

---

# GENERADORES DE NÚMEROS PSEUDOALEATORIOS DE DISTINTAS DISTRIBUCIONES DE PROBABILIDAD

---

**Juan Cruz Vazquez**

Simulación  
UTN - FRRo

Zeballos 1341, S2000 Rosario, Santa Fe  
juancruz.vazquez87@hotmail.com

**Alexis Jose Tomas**

Simulación  
UTN - FRRo

Zeballos 1341, S2000 Rosario, Santa Fe  
alexisjose Tomas@gmail.com

29 de mayo de 2023

## ABSTRACT

En este estudio realizamos un análisis detallado de las distintas distribuciones de probabilidad con el fin de comprender varios fenómenos de interés para la simulación. Para esto utilizamos el generador de números pseudoaleatorios de Python para poder generar secuencias con distintas distribuciones de probabilidad y testear la generación de valores de la forma más conveniente.

## 1. Introduction

Cuando se trata de tomar decisiones, a menudo nos encontramos con información incompleta que dificulta la elección óptima. Es en estos casos donde la inferencia estadística desempeña un papel crucial. Mediante el análisis de una muestra representativa de una población, podemos inferir características de dicha población y utilizar la teoría de probabilidades para evaluar riesgos, minimizarlos y maximizar ganancias. Los beneficios de la inferencia estadística se extienden a diversos campos como la física, las matemáticas, las ciencias sociales, la investigación médica, las finanzas, la economía, la filosofía, entre otros. En nuestro caso, emplearemos esta herramienta para generar secuencias de números pseudoaleatorios con distribuciones de probabilidad específicas, lo cual resulta sumamente útil.

## 2. Marco teórico y conceptual

### 2.1. Probabilidad

La probabilidad asociada a un suceso o evento aleatorio es una medida del grado de certidumbre de que dicho suceso pueda ocurrir. Se suele expresar como un número entre 0 y 1 (o un porcentaje entre 0 y 100), donde un suceso imposible tiene probabilidad cero y un suceso seguro tiene probabilidad uno. La probabilidad constituye un importante parámetro en la determinación de las diversas casualidades obtenidas tras una serie de eventos esperados dentro de un rango estadístico.

### 2.2. Variable aleatoria

Una variable aleatoria es una función que asigna a cada elemento del espacio muestral un número real. Puede ser discreta o continua. Una variable aleatoria es discreta si su recorrido es un conjunto finito o infinito numerable (susceptible de ser contado). Si el recorrido es un intervalo real, la variable aleatoria se dice continua.

### 2.3. Distribución de probabilidad

La distribución de probabilidad de una variable aleatoria es una función que asigna a cada suceso definido sobre la variable la probabilidad de que dicho suceso ocurra. La distribución de probabilidad está definida sobre el conjunto de todos los sucesos y cada uno de los sucesos es el rango de valores de la variable aleatoria. También puede decirse que

tiene una relación estrecha con las distribuciones de frecuencia. De hecho, una distribución de probabilidades puede comprenderse como una frecuencia teórica, ya que describe cómo se espera que varíen los resultados.

La distribución de probabilidad está completamente especificada por la función de distribución, cuyo valor en cada  $x$  real es la probabilidad de que la variable aleatoria sea menor o igual que  $x$ .

Existen diversas distribuciones de probabilidad, en este trabajo práctico hablaremos sobre 9 de ellas y a partir de sus transformadas inversas u otros métodos, podremos generar secuencias de números pseudoaleatorios de estas distribuciones.

## 2.4. El método de la transformada inversa

Es un método para la generación de números aleatorios de cualquier distribución de probabilidad continua cuando se conoce la inversa de su función de distribución.

### 2.4.1. Teorema

Sea  $X$  una variable aleatoria con función de probabilidad acumulada  $F$ , continua e invertible, y sea  $F^{-1}$  su función inversa. Entonces, la variable aleatoria  $U = F(X)$  tiene distribución uniforme en  $(0, 1)$ . Como consecuencia, si  $U$  es una variable aleatoria uniforme en  $(0, 1)$  entonces la variable aleatoria  $X = F^{-1}(U)$  satisface la distribución  $F$ .

### 2.4.2. Método

El método de la transformada inversa funciona de la siguiente manera:

- Se genera un número aleatorio a partir de la distribución uniforme estándar; se lo llama  $r$ .
- Se calcula el valor  $x$  tal que  $F(x) = r$  y se lo llama  $x$  elegido.
- Se toma  $x$  elegido como el número aleatorio extraído de la distribución caracterizada por  $F$ .

### 2.4.3. Demostración

Si deseamos generar los valores  $x_i$  de las variables aleatorias a partir de cierta estadística de población cuya función de densidad esté dada por  $f(x)$ , debemos, en primer lugar obtener la función de distribución acumulativa  $F(x)$ . Puesto que  $F(x)$  se define sobre el rango  $(0, 1)$ , podemos generar números aleatorios distribuidos uniformemente y además hacer  $F(x) = r$ . Resulta claro, entonces, cómo queda  $x$  determinada unívocamente por  $r = F(x)$ . Sigue que para cualquier valor particular de  $r$  que generemos, por ejemplo  $r_0$ , siempre es posible encontrar el valor de  $x$ ; en este caso  $x_0$ , que corresponde a  $r_0$  debido a la función inversa de  $F$ , si es conocida. Esto es:

$$x_0 = F^{-1}(r_0) \quad (1)$$

donde  $F^{-1}(x)$  es la transformación inversa (o mapeo) de  $r$  sobre el intervalo unitario en el dominio de  $x$ . Si generamos números aleatorios uniformes correspondientes a una  $F(x)$  dada, podemos resumir matemáticamente, este método como sigue:

$$r = F(x) = \int_{-\infty}^x f(t)dt \quad (2)$$

entonces

$$P(X \leq x) = F(x) = P[r \leq F(x)] = P[F^{-1}(r) \leq x] \quad (3)$$

y consecuentemente  $F^{-1}(r)$  es una variable que tiene  $f(x)$  como función de densidad de probabilidad. Este criterio equivale a resolver la ecuación 2 en  $x$ , en términos de  $r$ .

## 2.5. Metodo del Rechazo

El muestreo por rechazo es un método para generar variables aleatorias de una distribución objetivo  $f$ . La idea es generar variables aleatorias de una distribución propuesta  $g$ , de la cual es fácil obtener muestras, y luego aceptar o rechazar estas muestras en función de su probabilidad bajo  $f$ .

El algoritmo de muestreo por rechazo funciona de la siguiente manera:

1. Generar una variable aleatoria  $a \leq x \leq b$  a partir de la distribución propuesta  $g$  y un número aleatorio distribuido uniformemente  $(0, 1)$ .
2. Escalar el rango de  $f$  mediante un factor de escala  $c$  igual a la inversa del máximo valor que alcanza la función.
3. Generar una variable aleatoria  $r_2$  a partir de la distribución uniforme en el intervalo  $(0, 1)$ .
4. Si  $r_2 \leq cg(x)$ , entonces aceptar  $x$  como una muestra de  $f$ . De lo contrario, rechazar  $x$  y volver al paso 1.

El algoritmo de muestreo por rechazo garantiza generar variables aleatorias de la distribución objetivo  $f$ , siempre y cuando la distribución propuesta  $g$  tenga soporte en el mismo conjunto que  $f$  y es la razón para que  $g(x)$  esté acotada por una constante  $c$ . Esta constante  $c$  se llama "tasa de aceptación" del algoritmo de muestreo por rechazo.

El algoritmo de muestreo por rechazo es un método general que se puede utilizar para generar variables aleatorias de cualquier distribución objetivo. Sin embargo, puede ser ineficiente, ya que la tasa de aceptación puede ser muy baja. En la práctica, el algoritmo de muestreo por rechazo se utiliza a menudo en conjunto con otros métodos, como el muestreo de importancia, para mejorar la eficiencia del algoritmo.

## 2.6. Distribuciones de probabilidad

### 2.6.1. Distribución Uniforme

Quizá la función de densidad de probabilidad más simple es aquella que se caracteriza por ser constante, en el intervalo  $(a, b)$  y cero fuera de él. Esta función de densidad define la distribución conocida como uniforme. La distribución uniforme surge cuando se estudian las características de los errores por redondeo al registrar un conjunto de medidas sujetas a cierto nivel de precisión. Por ejemplo, si se registran medidas de peso con una aproximación determinada en gramos, puede suponerse que la diferencia en gramos entre el peso real y el peso registrado corresponde a un cierto número entre  $-0,5$  y  $+0,5$ , y que el error se encuentra distribuido uniformemente en este intervalo. El valor más sobresaliente que puede tener la distribución uniforme respecto a las técnicas de simulación radica en su simplicidad y en el hecho de que tal distribución se puede emplear para simular variables aleatorias a partir de casi cualquier tipo de distribución de probabilidad.

Matemáticamente, la función de densidad uniforme se define como sigue:

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{fuera del intervalo } (a, b) \end{cases} \quad (4)$$

En esta expresión,  $X$  es una variable aleatoria definida en el intervalo  $(a, b)$ . La gráfica de la distribución uniforme queda ilustrada en la figura 1.

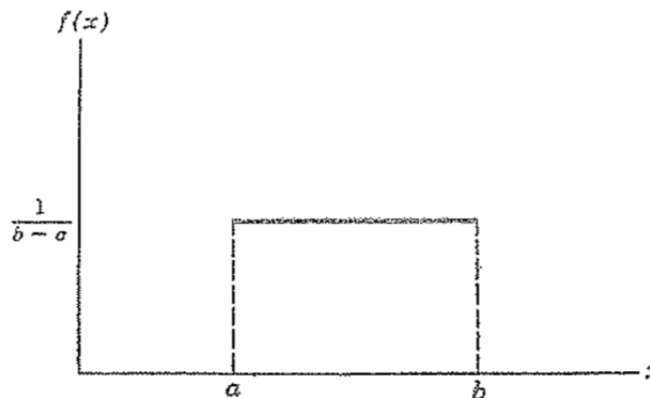


Figura 1: Gráfica de la distribución uniforme

La función de la distribución acumulativa  $F(x)$ , para una variable aleatoria  $X$  uniformemente distribuida, se puede representar por:

$$F(x) = \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a} \quad 0 \leq F(x) \leq 1 \quad (5)$$

El valor esperado y la variancia de una variable aleatoria uniformemente distribuida están dados por las siguientes expresiones:

$$EX = \int_a^b \frac{1}{b-a} x dx = \frac{b+a}{2} \quad (6)$$

$$VX = \int_a^b \frac{(x-EX)^2}{b-a} dx = \frac{(b-a)^2}{12} \quad (7)$$

Al efectuar aplicaciones de esta función, los parámetros de la función de densidad uniforme, esto es, los valores numéricos de  $a$  y de  $b$ , no necesariamente deben ser conocidos en forma directa. En casos típicos, aunque esto no sucede en todas las distribuciones uniformes, solamente conocemos la media y la variancia de la estadística que se va a generar. En estos casos, los valores de los parámetros se deben derivar al resolver el sistema que consta de las ecuaciones 6 y 7, para  $a$  y para  $b$ , pues se supone que  $EX$  y  $VX$  son conocidos. Este procedimiento, semejante a una técnica de estimación conocida en la literatura estadística como método de momentos, proporciona las siguientes expresiones:

- $a = EX - \sqrt{3VX}$
- $b = 2EX - a$

Para simular una distribución uniforme sobre cierto intervalo conocido ( $a, b$ ) deberemos, en primer lugar, obtener la transformación inversa para la ecuación 5.

$$x = a + (b-a)r \quad 0 \leq r \leq 1 \quad (8)$$

Enseguida 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 0 a 1. Cada número aleatorio determina, de manera única, un valor de la variable aleatoria  $x$  uniformemente distribuida.

