p)$  es la repetición de  $n$  experimentos de Bernoulli de probabilidad de éxito  $p$ , los valores de la variable aleatoria con distribución Binomial se pueden generar de muy diversos modos, aunque uno de los métodos mas simples, que en el caso de que el valor de  $n$  sea moderado resulta uno de los métodos mas eficientes, es el basado en la reproducción de ensayos de Bernoulli, siguiente el método de rechazos.

Este método empieza con valores conocidos de  $p$  y de  $n$  y consiste en generar  $n$  números aleatorios después de fijar  $x_0 = 0$ .

Para cada numero aleatorio  $r_i$  ( $1 \leq i \leq n$ ) se efectúa una prueba y la variable  $x_i$  se incrementa de acuerdo con el siguiente criterio:

$$x_i = x_{i-1} + 1 \quad \text{si } r_i \leq p \quad (13)$$

$$x_i = x_{i-1} \quad \text{si } r_i > p \quad (14)$$

Después de haberse generado  $n$  números aleatorios, el valor de  $x_n$  sera igual al valor de la variable aleatoria con distribución binomial  $x$ . Este procedimiento se puede repetir tantas veces como valores binomiales se requieran. Por lo tanto, podemos generar valores de esta distribución del siguiente modo:

1. Hacer  $x = 0, i = 1$ .
2. Si  $i \leq n$  ir al paso 3. En caso contrario ir al paso 5
3. Generar un numero aleatorio  $r_i$ . Si  $r_i \leq p$ , hacer  $x = x + 1$ .
4. Hacer  $i = i + 1$ . Ir al paso 2
5.  $x$  es el valor generado de la variable  $X$

Esto codificado en Python genera una función de la siguiente forma:

```
1 from generator import nro_random
2
3 def binomial(n, p): # n y p son par metros de la
4     x = 0           # distribucion
5     for i in range(n):
6         r = nro_random() # Genera nro pseudoaleatorio
7         if r <= p:
8             x += 1
9     return x
```

Listing 2: Distribución Binomial

Gráficamente corriendo el código anterior para varias muestras con diferentes parámetros (similares a los de la imagen anterior) se obtuvo un resultado bastante bueno, ya que no se pueden distinguir y no se muestran diferencias entre las funciones esperadas y las generadas. Hay que aclarar que incrementar el numero total de la cantidad de números generados no influye en que tan bueno sea el método de generación de variables aleatorias. Ya que para cualquier cantidad debería asemejarse a lo esperado.

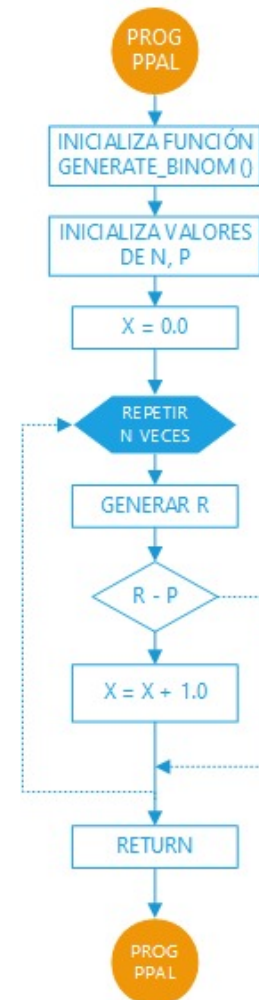


Figura 8: Diag. de flujo.



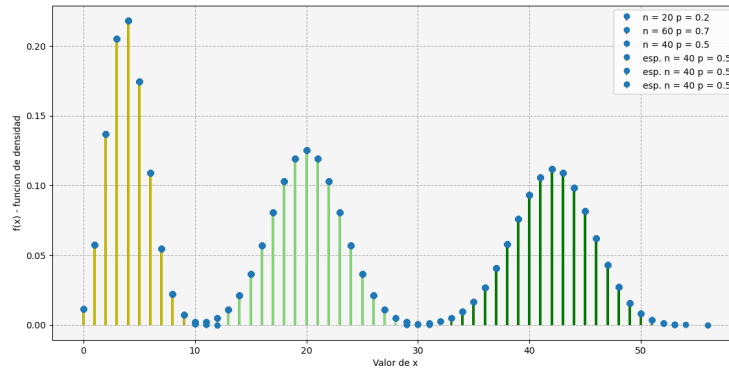


Figura 9: Múltiples muestras de dist. Binomial.

Igualmente, puesto que una variable  $X \sim Bi(n, p)$  es la suma de  $n$  variables aleatorias independientes  $B(p)$ , se podrían generar  $n$  valores independientes de una distribución  $B(p)$ , y posteriormente sumar tales valores. En cualquier caso, el método anterior requiere la generación de  $n$  números aleatorios y  $n$  comparaciones. Por ello, en general, es más eficiente la aplicación del método de inversión de la función de distribución teniendo en cuenta la siguiente relación recursiva que verifica la función puntual de probabilidad de un variable  $Bi(n, p)$ :

$$P(X = i + 1) = \frac{(n - i)p}{(i + 1)(1 - p)} P(X = i) \quad (15)$$

Entonces los pasos a seguir para generar un número pseudoaleatorio con esta distribución resulta:

1. Generar un número aleatorio  $r$ .
2. Hacer  $x = 0$ ,  $P = F = (1 - p)^n$ .
3. Si  $r < F$ , ir al paso 5.
4. Hacer  $P = \frac{(n-i)p}{(i+1)(1-p)} P$ ,  $F = F + p$ ,  $x = x + 1$ . Ir al paso 3
5.  $x$  es el valor generado de  $X$ .

El código en Python correspondiente a los pasos mencionados anteriormente se encuentra definido a continuación:

```

1 from generator import nro_random
2
3 def binomial(n, p):
4     r = nro_random()
5     x = 0
6     pr = (1 - p) ** n
7     F = pr
8     # Por el metodo de la transf. inversa
9     while r >= F:
10         pr *= ((p / (1 - p)) * (n - x) / (x + 1))
11         F += pr
12         x += 1
13     return x

```

Listing 3: Distribución Binomial

Generaremos variables aleatorias (variando los parámetros de la distribución), utilizando el método descrito anteriormente, que conformaran cada una una muestra de tamaño  $n$ . Estas muestras son graficadas a continuación.

Función de probabilidad puntual y acumulada:

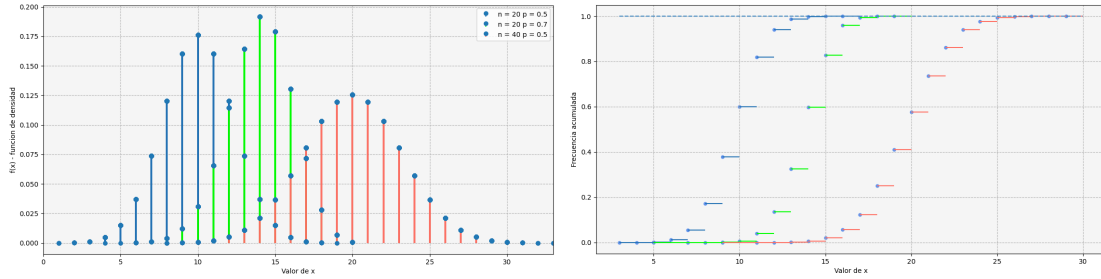


Figura 10: Funciones de la dist. Binomial.

Puede verse su fundamental similitud con la figura mostrada en la introducción de esta distribución. Se puede decir que el método utilizado (transformada inversa) para generar variables aleatorias con distribución Binomial resulta óptimo.

A continuación procedemos a evaluar una de estas muestras con el test de Chi-cuadrado para corroborar estos resultados (previamente habiendo incrementado el numero de iteraciones). La muestra a evaluar sera la siguiente:

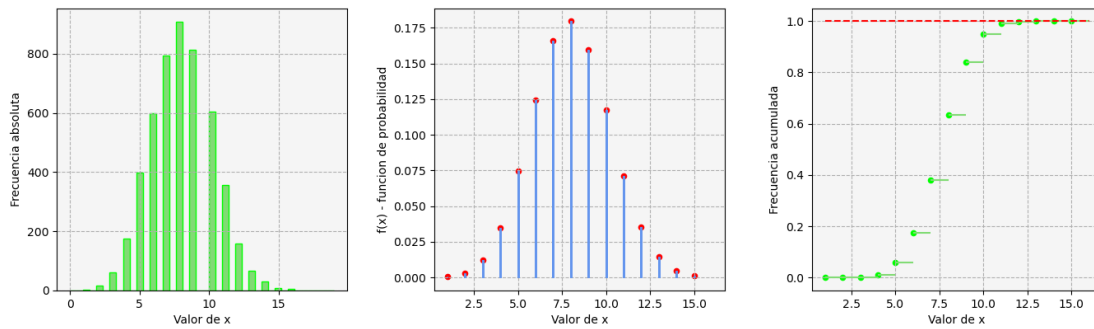


Figura 11: Muestra de distr. Binomial.

En el gráfico se presenta un diagrama de barras donde se visualiza el numero de ocurrencias de cada valor de x en la muestra. Luego se presenta la función de densidad correspondiente a cada valor obtenido en la muestra y finalmente se visualiza la frecuencia acumulada de dicha muestra.

Evaluéndola con el test Chi-Cuadrado se obtiene que esta muestra efectivamente pasó el test por presentar una eventual similitud con la curva de distribución esperada.

DISTRIBUCIÓN	CHICUADRADO	¿PASÓ EL TEST?
BINOMIAL	Val.Chi(19.45) < Val.table(25.0)	PASÓ

### 3.8. La distribución hipergeométrica

La distribución hipergeométrica es una distribución discreta que modela el número de eventos en una muestra de tamaño fijo cuando se conoce el número total de elementos en la población de la cual proviene la muestra. Cada elemento de la muestra tiene dos resultados posibles (es un evento o un no evento).

Las muestras no tienen reemplazo, por lo que cada elemento de la muestra es diferente. Cuando se elige un elemento de la población, no se puede volver a elegir. Por lo tanto, la probabilidad de que un elemento sea seleccionado aumenta con cada ensayo, presuponiendo que aún no haya sido seleccionado. La distribución hipergeométrica esta dada por la siguiente función de probabilidad:

$$p_X(x) = P(X = x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}} \quad (16)$$

Donde:

- $n$ : Representa que se eligen al azar y sin reposición  $n$  cantidad de elementos. Esto es llamado una muestra de tamaño  $n$ .
- $N$ : Representa la cantidad total de elementos de donde se obtiene la muestra de tamaño  $n$ .
- $r$ : Simboliza la cantidad de éxitos.
- $N - r$ : Simboliza la cantidad de fracasos.

