Clase 10

Repaso de Probabilidades y Estadística

Basado en repaso por Licesio J. Rodriguez-Aragón
U. de Castilla La Mancha

Probabilidades

- Definición de Probabilidad: es la razón entre el número de casos favorables de un suceso y el número de casos posibles. Dado un suceso **A** la probabilidad **P(A)** de que suceda **A** es:

$$P(A) = \frac{\text{numero de casos favorables a A}}{\text{numero de casos posibles}}$$

- Por definición se cumple que:

$$0 \leq P(A) \leq 1$$

Probabilidad

• Interpretación Frecuentista: la probabilidad de que durante un experimento suceda **A** es la frecuencia relativa con la que sucede **A** si uno realiza un número infinito de repeticiones independientes del experimento. Si **A** sucede un número $\mathbf{n_A}$ de veces en **n** repeticiones del experimento entonces:

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

• Interpretación Bayesiana: la probabilidad de un evento es una medida del grado de creencia que se tiene de que éste suceda, dada la información disponible. **P(A)** puede cambiar al recibirse mayor información.

Probabilidad Condicional

- La probabilidad de un suceso **A** puede verse afectada por la ocurrencia de un suceso **B**. Si la probabilidad de **A** no depende de la ocurrencia de **B** y viceversa decimos que los sucesos **A** y **B** son independientes.

- En general la probabilidad condicional de **A** dado **B** es:

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$

- Por lo tanto **A** y **B** son independientes si y solo si:

$$P(A \cap B) = P(A) \cdot P(B)$$

Probabilidad Total

- Si **{A₁,...,Aₙ}** son un conjunto de sucesos exhaustivos y mutuamente excluyentes:

$$\boxed{\sum_{i=0}^{n} P(A_i) = 1} \quad \wedge \quad \boxed{A_i \cap A_j = \emptyset \quad \forall i \neq j}$$

entonces se cumple que la probabilidad total de un suceso **B** es:

$$\boxed{P(B) = \sum_{i=0}^{n} P(B|A_i) \cdot P(A_i)}$$

Ejemplo

- Tres máquinas **M_1**, **M_2** y **M_3** fabrican piezas con una producción de 300, 450 y 600 piezas por hora, las cuales se almacenan de forma conjunta en el mismo almacén. Los porcentajes de piezas defectuosas que fabrica cada máquina son 2%, 3.5% y 2.5% respectivamente. ¿Cual es la probabilidad **P(D)** de elegir al azar una pieza defectuosa desde el almacén?

$$P(D) = \sum_{i=1}^{3} P(D|M_i) \cdot P(M_i)$$

- La probabilidades que necesitamos saber son:

$$P(M_1) = \frac{300}{300+450+600} = \frac{300}{1350} \;\; ; \;\; P(M_2) = \frac{450}{1350} \;\; ; \;\; P(M_3) = \frac{600}{1350}$$

$$P(D|M_1) = 0.02 \;\; ; \;\; P(D|M_2) = 0.035 \;\; ; \;\; P(D|M_3) = 0.025$$

- Por lo tanto:

$$P(D) = \frac{300 \cdot 0.02 + 450 \cdot 0.035 + 600 \cdot 0.025}{1350} = 0.0272$$

Teorema de Bayes

- Dada la definición de probabilidad condicional para dos sucesos **A** y **B** se cumple que

$$P(A|B) \cdot P(B) = P(A \cap B) = P(B \cap A) = P(B|A) \cdot P(A)$$

- Por lo tanto:

$$\boxed{P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}}$$

Ejemplo

• Una farmaceutica quiere vender un test de VIH. Los estudios clínicos muestran que si una persona está infectada (lo que ocurre en un 1% de la población) el test da positivo en un 99% de los casos, y que si una persona no está infectada el test da positivo en un 5% de los casos. ¿Es razonable vender este test al público?

