

CURSO PROPEDÉUTICO PROBABILIDAD

Marzo 2021

1.- INTRODUCCIÓN

- En 1933, el matemático ruso Andrey Nikolaevich Kolmogorov (1903-1987) propone ciertos axiomas que a la postre resultaron adecuados para la construcción de la teoría de la probabilidad.
- Esta teoría prevalece hoy en día y ha adquirido el calificativo de **teoría clásica**. Su importancia es la utilidad para resolver problemas puramente matemáticos, pero principalmente, para **modelar situaciones donde el azar es relevante**.

Probabilidad axiomática. En la definición axiomática de la probabilidad no se establece la forma explícita de calcular las probabilidades sino únicamente se proponen las reglas que el cálculo de probabilidades debe satisfacer.



A. N. Kolmogorov

Axiomas de la probabilidad

1. $P(A) \geq 0$. La probabilidad de cualquier evento aleatorio $A \in \Omega$ es un real finito no negativo.
2. $P(\Omega) = 1$. La probabilidad de ocurrencia de los sucesos elementales Ω es 1.
3. $P(A \cup B) = P(A) + P(B)$, cuando $A \cap B = \emptyset$. La probabilidad de la unión de sucesos disjuntos es la suma de sus respectivas probabilidades.

Life is just a long random walk

CONJUNTOS

UNIÓN DE CONJUNTOS. Sean **A** y **B** dos subconjuntos cualesquiera del conjunto universal. La unión de **A** y **B**, expresada por $A \cup B$, es el conjunto de todos los elementos que pertenecen a **A** o pertenecen a **B**.

$$A \cup B = \{x \mid x \in A \text{ o } x \in B\}$$

INTERSECCIÓN DE CONJUNTOS. Sean **A** y **B** dos conjuntos cualesquiera del conjunto universal. La intersección de **A** y **B**, expresada por $A \cap B$, es el conjunto de todos los elementos que pertenecen a **A** y a **B** simultáneamente, es decir:

$$A \cap B = \{x \mid x \in A \text{ y } x \in B\}$$

DIFERENCIA DE CONJUNTOS O COMPLEMENTO RELATIVO. Sean **A** y **B** dos conjuntos cualesquiera del conjunto universal. La diferencia o complemento relativo de **B** con respecto a **A**, es el conjunto de los elementos que pertenecen a **A**, pero no pertenecen a **B**.

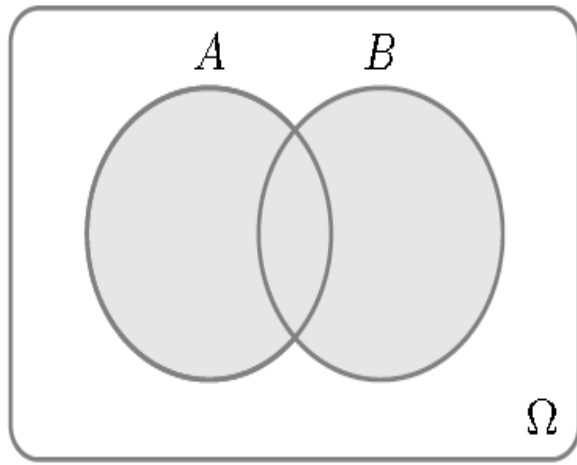
$$A - B = \{x \mid x \in A, x \notin B\}$$

Nota: $A - B \neq B - A$

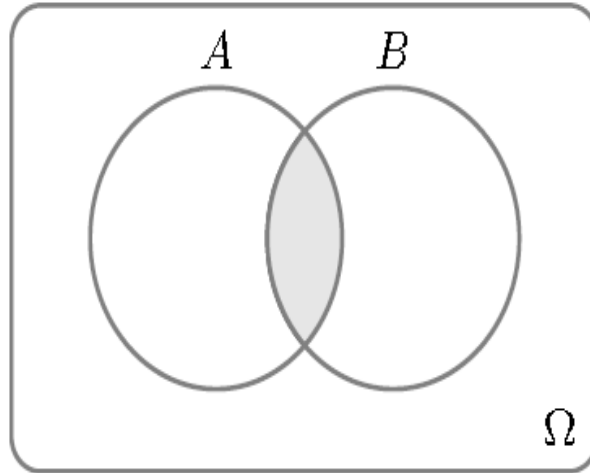
COMPLEMENTO ABSOLUTO O SIMPLEMENTE COMPLEMENTO. Sea **A** un subconjunto cualesquiera del conjunto universal. El complemento de **A** es el conjunto de elementos que perteneciendo al universo y no pertenecen al conjunto **A**, denotado por A' o A^c .

$$A' = \{x \mid x \in U, x \notin A\}$$

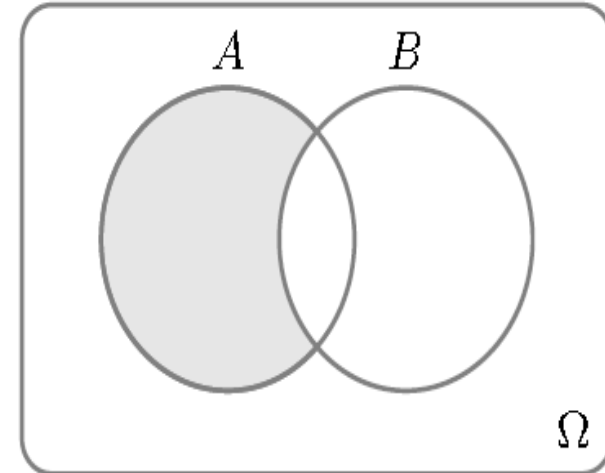
Nota: $A' = U - A$



$$A \cup B$$



$$A \cap B$$



$$A - B$$

RELACIONES ENTRE CONJUNTOS

IGUALDAD DE CONJUNTOS

Considerando el conjunto A y el conjunto B, si ambos tienen los mismos elementos, es decir, si cada elemento que pertenece a A también pertenece a B y si cada elemento que pertenece a B pertenece también a A.