Para aclarar estas afirmaciones es probable que lo mejor sea presentar una explicación gráfica. La figura 2 nos ilustra cómo cada valor generado de  $r$  está asociado con uno y sólo un valor de  $x$ . Por ejemplo, el valor específico de la función de distribución acumulativa en  $r_0$  determina el valor de  $x$  en  $x_0$ . Obviamente, este procedimiento se puede repetir tantas veces como se desee y en cada repetición se generará un nuevo valor de  $x$ . En este capítulo se observará cómo la generación de valores de variables aleatorias se puede seguir, mediante el uso de probabilidades acumulativas de muchas otras distribuciones.

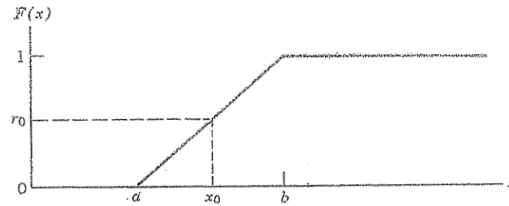


Figura 2: Explicación gráfica

### 2.6.2. Distribución Exponencial

La distribución exponencial es una distribución de probabilidad continua con un parámetro  $\lambda > 0$ . Esta ley de distribución describe procesos en los que interesa saber el tiempo hasta que ocurre determinado evento.

#### Función de densidad

Se dice que una variable  $X$  tiene una distribución exponencial, si se puede definir a su función densidad como:

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (9)$$

### Función de distribución acumulada

La función de distribución acumulada se calcula como sigue:

$$\begin{aligned} F(x) &= P(X \leq x) = \int_0^x \lambda e^{-\lambda t} dt = \\ &= \lambda \int_0^x e^{-\lambda t} dt = \\ &= \lambda \int_0^{\lambda x} -\frac{1}{\lambda} e^u du = \\ &= - \int_0^{\lambda x} e^u du = \\ &= -(e^{-\lambda x} - 1) = \\ &= 1 - e^{-\lambda x} \end{aligned} \quad (10)$$

### Estadísticos

La media se puede expresar como:

$$EX = \int_0^{\infty} x \lambda e^{-\lambda x} dx = \frac{1}{\lambda} \quad (11)$$

La varianza como:

$$VX = \int_0^{\infty} (x - \frac{1}{\lambda})^2 \lambda e^{-\lambda x} dx = \frac{1}{\lambda^2} \quad (12)$$

### Parámetros

Obsérvese que, como la distribución exponencial solamente tiene un parámetro  $\lambda$ , es posible expresarlo como:

$$\lambda = \frac{1}{EX} \quad (13)$$

### Gráficas

En la siguiente figura podemos observar la gráfica de la función densidad de probabilidad, asignándole a  $\lambda$  algunos valores:

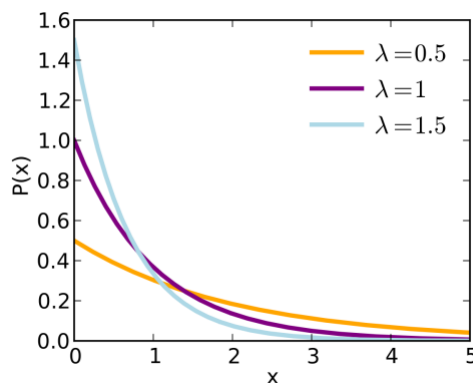


Figura 3: Gráfica de la distribución Exponencial

A su vez repitiendo la asignación de valores a  $\lambda$ , en la siguiente gráfica podemos ver la función de probabilidad acumulada:

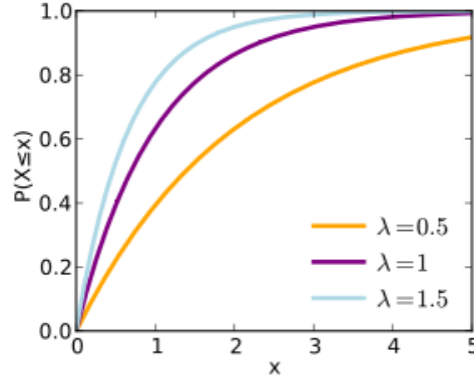


Figura 4: Gráfica de la función de probabilidad acumulada de la distribución exponencial

### Método para la generación de valores

Existen muchas maneras de lograr la generación de valores de variables aleatorias exponenciales. Puesto que  $F(x)$  existe explícitamente, la técnica de la transformación inversa nos permite desarrollar métodos directos para dicha generación. Debido a la simetría que existe entre la distribución uniforme sigue que la intercambiabilidad de  $F(x)$  y  $1 - F(x)$ . Por lo tanto:

$$\begin{aligned} e^{-\lambda x} &= r \\ -\lambda x &= \ln r \\ x &= -\frac{1}{\lambda} \ln r \end{aligned}$$

y como reemplazando  $\lambda$ :

$$x = -EX \ln r \quad (14)$$

Por consiguiente, para cada valor del número pseudoaleatorio  $r$  se determina un único valor para  $x$ . Los valores de  $x$  toman tan sólo magnitudes no negativas, debido a que  $\ln r \leq 0$  para  $0 \leq r \leq 1$ , y además se ajustan a la función de densidad exponencial (ecuación 9) con un valor esperado  $EX$ .

### 2.6.3. Distribución Gamma

Si un determinado proceso consiste de  $k$  eventos sucesivos y si el total del tiempo transcurrido para dicho proceso se puede considerar igual a la suma de  $k$  valores independientes de la variable aleatoria con distribución exponencial, cada uno de los cuales tiene un parámetro definido  $\alpha$ , la distribución de esta suma coincidirá con una distribución gamma con parámetros  $\alpha$  y  $k$ . La suma de los  $k$  (donde  $k$  es un entero positivo) valores de variable aleatoria con distribución exponencial con un mismo parámetro  $\alpha$ , también recibe el nombre de distribución de Erlang. Desde un punto de vista matemático la distribución de Erlang no es otra cosa que la convolución de  $k$  distribuciones exponenciales, o sea la distribución de la suma de  $k$  variables exponenciales. Aún más, si  $k$  se adecúa a la distribución binomial negativa o a la distribución geométrica, entonces la suma de  $k$  valores de variable aleatoria, con la misma  $\alpha$ , adoptará la forma de una distribución gamma. La forma más general de una distribución gamma se obtiene al considerar a  $k$  con valores positivos pero sin que estén restringidos a ser enteros. Casi siempre resulta posible ajustar alguna de las formas de la distribución gamma a un buen número de distribuciones de datos estadísticos sesgados en forma positiva.

La función gamma está escrita mediante la siguiente función de densidad:

$$f(x) = \frac{\alpha^k x^{k-1} e^{-\alpha x}}{(k-1)!} \quad (15)$$

donde  $\alpha > 0$ ,  $k > 0$  y  $x$  se considera no negativo. Pese a que no existe una forma explícita para describir la función acumulativa de la distribución gamma, Pearson ha logrado presentar, en forma tabular, los valores de la llamada

función gamma incompleta. Respecto a la media y la variancia de esta distribución, sus correspondientes expresiones están formuladas como sigue:

$$EX = \frac{k}{\alpha} \quad (16)$$

$$VX = \frac{k}{\alpha^2} \quad (17)$$

Si  $k = 1$ , entonces la distribución gamma resulta ser idéntica a la distribución exponencial; mientras que si  $k$  es un entero positivo, la distribución gamma coincide con la distribución de Erlang. Conviene también anotar que, a medida que  $k$  se incrementa, la distribución gamma tiende, en forma asintótica, a una distribución normal.

Para generar valores de variable aleatoria con distribución gamma y con un valor esperado y variancia dados, se pueden emplear las siguientes fórmulas a fin de determinar los parámetros de  $f(x)$  en la ecuación 15:

$$\alpha = \frac{EX}{VX} \quad (18)$$

$$k = \frac{(EX)^2}{VX} \quad (19)$$

Debido a que para una distribución gamma no se puede formular explícitamente una función de distribución acumulativa, debemos considerar un método alternativo para generar valores de variable aleatoria con distribución gamma. En relación a tales valores que satisfacen la distribución Erlang, estos se pueden generar con sólo reproducir el proceso aleatorio sobre el cual se basa la distribución de Erlang. Para lograr este resultado se debe tomar la suma de los  $k$  valores de variable aleatoria con distribución exponencial  $x_1, x_2, \dots, x_k$ , cuyo valor esperado es el mismo e igual a  $\frac{1}{\alpha}$ . En consecuencia, el valor de la variable aleatoria (Erlang)  $x$  se puede expresar como:

$$x = \sum_{i=1}^k x_i = -\frac{1}{\alpha} \sum_{i=1}^k \log r_i \quad (20)$$

El problema de generar valores de variable aleatoria con distribución gamma cuando el parámetro  $k$  no es un entero, queda aún en la actividad como un problema por resolverse, ya que en este caso todavía se formula un método estocástico satisfactorio, sin lo cual no puede justificarse algún proceso de simulación correspondiente.