• Dado que un test sale positivo, calculemos la probabilidad de tener VIH (**$P(I|T^+)$**) y de no tener VIH (**$P(I^c|T^+)$**) usando el teorema de Bayes y la definición de probabilidad total:

$$P(I|T^+) = \frac{P(T^+|I) \cdot P(I)}{P(T^+)} = \frac{P(T^+|I) \cdot P(I)}{P(T^+|I) \cdot P(I) + P(T^+|I^c) \cdot P(I^c)}$$

$$P(I^c|T^+) = \frac{P(T^+|I^c) \cdot P(I^c)}{P(T^+)} = \frac{P(T^+|I^c) \cdot P(I^c)}{P(T^+|I) \cdot P(I) + P(T^+|I^c) \cdot P(I^c)}$$

Ejemplo

- Las probabilidades que necesitamos son:

$$P(I) = 0.01 \;\; ; \;\; P(I^c) = 1 - P(I) = 0.99$$

$$P(T^+|I) = 0.99 \;\; ; P(T^+|I^C) = 0.05$$

- Reemplazando estos valores obtenemos que:

$$P(I|T^+) = \frac{1}{6} \quad ; \quad P(I^c|T^+) = \frac{5}{6}$$

- Una persona al azar que se hace el test y le sale positivo tiene un 83% de probabilidad de estar sano!!!!!!!!

Estadística

- Variables Aleatorias: en todo proceso de observación o experimento podemos definir una variable aleatoria asignando a cada resultado del experimento un número:

    - Resultado Numérico.
    - Etiqueta numérica asignada a casa resultado.

- Variable Aleatoria Discreta: puede tomar un número finito o infinito numerable de valores distintos.

    - e.g. lanzamiento de un dado o número que sale en una ruleta.

- Variable Aleatoria Continua: puede tomar un número infinito o no numerable de valores distintos.

    - e.g. distancia del centro del tablero a la que cae un dardo.

Función de Distribución de Probabilidad

- Dada una variable aleatoria discreta **X**, se define su función de distribución de probabilidad **f(x$_i$)** como la probabilidad de que **X** tome el valor **x$_i$**.

$$f(x_i) = P(X = x_i)$$

- La función de distribución de probabilidad cumple con:

$$0 \leq f(x_i) \leq 1$$

$$\sum_i f(x_i) = 1$$

- Se define la distribución cumulativa de probabilidad **F(x$_i$)** como:

$$F(x_i) = \sum_{j \leq i} f(x_j) = P(X \leq x_i)$$

Función de Distribución de Probabilidad

- Momentos de una Distribución de Probabilidad:

  Media o Esperanza (valor esperado de una distribución de probabilidad):

$$\mu = \langle X \rangle = \sum_i x_i f(x_i)$$

  Varianza (ancho de una distribución de probabilidad):

$$\sigma^2 = \langle (X - \langle X \rangle)^2 \rangle = \sum_i (x_i - \mu)^2 f(x_i)$$

Función de Distribución de Probabilidad

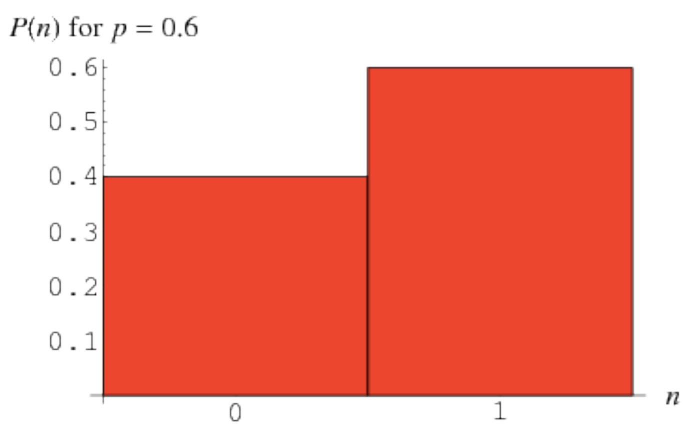
- Ejemplo: Distribución de Bernoulli

Resultado de un experimento con dos resultados posibles (**X={0,1}**).

$$P(X=1)=p \quad ; \quad P(X=0)=1-p=q$$

$$\boxed{f(x)=p^x q^{1-x}}$$

$$\mu=p \quad ; \quad \sigma^2=p(1-p)=pq$$



Función de Densidad de Probabilidad

- Dada una variable aleatoria continua **X**, se define su función de densidad de probabilidad **f(x)** como la probabilidad de que **X** tome el valor **x**.