$$A = B$$

SUBCONJUNTO

Si todo elemento de un conjunto A es también elemento de un conjunto B, entonces se dice que A es un subconjunto de B. Esto se representa mediante: $A \subset B$

LEYES DE CONJUNTOS

DE IDEMPOTENCIA

$$\mathbf{A \cup A = A} \quad \mathbf{A \cap A = A}$$

ASOCIATIVA

$$\mathbf{(A \cup B) \cup C = A \cup (B \cup C)}$$

$$\mathbf{(A \cap B) \cap C = A \cap (B \cap C)}$$

CONMUTATIVA

$$\mathbf{A \cup B = B \cup A} \quad \mathbf{A \cap B = B \cap A}$$

DISTRIBUTIVA

$$\mathbf{A \cup (B \cap C) = (A \cup B) \cap (A \cup C)}$$

$$\mathbf{A \cap (B \cup C) = (A \cap B) \cup (A \cap C)}$$

DE IDENTIDAD

$$\mathbf{A \cup U = U} \quad \mathbf{A \cap U = A}$$

$$\mathbf{A \cup \emptyset = A} \quad \mathbf{A \cap \emptyset = \emptyset}$$

DE INVOLUCIÓN

$$\mathbf{(A')' = A}$$

DE COMPLEMENTO

$$\mathbf{A \cup A' = U} \quad \mathbf{A \cap A' = \emptyset}$$

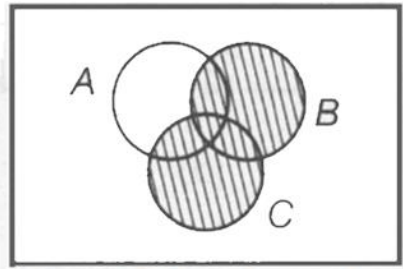
$$\mathbf{U' = \emptyset} \quad \mathbf{\emptyset' = U}$$

D'MORGAN

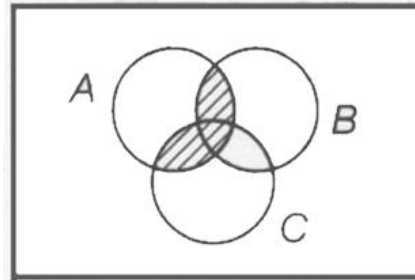
$$\mathbf{(A \cup B)' = A' \cap B'} \quad \mathbf{(A \cap B)' = A' \cup B'}$$

Propiedad distributiva

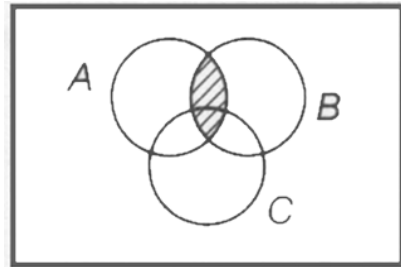
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$



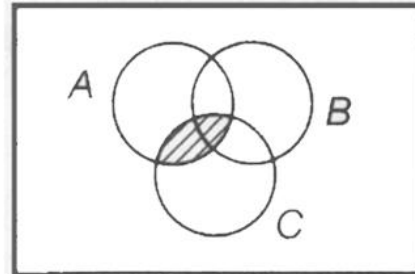
(a) Shaded region: $B \cup C$



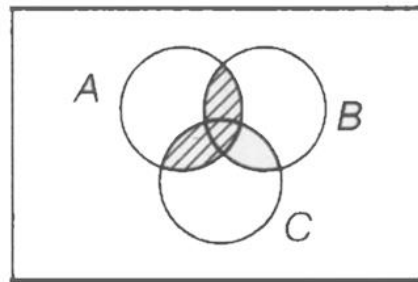
(b) $A \cap (B \cup C)$



(c) $A \cap B$



(d) $A \cap C$



(e) $(A \cap B) \cup (A \cap C)$

- Un **experimento** es el proceso por medio del cual se hacen observaciones
- El **espacio muestra** asociado con un experimento es el conjunto formado por todos los posibles puntos muestrales. Un espacio muestral o conjunto universo se denota por S .
- Cuando se realiza un experimento, se puede obtener uno o más resultados que se llaman **eventos**. Un evento en un espacio de puntos muestrales, es decir, es un subconjunto de S

Sea un **experimento** lanzar un dado honesto para anotar el número de la cara resultante, con $S = \{1, 2, 3, 4, 5, 6\}$. **Eventos** de interés podrían ser

- a) *Que caiga 1*
- b) *Que el número sea impar*

Sea un **experimento** para determinar el conjunto de personas que esperan en el andén de una estación del metro. Algunos **eventos** de interés podrían ser:

- a) *Que haya exactamente 240 personas.*
- b) *Que haya más de 200 personas.*
- c) *El número de personas que esperan sea de entre 240 y 780.*

Algunas propiedades de la probabilidad

a) $P(A^c) = 1 - P(A)$.

b) $P(\emptyset) = 0$.

c) Si $A \subseteq B$, entonces $P(A) \leq P(B)$.

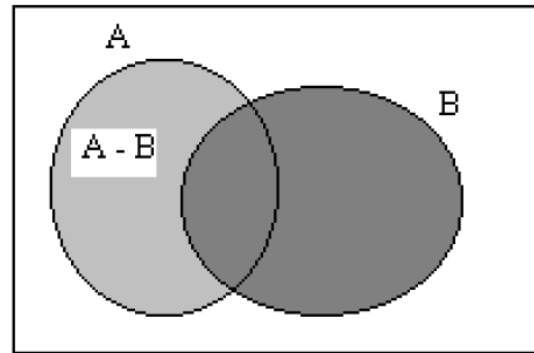
d) Si $A \subseteq B$, entonces $P(B - A) = P(B) - P(A)$.

