etsinf

UNIDAD DIDÁCTICA 4

DISTRIBUCIONES DE PROBABILIDAD – PARTE 2

Objetivo: El objetivo de esta Unidad Didáctica es introducir de forma resumida los conceptos elementales sobre las distribuciones de probabilidad y la esperanza matemática indispensables para el resto de la asignatura, así como presentar sucintamente los modelos más importantes en la práctica para variables aleatorias discretas (Binomial y Poisson) y continuas (Uniforme, Exponencial y Normal).

De acuerdo con el enfoque general de la asignatura, el tratamiento de los conceptos expuestos se realiza a un nivel elemental e intuitivo, pero suficiente para que un ingeniero pueda entenderlos y aplicarlos a los problemas reales que puedan aparecer en su ejercicio profesional.

Contenido

1. INTRODUCCIÓN Y CONCEPTOS BÁSICOS

    1.1 Variables aleatorias. Distribuciones de probabilidad

    1.2 Distribuciones de probabilidad discretas

    1.3 Distribuciones de probabilidad continuas

    1.4 Esperanza matemática

    1.5 Valor medio: concepto y propiedades

    1.6 Varianza: concepto y propiedades

2. PRINCIPALES DISTRIBUCIONES DE PROBABILIDAD.

    2.1 La distribución Binomial

    2.2 La distribución de Poisson

    2.3 La distribución de Uniforme

    2.4 La distribución Exponencial

    2.5 La distribución Normal

Fuente: Romero, R.; Zúnica, L. R. *Métodos estadísticos en ingeniería*. I.S.B.N 84-9705-727-9

2. Principales distribuciones de probabilidad (continuación)

En esta segunda parte de la Unidad Didáctica se presentan dos de los modelos de distribuciones para variables continuas más relevantes por su aplicación en el ámbito de la Informática: la distribución **Uniforme** y la distribución **Exponencial**.

2.3. La distribución de Uniforme

Una variable aleatoria **X** continua cuya función de densidad es constante en un intervalo (a,b) y nula fuera de dicho intervalo se dice que tiene una **Distribución Uniforme en (a,b)** y se simboliza como:

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

Por ejemplo, cualquier calculadora científica tiene una función (RAN o RND) que genera números aleatorios entre 0 y 1. Intuitivamente, la probabilidad que tiene cualquier número de ese intervalo en mostrarse cuando se llama a dicha función es la misma, esto es, la probabilidad se distribuye uniformemente entre 0 y 1.

Los números generados por esta función son valores de una variable aleatoria que sigue una **distribución Uniforme en (0,1)**

La distribución uniforme puede utilizarse también para modelar la pauta de variabilidad del tiempo de acceso a datos en unidades de discos, tiempo que será mayor o menor según la posición de la cabeza de lectura respecto a la situación del dato buscado.

También puede usarse para caracterizar el comportamiento del tamaño de un determinado tipo de fichero (entre 100 y 1000 Kb por ejemplo) generado por una aplicación.

Adicionalmente, la distribución uniforme es especialmente utilizada en simulación.

Se demuestra que la **función de densidad** que caracteriza a este tipo de variables (**Figura 6**) viene dada por la expresión:

$$\begin{aligned} f(x) &= K = \frac{1}{b-a} \quad x \in [a,\, b] \\ f(x) &= 0 \quad x \notin [a,\, b] \end{aligned}$$

La **probabilidad P(X≤x)** puede calcularse integrando la función de densidad f(x).

La **media** y la **varianza** de esta distribución se obtienen a partir de las expresiones de la Esperanza Matemática para variables continuas expuestas en los apartados 1.5 y 1.6 de esta Unidad Didáctica:

$$E(X) = \frac{a+b}{2} \qquad \sigma^2(X) = \frac{(b-a)^2}{12}$$

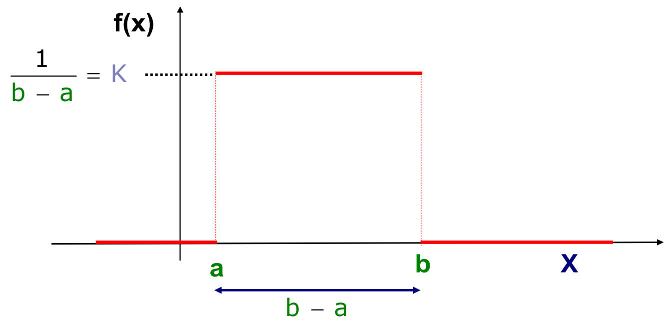

Figura 6. Función de densidad de una variable aleatoria X ~ U(a,b)

Ejercicio 11: el tiempo de acceso o búsqueda de un fichero en una antigua unidad de disco (HD) fluctúa **uniformemente** entre 0,1 y 0,5 segundos.

a) ¿Cuál es la variable aleatoria implicada en el estudio? ¿Qué distribución sigue?

b) ¿Cuál es la probabilidad de que el tiempo de acceso a un fichero sea exactamente 0,125 segundos?

c) ¿Cuál es la media del tiempo de acceso a un fichero? ¿Y la desviación típica?

d) ¿Cuál será el tiempo medio de acceso a 100 ficheros consecutivos? ¿Y la desviación típica?