#### 2.6.4. Distribución Binomial

Las variables aleatorias definidas por el número de eventos exitosos en una sucesión de ensayos independientes de Bernoulli, para los cuales la probabilidad de éxito es  $p$  en cada ensayo, siguen una distribución binomial. Este modelo estocástico también se puede aplicar al proceso de muestreo aleatorio con reemplazo, cuando los elementos muestreados tienen sólo dos tipos de atributos. El diseño de una muestra aleatoria de  $n$  elementos es análoga a  $n$  ensayos independientes de Bernoulli, en los que  $x$  es un valor binomial que está denotando al número de elementos de una muestra de tamaño  $n$  con atributos idénticos. Es ésta la analogía que sitúa la distribución binomial como uno de los modelos más importantes en las áreas del muestreo estadístico y del control de calidad.

La distribución binomial proporciona la probabilidad de que un evento o acontecimiento tenga lugar  $x$  veces en un conjunto de  $n$  ensayos, donde la probabilidad de éxito está dada por  $p$ . La función de probabilidad para la distribución binomial se puede expresar de la manera siguiente:

$$f(x) = \binom{n}{x} p^x q^{n-x} \quad (21)$$

donde  $x$  se toma como un entero definido en el intervalo finito  $0, 1, 2, \dots, n$ , al que se le asocia el valor  $q = (1 - p)$ .

El valor esperado y la variancia de la variable binomial  $X$  son:

$$EX = nx \quad (22)$$

$$VX = npq \quad (23)$$

La segunda expresión implica que la variancia de las variables binomiales siempre tiene un valor menor al de la media. Aún más, nótese cómo se define en la ecuación (4-136) la distribución de  $(n - x)$  con un valor esperado correspondiente de  $nq$ .

Cuando se conocen la media y la variancia, resulta inmediata la determinación de  $p$  y de  $n$ , las cuales pueden calcularse como sigue:

$$p = \frac{EX - VX}{EX} \quad (24)$$

$$n = \frac{(EX)^2}{EX - VX} \quad (25)$$

La distribución normal proporciona, cuando  $n$  es muy grande, una buena aproximación para la distribución binomial. Puesto que con la distribución normal resulta posible manipular valores negativos, a fin de hacer un buen uso de tal distribución la probabilidad de registrar observaciones negativas deberá ser despreciablemente pequeña. En la práctica esto significa que el valor esperado deberá ser por lo menos tres veces mayor que la desviación estándar, o sea

$$np \geq 3(npq)^{\frac{1}{2}} \quad (26)$$

lo cual implica que  $n \geq 9q/p$ . Los valores de variable aleatoria con distribución binomial se pueden generar de muy diversos modos, aunque uno de los métodos más simples, que en el caso de que el valor de  $n$  sea moderado resulta uno de los métodos más eficientes, es el basado en la reproducción de ensayos de Bernoulli, siguiendo 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  $x_0$  igual a cero. Para cada número 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:

$$x_i = x_{i-1} + 1 \quad \text{si } r_i \leq p \quad (27)$$

$$x_i = x_{i-1} \quad \text{si } r_i > p \quad (28)$$

Después de haberse generado  $n$  números aleatorios, el valor de  $x_n$  será igual al valor de la variable aleatoria con distribución binomial  $x$ . Este procedimiento se puede repetir tantas veces como valores binomiales se requieran.

Un segundo método para generar valores binomiales es el que se basa en las sumas aleatorias de valores de variables aleatorias con distribución geométrica, con lo cual se obtiene el número de éxitos en  $n$  ensayos. En el caso de que  $p$  sea pequeño, este método puede ser mucho más rápido.

### 2.6.5. Distribución Normal

La más conocida y más ampliamente utilizada distribución de probabilidad es sin duda la distribución normal y su popularidad se debe cuando menos a dos razones que presentan sus propiedades generales. "Las pruebas matemáticas nos señalan, que bajo ciertas condiciones de calidad, resulta justificado que esperemos una distribución normal mientras que la experiencia estadística muestre que, de hecho, muy a menudo las distribuciones se aproximan a la normal".

La distribución normal basa su utilidad en el teorema del límite central. Este teorema postula que, la distribución de probabilidad de la suma de  $N$  valores de variable aleatoria  $x_i$ ; independientes pero idénticamente distribuidos, con medias respectivas  $\mu_i$ ; y varianzas  $\sigma^2$  se aproxima asintóticamente a una distribución normal, a medida que  $N$  se hace muy grande, y que dicha distribución tiene como media y varianza respectivamente, a

$$\mu = \sum_{i=1}^N \mu_i \quad (29)$$

$$\sigma = \sum_{i=1}^N \sigma_i^2 \quad (30)$$



En consecuencia, el teorema del límite central permite el empleo de distribuciones normales para representar medidas globales operadas sobre los efectos de causas (errores) aditivas distribuidas en forma independiente, sin importar la distribución de probabilidad a que obedezcan las mediciones de causas individuales. El carácter que se atribuye a esta forma de interpretación del teorema del límite central resulta de particular importancia, ya que de esta manera se provee una justificación matemática para manipular la evidencia empírica, que tan frecuentemente aparece en los datos distribuidos en forma aproximadamente normal, obtenidos en una gran mayoría de problemas de investigación.

Otro rasgo valiosamente práctico de la distribución normal, es su utilidad para aproximar distribuciones de Poisson y binomiales para mencionar sólo dos entre muchas. A partir de la distribución normal, se pueden derivar otras muchas de las distribuciones existentes que juegan un papel muy importante en la estadística moderna, por ejemplo la Chi cuadrada, la t, y la distribución F, las cuales se originan a partir de consideraciones hechas sobre la distribución de probabilidad de la suma de los cuadrados de un número específico de valores de variables aleatoria con una distribución normal estándar [8, p. 233].

Si la variable aleatoria  $X$  tiene una función de densidad  $f(x)$  dada como

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-1/2 \left( \frac{x - \mu_x}{\sigma_x} \right)^2} \quad (31)$$

La función inversa de la distribución normal está definida matemáticamente de la siguiente manera:

Para una distribución normal estándar (media 0 y desviación estándar 1), la función inversa se denota como  $Z(p)$ , donde  $p$  es la probabilidad deseada. La fórmula para calcular la función inversa es:

$$Z(p) = \sqrt{2} * \text{erfinv}(2p - 1) \quad (32)$$

Donde  $\text{erfinv}$  es la función inversa de la función de error.

Para una distribución normal con una media  $\mu$  y una desviación estándar  $\sigma$ , la función inversa se denota como  $X(p, \mu, \sigma)$ . La fórmula para calcular la función inversa en este caso es:

$$X(p, \mu, \sigma) = \mu + \sigma * Z(p) \quad (33)$$

Donde  $Z(p)$  es la función inversa de la distribución normal estándar.

### 2.6.6. Distribución de Pascal

Entre las primeras y probablemente más simples de las formulaciones matemáticas de procesos estocásticos, se encuentra la llamada de ensayos de Bernoulli. Estos ensayos son experimentos independientes al azar, en los que el resultado de cada ensayo queda registrado, ya sea como un éxito o un fracaso. La probabilidad de éxito se denota por  $p$  ( $0 \leq p \leq 1$ ) y se supone que  $p$  es constante para cualquier sucesión particular de ensayos.

La probabilidad de un fracaso se denota por  $q$ , donde:

$$q = 1 - p \quad (34)$$

Una sucesión de ensayos Bernoulli, combinada con cierto proceso del conteo, viene a constituir la base conceptual para una gran familia de distribuciones discretas de probabilidad, incluyendo la geométrica, binomial negativa, Poisson y otras distribuciones binomiales. Los valores de variable aleatoria que se generan al contar el número de fracasos en una sucesión de ensayos (o eventos) antes de que ocurra el primer éxito, son valores de variable aleatoria que se ajustan a una distribución geométrica.

La distribución geométrica de probabilidad tiene un gran valor y utilidad en el área del control estadístico de calidad, así como también para las distribuciones de rezagos y movimientos en modelos econométricos.

La distribución geométrica queda descrita por la siguiente función de probabilidad:

$$f(x) = pq^x \quad x = 0, 1, 2, \dots \quad (35)$$

y la función de distribución acumulativa está definida por:

$$F(x) = \sum_{X=0}^x pq^x \quad x = 0, 1, 2, \dots \quad (36)$$

Puesto que por definición se tiene  $F(x) = P(X \leq x)$ , y como  $P(X = 0) = F(0) = p$ , el rango de  $F(x)$  es  $p \leq F(x) \leq 1$ . Por otra parte,  $P(X > x) = 1 - F(x)$ , lo que implica que  $P(X > 0) = q$  y además

$$1 - F(x) = q^{x+1} \quad (37)$$

El valor esperado y la variancia de la variable geométrica, están dados por:

$$EX = \frac{q}{p} \quad (38)$$

$$VX = \frac{q}{p^2} = \frac{EX}{p} \quad (39)$$

Esta última expresión implica que la variancia difiere de la media en un factor de  $\frac{1}{p}$ . De la ecuación (30) resulta claro cómo la distribución geométrica siempre tendrá el perfil de una J con modo en el punto  $x = 0$ .

La distribución geométrica tiene sólo un parámetro  $p$ , el cual se puede expresar como una función de la media  $EX$

$$p = \frac{1}{1 + EX} \quad (40)$$

Para generar en una computadora valores de variable aleatoria con distribución geométrica se emplea la técnica de la transformación inversa y la fórmula que aparece en la ecuación (32). Al observar que el rango de la expresión  $[1 - F(x)/q]$  es unitario, resulta que

$$r = q^x \quad (41)$$

y consecuentemente

$$x = \frac{\log r}{\log q} \quad (42)$$

donde al valor  $x$  y siempre se le redondea al entero que sea menor. Esto se puede lograr de manera muy simple con sólo retener los números hasta antes del punto decimal, o bien convirtiendo los números en modo de punto flotante al de punto fijo. Para generar un valor de variable aleatoria con distribución geométrica haciendo uso de esta técnica, se requiere únicamente el empleo de un número aleatorio uniforme, tal como lo muestra la ecuación (37).