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

f) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

g) $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$.

h) $P(A - B) = P(A) - P(A \cap B)$

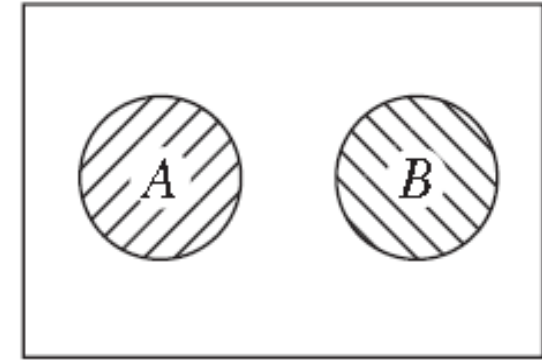


$$P(A - B) = P(A) - P(A \cap B)$$

$$P\left[\bigcup_{k=1}^n A_k\right] = \sum_{j=1}^n P[A_j] - \sum_{j < k} P[A_j \cap A_k] + \dots + (-1)^{n+1} P[A_1 \cap \dots \cap A_n].$$

EVENTOS MUTUAMENTE EXCLUYENTES:

- Llamados también **disjuntos** o **ajenos**.
- Para dos eventos A y B , la ocurrencia de uno excluye la ocurrencia del otro.
- No tienen elementos en común: $A \cap B = \emptyset$.



Ejemplo 1: Sean A y B eventos **ajenos** tales que $P(B) = 0.3$ y $P(A \cap B^c) = 0.2$.
Compruebe que $P(A \cup B) = 0.5$. •

EVENTOS INDEPENDIENTES:

- Cuando la ocurrencia o no ocurrencia de un evento **no afecta** la ocurrencia del otro evento.
- Los eventos **A y B** son independientes si $P(A \cap B) = P(A) P(B)$
- Los eventos **A, B, C** son independientes si se cumplen:
 - $P(A \cap B \cap C) = P(A) P(B) P(C)$,
 - $\{A, B\}$ son independientes, así como $\{A, C\}$ y $\{B, C\}$.

Ejemp 2.- En una caja hay 5 libros de Arte, 3 de Química y 4 Comics. Aleatoriamente se extrae un libro, se anota su tipo y se **regresa** a la caja.

Al considerar **independencia** en la selección de libros,

¿cuál es la probabilidad de que al extraer 3 libros en **forma consecutiva** los 3 sean Comics?

EVENTOS DEPENDIENTES:

Si los eventos A y B **están relacionados** de tal modo que la ocurrencia de B **depende** de la ocurrencia de A.

Regla del producto

- $P(A \cap B) = P(A, B) = P(B / A) P(A)$, y
- $P(B \cap A) = P(B, A) = P(A / B) P(B)$

Ejemp 3.- En una caja hay 5 libros de Arte, 3 de Química y 4 Comics.
Si se extraen al azar 2 libros en **forma consecutiva y sin reemplazo**.

Sin considerar independencia,

¿Cuál es la probabilidad de que ambos sean Comics?

Ejemplo 4:

En la fabricación de circuitos integrados resultan algunos defectos:

- el **5%** de la producción no lleva impreso el logotipo a color,
- el **7%** tiene incompleto el logotipo,
- el **2%** tiene ambos defectos.

Considera los eventos: $A = \{\text{circuito con defecto de color}\}$, y $B = \{\text{circuito con logotipo incompleto}\}$.

Si se elije un circuito al azar, Cuál es la probabilidad de que

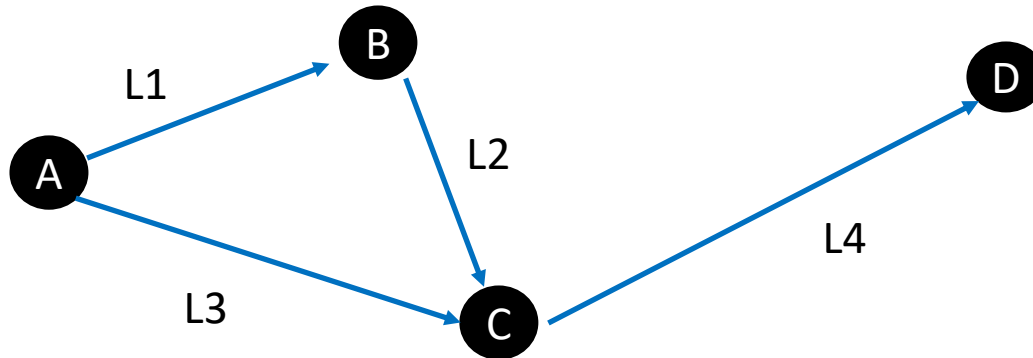
- a) Tenga **al menos** uno de los defectos
- b) Tenga **solo** el defecto A
- c) Tenga **solo uno** de entre los **dos defectos**
- d) **No** tenga defectos

INDEPENDENCIA Y EXCLUSIVIDAD

Ejemplo 6:

Sean los **nodos** A, B, C y D de una red de comunicaciones, con las **líneas de enlace** L1, L2, L3 y L4.

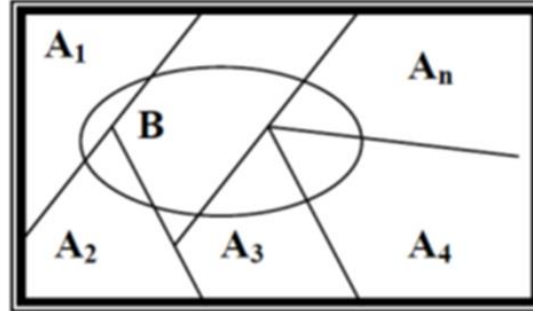
- La probabilidad de disponibilidad de cada enlace es $p(L_i) = p$.
- Para enviar un mensaje de A a D se dispone de **dos** posibles trayectorias, T1 Y T2.
- Considera **independencia** en la disponibilidad de cada enlace.
- Hay **independencia** entre T1 y T2.



