Distribuciones

Ejercicio 1)

Imaginemos que tenemos los siguientes experimentos, razona qué distribución de tipo discreta los representa:

1. Jugar a cara o cruz.
2. El nacimiento de un bebé (niño o niña).
3. Fabricación de piezas en una factoría (aceptables o defectuosas).
4. El resultado de una operación médica (éxito o fracaso).
5. El lanzamiento a una canasta (encestar o fallar).
6. El resultado de un examen (aprobar o suspender).

Ejercicio 2)

Tiramos una moneda 6 veces y contamos el número de caras X.

a) ¿Qué distribución representa a este experimento? ¿Cuáles son sus parámetros?
b) Hallar la función de probabilidad.
c) Representarla en un diagrama de barras.
d) Hallar la media y la varianza.

Ejercicio 3)

Se extraen cinco bolas con reemplazamiento de una urna que contiene 6 bolas rojas y 8 negras. ¿Qué es más probable: sacar 2 bolas rojas o sacar 3 bolas rojas?

Ejercicio 4)

Supongamos que el número de imperfecciones en un alambre delgado de cobre sigue una distribución de Poisson con una media de 2.3 imperfecciones por milímetro.

a) Determine la probabilidad de 2 imperfecciones en un milímetro de alambre.
b) Determine la probabilidad de 10 imperfecciones en 5 milímetros de alambre.
c) Determine la probabilidad de al menos una imperfección en 2mm de alambre.

Ejercicio 5)

Un examen de 100 preguntas admite como respuesta en cada una de ellas dos posibilidades, verdadero o falso. Si un alumno contesta al azar calcula la probabilidad de que acierte más de 60 respuestas.

Ejercicio 6)

La contaminación constituye un problema en la fabricación de discos de almacenamiento óptico. El número de partículas de contaminación que ocurre en un disco óptico tiene una distribución de Poisson y el número promedio de partículas por centímetro cuadrado de superficie del disco es 0.1. El área de un disco bajo estudio es 100 centímetros cuadrados.

a) Encuentre la probabilidad de que ocurran 12 partículas en el área del disco bajo estudio.
b) La probabilidad de que ocurran cero partículas en el área del disco bajo estudio.
c) Determine la probabilidad de que 12 o menos partículas ocurran en el área del disco bajo estudio.

Ejercicio 7)

Supóngase que la concentración que cierto contaminante se encuentra distribuida de manera uniforme en el intervalo de 0 a 20 pares de millón. Si se considera tóxica una concentración de 8 o más.

a) ¿Cuál es la probabilidad de que al tomarse una muestra la concentración de esta sea tóxica?
b) Halle la concentración media y la varianza.
c) Calcula la probabilidad de que la concentración sea exactamente 10.

Ejercicio 8)

Una empresa electrónica observa que el número de componentes que fallan antes de cumplir 100 horas de funcionamiento es una variable aleatoria de Poisson. Si el número promedio de estos fallos es 8:

a) ¿Cuál es la probabilidad de que falle un componente en 25 horas?
b) ¿Y de qué fallen no más de dos componentes en 50 horas?
c) ¿Cuál es la probabilidad de que fallen por lo menos 10 en 125 horas?

Ejercicio 9)

Supongamos que el consumo familiar de un cierto producto se distribuye como una variable aleatoria de distribución uniforme, con esperanza igual a 10 y varianza unidad. Determina la probabilidad de que dicho consumo esté comprendido entre 8 y 12 unidades.

Ejercicio 10)

El personal de una compañía usa una Terminal para realizar sus pedidos internacionales. Si el tiempo que cada comercial gasta en una sesión en la Terminal tiene una distribución exponencial con media 36 minutos, encontrar:

a) La probabilidad de que un comercial utilice la Terminal 30 minutos o menos.
b) Si un comercial ha estado 30 minutos en la Terminal, ¿cuál es la probabilidad de que pase al menos una hora o más en la Terminal?
c) El 90% de las sesiones terminan en menos de x_T minutos. ¿Cuánto vale x_T?

Ejercicio 11)

Se ha comprobado que el tiempo de vida de cierto tipo de marcapasos sigue una distribución exponencial con media de 16 años. ¿Cuál es la probabilidad de que a una persona a la que se le ha implantado este marcapasos se le deba reimplantar otro entre los 10 y los 20 años?