En símbolos, con sus parámetros de entrada, se representa como  $X \sim HG(N, r, n)$ . Se tiene entonces que el valor esperado y la varianza se caracterizan como sigue:

$$E(X) = np \quad V(X) = np(1 - p)\left(\frac{N - n}{N - 1}\right) \quad (17)$$

Si  $n$  es pequeño, con relación a  $N$  ( $n < N$ ), la probabilidad de un éxito varía muy poco de una prueba a otra, así pues, la variable, en este caso, es esencialmente binomial; en esta situación,  $N$  suele ser muy grande y los números combinatorios se vuelven prácticamente inmanejables, así pues, la probabilidades se calculan más cómodamente aproximando por las ecuaciones de una binomial con  $p = r / N$ .

Gráficamente la función de probabilidad puntual y su respectiva función acumulada quedan determinadas como se muestra en la figura que esta debajo. Esta gráfica de la distribución sera la que deberemos generar mediante métodos matemáticos y con ayuda de nuestro generador.

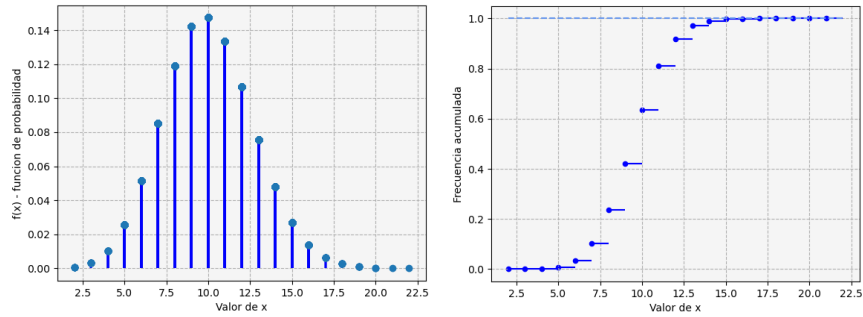


Figura 12: Muestra de distribución hipergeometrica.

Variando los parámetros la distribución hipergeometrica puede comportarse de diversos modos. La función de probabilidad puntual y la función acumulada para parámetros  $n = 3$ ,  $N = 25$ ,  $r = 3$  se obtiene que la distribución se encuentra representada como sigue:

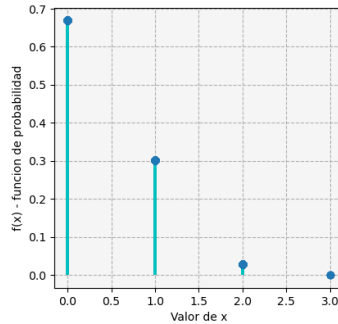


Figura 13: Muestra de distribución hipergeometrica.

### 3.8.1. Generación de la distribución hipergeométrica

La generación de valores hipergeométricos involucra la simulación de experimentos de muestreo sin reemplazo. Bastara sencillamente con que alteremos el método de ensayos de Bernoulli para generar valores binomiales, con objeto que  $N$  y  $p$  varíen en forma dependiente respecto al numero total de elementos que previamente se han obtenido entre la población y el numero de elementos  $r$  que se han extraído. A medida que se extrae un elemento de una muestra de  $n$  elementos, se reduce el valor de  $N = N_0$  de acuerdo a la formula:

$$N_i = N_{i-1} - 1 \quad i = 1, 2, ..n \quad (18)$$

De manera similar, el valor de  $p = p_0$  se transforma según:

$$p_i = \frac{(N_{i-1}p_{i-1}) - S}{N_{i-1} - 1} \quad (19)$$

a medida que se saca el  $i$ -ésimo elemento de la muestra de  $n$  elementos, donde  $S = 1$  cuando el elemento de muestra  $(i - 1)$  pertenece al numero de éxitos  $r$  y  $S = 0$  cuando el elemento de la muestra  $(i - 1)$  pertenece a  $N - r$  fracasos. Los valores iniciales de  $N_0$  y  $P_0$  corresponden: a  $N$  el tamaño inicial de la población y a  $p$  la proporción de la población total que consta de elementos de  $r$ . Cabe aclarar que el valor de  $x$  sera incrementado cada vez que se hace  $S = 1$  porque se están contabilizando la cantidad de porque esta representando la cantidad de elementos que pertenecen a ese suceso. Con ayuda de esto generaremos números que simularían provenir de una distribución hipergeométrica. Para esto, y resumiendo, la secuencia de pasos que vamos a implementar va a ser la siguiente:

1. Hacer  $x = 0, p = p_0, i = 0$ .
2. Si  $i < n$ , generar un numero pseudoaleatorio  $r$  sino terminar.
3. Si  $r > p$  hacer  $S = 0$  e ir al paso 5 sino ir al paso 4.
4. Hacer  $S = 1$  y  $x = x + 1$
5. Hacer  $p = \frac{(N * p - S)}{(N - 1)}$ ,  $N = N - 1, i = i + 1$  y volver al paso 2.

Ejecutando los pasos descriptos anteriormente en Python generaremos la muestra con la distribución deseada. Para esto el código en Python que vamos a utilizar va a ser el que se encuentra a continuación, al cual se le agrego una validación previa para verificar que el numero que se genere no sea mayor al tamaño de la muestra, ya que esto resulta imposible.

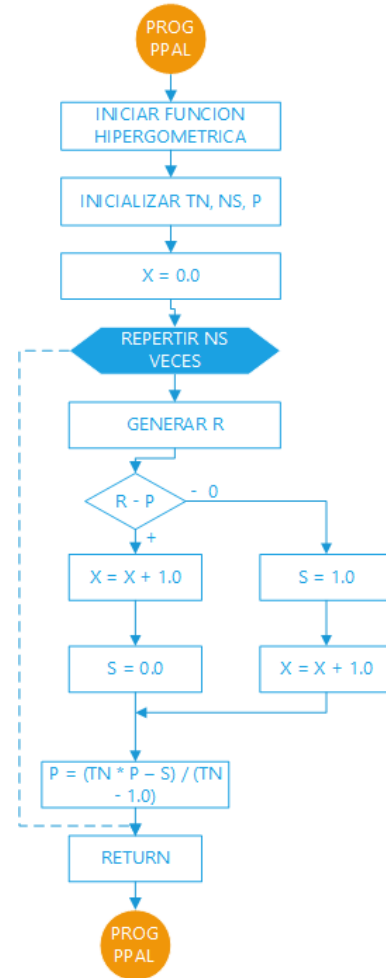


Figura 14: Diag. de flujo.

```

1 def hipergeometrica(p, N, n):
2     #Tamaño N, se toma una muestra de tamaño n
3     x = n + 1
4     while x > n: #x no puede ser mayor al tamaño de la muestra
5         x = 0
6         for i in range(n):
7             r = nro_random()
8             if r > p:
9                 s = 0
10            else:
11                s = 1
12                x += 1
13            p = (N * p - s) / (N - 1)
14    return x
  
```

Listing 4: Distribución Hipergeométrica

El diagrama de barras, la función de probabilidad puntual y función acumulada de la muestra generada con el código mencionado se encuentra a continuación:

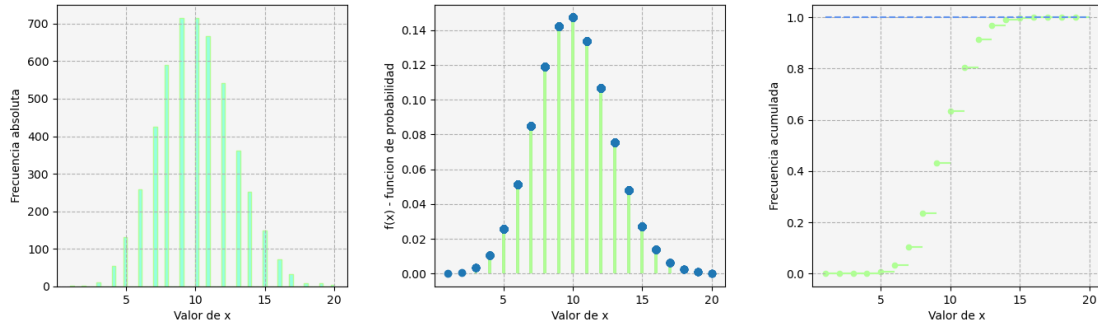


Figura 15: Muestra de distribución Hipergeometrica.

Los parámetros utilizados fueron  $n = 100$ ,  $N = 500$  y  $r = 50$ . Gracias a esto puede sacarse la probabilidad  $p$  mediante:

$$p = \frac{r}{N} = \frac{50}{500} = 0,1 \quad (20)$$

Con estos valores se puede obtener el valor de la esperanza y la varianza de la distribución, entonces se tiene que:

$$E(X) = np = 100 * 0,1 = 10 \quad V(X) = 100 * 0,1 * 0,9 \left( \frac{500 - 100}{499} \right) = 7,2144 \quad (21)$$

Se observaba como la muestra se adapta correctamente a los valores esperados. Es decir, la esperanza de la función de probabilidad coincide con la calculada previamente.

### 3.9. La distribución Poisson

Es una distribución de probabilidad discreta que expresa, a partir de una frecuencia de ocurrencia media, la probabilidad de que ocurra un determinado número de eventos durante cierto período de tiempo. Concretamente, se especializa en la probabilidad de ocurrencia de sucesos con probabilidades muy pequeñas, o sucesos raros. Fue propuesta por Siméon-Denis Poisson, que la dio a conocer en 1838.

Ejemplos de esta distribución son:

- El número de autos que pasan a través de un cierto punto en una ruta (suficientemente distantes de los semáforos) durante un periodo definido de tiempo.
- El número de errores de ortografía que uno comete al escribir una única página.
- El número de llamadas telefónicas en una central telefónica por minuto.
- El número de servidores web accedidos por minuto.