¿Cuál es la probabilidad de que la red esté disponible para enviar un mensaje en un instante dado? $P(A \rightarrow D) = ?$

BAYES

Si los eventos $A_1, A_2, A_3, \dots, A_n$, definen a un conjunto universo S ; y si estos eventos A_i son mutuamente excluyentes (su unión es S). Sea B otro evento, tal que:



$$B = S \cap B = (A_1 \cup A_2 \cup A_3 \dots \cup A_n) \cap B$$

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) + \dots + P(A_n \cap B)$$

Luego por la regla de multiplicación:

$$P(B) = P(A_1) P(B/A_1) + P(A_2) P(B/A_2) + P(A_3) P(B/A_3) \dots P(A_n) P(B/A_n)$$

Si $A_1, A_2, A_3, \dots, A_n$ es una partición de S , y B es cualquier evento. Entonces para cualquier i ,

$$P(A_i / B) = \frac{P(A_i)P(B/A_i)}{P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + \dots + P(A_n)P(B / A_n)}$$

Es decir:

$$P(A_i / B) = \frac{P(A_i)P(B/A_i)}{\sum P(A_i)P(B/A_i)}$$

PROBABILIDAD TOTAL

Si los sucesos A_1, A_2, \dots, A_n constituyen una partición del Ω , tal que $P(A_i) > 0, \forall i$, tendremos que cualquier suceso B podrá particionarse de la forma, $B = \cup_{i=1}^n B \cap A_i$ y tratándose de una unión disjunta podremos escribir

$$P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i)P(A_i).$$

Este resultado se conoce con el nombre de *teorema de la probabilidad total*.

Ejemplo 7:

Una gran población se divide entre quienes son alumnos (A) y quienes no lo son (\bar{A}). Sean las **probabilidades conjuntas**:

P(A, G)	Con gripa (G)	Sin gripa (\bar{G})
Alumno (A)	0.5	0.2
No alumno (\bar{A})	0.1	0.2

Determina:

- $P(A \cap G)$
- $P(A)$ y $P(G)$
- ¿Son A y G eventos independientes?

Bayes

$$\begin{aligned} P(A_i | B) &= \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i \cap B)}{\sum_j P(A_j \cap B)} \\ &= \frac{P(B | A_i)P(A_i)}{P(B)} = \frac{P(B | A_i)P(A_i)}{\sum_j P(A_j \cap B)} \end{aligned}$$

Ejemp 8.- Desarrola las siguientes expresiones

- a) $P(B \cap A)=?$
- b) $P(C \cap B \cap A)=?$
- c) $P(D \cap C \cap B \cap A)=?$
- d) $P[(A \cup C) | B]= ?$

Ejemp 8.- Se tienen dos cajas que contienen smartphones en la proporción siguiente

	Caja A	Caja B	Total
Iphone	3	2	5
Samsung	2	8	10
Total	5	10	15

Se arroja una moneda.

- Si se obtiene **águila** se extrae aleatoriamente un teléfono de la **caja A**;
- Si se obtiene **sol** se extrae un teléfono de la **caja B**.

Dado que

- **R** indica el evento “extraer un Iphone”
- **A** y **B** indican los eventos escoger **caja A** y **caja B**, respectivamente.

Determina **la probabilidad** de extraer **un Iphone** en los siguientes casos:

- Caso de una moneda honesta;
- Cuando $P(\text{águila})= 0.2$ y $P(\text{sol})= 0.8$;
- Al emplear una moneda honesta, ¿Cuál es la probabilidad de **seleccionar la caja A** dado que se ha extraído un **teléfono R**?

Ejemplo 9:

En una gran población con **igual distribución** de mujeres y hombres,

- El 4% de los hombres tiene astigmatismo (la probabilidad de padecer astigmatismo es de 0.04 en los hombres)
- El 1% de la mujeres tiene astigmatismo (la probabilidad de padecer astigmatismo es de 0.01 en las mujeres)

Sean los eventos

$M = \{\text{mujer}\}$,

$H = \{\text{hombre}\}$, y

$D = \{\text{la persona padece de astigmatismo}\}$.

Si una persona es elegida aleatoriamente, determina

- a) La probabilidad de que padezca astigmatismo, dado que se ha escogido a una mujer.
- b) $P(H)$ y $P(M)$
- c) La probabilidad de que padezca astigmatismo

BAYES

Ejemp 10.-

La probabilidad de que un alumno estudie para un examen es de 0.2.

Dado que el alumno ha estudiado la probabilidad de que apruebe es de 0.8, y si no estudia la probabilidad de que apruebe es de 0.5.

Sean los eventos

$E = \{\text{El alumno ha estudiado}\},$

$\bar{E} = \{\text{El alumno no ha estudia}\},$

$A = \{\text{Aprobar el examen}\}.$

Determina

a) $P(A \cap E)$

b) La probabilidad de que el alumno apruebe el examen

c) Dado que el alumno aprobó el examen, ¿Cuál es la probabilidad de que haya estudiado?

Ejemplo 11: Probabilidad condicional y conjunta

En un aeropuerto, la probabilidad

- De que un avión de línea económica despegue a tiempo es de 0.83,
- De que llegue a tiempo es de 0.82, y
- De que despegue y que además llegue a tiempo es de 0.78.