Ejercicio 12)

Demostrar que la tipificación de una variable aleatoria Normal genérica, la convierte en una Normal estándar:

$$X \sim N(\mu, \sigma^2) \quad \rightarrow \quad Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

a) Si $X \sim N(\mu = 2, \sigma^2 = 2^2)$, hallar $P(X < 8)$.

Ejercicio 13)

Si $Z \sim N(0,1)$, hallar el valor de z_0 en cada caso:

a) $P(z < z_0) = 0.5$
b) $P(z < z_0) = 0.8729$
c) $P(z > z_0) = 0.9015$

Ejercicio 14)

Los pesos de 2000 soldados presentan una distribución Normal de media 65kg y desviación típica 8kg. Calcula la probabilidad de que un soldado elegido al azar pese:

a) Más de 61kg.
b) Entre 63 y 69kg.
c) Menos de 70kg.
d) Más de 75kg.

Ejercicio 15)

Imagina que en una tienda, en promedio entran 8 clientes cada hora y llama N a la variable que mide el número de clientes que entran en nuestra tienda cada hora.

a) ¿Qué distribución tiene la variable aleatoria N?
b) ¿Qué distribución tiene el tiempo de espera hasta el cliente siguiente?
c) Calcula la probabilidad de que tengamos que esperar 15 minutos hasta que un nuevo cliente entre en la tienda.

Tareas

1*

El departamento de control de calidad de una empresa que fabrica pañuelos sabe que el 5% de su producción tiene algún tipo de defecto. Los pañuelos se empaquetan en cajas de 15. Calcular la probabilidad de que una caja contenga:

a) 2 elementos defectuosos.
b) Menos de 3 elementos defectuosos.
c) Entre 3 y 5 elementos defectuosos (ambos incluidos).

2*

La duración de un láser semiconductor a potencia constante tiene una distribución Normal con media 7000 horas y desviación típica de 600 horas.

a) ¿Cuál es la probabilidad de que el láser falle antes de 5000 horas?
b) ¿Cuál es la duración en horas excedida por el 95% de los láseres?

3*

Aproximando con una distribución Normal, calcular la probabilidad de que al lanzar una moneda 100 veces, el número de caras obtenido esté comprendido entre 46 y 55.

4*

El tiempo de reparación de unas máquinas de escribir tiene una distribución aproximadamente exponencial, con media de 22 minutos.

a) ¿Cómo quedaría la distribución de la exponencial?
b) ¿Cómo quedaría la función de distribución?
c) Determina la probabilidad de que el tiempo de reparación sea menor que quince minutos.

5*

En un libro de 500 páginas las erratas se distribuyen aleatoriamente con una distribución de Poisson, con un promedio de 300 erratas en 500 páginas. Hallar la probabilidad de que en una página haya menos de dos erratas.

RESUMEN TEÓRICO Y FÓRMULAS:

DISTRIBUCIONES DISCRETAS

Distribución Bernoulli:
Representa un experimento donde solo hay dos resultados posibles, uno que fijamos como “éxito” y lo contrario que sería el “fracaso”.
Ejemplos:

7. Jugar a cara o cruz.
8. El nacimiento de un bebé (niño o niña).
9. Fabricación de piezas en una factoría (aceptables o defectuosas).
10. El resultado de una operación médica (éxito o fracaso).
11. El lanzamiento a una canasta (encestar o fallar).
12. El resultado de un examen (aprobar o suspender).

Notación

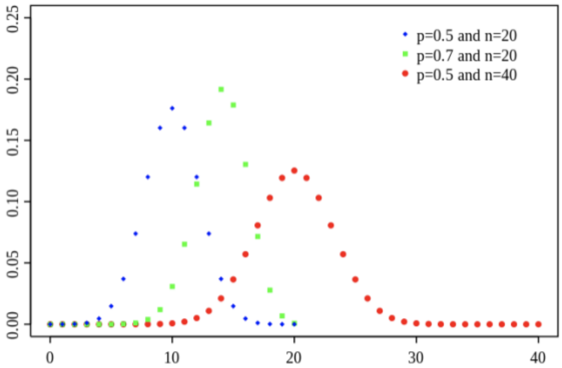
$$X \sim Bernoulli(p)$$

Donde p es la probabilidad de "éxito", es decir, $p = P(X = 1)$.