### 2.6.7. Distribución hipergeométrica

Considérese una población que consta de  $N$  elementos tales que cada uno de ellos pertenece a la clase I o a la II. Denotemos por  $N_p$  al número de elementos que pertenecen a la clase I y por  $N_q$  al número de elementos que son miembros de la clase II, donde  $p+q=1$ . Si en una población de  $N$  elementos se toma una muestra aleatoria que conste de  $n$  elementos ( $n < N$ ) sin que tenga lugar algún reemplazo, entonces el número de elementos  $x$  de la clase I en la muestra de  $n$  elementos, tendrán una distribución de probabilidad hipergeométrica. En las áreas de control de calidad y en la de control de producción, con mayor frecuencia se encuentran las aplicaciones de la distribución hipergeométrica. Por ejemplo, si el intervalo entre la llegada de órdenes sucesivas de los clientes relativas a cierto producto de la compañía, se distribuye geoméricamente, entonces la demanda total en cualquier periodo dado de tiempo tendrá una distribución hipergeométrica. La distribución hipergeométrica está descrita por la siguiente función de probabilidad:

$$f(x) = \frac{\binom{N_p}{x} \binom{N_q}{n-x}}{\binom{N}{n}}$$

$$0 \leq x \leq N_p$$

$$0 \leq n - x \leq N_q$$

donde  $x$ ,  $n$  y  $N$  son enteros. El valor esperado y la variancia se caracterizan como sigue:

$$EX = np$$

$$VX = npq \left( \frac{N-n}{N-1} \right)$$

La generación de valores hipergeométricos involucra, sustancialmente, la simulación de experimentos de muestreo sin reemplazo. En otras palabras, bastará 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 número total de elementos que previamente se han obtenido entre la población y el número de elementos de la clase I 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 con la fórmula:  $N_i = N_{i-1} - 1$   $i = 1, 2, \dots, n$ . De manera similar, el valor de  $p = p_0$  se transforma según:

$$p_i = \frac{N_{i-1}p_{i-1} - s}{N_{i-1} - 1}$$

$$i = 1, 2, \dots, n$$

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 a la clase I y  $S=0$  cuando el elemento de muestra ( $i-1$ ) pertenece a la clase II. Ciertamente, los valores iniciales de  $N_0$  y  $P_0$  corresponden: a  $N$  el tamaño inicial de la población y a la proporción de la población total que consta de elementos de la clase I.

### 2.6.8. Distribución de Poisson

El nombre de esta distribución proviene de su creador, Siméon-Denis Poisson (1781-1840), un matemático y filósofo francés, que quería modelar la frecuencia de eventos durante un intervalo de tiempo fijado. También participó en perfeccionar la ley de los grandes números.

La distribución de Poisson es una distribución de probabilidad discreta que modeliza la frecuencia de eventos determinados durante un intervalo de tiempo fijado a partir de la frecuencia media de aparición de dichos eventos. En otras palabras, la distribución de Poisson es una distribución de probabilidad discreta que, tan solo conociendo los eventos y su frecuencia media de ocurrencia, podemos saber su probabilidad.

Esta distribución se utiliza en el campo de riesgo operacional con el objetivo de modelar las situaciones en que se produce una pérdida operacional. En riesgo de mercado se emplea el proceso de Poisson para los tiempos de espera entre transacciones financieras en bases de datos de alta frecuencia. También, en riesgo de crédito se tiene en cuenta para modelar el número de quiebras.

Dada una variable aleatoria discreta  $X$  decimos que su frecuencia se puede aproximar satisfactoriamente a una distribución de Poisson, tal que:

$$X \sim P(\mu)$$

A diferencia de la distribución normal, la distribución de Poisson solo depende de un parámetro,  $\mu$ .

$\mu$  informa del número esperado de eventos que ocurrirán en un intervalo de tiempo fijado. Cuando se habla de algo “esperado” tenemos que redirigirlo a pensar en la media. Por tanto,  $\mu$  es la media de la frecuencia de los eventos. Tanto la media como la varianza de esta distribución son  $\mu$ , estrictamente positiva.

### Función de densidad de probabilidad

$$P(X = x) = \frac{\mu^x e^{-\mu}}{x!}$$

$$\forall x = 0, 1, 2, \dots, n$$

Esta función se entiende como la probabilidad de que la variable aleatoria  $X$  tome un valor concreto  $x$ . Es la exponencial de la media negativa multiplicada por la media elevada a la observación y todo dividido por el factorial de la observación. Como está indicado, para conocer la probabilidad de cada observación, tendremos que sustituir en la función todas las observaciones. En otras palabras,  $x$  es un vector de dimensión  $n$  que contiene todas las observaciones de la variable aleatoria  $X$ . La media también sería un vector pero de una dimensión, tal que:

$$x = 0, 1, 2, \dots, n$$

$$\mu = \text{media}$$

Una vez ya tenemos las probabilidades calculadas, junto con las observaciones ya podemos dibujar la distribución de densidad de probabilidad.

## 2.7. Distribución Empírica

La distribución empírica, también conocida como función de distribución empírica (FDE), es una herramienta estadística utilizada para describir la distribución de un conjunto de datos observados. Proporciona una estimación acumulada de la probabilidad de que una variable aleatoria tome un valor menor o igual a un cierto punto.

## Representación matemática

$$FDE(x) = \frac{1}{N} \sum_{i=1}^N I(x \leq x_i) \quad (43)$$

El desarrollo matemático de la distribución empírica implica los siguientes pasos://

1. **Obtención de datos:** Primero, se recopila un conjunto de datos observados. Por ejemplo, supongamos que tenemos una muestra de tamaño  $n$ , y los datos se denotan como  $x_1, x_2, \dots, x_n$ .
2. **Ordenamiento de los datos:** A continuación, se ordenan los datos de manera ascendente. Si denotamos los datos ordenados como  $x(1), x(2), \dots, x(n)$ , entonces tenemos  $x(1) \leq x(2) \leq \dots \leq x(n)$ .
3. **Cálculo de la función de distribución empírica:** La función de distribución empírica se define como la proporción de observaciones que son menores o iguales a un cierto valor. Para calcular la función de distribución empírica en un punto específico  $x$ , se utiliza la siguiente fórmula:

$$F(x) = (\text{número de observaciones } \leq x) / n \quad (44)$$

Esta fórmula indica que se cuentan las observaciones que son menores o iguales a  $x$  y se divide por el tamaño total de la muestra.

4. **Representación gráfica:** Finalmente, se puede representar gráficamente la función de distribución empírica utilizando un gráfico de escalones. En el eje  $x$  se colocan los valores de los datos ordenados, y en el eje  $y$  se representan los valores de  $F(x)$  correspondientes.

La distribución empírica es una aproximación de la función de distribución teórica desconocida subyacente de la cual se obtuvieron los datos observados. Proporciona una forma visual de entender cómo están distribuidos los datos y se utiliza a menudo como punto de partida para realizar inferencias estadísticas o para comparar la distribución de los datos observados con una distribución teórica específica.

## 3. Testeos

### 3.1. Test de Kolmogorov-Smirnov

Se usará este test para probar luego, la generación de variables aleatorias de las distintas distribuciones a partir de los métodos de la transformada y de rechazo.

El test de Kolmogorov-Smirnov es una prueba estadística utilizada para determinar si una muestra proviene de una distribución específica. Compara la función de distribución empírica acumulada (FDE) de la muestra con la función de distribución acumulada (FDA) teórica esperada. La estadística del test se calcula como la mayor discrepancia absoluta entre la FDE y la FDA. La fórmula para la estadística del test es:

$$D_n = \max_i (\max(|FDE(X_i) - FDA(X_i)|, |FDE(X_i - 1) - FDA(X_i)|)) \quad (45)$$

donde  $D_n$  es la estadística del test,  $FDE$  es la función de distribución empírica acumulada,  $FDA$  es la función de distribución acumulada teórica, y  $X_i$  es el  $i$ -ésimo valor ordenado de la muestra. El valor  $p$  del test se obtiene comparando la estadística del test con una distribución de referencia apropiada. Si el valor  $p$  es menor que un umbral predefinido, se rechaza la hipótesis nula de que la muestra sigue la distribución teórica.

### 3.2. Prueba de bondad de ajuste $\chi^2$

La prueba de bondad de ajuste  $\chi^2$  es una prueba de hipótesis estadística que se usa para averiguar si es probable que una variable provenga o no de una distribución específica. En otras palabras, nos dice si la muestra disponible representa (o ajusta) razonablemente los datos que uno esperaría encontrar en la población.

#### Hipotesis:

$H_0$  Los datos son muestra de la distribución uniforme.

$H_1$  Los datos no son muestra de la distribución uniforme.

#### Prueba:

1. Partir la distribución en n celdas que son exhaustivas y mutuamente excluyentes.
2. Contar el número de observaciones  $O_i$  encontrados en cada celda.
3. Calcular el valor esperado en cada celda

$$e_i = np_i \quad (46)$$

4. Calcular la sumatoria de las estadísticas de prueba

$$X^2 = \sum_{i=1}^n \frac{(oi - ei)^2}{ei} \quad (47)$$

5. Si  $\chi^2 < \chi_{p,q}^2$ , se acepta la hipótesis  $H_0$

p : Representa el intervalo de confianza. q : Representa los grados de libertad.

## 4. Códigos y Testeos

### 4.1. Distribución Uniforme

Para este experimento tomamos:

- $A = 2$
- $B = 5$
- $n = 10000$

Donde n es la cantidad de valores de la lista.

Con esta lista generamos un histograma donde se puede ver la cantidad de veces que se obtuvo cada uno de los 4 valores. Y se puede ver el esperado en la línea roja.

**Código en Python 3.11 para generar números con distribución uniforme**

```
1 def Uniforme(size): # Dist. Uniforme
2     ListUni = []
3     for i in range(size):
4         nr = random.random()
5         vu = round(a + nr * (b - a), 4)
6         ListUni.append(vu)
7     return ListUni
```

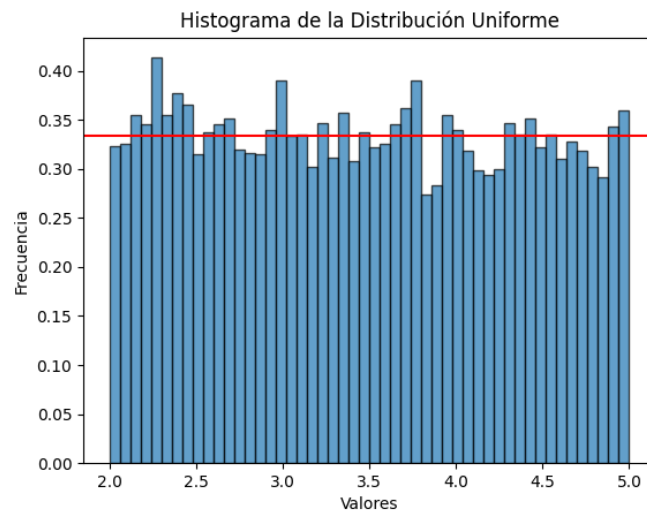


Figura 5: Distribucion Uniforme

En la gráfica se puede ver la línea roja que es lo esperado, y como los valores son muy cercanos a este.