Sean los eventos

$D = \{\text{despegar a tiempo}\}$

$A = \{\text{llegar a tiempo}\}$

Determina la probabilidad de que el avión:

- a) Llegue a tiempo, dado que despegó a tiempo.
- b) Haya despegado a tiempo, dado que llegó a tiempo

Ejemplo 13.-

En una ciudad que tiene **igual número** de personas que trabajan en sectores de administración, comercio, salud y servicio municipal, se encontró que son **mujeres** el **35 % de los administrativos**, el **25 % de los comerciantes**, el **20 % del servicio de salud** y el **15 % del servicio municipal**.

Sean los eventos:

$A = \{\text{Administrativo}\}$,

$B = \{\text{Comercio}\}$

$C = \{\text{Salud}\}$,

$D = \{\text{Servicio municipal}\}$

$M = \{\text{Mujer}\}$,

$\bar{M} = \{\text{Hombre}\}$

Determina

a) $P(A)$, $P(B)$, $P(C)$ y $P(D)$

b) $P(M | A)$, $P(M | B)$

c) La probabilidad de que una persona escogida aleatoriamente sea **mujer**

d) La probabilidad que una mujer escogida al azar, sea administrativa

e) La probabilidad que una persona escogida aleatoriamente, sea hombre

2.- VARIABLE ALEATORIA

Sea la variable aleatoria X , se define a la función de distribución de probabilidad como $P(X \leq x) = F(x)$, para $-\infty < x < \infty$, siendo $F(x)$ una función monótona no decreciente.

Toda función de distribución $F(x)$ satisface las siguientes propiedades:

a) $0 \leq F(x) \leq 1.$

b) $\lim_{x \rightarrow \infty} F(x) = 1.$

c) $\lim_{x \rightarrow -\infty} F(x) = 0.$

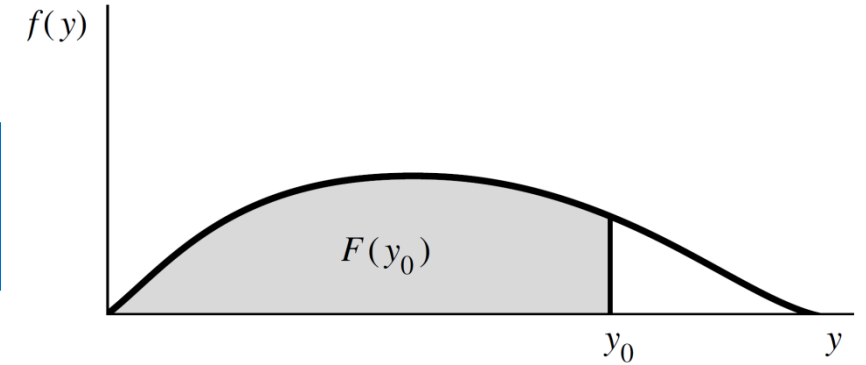
d) Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2).$

e) Si $x_1 \leq x_2$, entonces $P(x_1 < X \leq x_2) = F(x_2) - F(x_1).$

Para calcular $P(Y \leq y) = F(y)$:

$$F(y) = \int_{-\infty}^y f(t) dt$$

$$f(y) = \frac{dF(y)}{dy}$$



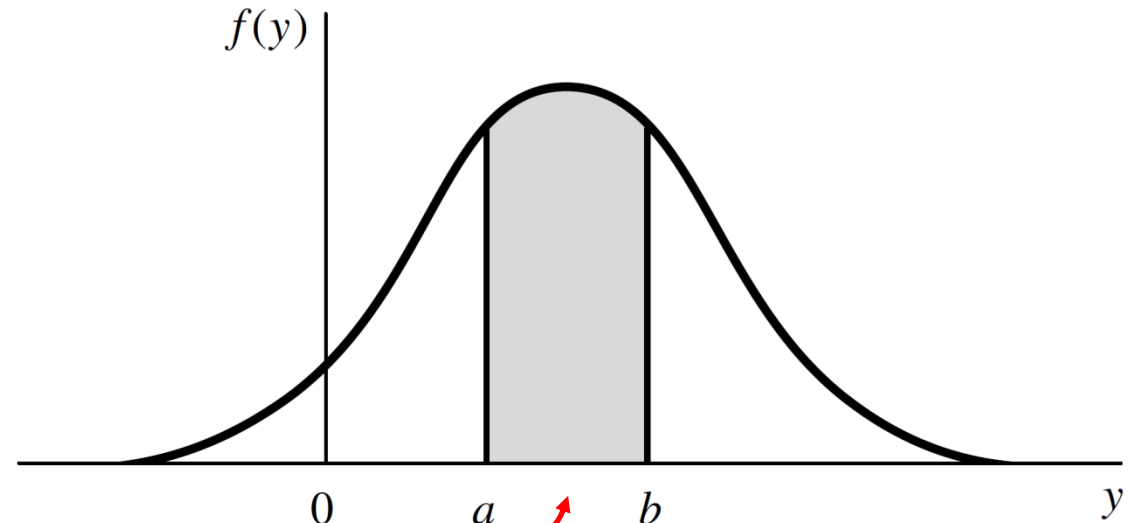
Propiedades de una función de densidad

1. $f(y) \geq 0$ para toda y , $-\infty < y < \infty$
2. $\int_{-\infty}^{\infty} f(y) dy = 1$.