2.4. La distribución Exponencial

Esta distribución describe la pauta de variabilidad de variables aleatorias continuas que hacen referencia a la duración o tiempo transcurrido entre dos sucesos consecutivos o al tiempo que transcurre hasta que se produce un determinado suceso o evento. Por ejemplo, el tiempo transcurrido entre dos fallos consecutivos de un sistema informático puede modelarse como una variable exponencial. Los valores que puede tomar este tipo de variables son siempre mayores o iguales a 0 (R^+).

La distribución exponencial depende del parámetro α. Nos referimos a ella como:

$$\boxed{X \sim Exp(\alpha)}$$

La distribución Exponencial es frecuentemente utilizada para modelar el comportamiento de variables que representan:

- Vida o duraciones de equipos y/o componentes **(Fiabilidad)** como el tiempo hasta el fallo en equipos industriales o el tiempo hasta el fallo de componentes electrónicos, …
- Tiempo que se tarda en realizar un proceso **(Sistemas de Colas)** como tiempo de espera en cola hasta que un proceso es atendido, tiempo de ejecución de una instrucción, tiempo que puede transcurrir en un servicio de urgencias hasta que es atendido un paciente,…

Ejemplo: La duración (X) hasta el fallo de un componente o un equipo es tal que la P(X>t) decrece exponencialmente a medida que aumenta t, como se muestra en la siguiente figura:

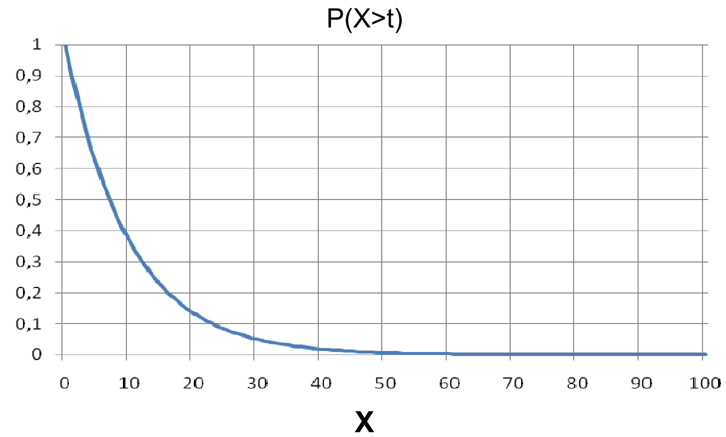

Figura 7. Probabilidad de que una variable aleatoria Exp(α) tome valores mayores que t

Así, se tiene que:

$$\mathbf{P(X > t) = e^{-\alpha t} \quad X \geq 0}$$

Por lo que:

$$\mathbf{P(X \leq t) = 1 - e^{-\alpha t} \quad X \geq 0 \quad (1)}$$

La expresión de la **función de densidad** de la **distribución Exponencial** es la siguiente (**Figura 8**):

$$\begin{array}{ll} \mathbf{f(x) = \alpha e^{-\alpha x}} & \mathbf{x \geq 0} \\ \mathbf{f(x) = 0} & \mathbf{x < 0} \end{array}$$

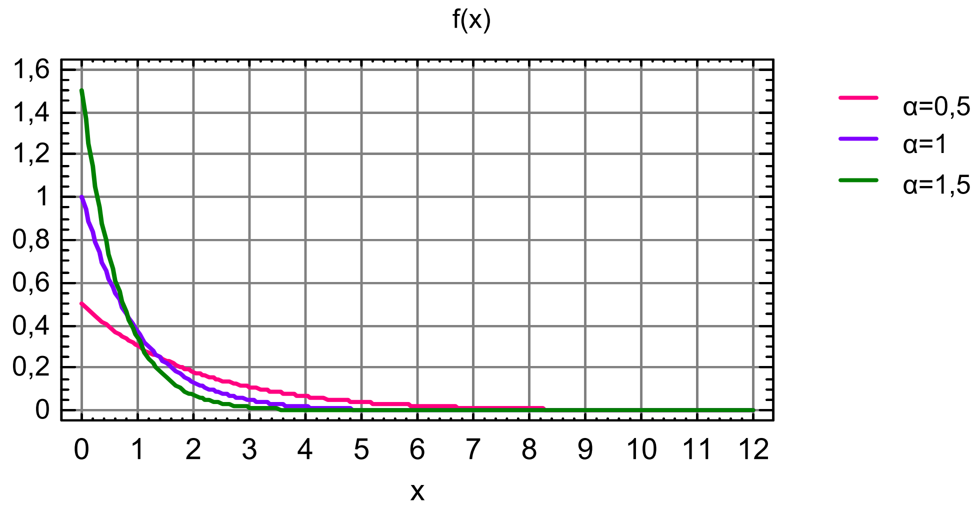

Figura 8. Función de densidad de una variable aleatoria exponencial para distintos valores de α

La probabilidad acumulada **P(X≤x)** (1), puede obtenerse sin más que integrar la función de densidad **f(x)** entre **0** y **x**.