#### Números generados con metodo del rechazo

```

1 def runiforme(size, a, b):
2     ListUni = []
3     while len(ListUni) < size:
4         u = random.random()
5         x = a + u * (b - a)
6         y = random.random()
7         if y <= 1 / (b - a):
8             ListUni.append(x)
9     return ListUni

```

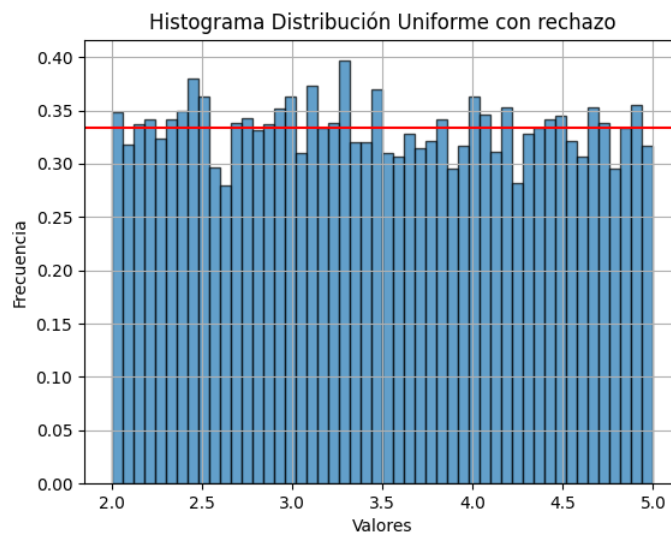


Figura 6: Distribucion Uniforme con rechazo

En la gráfica se puede ver la línea roja que es lo esperado, y como los valores son más cercanos que el caso anteriormente mostrado.

## 4.2. Distribución Exponencial

### Código en Python 3.11 para generar números con distribución exponencial

```
1 def generador_exponencial(media, n):
2     valores=[]
3     for i in range(0, n):
4         R = random.random()
5         X = -media * math.log(R)
6         valores.append(X)
7     return valores
```

- $media = 1$

- $n = 1000$

Donde n es la cantidad de valores de la lista.

Con esta lista generamos un histograma donde se puede ver la cantidad de veces que se obtuvieron los distintos valores. Si comparamos la figura 7 con la figura 8 podemos notar que

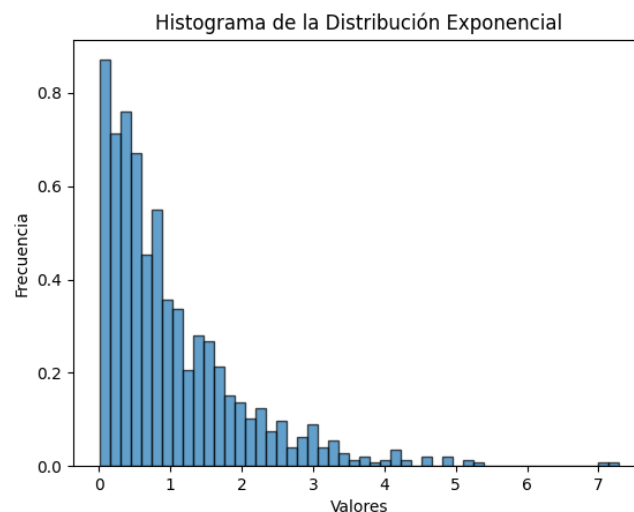


Figura 7: Histograma distribución exponencial

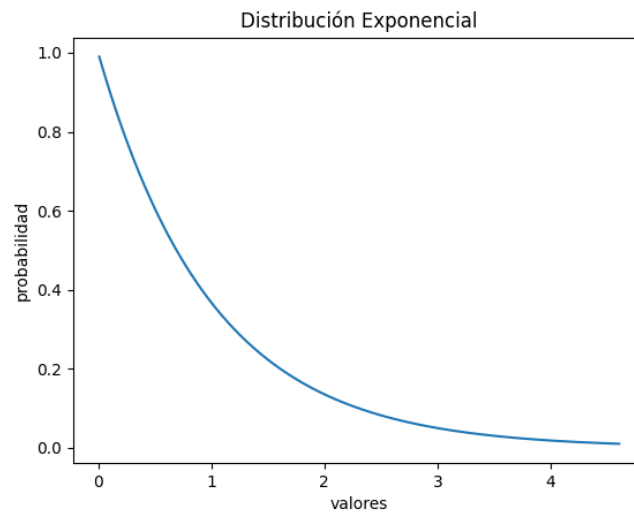


Figura 8: Funcion densidad distribucion exponencial

**Test Kolmogorov-Smirnov**

0.04014 &lt; 0.07756

Pasa el Test

**Números generados con metodo de transformada inversa**■  $\lambda = 0,5$ 

```

1 def t_inversa_exp(lam, n):
2     u = np.random.uniform(0, 1, n) # Genera n números aleatorios uniformes
3     x = -np.log(1 - u) / lam # Aplica la transformada inversa
4     return x

```

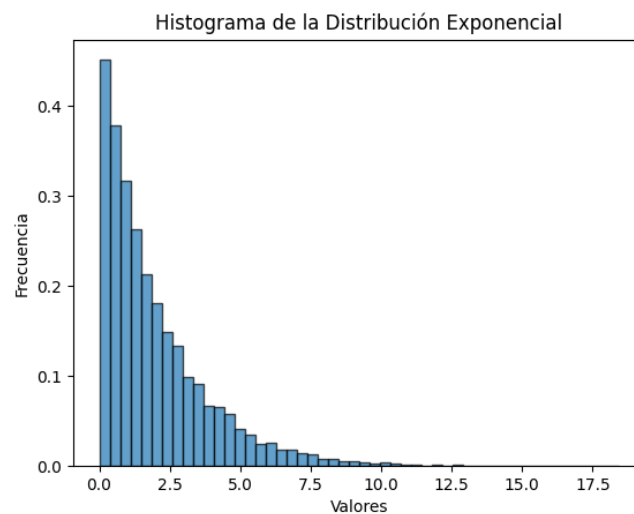


Figura 9: Funcion densidad distribucion exponencial



**Test Kolmogorov-Smirnov**

0.05159 &gt; 0.00940

No pasa el Test

**Números generados con metodo del rechazo**

```

1 def RExponencial(media, size): # Método de rechazo para distribución exponencial
2     ListExp = []
3     while len(ListExp) < size:
4         u = random.random()
5         v = -media * math.log(u)
6         if u <= math.exp(-v / media):
7             ListExp.append(v)
8     return ListExp

```

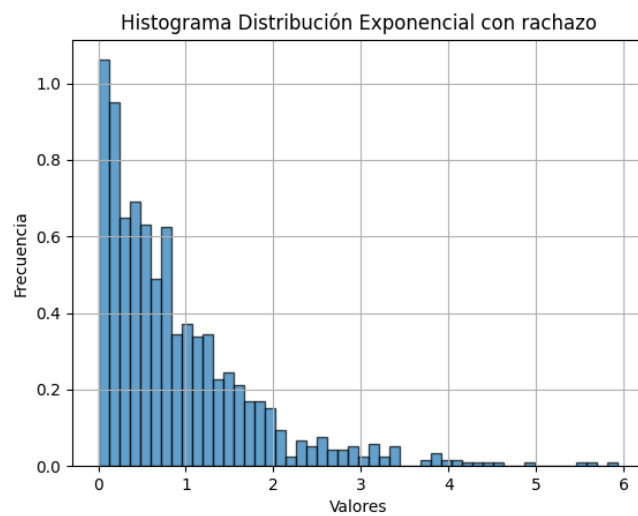


Figura 10: Histograma distribución exponencial - metodo del rechazo

**Test Kolmogorov-Smirnov**

0.01723 &lt; 0.92278

Pasa el Test

**4.3. Distribución Gamma****Código en Python 3.11 para generar números con distribución gamma** Para este experimento tomamos:

- $A = 2,6$

```

1 def gammaSR(VA, A):
2     TR = 1.0
3     for _ in range(VA):
4         R = random.random()
5         TR *= R
6     X = -math.log(TR) / A
7     return X

```

La Distribución Gamma esta dada por la formula:

$$p(x; a, b) = a(ax)^{b-1} e^{-ax\gamma(b)} \quad (48)$$

En dónde los parámetros  $a$  y  $b$  y la variable  $x$  son números reales positivos y  $\gamma(b)$  es la función gamma. La Distribución Gamma comienza en el origen de coordenadas y tiene una forma bastante flexible. Otras distribuciones son casos especiales de ella.

Las gráficas que obtuvimos fueron:

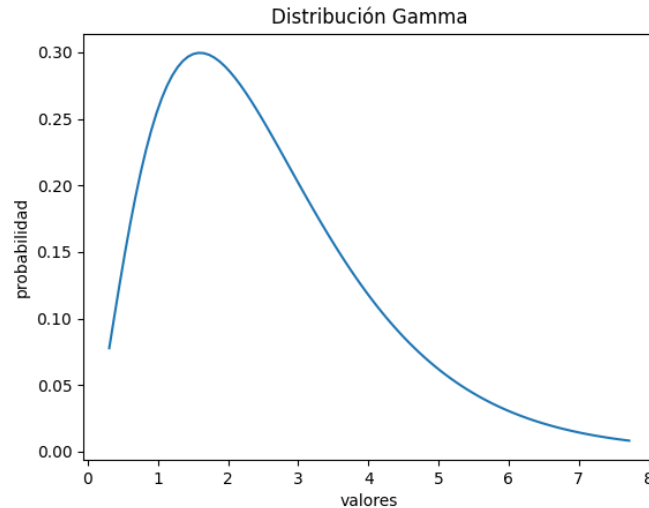


Figura 11: Función de distribución gamma

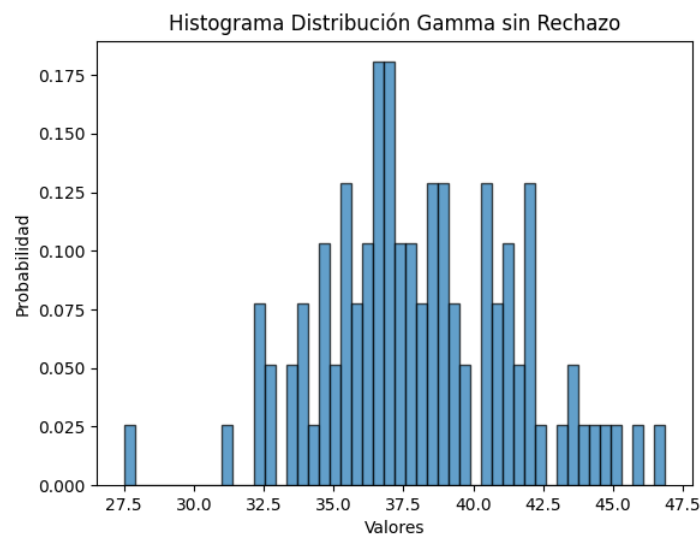


Figura 12: Distribucion gamma

### Números generados con metodo del rechazo

```
1 def GammaR(A,K):
2     c = (1 / (A**A * math.exp(-A)))
3     M = K.pdf(A) / (c * stats.expon.pdf(A, scale=1))
4     x = np.linspace(K.ppf(0.01), K.ppf(0.99), 100)
5     fp = K.pdf(x)
6     aleatorios = []
7     i = 0
```

```

8  while i < 100:
9      y = random.expovariate(1)
10     u = random.uniform(0, M)
11     propuesta = y + A
12     factor_aceptacion = K.pdf(propuesta) / (c * stats.expon.pdf(y, scale=1))
13     if u <= factor_aceptacion:
14         aleatorios.append(propuesta)
15         i += 1
16  return x, fp, aleatorios