Probabilidad en un rango:

$$\begin{aligned} P(a < Y \leq b) &= P(Y \leq b) - P(Y \leq a) \\ &= F(b) - F(a) = \int_a^b f(y) dy \end{aligned}$$

$$P(a \leq Y \leq b) = \int_a^b f(y) dy$$



Ejemp 20.-Dada la función de densidad continua $f(y)$ determina el valor de c para el cual $f(y)$ es una función de densidad válida.

$$f(y) = \begin{cases} cy^2, & 0 \leq y \leq 2 \\ 0, & \text{caso contrario} \end{cases}$$

$$\begin{aligned} F(\infty) &= \int_{-\infty}^{\infty} f(y) dy = 1 \\ &= \int_0^2 cy^2 dy = \left. \frac{cy^3}{3} \right]_0^2 = \left(\frac{8}{3} \right) c \\ & \quad (8/3)c = 1 \\ & \quad c = 3/8 \end{aligned}$$

VALOR ESPERADO, MEDIA O ESPERANZA

El valor esperado se usa como una medida de agrupamiento o de tendencial central de una distribución de probabilidad de una v.a.

Sea X una v.a con función de probabilidad o función de densidad $f(x)$. Sea $g(X)$ una función de la v.a X . El valor esperado de X , simbolizado por $E(X)$ es:

$$E[g(X)] = \begin{cases} \sum_x h(x) \cdot f(x) & \text{si } X \text{ es una v.a discreta} \\ \int_{-\infty}^{+\infty} h(x) \cdot f(x) dx & \text{si } X \text{ es una v.a continua} \end{cases}$$

MOMENTOS

Respecto al origen

$$E(X^n) = \begin{cases} \int_{-\infty}^{\infty} x^n f(x) dx, \\ \sum_x x^n f(x). \end{cases}$$

Respecto a la media

$$E[(X - \mu)^n] = \begin{cases} \int_{-\infty}^{\infty} (x - \mu)^n f(x) dx, \\ \sum_x (x - \mu)^n f(x). \end{cases}$$

1) Si $g(X) = X$, entonces se está calculando la esperanza de la v.a X

$$E(X) = \begin{cases} \sum_x x \cdot f(x) & \text{si } X \text{ es una v.a discreta} \\ \int_{-\infty}^{+\infty} x \cdot f(x) dx & \text{si } X \text{ es una v.a continua} \end{cases}$$

$$E(X) = \mu$$

2) Si $g(x) = (X - \mu)^2$, entonces $E[g(x)]$ se llama **varianza** de la v.a X y se simboliza como σ^2

$$\sigma^2 = E[g(X)] = \begin{cases} \sum_x (X - \mu)^2 \cdot f(x) & \text{si } X \text{ es una v.a discreta} \\ \int_{-\infty}^{+\infty} (X - \mu)^2 \cdot f(x) dx & \text{si } X \text{ es una v.a continua} \end{cases}$$

$\sigma = \sqrt{Var(X)}$ se conoce como desviación estándar.

σ mide la dispersión de los valores de la v.a X con respecto a su media (μ)

Propiedades de la esperanza

Sea X una v.a, y c una constante, entonces

$$1) E(c) = c$$

$$2) E(cX) = c \cdot E(X)$$

$$3) E[h(X) \pm g(X)] = E[h(X)] \pm E[g(X)]$$

$$\begin{aligned}\sigma^2 &= E(X - \mu)^2 \\ &= E(X^2 - 2X\mu + \mu^2) \\ &= E(X^2) - E(2X\mu) + E(\mu^2) \\ &= E(X^2) - 2\mu E(X) + \mu^2 && \text{pero } E(X) = \mu \\ &= E(X^2) - 2E(X)E(X) + [E(X)]^2 \\ &= E(X^2) - [E(X)]^2\end{aligned}$$

$$\text{Luego, } Var(X) = \sigma^2 = E(X^2) - [E(X)]^2$$

Así,

$$Var(X) = \begin{cases} \sum_x x^2 \cdot f(x) - \left[\sum_x x \cdot f(x) \right]^2 & \text{si } X \text{ es una v.a discreta} \\ \int_{-\infty}^{+\infty} x^2 \cdot f(x) - \left[\int_{-\infty}^{+\infty} x \cdot f(x) \right]^2 & \text{si } X \text{ es una v.a continua} \end{cases}$$

Sean X y Y dos variables aleatorias, y sea c una constante. Entonces

a) $\text{Var}(X) \geq 0$.

b) $\text{Var}(c) = 0$.

c) $\text{Var}(cX) = c^2 \text{Var}(X)$.

d) $\text{Var}(X + c) = \text{Var}(X)$.

e) $\text{Var}(X) = E(X^2) - E^2(X)$.

f) En general, $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$.

Ejemp 21: Caso continuo

Sean dos v.a X e Y, determina $E[X + Y]$

Ejemp. 22: Caso continuo

Sea la variable aleatoria continua X con función de densidad $f(x) = 2x$, para $0 \leq x \leq 1$. Determina

- a) la esperanza de X.
- b) $\text{Var}[X]$

$$\text{a) } E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x 2x dx = \frac{2}{3} x^3 \Big|_0^1 = \frac{2}{3}$$

$$\begin{aligned} \text{b) } \text{Var}(X) &= \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx \\ &= \int_0^1 (x - 2/3)^2 2x dx \\ &= \int_0^1 (2x^3 - 8x^2/3 + 8x/9) dx \\ &= (x^4/2 - 8x^3/9 + 4x^2/9) \Big|_0^1 \\ &= 1/18. \end{aligned}$$

Ejemplo 23: Caso variable discreta

Se lanza una moneda tres veces, si las tres veces aparece Águila o las tres veces aparece Sol, un jugador gana \$5, pero si no es así pierde \$3. Sea **X la v.a.** que denota **la ganancia** del jugador.

¿Cuál es el valor esperado de este juego?

$S = \{AAA, AAS, ASA, ASS, SAA, SAS, SSA, SSS\}$

X	5	-3
$p_X(x)$	2/8	6/8

$$E(X) = \sum_k x_k p_X(x_k) \quad X: \text{discrete}$$

$$E(X) = 5 \left(\frac{1}{4} \right) - 3 \left(\frac{3}{4} \right) = -1$$

Ejemplo 24: Caso discreto

En una agencia de teléfonos celulares, tres modelos de teléfonos suelen ser vendidos en número de 20, 21 y 22, diariamente, con probabilidades de selección de 0.3, 0.5 y 0.2, respectivamente.