Entonces, decimos que una variable aleatoria  $X$  tiene una distribución de Poisson con parámetros  $\lambda > 0$  y notamos  $X \sim P(\lambda)$  cuando:

$$P_X(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad x = 0, 1, 2, \dots \quad (22)$$

Siendo:

- $x$  es el número de ocurrencias del evento o fenómeno (la función nos da la probabilidad de que el evento suceda precisamente  $x$  veces).
- $\lambda$  es un parámetro positivo que representa el número de veces que se espera que ocurra el fenómeno durante un intervalo dado.
- $e$  es la base de los logaritmos naturales ( $e = 2,71828\dots$ )

De la definición de esperanza y varianza de una variable aleatoria discreta, resulta para una variable aleatoria  $X$  con distribución de Poisson con parámetro  $\lambda$ :

$$E(X) = \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} \quad (23)$$

$$V(X) = \sum_{x=0}^{\infty} (x - \lambda)^2 * \frac{\lambda^x e^{-\lambda}}{x!} \quad (24)$$

De ambas formulas se puede probar que  $E(X) = V(X) = \lambda$ . El aspecto de la distribución depende muchísimo de la magnitud de la media. Como ejemplo, mostramos tres casos con  $\lambda = 0,5$ ,  $\lambda = 1,5$ ,  $\lambda = 15$  y  $\lambda = 40$ . Obsérvese que la asimetría de la distribución disminuye al crecer  $\lambda$  y que, en paralelo, la gráfica empieza a tener un aspecto acampanado.

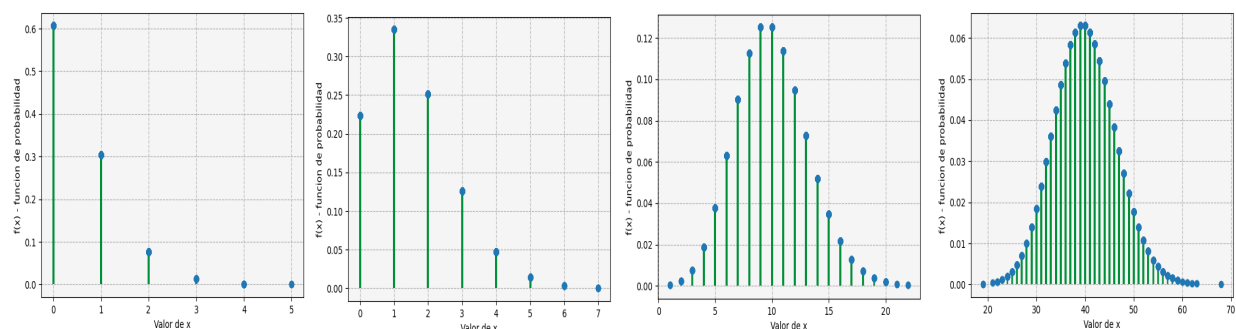


Figura 16: Distribucion de Poisson.

A su vez, para un  $\lambda = 30$  se tiene que la función de probabilidad de esta distribución junto con su función de probabilidad acumulada se encuentran determinadas abajo. Esta sera la distribución que compararemos posteriormente para evaluar nuestra generación de una variable aleatoria de Poisson.

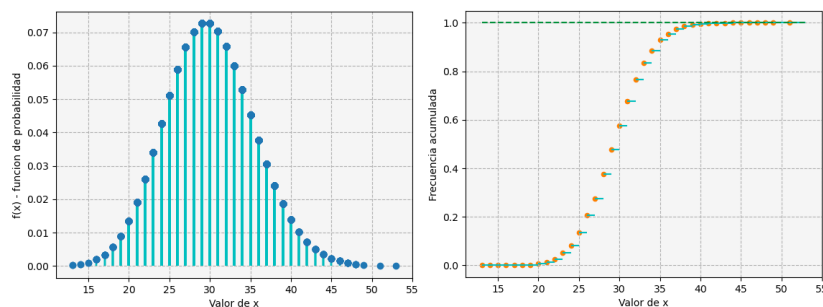


Figura 17: Distribución de Poisson para  $\lambda = 30$ .

### 3.9.1. Generación de la distribución de Poisson

La distribución de Poisson esta relacionada con la distribución exponencial, de modo que el numero de veces que ocurre un determinado suceso en un intervalo de tiempo de longitud unidad sigue una distribución de Poisson de parámetro  $\lambda$  si y solo si los tiempos entre sucesos son independientes y se distribuyen según una distribución  $E(\lambda)$ . Haciendo uso de esta propiedad, podemos generar valores de una distribución  $P(\lambda)$ . Para esto, un método para generar valores de variable aleatoria con distribución de Poisson deberá considerar la generación de intervalos  $t_1, t_2, t_3, \dots$  distribuidos en forma exponencial con un valor esperado igual a 1. Una vez generados estos intervalos aleatorios, se acumulan hasta que su suma exceda al valor de  $\lambda$ . En términos matemáticos el valor poissoniano  $x$  se determina haciendo uso de la siguiente desigualdad:

$$\sum_{i=0}^x t_i \leq \lambda \leq \sum_{i=0}^{x+1} t_i \quad (25)$$

donde los valores de una variable aleatoria  $t_i$  se generan por medio de la formula  $t_i = -\log(r_i)$  con una media unitaria. Un método mas rápido para generar valores poissonianos  $x$  es el que consiste en reformular la ecuación anterior de la siguiente manera:

$$\prod_{i=0}^x r_i \leq e^{-\lambda} \leq \prod_{i=0}^{x+1} r_i \quad (26)$$

Mediante este esquema generaremos en Python una función que nos genere la distribución de Poisson deseada. Esto se encuentra definido a continuación:

```

1 from math import exp
2
3 def poisson(lamda):
4     x = 0
5     b = exp(- lamda)
6     producto = 1
7     while producto >= b:
8         r = nro_random()      # Genera nro
9         pseudoaleatorio
10        producto *= r
11
12        if producto >= b:
13            x += 1
14        else:
15            return x

```

Listing 5: Distribución de Poisson

Sin embargo, al igual que en la distribución anterior, podemos aplicar el método de inversión de la función de distribución teniendo en cuenta la siguiente relación recursiva que verifica la función puntual de probabilidad de un variable  $Po(\lambda)$ :

$$P(X = i + 1) = \frac{\lambda}{i + 1} P(X = i) \quad (27)$$

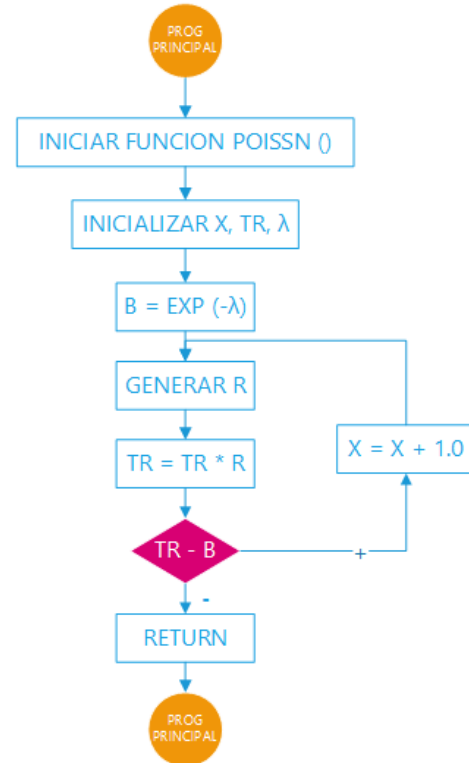


Figura 18: Diag. de flujo.

Gracias a esto, el esquema de generación será el siguiente:

1. Generar un número aleatorio  $r_i$ .
2. Hacer  $x = 0, P = F = e^{-\lambda}$ .
3. Si  $r_i < F$ , ir al paso 5.
4. Hacer  $P = \frac{\lambda}{x+1}P, F = F + p, x = x + 1$ . Ir al paso 3.
5.  $x$  es el valor generado de  $X$

El diagrama de flujo de dichos pasos se encuentra representado a continuación:

La ventaja de este esquema es que es necesario la generación de un único número aleatorio y, en promedio, serán necesarias  $1 + \lambda$  comparaciones. Basándonos en esta secuencia de pasos generamos el correspondiente código en Python de la siguiente manera:

```
1 from math import exp
2
3 def poisson(lamda):
4     r = nro_random() # Genera nro pseudoaleatorio
5     x = 0
6     p = exp(- lamda)
7     F = p
8
9     while r >= F:
10         p = (lamda * p) / (x + 1)
11         F += p
12         x += 1
13     return x
```

Listing 6: Distribución de Poisson

Graficando una muestra obtenida con el código mencionado anteriormente y con parámetro  $\lambda = 30$  se obtuvo el siguiente resultado:

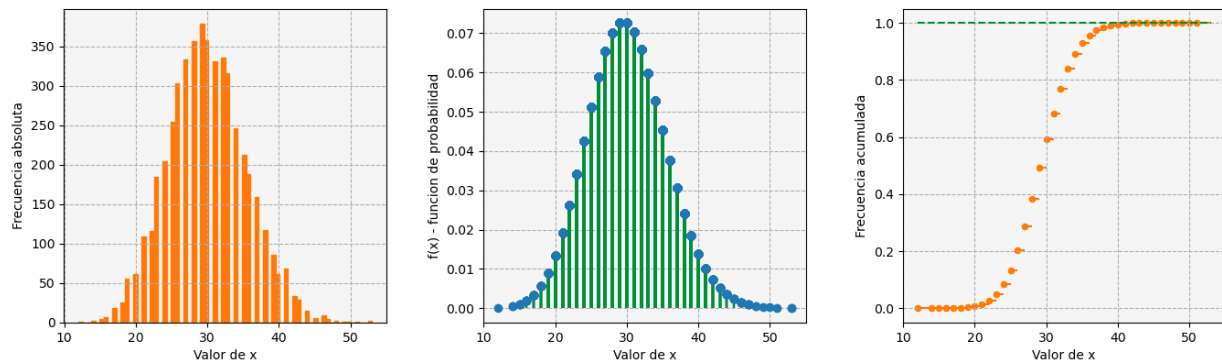


Figura 19: Distribución de Poisson para  $\lambda = 30$ .