- NO ES UNA PROBABILIDAD!!!! ES UNA DENSIDAD DE PROBABILIDAD. Solo al integrarla obtenemos probabilidades.

$$f(x)dx = P(x - \frac{dx}{2} < X < x + \frac{dx}{2})$$

- La función de densidad de probabilidad cumple con:

$$0 \leq f(x) \leq 1$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

- Se define la distribución cumulativa de probabilidad **F(x_i)** como:

$$F(x) = \int_{-\infty}^{x} f(x')dx' = P(X \leq x)$$

Función de Distribución de Probabilidad

- Momentos de una Distribución de Probabilidad:

  Media (valor esperado de una densidad de probabilidad):

$$\mu = \langle X \rangle = \int_{-\infty}^{\infty} x f(x) dx$$

  Varianza (ancho de una densidad de probabilidad):

$$\sigma^2 = \langle (X - \langle X \rangle)^2 \rangle = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

Función de Densidad de Probabilidad

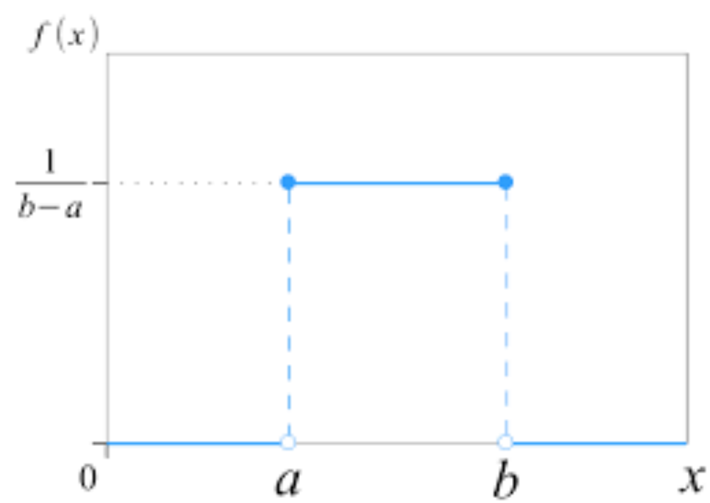
- Ejemplo: Distribución Uniforme

La variable aleatoria X tiene una probabilidad constante en el rango a<x<b y cero fuera de este rango.

$$\boxed{f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{en el resto} \end{cases}}$$

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

$$\mu = \frac{a+b}{2}; \quad \sigma^2 = \frac{(b-a)^2}{12}$$

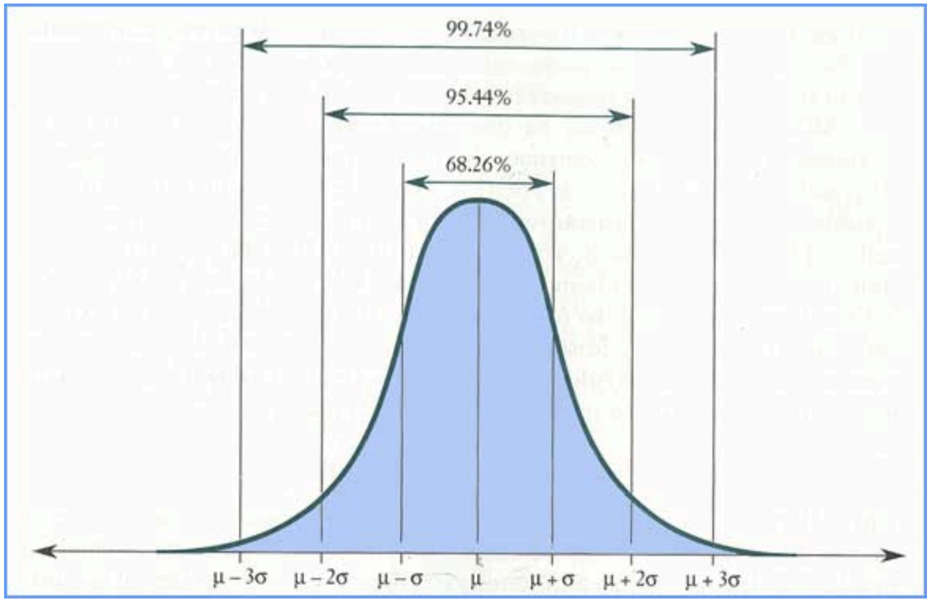


Función de Densidad de Probabilidad

- Ejemplo: Distribución Gaussiana o Normal

Una variable X tiene una distribución Normal ***N*(μ,σ)** si tiene una función de densidad de probabilidad:

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}$$



Descubierta por Gauss como la función que describe la distribución de valores obtenidos de mediciones experimentales sujetas a errores de medición no sistemáticos. De especial importancia para la Termodinámica.