Sea X la v.a. que representa las ventas diarias de los modelos de teléfonos, de los datos anteriores se tiene la siguiente tabla:

X	20	21	22
$p(x)$	0.3	0.5	0.2

En un día cualquiera y respecto a la cantidad de teléfonos vendidos:

a) ¿Cuál es la media?

b) ¿Cuál es la varianza?

$$a) \quad E(X) = \sum_k x_k p_X(x_k)$$

$$E(X) = (20)(0,3) + (21)(0,5) + (22)(0,2)$$

$$E(X) = 20,9$$

$$b) \quad Var(X) = E(X^2) - [E(X)]^2$$

$$E(X^2) = (400)(0,3) + (441)(0,5) + (484)(0,2)$$

$$E(X^2) = 437,3$$

$$Var(X) = 437,3 - 436,81$$

$$Var(X) = 0,49$$

Ejemplo 25: Caso continuo

Sea X una v.a. con **media 5** y **varianza 9**. Para la v.a. $Y = \frac{1}{3}(X - 5)$, determina:

- a) El valor esperado
- b) La varianza de Y :

b)

$$\text{Var}(cX) = c^2 \text{Var}(X)$$

$$\text{Var}(X + c) = \text{Var}(X)$$

$$\begin{aligned} \text{Var}\left[\frac{1}{3}(X - 5)\right] &= \left(\frac{1}{3}\right)^2 \cdot \text{Var}(X) \\ &= \frac{1}{9} \cdot 9 \\ &= 1 \end{aligned}$$

Ejemplo 26:

Sea X una v.a con la función de densidad:

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2 \\ 0, & \text{Caso contrario} \end{cases}$$

Determina:

a) $E[X]$

b) $X[g(X)]$, siendo $g(X) = 4X + 3$

$$\text{a) } E(X) = \int_{-1}^2 X \cdot \frac{X^2}{3} dX$$

$$E(X) = \frac{1}{3} \cdot \frac{1}{4} X^4 \Big|_{-1}^2$$

$$E(X) = \frac{15}{12}$$

$$\text{b) } E(4X + 3) = 4E(X) + 3$$

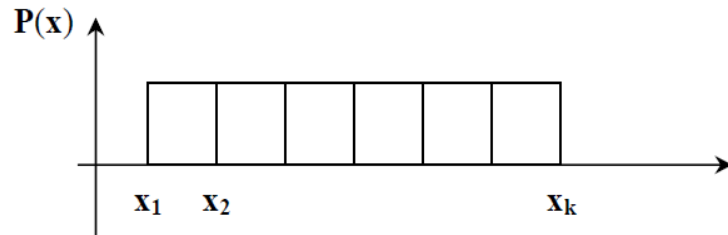
$$= 4 \cdot \frac{15}{12} + 3$$
$$= 8$$

3.- Funciones de densidad

FUNCIÓN UNIFORME DISCRETA

Si la variable aleatoria X asume valores de X_1, X_2, \dots, X_k con iguales probabilidades, entonces la distribución uniforme es:

$$f(x, k) = \frac{1}{k}$$



$$\mu = \frac{\sum_{i=1}^k x_i}{k}$$

La distribución de probabilidad del lanzamiento de un dado es:

$$S = \{1, 2, 3, 4, 5, 6\}$$

$$P(x = 1, 2, \dots, 6) = \frac{1}{6}$$

$$\sigma^2 = \frac{\sum_{i=1}^k (x_i - \mu)^2}{k}$$

$$\mu = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3.5$$

$$\sigma^2 = \frac{(1 - 3.5)^2 + (2 - 3.5)^2 + \dots + (6 - 3.5)^2}{6} = 2.91$$

FUNCIÓN UNIFORME CONTINUA

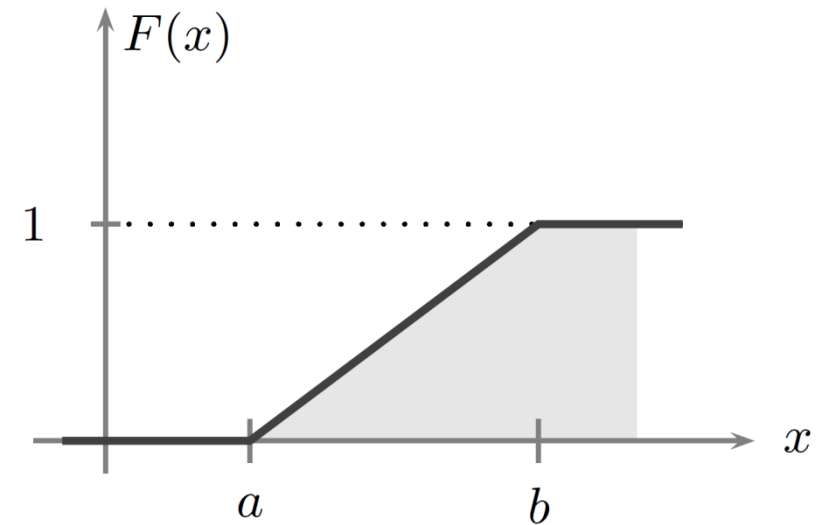
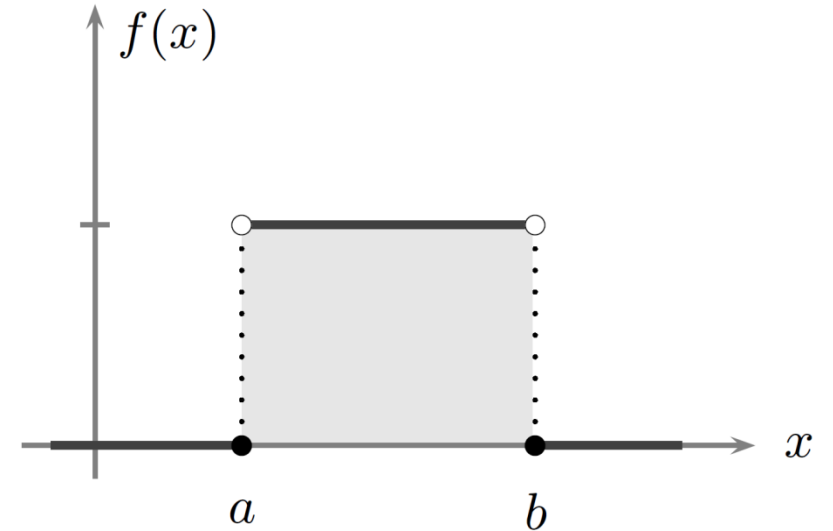
$X \sim U[a, b]$