Puede verse como el centro de la distribución se encuentra en  $\lambda = 30$  y esto es correcto debido a que la esperanza de Poisson es igual a  $\lambda$ . Además se ve rápidamente como esta gráfica tiende a ser exactamente igual a la presentada en la sección anterior. A partir de este último código también se generaron diferentes muestras variando el único parámetro que presenta la distribución. El resultado fue el siguiente:



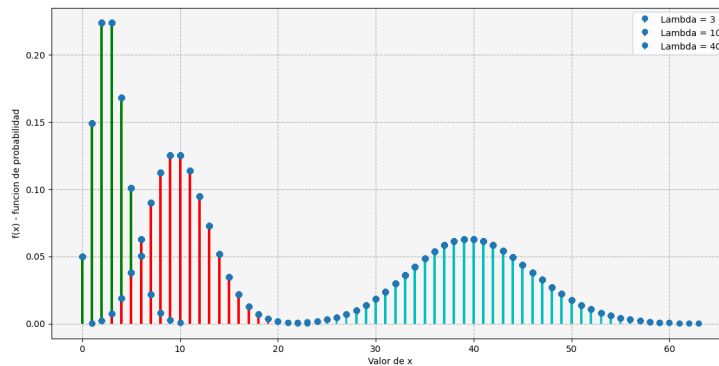


Figura 20: Muchas muestras de distribución de Poisson.

Se ve como nuevamente se adaptan las muestras generadas a los valores esperados. Cabe aclarar que podrían haberse generado múltiples corridas para un mismo valor del parámetro lamda, pero como las muestras obtenidas resultan ser todas extremadamente similares cuando  $n$  tiende a infinito (como se vio en el apartado de distribuciones continuas) preferimos representar múltiples gráficas con diferentes parámetros para tener una mejor comprensión de como se comporta la muestra en diferentes circunstancias.

Sometimos diferentes muestras al test de Chi-cuadrado y como era de esperarse (debido al previo análisis gráfico), todas las muestras han pasado la prueba sin dificultad. Los valores obtenidos de Chi-cuadrado no se encuentran forzados, es decir, no se encuentran cerca del valor máximo admisible, sino que se encuentran en un rango factible que permite considerar buenas a las respectivas muestras.

DISTRIBUCIÓN	VALOR DE $\lambda$	CHICUADRADO	¿PASÓ EL TEST?
POISSON	$\lambda = 0,4$	$Val.Chi(8,02) < Val.table(9,49)$	PASÓ
POISSON	$\lambda = 10$	$Val.Chi(24,09) < Val.table(35,17)$	PASÓ
POISSON	$\lambda = 20$	$Val.Chi(29,78) < Val.table(33,92)$	PASÓ
POISSON	$\lambda = 40$	$Val.Chi(73,1) < Val.table(89,39)$	PASÓ

### 3.10. La distribución Empírica Discreta

En algunos casos es necesario usar los datos observados directamente para especificar una distribución llamada distribución empírica en lugar de una distribución teórica. Esto se hará cuando los datos no se ajustan a ninguna distribución de probabilidad conocida. Este método puede utilizarse para simular tanto distribuciones empíricas, discretas y continuas que puedan ser aproximadas mediante una distribución discreta. Sin embargo, es recomendable utilizar este método solo cuando no exista una distribución que se ajuste a lo que nosotros queremos, ya que, como se vio, existen métodos mucho mas eficientes para generar distribuciones ya conocidas y que fueron previamente testeados.

En este apartado lo que vamos a hacer es determinar una distribución empírica que se ajuste a una distribución Binomial. En realidad como dijimos, este método se utiliza cuando no se conoce la distribución de una muestra, pero para hechos prácticos vamos a simular una determinada distribución. Podríamos haber tomado cualquier otra distribución para nuestro caso de estudio, fue mera elección al azar el uso de esta, pudiendose aplicar a cualquier otra. Previamente ya se definieron los parámetros de la distribución binomial y sus respectivas funciones de probabilidad puntual y acumulada. Es por esto que pasaremos directamente a mencionar como obtendremos una aproximación de la misma.

Los pasos a seguir van a ser:

1. Asignamos a  $n$  números su respectiva función de probabilidad puntual y sacamos la probabilidad de cada número con respecto al total.
2. Una vez que tenemos cada numero con su respectiva probabilidad, calculamos la frecuencia acumulada de dichos números y los guardamos en un arreglo.
3. Generamos un numero random  $r_i$  y recorremos el arreglo que contiene la frecuencia acumulada de cada numero. Para cada valor de la frecuencia acumulada comparamos si el mismo es mayor al numero  $r_i$  generado.

- Si el valor de la frecuencia acumulada en un numero es mayor a  $r_i$  entonces el valor que representa a esa frecuencia acumulada es el valor de  $x$  sino se continúa recorriendo el mismo hasta encontrar el valor mayor, el ultimo valor que se puede obtener será igual a la unidad.

Haciendo uso de los pasos mencionados anteriormente se procedió a programarlo en Python. El código generado se encuentra a continuación:

```

1 def discreta_empirica(xs, frec_acum):
2     r = nro_random() # Genero un nro random.
3     for i in range(len(xs)):
4         if(frec_acum[i] > r): # Si el r es < a frec.acum le corresponde ese valor a x
5             return xs[i]
6
7 def calcula_frec_acum():
8     # Elijo una distribucion cualquiera
9     # En este caso utilizamos la dist. Binomial
10    nb = 20
11    p = 0.4
12    func = []
13    probs = []
14    xs = np.arange(0, nb)
15
16    func = [f_den_binomial(x, nb, p) for x in xs]
17    probs = [f / sum(func) for f in func] # Calculo la probabilidad de cada nro
18
19    frec_acum = []
20    frec_acum.append(probs[0])
21    for i in range(1, len(xs)):
22        acumulado = frec_acum[i - 1] + probs[i]
23        frec_acum.append(acumulado) # Calculo la frec.acumulada para cada nro.
24
25    return xs, frec_acum
26
27 def generate_discreta_empirica(n):
28     xs, frec_acum = calcula_frec_acum()
29     corrida = []
30     for i in range(n):
31         x = discreta_empirica(xs, frec_acum)
32         corrida.append(x)
33     return corrida

```

Listing 7: Distribución Empírica Discreta

Nuestro objetivo como se dijo es asemejarnos a una distribución binomial. Para esto ya conocemos como debería ser la función de probabilidad puntual y acumulada esperada. Por lo tanto generamos una muestra con esta distribución a partir del método presentado anteriormente. El resultado que obtuvimos fue generado para los parámetros  $X \sim Bi(n = 20, p = 0,4)$ :

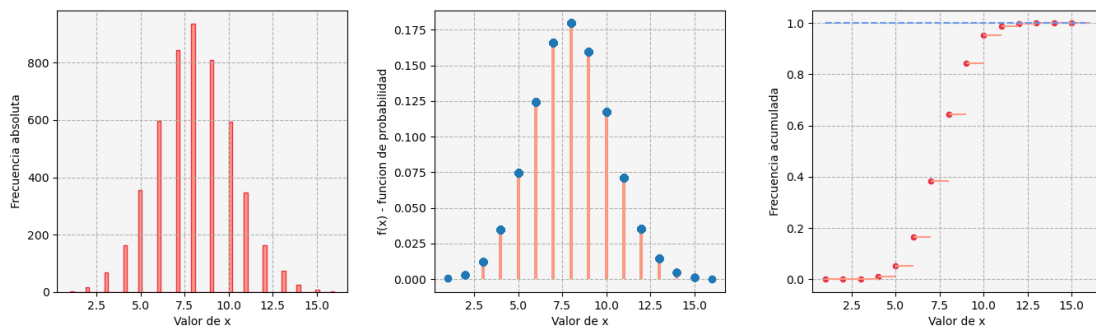


Figura 21: Distribución empírica discreta para la Binomial.

El resultado es similar a lo esperado. Consideramos esto porque las gráficas son similares a las vistas anteriormente y además la esperanza esperada debería ser:

$$E(X) = np = 20 * 0,4 = 8 \quad (28)$$

que es exactamente el valor que corresponde a la esperanza de la muestra generada. Probando para múltiples muestras el resultado obtenido vuelve a ser contundente y nos vuelve a informar que, por lo percibido, estamos ante un buen método de generación de variables aleatorias. El inconveniente de este método es que, desde el punto de vista de la velocidad de computación, facilidad de programación y requisitos de memoria otros métodos pueden hacer mejor uso de ellos, y de esta forma puede volverse ineficiente comparándolo con los demás.

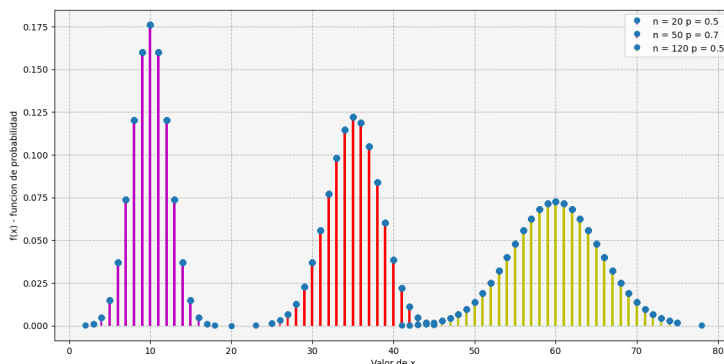


Figura 22: Distribución empírica discreta para la Binomial.

Evaluaremos ahora, mediante el test de Chi-Cuadrado, los resultados obtenidos. La respuesta de si paso o no el test se encuentra descripto a continuación para diferentes parámetros de la distribución binomial.

DISTRIBUCIÓN	VALOR DE n y p	CHICUADRADO	¿PASÓ EL TEST?
EMP.DISCRETA	$n = 20, p = 0,5$	$Val.Chi(22,84) > Val.table(23,68)$	PASÓ
EMP.DISCRETA	$n = 50, p = 0,7$	$Val.Chi(25,12) > Val.table(30,68)$	PASÓ
EMP.DISCRETA	$n = 120, p = 0,5$	$Val.Chi(68,1) < Val.table(70,19)$	PASÓ

## 4. Generación de variables aleatorias continuas

### 4.1. Variables aleatorias continuas.

Hemos introducido a las variables aleatorias como funciones que asignan a cada elemento de un espacio muestral un numero real. Cuando la variable aleatoria es continua, es decir, cuando el recorrido de la variable aleatoria es un intervalo real, matemáticamente no es factible asignar una probabilidad positiva a cada elemento del recorrido de la variable aleatoria y al mismo tiempo verificar que la suma de las probabilidades de los distintos valores sea igual a uno. Esto nos obliga a desarrollar un método diferente para describir la distribución de probabilidad para una variable aleatoria continua.