Función de probabilidad de una Bernoulli:
$P(X = x) = p^x(1-p)^{1-x}$

Valor esperado y Varianza:
$E(X) = p$
$Var(X) = p(1-p)$

Distribución Binomial:



Representa un experimento que consiste en n repeticiones (independientes) de un experimento Bernoulli.
Ejemplos:

1. Lanzar una moneda 10 veces y contar el número de caras.
2. Observar 30 nacimientos de bebés en un hospital y contar el número de niñas.
3. Fabricación de 100 piezas y contar las defectuosas.
4. Lanzamiento de n veces a canasta y contar los éxitos.

Números combinatorios

- El número combinatorio $\binom{n}{k}$ se lee "número combinatorio de n sobre (o en) k" y se calcula de la siguiente forma:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

- El factorial es el producto sucesivo decreciente: $n! = n * (n-1) * (n-2) * \ldots * 2 * 1$

- $0! = 1$

- Ejemplo:

$$\binom{5}{2} = \frac{5!}{2!\,(5-2)!} = \frac{5*4*3*2*1}{(2*1)(3*2*1)} = 10$$

Una variable aleatoria X se llama binomial si su valor representa el número de éxitos que ocurren en n pruebas independientes, teniendo todas ellas la misma probabilidad de éxito, que designamos como p. Su función de probabilidad es:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{para } k = 0, 1, \dots, n$$

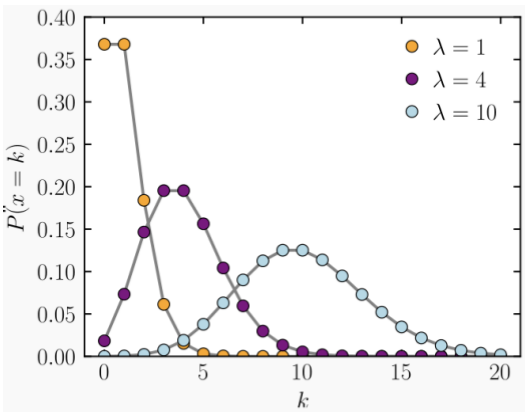
Notación:

$$X \sim B(n, p)$$

Las constantes n y p son los parámetros de la distribución, donde $n > 0$ y un número entero, y $0 \leq p \leq 1$ ya que es una probabilidad (de tener éxito). La probabilidad de fracaso sería $1 - p$ y se suele designar con la letra q, con lo cual $p + q = 1$ por la ley de probabilidad.

Media: $n * p$
Varianza: $n * p * q$
Desviación típica: $\sqrt{n * p * q}$

Distribución Poisson:



Expresa, a partir de una frecuencia de ocurrencia media, la probabilidad de que ocurra un determinado número de eventos durante cierto período de tiempo.

Notación:

$$X \sim Poisson(\lambda)$$

Función de probabilidad:

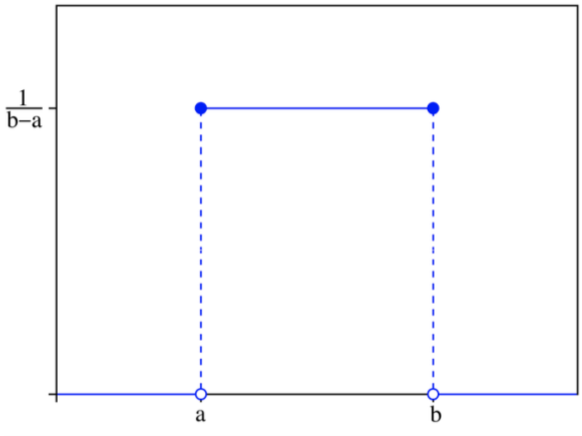
$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

donde $k = 0,1,2, \dots$ es el número de ocurrencias del evento.