La **media** y la **varianza** de esta distribución se obtienen a partir de las expresiones de la Esperanza Matemática para variables continuas expuestas en los apartados 1.5 y 1.6 de esta Unidad Didáctica:

$$\boxed{E(X) = \frac{1}{\alpha}} \qquad \boxed{\sigma^2(X) = \frac{1}{\alpha^2}}$$

Ejercicio 12: La duración T (horas de funcionamiento hasta que fallan o dejan de funcionar correctamente) de las pantallas LCD de gama media fluctúa exponencialmente. Se sabe que la vida media de las pantallas es de 40 años, suponiendo un funcionamiento de 4 horas al día.

a) ¿Cuál es la variable aleatoria implicada en el estudio? ¿Qué distribución sigue?

b) ¿Cuál es la duración media de las pantallas? ¿Y su desviación típica?

c) ¿Qué porcentaje de las pantallas durarán más de 10 años?

d) Suponiendo que las pantallas tienen 1 año de garantía, ¿Qué porcentaje de las mismas tendrán que usar la garantía?

e) ¿Cuántos años tienen que transcurrir para que el 50% de las pantallas dejen de funcionar correctamente? ¿A qué parámetro de los estudiados corresponde el valor obtenido?

La **propiedad fundamental** de la distribución exponencial es que la distribución condicional de $X/X>x_0$ no depende del valor x_0 sino que es igual a la distribución marginal inicial de X. Por esta propiedad se dice que la distribución exponencial **no tiene memoria**.

$$\boxed{P\left(X > x_0 + x \,/\, X > x_0\right) = P(X > x) = e^{-\alpha x}}$$

ya que

$$P\left(X > x_0 + x \,/\, X > x_0\right) = \frac{P(X > x_0 + x)}{P(X > x_0)} = \frac{e^{-\alpha(x_0+x)}}{e^{-\alpha x_0}} = e^{-\alpha x}$$

Ejercicio 13. El tiempo transcurrido hasta el fallo de un componente electrónico sigue una distribución exponencial cuya mediana es 69,3 horas.

a) ¿Qué porcentaje de los componentes tendrán una duración superior a 90 horas?

b) ¿Cuál es el tiempo de vida media de un componente?

c) Sabiendo que ya han transcurrido 300 horas sin que se produzca ningún fallo, ¿cuál es la probabilidad de que en las próximas 50 horas el componente funcione correctamente?

Ejercicio 14. La duración T (horas de funcionamiento hasta que fallan) de ciertos componentes electrónicos fluctúa aleatoriamente siguiendo una distribución exponencial. Se sabe que las componentes duran en promedio 400 horas. Estos componentes se utilizan para el montaje de un dispositivo electrónico (D) que debe conectar dos bornes A y B. Se desea que el dispositivo tenga una fiabilidad de al menos 99% a las 400 horas.

Con el fin de garantizar esta fiabilidad se montan en paralelo N componentes del tipo estudiado. ¿Cuánto debe valer como mínimo N?

Ejercicios resueltos

Apartado 5.A.1 del Capítulo 5 del libro de R. Romero y L.R. Zúnica "Métodos Estadísticos en Ingeniería" SPUPV 637

Apartado 6.A.1 del Capítulo 6 del libro de R. Romero y L.R. Zúnica "Métodos Estadísticos en Ingeniería" SPUPV 637

Apartados 7.A.1 y 7.A.2 del Capítulo 7 del libro de R. Romero y L.R. Zúnica "Métodos Estadísticos en Ingeniería" SPUPV 637

Ver boletín correspondiente en PoliformaT (EST GII: Recursos / 04 | Ejercicios)

Para saber más

- Descartes: Distribución Uniforme. Instituto de Tecnologías Educativas (ITE) del Ministerio de Educación: http://descartes.cnice.mec.es/materiales_didacticos/distribuciones_probabilidad/ejemplo_cont.htm

- Descartes: Distribución Exponencial. Instituto de Tecnologías Educativas (ITE) del Ministerio de Educación: http://descartes.cnice.mec.es/materiales_didacticos/distribuciones_probabilidad/ejemplo_cont.htm#2

- The Gauss distribution: suma de n variables aleatorias uniformemente distribuidas (*Franz Embacher and Petra Oberhuemer;Maths Online*): http://www.univie.ac.at/future.media/moe/galerie/wstat1/wstat1.html

Fuentes

Métodos Estadísticos en Ingeniería (Romero Villafranca, Rafael)