#### 4.1.1. Función de densidad de probabilidad

Como se dijo, la función de distribución describe el comportamiento probabilístico de una variable aleatoria X asociada a un experimento aleatorio. Podemos formalizar a través de la siguiente definición: Sea X una variable aleatoria continua. Decimos que  $f_X$  es la función de densidad de probabilidad asociada a X cuando se verifica

1.  $f_X(x) \geq 0, \forall x \in \mathbb{R}$
2.  $\int_{-\infty}^{\infty} f_X(x) dx = 1$
3.  $P(a \leq X \leq b) = \int_a^b f_X(x) dx$

#### 4.1.2. Función de distribución acumulada

Sea  $X$  una variable aleatoria continua y  $f_X$  su función de densidad de probabilidad. La función

$$F_x : \mathbb{R} \rightarrow [0, 1] \quad (29)$$

$$x \rightarrow F_x(x) = \int_{-\infty}^x f_X(t) \quad (30)$$

se denomina función de distribución acumulada de la variable aleatoria  $X$ .

#### 4.1.3. Estadísticos: Esperanza y Varianza

- **Esperanza matemática:** Sea  $X$  una variable aleatoria continua y  $f_X$  su función de densidad de probabilidad. La esperanza matemática o media poblacional de  $X$ , que se nota como  $E(X)$  o  $\mu_X$  es el número:

$$E(X) = \mu_X = \int_{-\infty}^{+\infty} x f_X(x) dx \quad (31)$$

siempre que la integral exista y sea finita.

- **Varianza:** Sea  $X$  una variable aleatoria continua y  $f_X$  su función de densidad de probabilidad. La varianza de  $X$  que se nota  $V(X)$  o  $\sigma_X^2$  es el número no negativo

$$V(X) = \sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f_X(x) \quad (32)$$

#### 4.2. La distribución uniforme

La distribución Uniforme es el modelo continuo más simple. Corresponde al caso de una variable aleatoria que sólo puede tomar valores comprendidos entre dos extremos  $a$  y  $b$ , de manera que todos los intervalos de una misma longitud dentro de  $(a, b)$  tienen la misma probabilidad. La distribución surge cuando se estudia las características de los errores por redondeo al registrar un conjunto de medidas sujetas a cierto nivel de precisión. También puede expresarse como el modelo probabilístico correspondiente a tomar un número al azar dentro de un intervalo  $(a, b)$ .

Los **parámetros** de entrada para esta distribución son los valores  $a$  y  $b$ , que representan los extremos de la función, siendo  $a$  el extremo inferior y  $b$  el superior. Decimos que una variable aleatoria  $X$  se distribuye uniformemente en el intervalo  $[a, b]$  y notamos  $X \sim U(a, b)$ , cuando su función de densidad de probabilidad es:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{fuera del intervalo } (a, b) \end{cases} \quad (33)$$

El hecho de que una variable aleatoria tenga un comportamiento uniforme significa que intervalos de igual amplitud contenidos en el intervalo  $[a, b]$  tienen la misma probabilidad de ocurrir. La función de distribución acumulativa  $F(x)$  para una variable aleatoria  $X$  uniformemente distribuida se representa como:

$$F(X) = \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a} \quad 0 \leq F(x) \leq 1 \quad (34)$$

Entonces resulta que la función de distribución de probabilidad se puede calcular según:

$$F(x; a, b) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } b \leq x \end{cases} \quad (35)$$

Su función de densidad de probabilidad y de distribución acumulada quedan determinadas por:

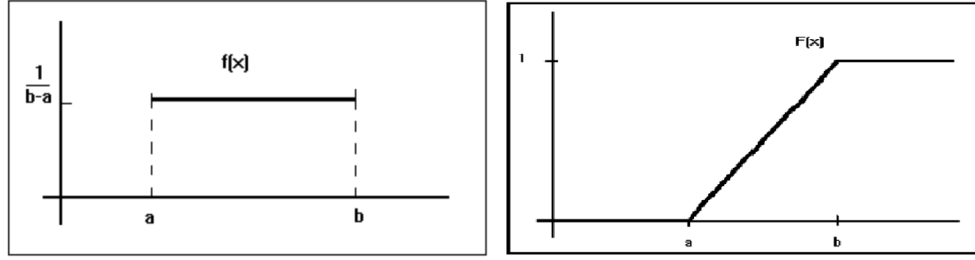


Figura 23: Funciones de la Distribución uniforme.

La esperanza de la distribución uniforme se puede calcular como:

$$E(X) = \mu_X = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2} \quad (36)$$

Y su varianza se puede calcular como

$$V(X) = \sigma_X^2 = \int_a^b \frac{(x - E(X))^2}{b-a} dx = \frac{(b-a)^2}{12} \quad (37)$$

Los parámetros de entrada para esta distribución, como dijimos previamente, son los valores  $a$  y  $b$ , que representan los extremos de la función, siendo  $a$  el extremo inferior y  $b$  el superior. En los casos típicos, aunque no siempre sucede, se suele conocer los valores de la esperanza y la varianza. En dicho caso, se puede calcular los valores de  $a$  y  $b$  como:

$$a = E(X) - \sqrt{3V(X)} \quad (38)$$

$$b = 2 * E(X) - a \quad (39)$$

#### 4.2.1. Generación de la distribución uniforme

Para simular una distribución uniforme sobre cierto intervalo conocido  $(a,b)$  deberemos, en primer lugar, obtener la transformación inversa para la ecuación de la función de distribución acumulada. Entonces la transformación inversa de dicha función queda determinada por:

$$x = a + (b-a)r \quad 0 \leq r \leq 1 \quad (40)$$

Con esto generamos un conjunto de números aleatorios correspondiente al rango de las probabilidades acumulativas, es decir, los valores de variables aleatorias uniformes definidas sobre el rango de 0 a 1. Cada número aleatorio  $r$  determina, de manera única, un valor de la variable aleatoria  $x$  uniformemente distribuida. Mediante el suministro de este método generaremos una muestra con determinados parámetros de  $a$  y  $b$  para ver que tan efectivo es el mismo. Utilizamos  $a = 0$  y  $b = 20$ . Con un  $n$  suficientemente grande podemos ver como se obtienen resultados óptimos y que corresponden con lo esperado.

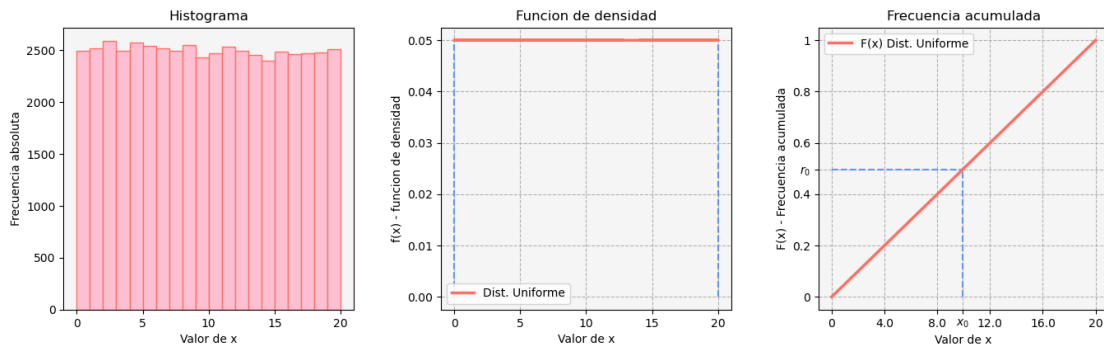


Figura 24: Muestra de la generación de la distr. uniforme.

En la gráfica se representa la correspondencia entre el valor de  $x$  generado y el numero aleatorio  $r_i$  a partir del cual se obtuvo. Es decir, muestra como cada valor generado de  $r$  esta asociado con uno y solo un valor de  $x$ . Por ejemplo, el valor específico de la función de distribución acumulativa en  $r_0$  determina el valor de  $x$  en  $x_0$ . Este procedimiento se puede repetir tantas veces como se desee y en cada repetición se generara un nuevo valor de  $x$ . De esta forma se puede comprender como actúa el método de la transformada inversa.

El código en Python utilizado para generar la muestra presentada anteriormente se encuentra definido abajo. A este fragmento de código lo iteramos  $n$  veces para generar una muestra de tamaño  $n$  de números aleatorios con distribución uniforme con parámetros  $a$  y  $b$ .

```
1 from generator import nro_random
2
3 def uniform(a, b):
4     r = nro_random() # Genera nro pseudoaleatorio
5     x = a + (b - a) * r # Transformada inversa
6     return x
```

Listing 8: Distribución Discreta

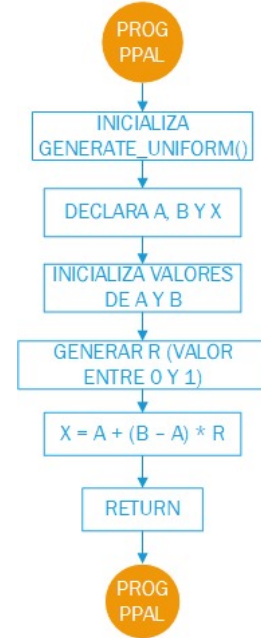


Figura 25: Diag. de flujo.

### 4.3. La distribución exponencial

Este modelo suele utilizarse para variables que describen el tiempo hasta que se produce un determinado suceso. Ejemplos de distribuciones exponenciales pueden ser el intervalo entre los accidentes en una fabrica, la llegada de pedidos a una compañía, el registro de pacientes en un hospital, el aterrizaje de aviones en un aeropuerto y más. Para los valores de variables aleatorias de tipo exponencial, se deben satisfacer las siguientes suposiciones:

1. La probabilidad de que ocurra un evento en el intervalo  $[t, (t + \Delta t)]$  es  $\alpha \Delta t$ .
2.  $\alpha$  es una constante que no depende de  $t$  o de algún otro factor.
3. La probabilidad de que durante un intervalo  $[t, (t + \Delta t)]$  ocurra mas de un evento, tiende a 0 a medida que  $\Delta t$  tiende a 0, y su orden de magnitud deberá ser menor que el de  $\alpha \Delta t$ .