Media: $E(X) = \lambda$
Varianza: $Var(X) = \lambda$

DISTRIBUCIONES CONTINUAS

Distribución Uniforme:



Caracteriza a un intervalo igualmente probable.
Notación:

$$X \sim U(a,b)$$

Función de densidad:

$$f_X(x) = \frac{1}{b-a} \quad \text{para } x \in [a,b]$$

Función de distribución:

$$F_X(x) = \begin{cases} 0 & ,x < a \\ \dfrac{x-a}{b-a} & ,a \leq x < b \\ 1 & ,x \geq b \end{cases}$$

Valor esperado:

$$E(X) = \frac{a+b}{2}$$

Varianza:

$$Var(X) = \frac{(b-a)^2}{12}$$

Distribución Uniforme estándar:

$$X_U \sim U(0,1)$$

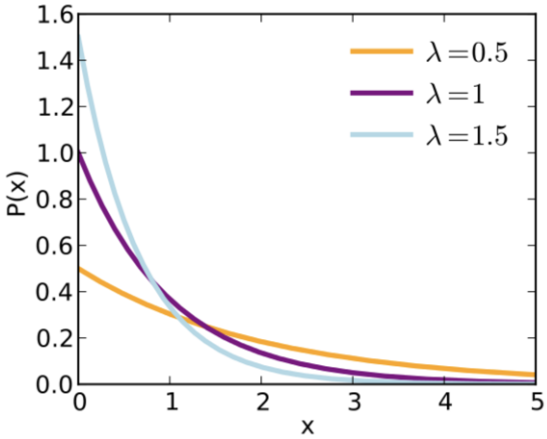
$$E(X_U) = 1/2$$

$$Var(X_U) = \frac{1}{12}$$

Relaciones con otras distribuciones, si $X \sim U(0,1)$:

- $Y = -\ln(X)/\lambda$ tiene una distribución exponencial con parámetro λ, es decir, $Y \sim Exp(\lambda)$

Distribución Exponencial:



Se utiliza para modelar tiempos de espera para la ocurrencia de un cierto evento.
Notación:

$$X \sim Exp(1/\lambda)$$

Función de densidad:

$$f_X(x) = \lambda e^{-\lambda x}, \text{ para } x \geq 0$$

Función de distribución:

$$F_X(x) = 1 - e^{-\lambda x}, \text{ para } x \geq 0$$

Valor esperado:

$$E(X) = \frac{1}{\lambda}$$

$$Var(X) = \frac{1}{\lambda^2}$$

<u>Relación con la distribución de Poisson:</u>

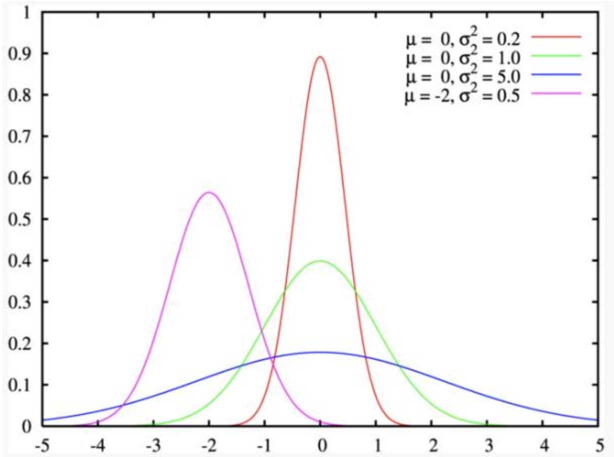
La exponencial se puede utilizar para modelar los tiempos de espera entre dos llegadas (o aciertos o eventos) de Poisson sucesivos, mientras que Poisson modela la probabilidad del número de llegadas (o eventos).