```

Las gráficas que obtuvimos fueron:

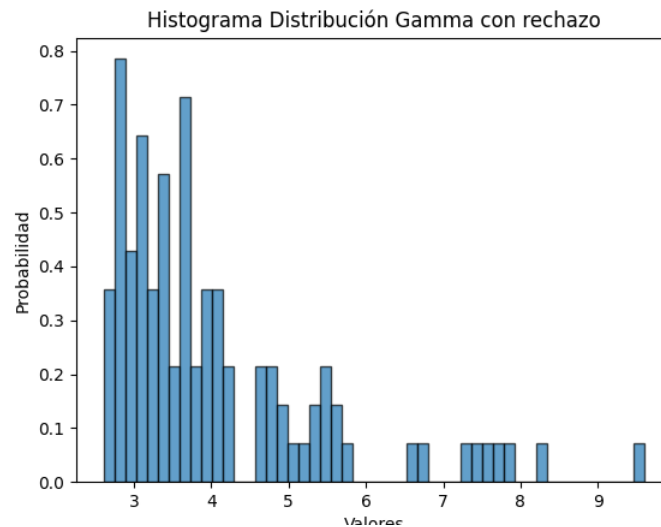


Figura 13: Distribucion gamma con rechazo

#### 4.4. Distribución Normal

**Código en Python 3.11 para generar números con distribución normal** Para este experimento tomamos:

- $\mu = 10$
- $\sigma = 1$

llenamos la lista con 1000 valores generados con la función random de python:

```

1  def generador_normal(ex, stdx, n):
2      valores=[]
3      for i in range(0, n):
4          sum = 0
5          for j in range(1, 13):
6              r = random.random()
7              sum = sum + r
8          x = ex + stdx * (sum - 6.0)
9          valores.append(x)
10  return valores

```

Las gráficas que obtuvimos fueron:

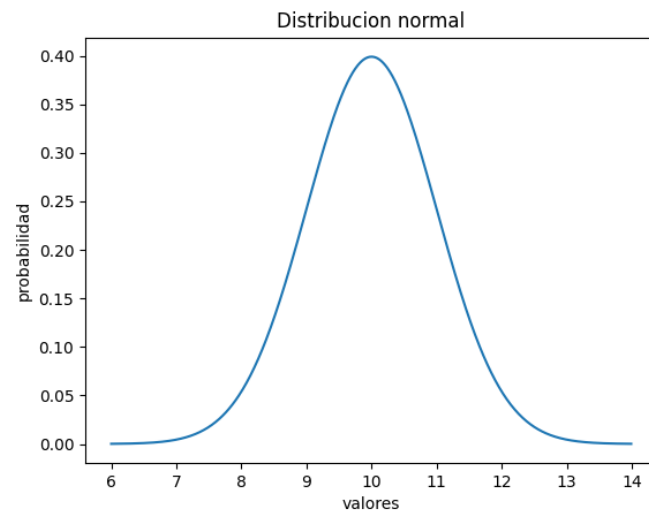


Figura 14: Distribucion normal

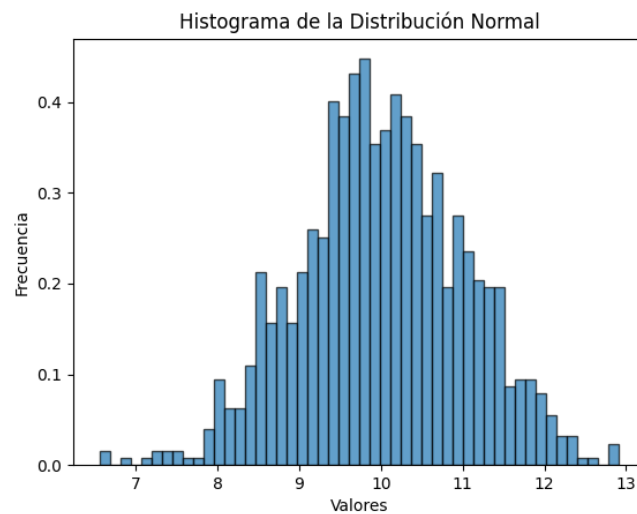


Figura 15: Histograma distribucion normal

**Test Kolmogorov-Smirnov**

0.04156 &lt; 0.06139

Pasa el Test

**Números generados con metodo de transformada inversa**

```

1 def t_inversa_normal(n):
2     u = np.random.uniform(0, 1, n) # Genera n números aleatorios uniformes
3     x = np.sqrt(2) * erfinv(2 * u - 1) # Aplica la transformación inversa
4     return x

```

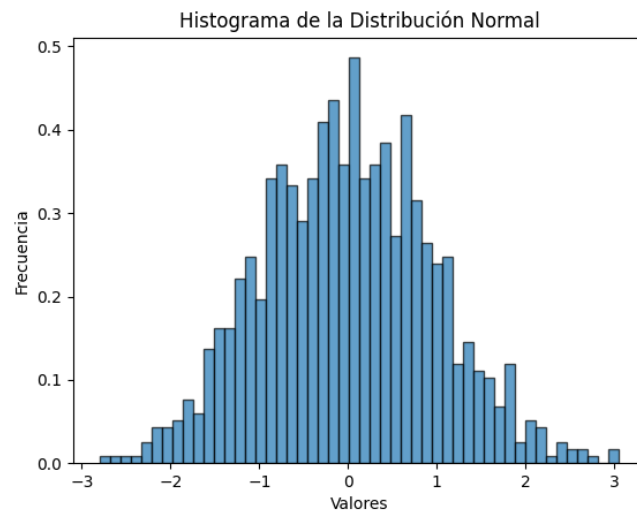


Figura 16: Histograma distribucion normal - Metodo de transformada inversa

### Test Kolmogorov-Smirnov

$0.10634 > 2.66222e^{-10}$

No pasa el Test

### Números generados con metodo del rechazo

```

1 def normalR(EX, STDY):
2     while True:
3         SUM = 0.0
4         for _ in range(12):
5             R = random.random()
6             SUM += R
7         X = EX + STDY * (SUM - 6.0)
8         # Generar un valor aleatorio U entre 0 y 1
9         U = random.random()
10        # Calcular la función de densidad de probabilidad (PDF)
11        # de una distribución normal estándar en X
12        PDF_X = (1 / (math.sqrt(2 * math.pi) * STDY)) * math.exp(-(X - EX)**2 / (2 * STDY**2))
13        # Comprobar si U está por debajo de la PDF_X
14        if U <= PDF_X:
15            return X

```

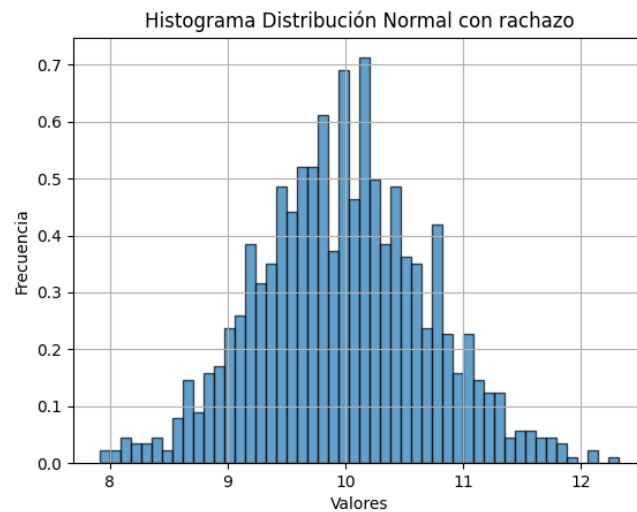


Figura 17: Histograma distribución normal - Metodo del rechazo

#### 4.5. Distribución Pascal

##### Código en Python 3.11 para generar números con distribución pascal

Para este experimento tomamos:

- $K = 5$
- $Q = 0,1$
- $Size = 1000$

llenamos la lista con 1000 valores generados con la función random de python:

```

1 def pascal(K, Q):
2     TR = 1.0
3     QR = math.log(Q)
4     for I in range(1, K + 1):
5         R = random.random()
6         TR = TR * R
7     NX = math.log(TR) / QR
8     X = NX
9     return X

```

Las gráficas que obtuvimos fueron:

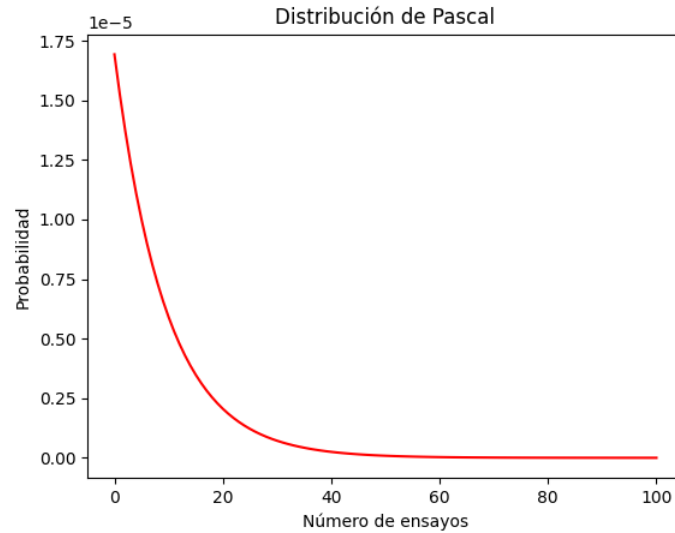


Figura 18: Distribución pascal

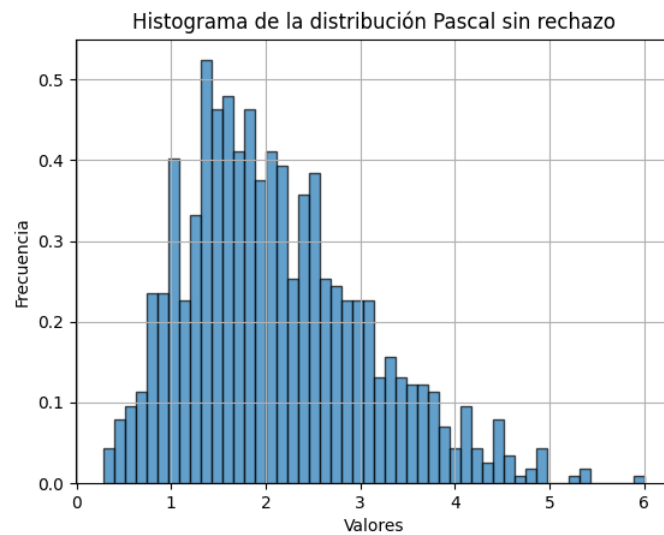


Figura 19: Histograma distribucion pascal

### Números generados con metodo del rechazo

```

1 def pascalR(K, Q):
2     while True:
3         Y = -math.log(random.random()) / Q
4         X = 0
5         U = random.random()
6         while U >= math.exp(-K * Q):
7             U *= random.random()
8             X += 1
9         if X == K:
10            return Y

```

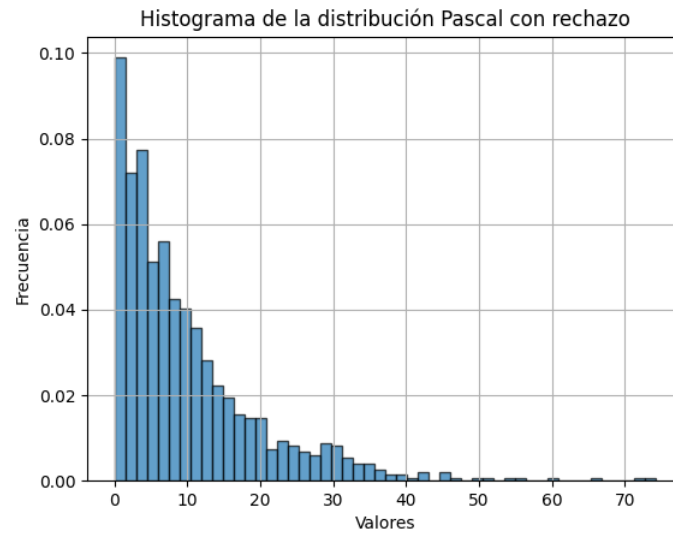


Figura 20: Histograma distribucion pascal - Metodo del rechazo

#### 4.6. Distribución Binomial

Para este experimento tomamos:

- $N = 100$
- $Q = 0,1$
- $Size = 1000$

**Código en Python 3.11 para generar números con distribución binomial**

```
1 def binom(N, P):
2     X = 0.0
3     for I in range(1, N+1):
4         R = random.random()
5         if R - P < 0:
6             X = X + 10
7     return X
```

#### Graficas de la distribucion Binomial

Para este experimento tomamos 100 ensayos(n) y con una probabilidad fija  $p=0.1$  de ocurrencia de éxito entre los ensayos.



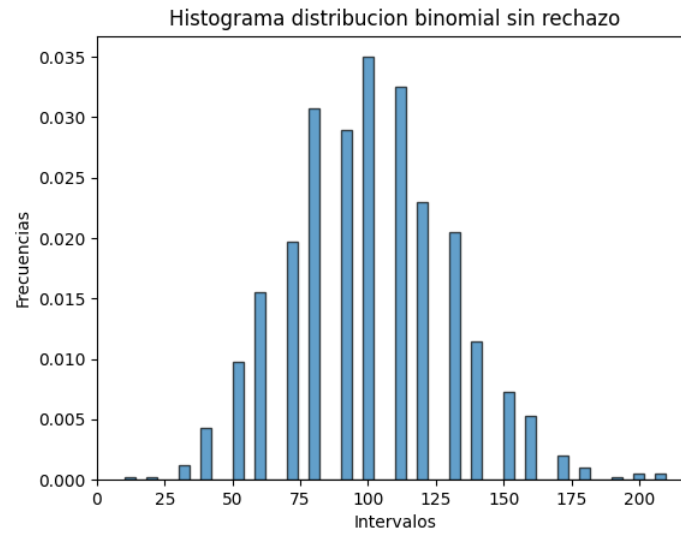


Figura 21: Histograma distribución binomial

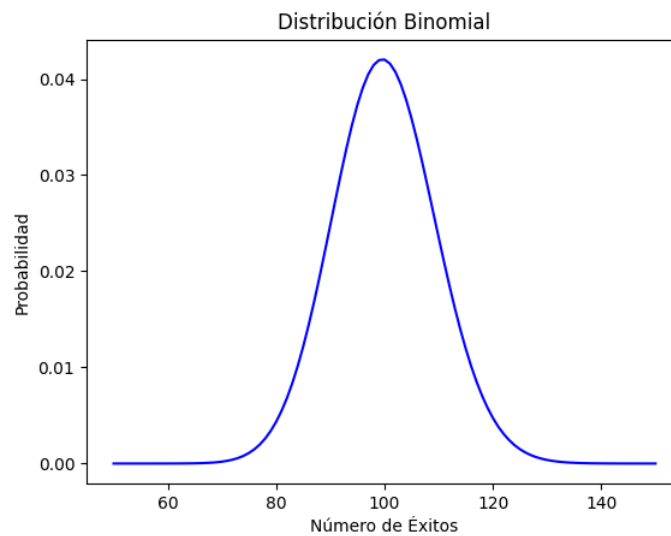


Figura 22: Distribucion binomial

**Test de chi cuadrado  $\chi^2$**

$0.05 > 4,01371e^{-201}$

No pasa el Test

**Números generados con metodo del rechazo**

```

1 def binomR(N, P):
2     X = 0
3     for _ in range(N):
4         R = random.random()
5         if R < P:
6             X += 1
7     return X

```

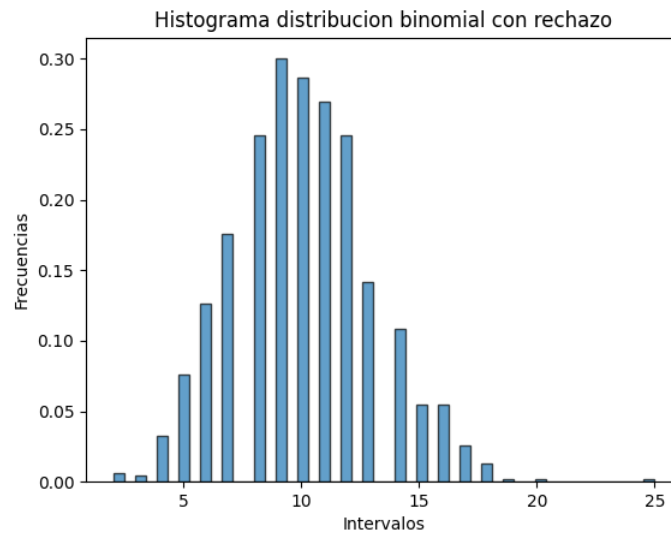


Figura 23: Histograma distribución binomial - Método del rechazo

**Test de chi cuadrado  $\chi^2$**

$0.05 > 6,80402^{-187}$

No pasa el Test

#### 4.7. Distribución Hipergeométrica

Para este experimento tomamos:

- $N = 500$

- $K = 50$

- $n = 100$

llenamos la lista con 1000 valores generados con la función random de python:

```

1  def hipergeometrica(TN, NS, P):
    while True:
        X = 0
        for _ in range(NS):
            R = random.random()
            if R < P:
                X += 1
        if X <= TN:
            return X

```

Las gráficas que obtuvimos fueron:

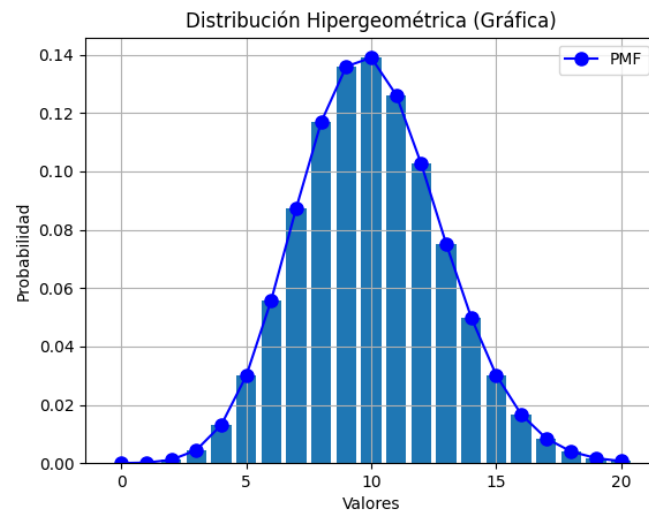


Figura 24: Distribucion Hipergeométrica

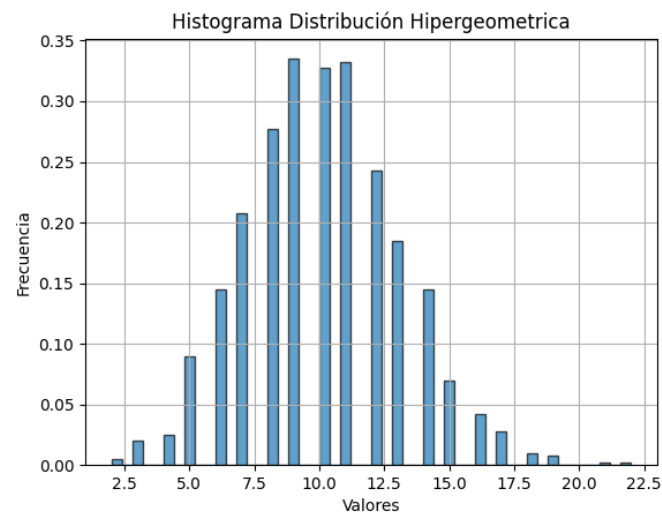


Figura 25: Histograma de la distribucion Hipergeométrica

### Números obtenidos por metodo del rechazo

```

1 def rechazo_hipergeometrica(TN, NS, P, n):
2     resultados = []
3     while len(resultados) < n:
4         Y = hipergeometrica(TN, NS, P)
5         U = random.random()
6         if U < Y / NS:
7             resultados.append(Y)
8     return resultados