Decimos por lo tanto que una variable aleatoria  $X$  tiene una distribución exponencial de parámetro  $\alpha > 0$  y notamos  $X \sim E(\alpha)$  cuando su función de densidad de probabilidad es:

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases} \quad (41)$$

Además, si  $x > 0$ , entonces su función de distribución acumulada

$$F_X(x) = \int_0^x \alpha e^{-\alpha t} dt = 1 - e^{-\alpha x} \quad (42)$$

y en consecuencia  $P(X > x) = e^{-\alpha x}$ . Haciendo uso de las ecuaciones descriptas en la introducción a las variables aleatorias continuas, la esperanza y la varianza para esta distribución son:

$$E(X) = \int_0^{\infty} x \alpha e^{-\alpha x} dx = \frac{1}{\alpha} \quad V(X) = \int_0^{\infty} (x - \frac{1}{\alpha})^2 \alpha e^{-\alpha x} dx = \frac{1}{\alpha^2} = E(X)^2 \quad (43)$$

La distribución exponencial solamente tiene el parámetro  $\alpha$ , el cual se puede expresar como  $\alpha = 1/E(X)$ . A partir de esto se tiene que gráficamente la distribución exponencial queda determinada de la siguiente manera:

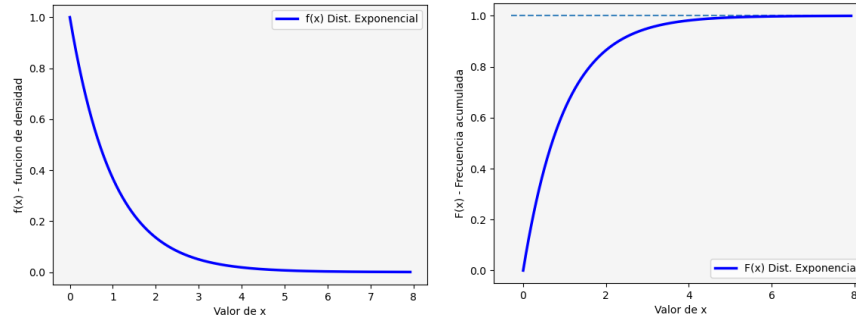


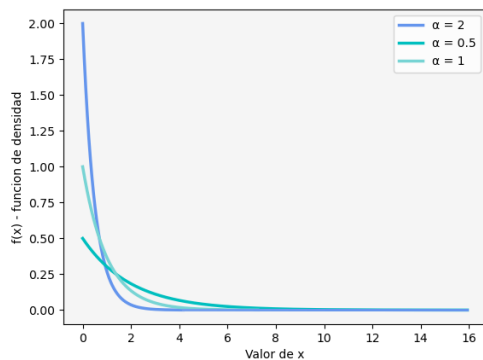
Figura 26: Distribucion exponencial.

#### 4.3.1. Generación de la distribución exponencial

Se desea simular una variable aleatoria con distribución exponencial; su función de densidad fue mencionada previamente:  $\alpha e^{-\alpha x}$  si  $x \geq 0$  y su función acumulada  $F_X(x) = 1 - e^{-\alpha x}$ . Igualando esta función acumulada al número aleatorio  $r_i$  y encontrando la transformada inversa (despejando  $x$ ) se obtiene:

$$x = -\frac{1}{\alpha} \log r_i \quad (44)$$

Por consiguiente, para cada valor del número pseudoaleatorio  $r$  se determina un único valor para  $x$ . Los valores de  $x$  toman solo magnitudes no negativas. Haciendo uso de esto generaremos diferentes muestras variando el parámetro  $\alpha$  y de esta forma obtendremos diferentes variables aleatorias. Esto se encuentra graficado a continuación:



1

Figura 28: Distribución exponencial.

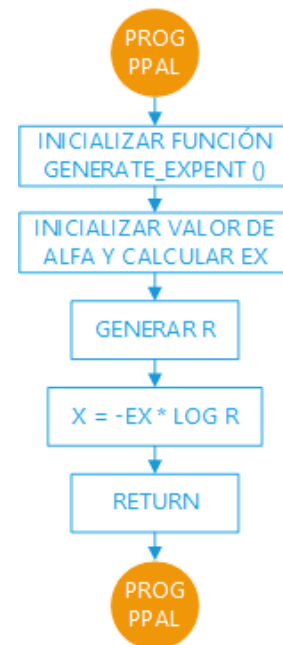


Figura 27: Diag. de flujo.

Variando los parámetros se obtuvieron resultados que reflejan como esta trabajando el método y como es capaz de generar una buena muestra que tiende a ser una distribución exponencial óptima. El código que estamos utilizando para generar los valores de las variables aleatorias con distribución exponencial se encuentra explicito a continuación.

```

1 from math import log
2 from generator import nro_random
3
4 def exponencial(alfa):
5     r = nro_random() # Genera nro pseudoaleatorio
6     x = ti_exponencial(alfa, r)
7     return x
8
9 def ti_exponencial(alfa, r):
10    e = 1 / alfa # Esperanza
11    x = - e * log( r ) # Transformada inversa
12    return x

```

Listing 9: Distribución Exponencial

Ahora tomaremos una muestra de una variable aleatoria con parametro  $\alpha = 1$  en donde mostraremos su función de distribución acumulada junto con su histograma de frecuencias. Luego tomaremos muestras similares a esta para poder evaluarlas mediante el test Chi-Cuadrado. La muestra se representa como sigue:

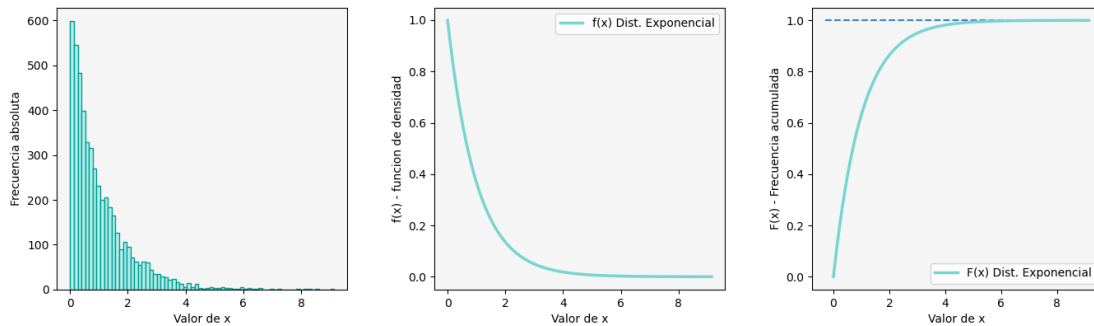


Figura 29: Distribución exponencial.

La primer figura define un histograma en donde se marcan la cantidad de veces que ocurrió el valor x, en la segunda figura se muestra la función de densidad de probabilidad para los mismos valores y finalmente en la ultima figura se muestra la función de distribución acumulada.

Sometiendo diferentes muestras al test de Chi-Cuadrado se obtiene que pasan la prueba. Esto era de esperarse debido a que las gráficas mostraban resultados que son muy similares a los resultados esperados. La tabla con el resultado del test:

DISTRIBUCIÓN	VALOR DE $\alpha$	CHICUADRADO	¿PASÓ EL TEST?
EXPONENCIAL	$\alpha = 0,1$	$Val.Chi(95,02) < Val.table(96,22)$	PASÓ
EXPONENCIAL	$\alpha = 0,4$	$Val.Chi(10,11) < Val.table(23,68)$	PASÓ
EXPONENCIAL	$\alpha = 0,7$	$Val.Chi(15,01) < Val.table(15,51)$	PASÓ
EXPONENCIAL	$\alpha = 1$	$Val.Chi(6,29) < Val.table(7,81)$	PASÓ



#### 4.4. La distribución gamma

Se trata de una familia de distribuciones que provee un modelo adecuado para histogramas que presentan cierto tipo de asimetría. Antes de presentar a las v.a. con distribución Gamma, es necesario recordar cómo se define la función Gamma o factorial, la cual cumple un rol importante en muchas ramas de la Matemática.

Como se mencionó anteriormente, es una distribución adecuada para modelizar el comportamiento de variables aleatorias continuas con asimetría positiva. Es decir, variables que presentan una mayor densidad de sucesos a la izquierda de la media que a la derecha. En su expresión se encuentran dos parámetros, siempre positivos,  $\alpha$  y  $\beta$  y de los que depende su forma y alcance por la derecha, y también la función gamma, responsable de la convergencia de la distribución. Podemos presentar las siguientes propiedades:

- Los valores de la esperanza y varianza, se determinan mediante  $E(X) = \alpha\beta$  y  $V(X) = \alpha\beta^2$ .
- El primer parámetro,  $\alpha$ , sitúa la máxima intensidad de probabilidad y por este motivo es denominada la forma de la distribución. Cuando se toman valores próximos a cero aparece entonces un dibujo muy similar al de la distribución exponencial. Cuando se toman valores grandes de  $\alpha$ , el centro de la distribución se desplaza a la derecha, por lo que va apareciendo la forma de la campana de Gauss con asimetría positiva.
- El segundo parámetro,  $\beta$  también llamado  $k$ , es el que determina la forma o alcance de la asimetría positiva desplazando la densidad de probabilidad en la cola de la derecha. Para valores elevados de  $\beta$  la distribución acumula más densidad de probabilidad en el extremo derecho de la cola, alargando mucho su dibujo y dispersando la probabilidad a lo largo del plano.

Al dispersar la probabilidad la altura máxima de densidad de probabilidad se va reduciendo; de aquí que se le denomine escala. Valores mas pequeños  $\beta$  conducen a una figura más simétrica y concentrada, con un pico de densidad de probabilidad más elevado. Una forma de interpretar  $\beta$  es “tiempo promedio entre ocurrencia de un suceso”.

Matemáticamente la función de densidad se define como:

$$f(x) = \frac{\alpha^k x^{(k-1)} e^{-\alpha x}}{(k-1)!} \quad (45)$$

Donde

- $\alpha > 0$ ,  $k > 0$  y  $x$  se considera no negativo.
- Si  $k = 1$ , la distribución gamma resulta ser idéntica a la exponencial.
- Si  $k > 0$  y entero, la distribución coincide con la de Erlang.