Dado un intervalo de tiempo $[0, T]$, la distribución exponencial mide el tiempo de espera (continuo) para la ocurrencia de eventos, cuyo número de ocurrencia (definido en un intervalo de tiempo $[0, T]$) está medido por una variable aleatoria $N \sim Poisson(\lambda)$:

Entonces, si N es el número de eventos que son probables de ocurrir en un intervalo de tiempo $[0, T], y\ X$ es el tiempo de espera hasta el siguiente evento, se tiene que:

$$N \sim Poisson(\lambda) \rightarrow \quad X \sim Exp\left(\frac{1}{\lambda}\right)$$

Distribución Normal:

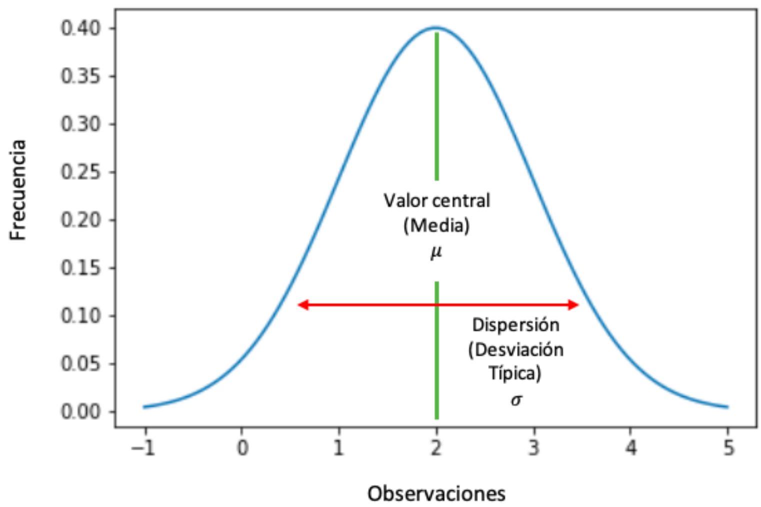


La función de distribución Normal juega un papel central en Estadística ya que, además de sus interesantes propiedades de reproductividad y de aproximación de otras distribuciones, sirve para modelizar una gran cantidad de situaciones prácticas.

Una v.a. X se dice que tiene una distribución de probabilidades Normal con media μ y varianza σ^2 y se denomina $N(\mu, \sigma^2)$ si su función de densidad es de la forma:

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

Cuando se dibuja la función de densidad en una gráfica, se observa que es simétrica respecto a μ.



La Normal estándar es la que tiene media 0 y varianza 1:

$$Z \sim N(0,1)$$

<u>Tipificación de la Normal:</u>

Cuando tenemos una v.a. $N(\mu, \sigma^2)$, para calcular las probabilidades y poder usar las tablas, se efectúa una transformación llamada tipificación (o estandarización) que la convierte en una Normal estándar $N(0,1)$:

$$X \sim N(\mu, \sigma^2) \quad \rightarrow \quad Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

A esta variable Z se le llama normalizada (o estandarizada o tipificada). De esta forma, solo es necesario disponer de la tabla correspondiente a la $N(0,1)$ para realizar los cálculos ya que:

$$P_X(X \leq x_0) = P_Z(Z \leq \frac{x_0 - \mu}{\sigma})$$

Teorema Central del Limite (TCL):

Sean $X_1, X_2, \dots, X_n$ un conjunto de variables aleatorias independientes e idénticamente distribuidas con media μ y varianza $\sigma^2 \neq 0$. Sea:

$$S_n = X_1 + \cdots X_n$$

Entonces el resultado de la suma de las n variables se puede aproximar por una distribución Normal con media $n\mu$ y varianza $n\sigma^2$ (desviación típica $\sigma\sqrt{n}$):

$$S_n \sim Normal\left(n\mu, \sigma\sqrt{n}\right)$$

Tipificando llegaríamos a que:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \sim N(0,1)$$

Aplicación del TCL: aproximación de la Binomial por la Normal

Supongamos que X es una v.a. con distribución de probabilidad binomial $B(n,p)$ y que el parámetro n tiende a infinito mientras que el parámetro p permanece constante. Entonces, como la binomial es el resultado de la suma de n Bernouillis independientes con probabilidad p, se puede aplicar el TCL. Es decir

$$X = \sum_{i=1}^{n} \delta_i, \text{ donde } \delta_i \sim Bernouilli\ (p)$$

Por lo cual, como $E(X) = np$ y $Var(X) = np(1-p)$, se tiene que la distribución de la variable aleatoria:

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

tiende a la función de distribución Normal estándar:

$$Z \sim N(0,1)$$

En la práctica se usa la aproximación cuando:

$$n \geq 30, \qquad np \geq 5, \qquad n(1-p) \geq 5$$