```

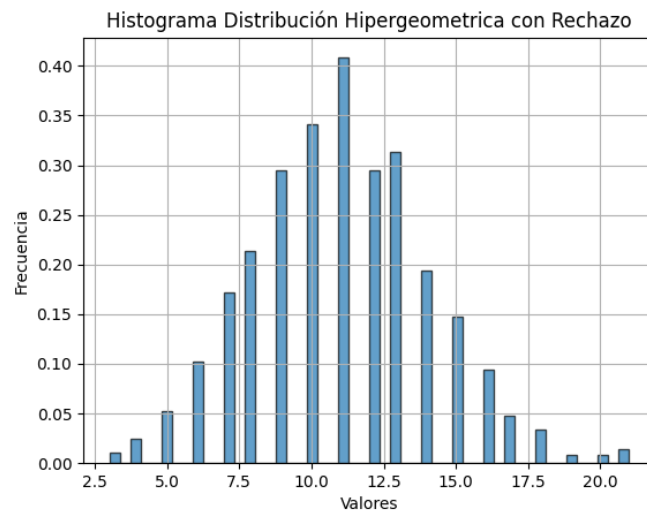


Figura 26: Histograma de la distribución Hipergeométrica - Método del rechazo

#### 4.8. Distribución de Poisson

Para esta prueba generamos 1000 valores utilizando  $\lambda = 3$ .

```

1 def poissn(p):
2     x = 0
3     L = math.exp(-p)
4     tr = 1.0
5     while True:
6         tr *= random.random()
7         if tr <= L:
8             return x
9         else:
10            x += 1

```

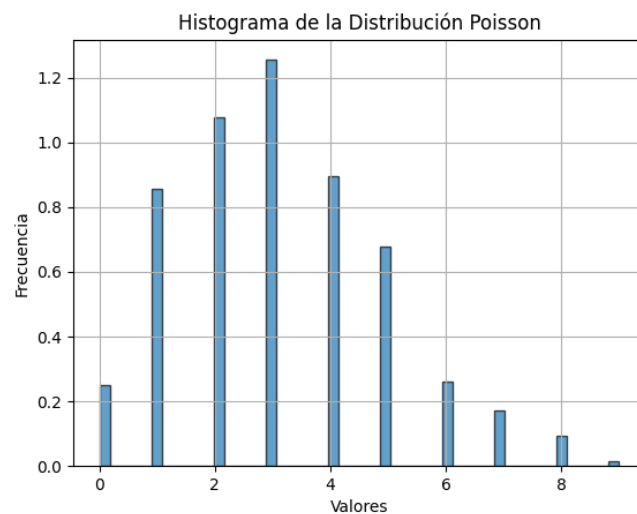


Figura 27: Gráfica de la distribución de Poisson

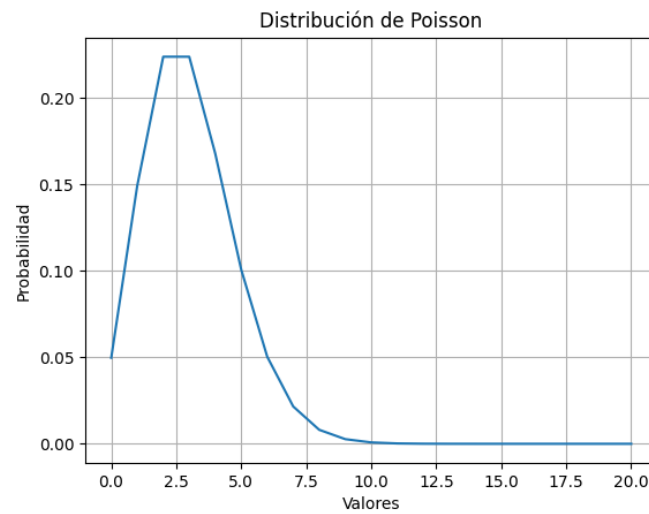


Figura 28: Histograma de valores generados

Como podemos observar, la mayor cantidad de valores generados se encuentran en el rango (2,3), y si nos fijamos en la función de densidad, la mayor probabilidad de aparición se encuentra en el mismo rango.

#### Números obtenidos por metodo del rechazo

```

1 def poissnR():
2     numeros = []
3     l = 3 # lambda
4     for i in range(1000):
5         u = random.random()
6         x = 0
7         P = F = math.exp(-1)
8         while u >= F:
9             P = (1 / (x + 1)) * P
10            F = F + P
11            x = x + 1
12            numeros.append(x)
13     return numeros

```

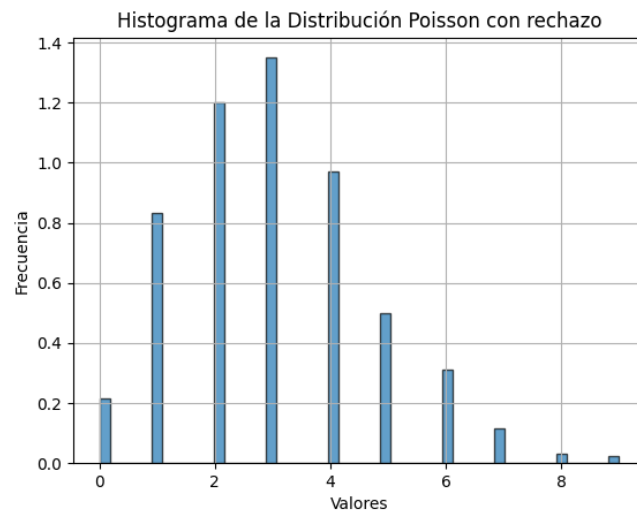


Figura 29: Histograma de valores generados con metodo del rechazo

### Test de chi cuadrado $\chi^2$

23.51814150606464 > 14.0

Se rechaza la hipótesis nula. Las muestras no siguen una distribución de Poisson.

## 4.9. Distribucion Empirica Discreta

```

1 def empirica_discreta(values, probabilities, n):
2     cumulative_probabilities = [sum(probabilities[:i+1]) for i in range(len(probabilities))]
3     random_values = []
4
5     for _ in range(n):
6         random_value = random.random()
7
8         for i, cumulative_prob in enumerate(cumulative_probabilities):
9             if random_value < cumulative_prob:
10                 random_values.append(values[i])
11                 break
12
13     return random_values

```

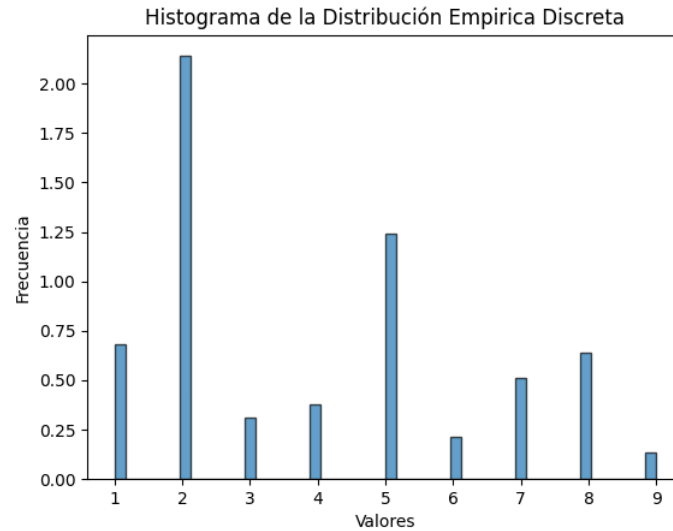


Figura 30: Histograma de la distribucion empirica discreta

### Números obtenidos por metodo del rechazo

```

1 def empirica_discretaCR(values, probabilities, n):
2     random_values = []
3     max_probability = max(probabilities)
4     while len(random_values) < n:
5         random_index = random.randint(0, len(values) - 1)
6         random_probability = random.uniform(0, max_probability)
7         if random_probability <= probabilities[random_index]:
8             random_values.append(values[random_index])
9     return random_values

```

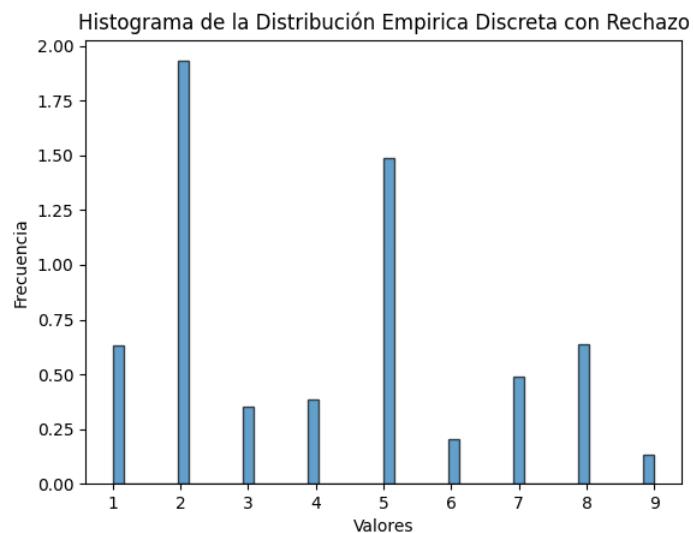


Figura 31: Histograma de la distribucion empirica discreta - Método del rechazo

## 5. Conclusiones

A lo largo del desarrollo fuimos generando las distintas distribuciones utilizando los diferentes métodos, los cuales son el **Metodo de la Transformada Inversa** y el **Metodo del Rechazo**. Fue común para todas las distribuciones generadas, la semejanza de la muestra con la función densidad ya conocida, sin tener en cuenta que aun así algunas no pasaron los Tests realizados. Esto, sumado a la verificación de varias de las distribuciones mediante la visualización gráfica de su comportamiento, nos lleva a concluir que las distribuciones generadas son eficientes y corresponden con su definición matemática.

## 6. Referencias

- Naylor, T.H. Técnicas de simulación en computadoras, 1982.
- Método de la transformada inversa: [https://es.wikipedia.org/wiki/M%C3%A9todo\\_de\\_la\\_transformada\\_inversa](https://es.wikipedia.org/wiki/M%C3%A9todo_de_la_transformada_inversa)
- Distribucion Empirica Discreta: [https://es.wikipedia.org/wiki/Distribuci%C3%B3n\\_emp%C3%ADrica](https://es.wikipedia.org/wiki/Distribuci%C3%B3n_emp%C3%ADrica)
- GitHub <https://github.com/AlexisTomas2000/TP-Simulacion/tree/main/TP-2.2>