Gráficamente la distribución gamma, variando los valores de sus parámetros se percibe de diferentes formas como se mencionó recientemente. Es por esto que se tiene:

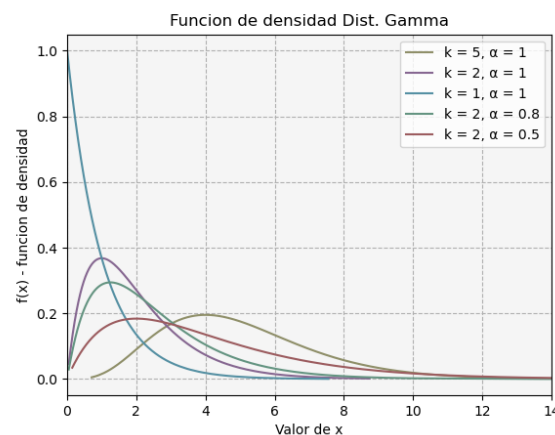


Figura 30: Distribución gamma.

#### 4.5. Generación de la distribución gamma

Para generar una variable aleatoria con distribución gamma debemos considerar un método alternativo al de la transformada inversa, ya que esta distribución no cuenta explícitamente con una función de distribución acumulativa. Para lograr este resultado se debe tomar la suma de los  $k$  valores de la variable aleatoria con distribución exponencial  $x_1, x_2, x_3, \dots, x_k$ , cuyo valor esperado es el mismo e igual a  $1/\alpha$ . En consecuencia el valor de la variable aleatoria  $x$  se puede expresar como:

$$x = \sum_{i=1}^k x_i = -\frac{1}{\alpha} \sum_{i=1}^k \log(r_i) = -\frac{1}{\alpha} \left( \log \prod_{i=1}^k r_i \right) \quad (46)$$

Mediante esta ecuación sacaremos una muestra de tamaño  $n$  para generar  $n$  números con distribución gamma que corresponderán a una variable aleatoria. Como dijimos que se puede expresar mediante la transformada inversa de la distribución exponencial resulta que el código en Python utilizado para generarlo se encuentra definido a continuación:

```
1 def gamma(alfa , k):
2     producto = 1
3     for i in range(k):
4         r = nro_random()           # Genera nro
5         pseudoaleatorio            # pseudoaleatorio
6         producto *= r
7     x = ti_exponencial(alfa , producto) # Numero de
8                                         Erlang con exponencial
9     return x
```

Listing 10: Distribución Gamma

En este caso decidimos ejecutar múltiples corridas para evaluar que tanto se asemejan entre si, y asegurarnos que no influye la corrida que hagamos todas corresponderán a una gráfica similar. De esta forma vimos la independencia con respecto al numero de corridas y puede verse debajo. Para un numero de corridas grande y un  $n$  grande también se obtuvo que:

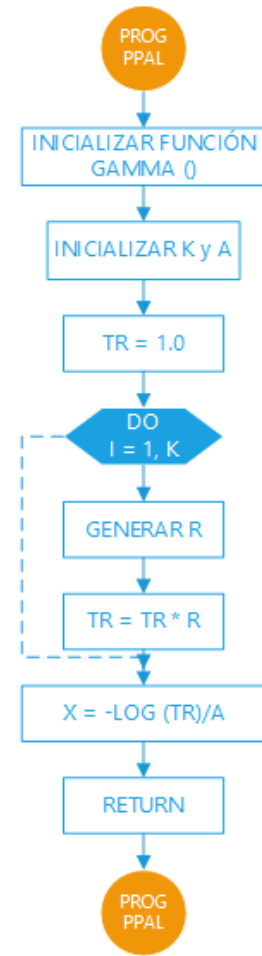


Figura 31: Diag. de flujo.

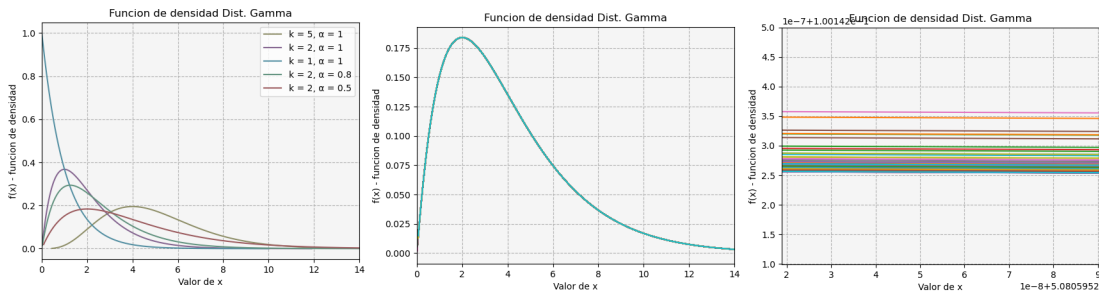


Figura 32: Múltiples corridas con distribución gamma.

Se ve como la similitud entre ellas es extrema, y no existen casi diferencias entre todas. A escala muy pequeña obviamente existirá un cierto desfasaje pero esto se debe a la naturalidad de la aleatoriedad en la generación de la variable aleatoria. Ya que las muestras por mas que asemejen ser similares, están compuestas por diferentes cantidades de números generados con dicha distribución. Además se nota que si bien este desfasaje existe tiende a nulo ya que la escala se encuentra a niveles exponenciales negativos. Luego tomamos una muestra haciendo tender  $n$  a un numero muy grande para poder evaluar en base a su frecuencia acumulada. El resultado fue el siguiente:

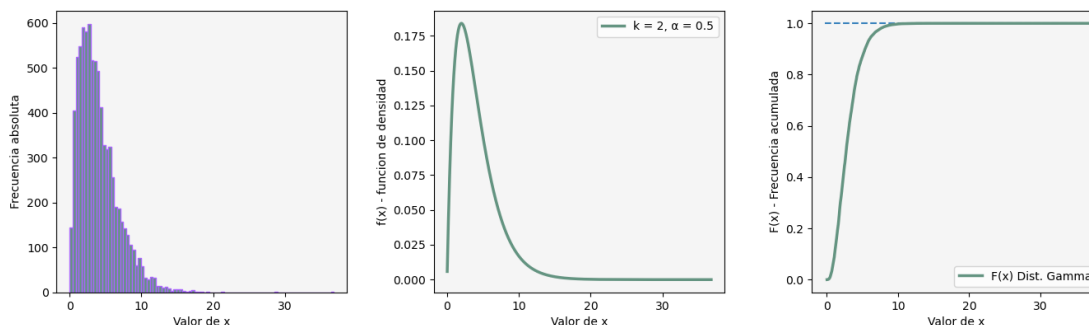


Figura 33: Muestra con distribución gamma.

La primera figura corresponde a un histograma en el que se contó la cantidad de veces que ocurrió cada valor de la variable  $X$ . Estos gráficos son similares a lo esperado y se puede decir que representan una buena generación de dicha variable aleatoria y para nuestro caso de estudio es deseable y recomendable utilizar esta forma de generar los números pertenecientes a la distribución Gamma.

#### 4.6. La distribución normal

En estadística y probabilidad, la distribución normal, también llamada distribución de Gauss (en honor a Carl F. Gauss), distribución gaussiana o distribución de Laplace-Gauss, refleja cómo se distribuyen los datos en una población.

Se trata de la distribución más frecuente en estadística, y se considera la más importante por la gran cantidad de variables reales que adoptan su forma. Así, muchas de las características en la población se distribuyen según una distribución normal. 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 en telecomunicaciones
- Errores cometidos al medir ciertas magnitudes

La densidad está concentrada en torno a la media y se hace muy pequeña conforme nos alejamos del centro por la derecha o la izquierda (las 'colas' de la distribución). Cuanto más alejado es el valor del centro de la función de densidad, menos probable es observar ese valor. Dos parámetros determinan una distribución normal: la media y la desviación estándar. Por tanto, puede ser adecuado hablar de las distribuciones normales, en plural, y decir que son una familia biparamétrica de distribuciones. Es por esto que la Esperanza y la Varianza quedan determinadas como sigue:

$$E(x) = \mu_x \quad V(x) = \sigma_x^2 \quad (47)$$

Alrededor del 95 % de las observaciones está dentro de 2 desviaciones estándar de la media. El 95 % de los valores se ubicará dentro de 1.96 desviaciones estándar con respecto a la media (entre 1.96 y +1.96). Aproximadamente el 68 % de las observaciones está dentro de una 1 desviación estándar de la media (-1 a +1), y alrededor del 99.7 % de las observaciones estarían dentro de 3 desviaciones estándar con respecto a la media (-3 a +3). Su correspondiente función de densidad es dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty \quad (48)$$

Propiedades:

- La función  $f_X(x)$  es mayor a 0 para todo  $x$  real y además  $\int_{-\infty}^{+\infty} f_X(x) dx = 1$ .
- La gráfica  $f_X$  es simétrica respecto de la recta de ecuación  $x = \mu$ . Es decir  $f_X(\mu - x) = f_X(\mu + x)$ .

- La función  $f_X$  tiene un máximo absoluto para  $x = \mu$  y presenta puntos de inflexión en los puntos de abscisa  $x = \mu \pm \sigma$

Sea  $X \sim N(\mu, \sigma)$ . Entonces la probabilidad de que una variable aleatoria con dicha distribución tome un valor en el intervalo  $[a, b]$  viene dada por la integral

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \quad (49)$$

En la práctica, estas integrales se calculan con ayuda de tablas. La función  $f(t) = e^{-\frac{(t-\mu)^2}{2\sigma^2}}$  "no se puede integrar". Para el cálculo a través de tablas resulta útil saber que si  $X \sim N(\mu, \sigma)$ , entonces para  $Z = \frac{X-\mu}{\sigma}$  resulta  $Z \sim N(0, 1)$ . En efecto esto permite determinar que  $F_X(x) = P(X \leq x) = P(Z \leq \frac{x-\mu}{\sigma})$ .

Gráficamente la función normal tiene la siguiente forma:

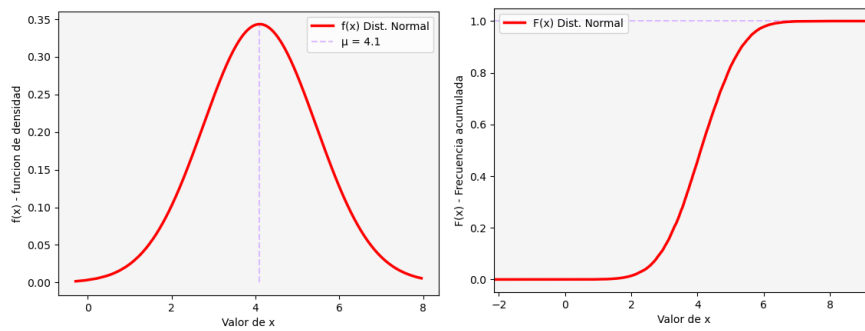


Figura 34: Distribución normal.

#### 4.6.1. Generación de la distribución normal

Existen varios métodos para generar en una computadora, valores de la variable aleatoria distribuidos en forma normal. En este apartado veremos el procedimiento llamado del límite central. Si  $r_1, r_2, \dots, r_n$  representan variables aleatorias independientes, cada una de las cuales posee la misma distribución de probabilidad. Entonces por el teorema del límite central se sabe que la esperanza de la suma de dichas variables aleatorias será la cantidad de variables  $N$  multiplicado por  $\sigma$ . Esto se debe a que todas las variables tienen la misma esperanza matemática igual a  $\theta$ . Entonces

$$E\left(\sum_{i=1}^N r_i\right) = N\theta \quad V\left(\sum_{i=1}^N r_i\right) = N\sigma^2 \quad (50)$$

De esto resulta  $z = \frac{\sum_{i=1}^N r_i - N\theta}{\sigma\sqrt{N}}$ . Se sigue que  $z$  es un valor de variable aleatoria con distribución normal estándar. El procedimiento para simular valores normales requiere el uso de la suma de  $K$  valores de variable aleatoria distribuidos uniformemente; esto es, la suma de  $r_1, r_2, \dots, r_k$  con cada  $r_i$  definida en el intervalo  $0 < r_i < 1$ . Aplicando la convención notacional de la forma matemática del teorema del límite central resulta  $U \sim E(r_i) = \theta = \frac{a+b}{2} = \frac{1}{2}$  y  $\sigma = \frac{b-a}{\sqrt{12}} = \frac{1}{\sqrt{12}}$ . Gracias a esto se puede escribir a  $z$  como

$$z = \frac{\sum_{i=1}^K r_i - \frac{K}{2}}{\sqrt{\frac{K}{12}}} \quad (51)$$

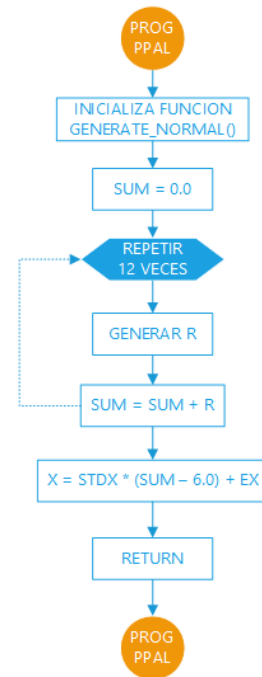


Figura 35: Diag. de flujo.

y esta sera la ecuación que utilizaremos para generar números con distribución normal. El código en Python que determina esta ecuación se encuentra a continuación:

```
1 def normal(mu, sigma):
2     k = 12
3     sum_r = 0
4     for i in range(k):
5         r = nro_random()
6         sum_r += r
7
8     x = sigma * ((12 / k) ** 0.5) * (sum_r - k / 2) + mu
9     return x
```

Listing 11: Distribución Normal

Generando diferentes variables aleatorias con diferentes parámetros para  $\sigma$  y un mismo  $\mu$  se obtuvieron las siguientes muestras:

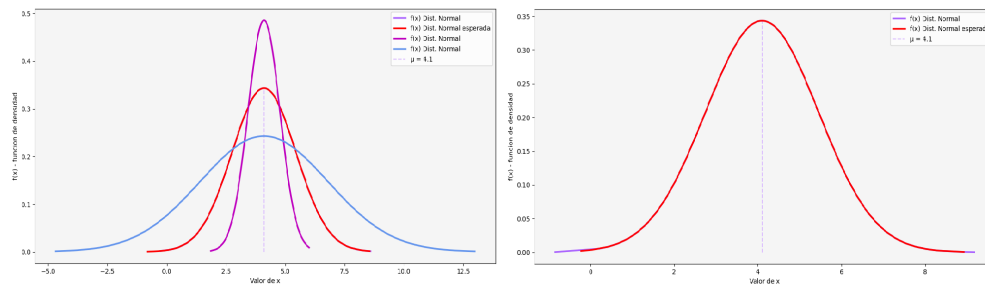


Figura 36: Múltiples muestras distr. normal.

En la segunda imagen se encuentra representado en rojo el valor que se esperaba obtener de esta distribución para un  $s = \sqrt{1,82}$ , puede verse que coincide de manera exacta con el valor obtenido en color violeta para un mismo valor de  $s$ . Ahora elegiremos una muestra para someterla al test de Chi-Cuadrado para evaluar su semejanza con una muestra esperada. La muestra seleccionada fue la siguiente:

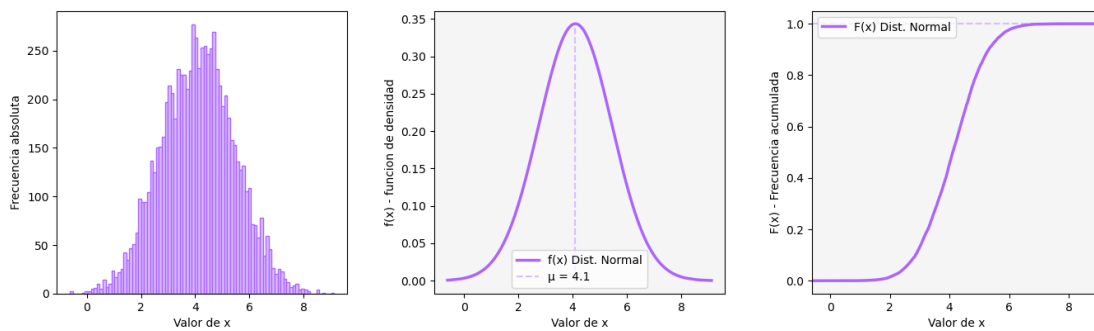


Figura 37: Muestra de distr. normal.

En la figura mostrada puede verse con mas detalle como se asemeja la distribución normal observada con lo esperado. Ahora sometiendola al test de Chi-Cuadrado se obtiene que efectivamente paso la prueba, además agregamos debajo otras muestras que fueron evaluadas con el mismo test pero variando el parámetro de la esperanza y la varianza de la distribución. Pasamos a considerar a este método como uno que sirve para nuestras necesidades. Ya que sabemos que si bien este método es muy conocido no podíamos asegurarnos de que tan eficiente era sin previamente hacer un muestreo del mismo. La tabla que contiene los resultados del test se encuentra a continuación.

DISTRIBUCIÓN	VALOR DE $\mu$ y $\sigma$	CHICUADRADO	¿PASÓ EL TEST?
NORMAL	$\mu = 4,1 \sigma = \sqrt{1,82}$	$Val.Chi(6,66) < Val.table(23,68)$	PASÓ
NORMAL	$\mu = 4,1 \sigma = \sqrt{4}$	$Val.Chi(13,1) < Val.table(26,3))$	PASÓ
NORMAL	$\mu = 7 \sigma = \sqrt{1,82}$	$Val.Chi(15,01) < Val.table(15,51)$	PASÓ
NORMAL	$\mu = 17 \sigma = \sqrt{1,82}$	$Val.Chi(16,37) < Val.table(31,41)$	PASÓ

## 5. Conclusión

El generador de números pseudoaleatorio utilizado para generar variables aleatorias con diferentes distribuciones de probabilidad, tanto continuas como discretas, usado entonces, para la obtención de resultados generan resultados esperados como predice la teoría. Demostrando que este generador reproduce exactamente un análisis de las distribuciones correspondientes se supone que se debe hacer uso de estos métodos. En las situaciones donde se utiliza este generador, las matemáticas necesarias para desarrollar la distribución de probabilidad son tan complejas que los métodos vistos son la única manera de lograr el entendimiento simple necesario para analizar los modelos desarrollados. Sin descartar la implementación de otros métodos existentes pudimos ver que los evaluados previamente corresponden a resultados contundentes en cuanto a lo necesario para comenzar a hacer uso de diferentes herramientas de simulación. Existen métodos que resultan mas populares pero no siempre pueden hacer un mejor uso de los elementos de computación. Lo ideal es tener un buen equilibrio entre estos dos aspectos para no generar problemas a la hora de querer generar una gran cantidad de números con una determinada distribución. Sin embargo, se vio que cuando no se cuenta con una definición formal de la distribución con la cual se está trabajando, se puede hacer uso de la distribución empírica quién no asegura hacer un buen uso de los elementos de hardware pero que puede salvarnos en situaciones donde no podamos contar con otros métodos y se demostró que produce resultados convincentes con lo esperado. Además si bien hicimos uso del generador de números pseudoaleatorios de Python queda implícito que con un generador GCL, quién fue previamente testado, podrían obtenerse los mismos resultados.

## Referencias

- [1] Detalle de números pseudoaleatorios <https://tereom.github.io/est-computacional-2018/numeros-pseudoaleatorios.html>,
- [2] Articulo de variables aleatorias <http://ve.scielo.org/pdf/uct/v16n64/art06.pdf>
- [3] Transformada inversa [http://www.est.uc3m.es/esp/nueva\\_docencia/leganes/ing\\_telecomunicacion/estadistica/](http://www.est.uc3m.es/esp/nueva_docencia/leganes/ing_telecomunicacion/estadistica/)
- [4] Metodos de simulacion [https://www.researchgate.net/publication/315663039\\_Acomparison\\_of\\_inverse\\_transform](https://www.researchgate.net/publication/315663039_Acomparison_of_inverse_transform)
- [5] Numeros pseudoaleatorios <https://es.slideshare.net/iorifoar/numeros-pseudoaleatorios>,
- [6] Numeros pseudoaleatorios <https://es.khanacademy.org/computing/computer-science/cryptography/crypt/v/random-vs-pseudorandom-number-generators>
- [7] Numeros pseudoaleatorios <http://www.sci.csueastbay.edu/~btrumbo/Stat3401/Hand3401/CongGenIntroB.pdf>
- [8] Variables aleatorias <https://webs.um.es/mpulido/miwiki/lib/exe/fetch.php?id=amiocache=cachemedia=wiki:simt3.pdf>,
- [9] Simulacion de variables aleatorias <https://tereom.github.io/est-computacional-2018/simulacion-de-variables-aleatorias.html>,
- [10] Distr. de probabilidad <https://www.ucm.es/data/cont/media/www/pag-54183/APUNTES20ESTADC38DSTICA202.pdf>