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in (a, b), \\ 0 & \text{otro caso.} \end{cases}$$

$$E[X] = (a+b)/2$$

$$\text{Var}[X] = (b-a)^2/12$$

$$F(x) = \begin{cases} 0 & \text{si } x < a, \\ (x-a)/(b-a) & \text{si } a \leq x < b, \\ 1 & \text{si } x \geq b. \end{cases}$$



FUNCIÓN DE DENSIDAD UNIFORME CONTINUA

Ejemplo 30: Una línea de autobuses, al concluir un estudio, considera que su **consumo mensual** de combustible se expresa mediante una **función de densidad uniforme**:

$$f(x) = \begin{cases} \frac{1}{40000}, & 10000 < x \leq 50000 \\ 0, & \text{Caso contrario} \end{cases}$$

Donde **x** es la v.a. que **representa el consumo de combustible**. Determina:

- El valor esperado del consumo mensual.
- La varianza del consumo mensual.
- La probabilidad de **consumir** un máximo de 40,000 litros.
- La probabilidad de que el **consumo** esté en el rango de 20,000 y 30,000 litros.

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} xf(x) dx = \int_{10,000}^{50,000} x \frac{1}{40,000} dx \\ &= \frac{1}{40,000} \left[\frac{x^2}{2} \right]_{10,000}^{50,000} = \frac{1}{8 \cdot 10^4} [50,000^2 - 10,000^2] = \frac{10^8}{8 \cdot 10^4} [5^2 - 1^2] = \frac{10^4}{8} 24 = 30,000 \end{aligned}$$

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{10,000}^{50,000} x^2 \frac{1}{40,000} dx \\ &= \frac{1}{40,000} \left[\frac{x^3}{3} \right]_{10,000}^{50,000} = \frac{1}{12 \cdot 10^4} [50,000^3 - 10,000^3] = \frac{10^{12}}{12 \cdot 10^4} [5^3 - 1^3] = \frac{10^8}{12} 124 = \frac{31}{3} \cdot 10^8 \\ \text{Var}[X] &= E[X^2] - (E[X])^2 = \frac{31}{3} \cdot 10^8 - 9 \cdot 10^8 = \frac{4}{3} \cdot 10^8 \end{aligned}$$

GN4480100S8

Deutsche Bundesbank

Wolfgang Krauß

Frankfurt am Main
1. September 1999



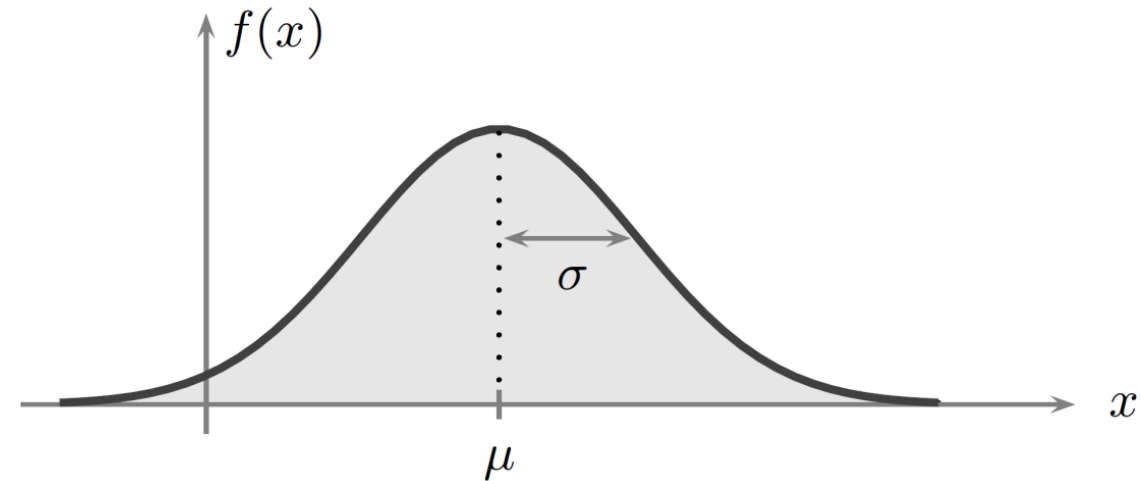
DISTRIBUCIÓN NORMAL O GAUSSIANA

$$X \sim N(m, \sigma^2)$$

$$X \sim N(m, \sigma^2): \quad f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$$

V.a. X con distribución normalizada o estándar:

$$X \sim N(0, 1): \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$



Para calcular el área de la función de densidad normal correspondiente a $P(a \leq Y \leq b)$ se debe evaluar la integral:

$$\int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)} dy$$

Esta integral no tiene forma cerrada de solución, en consecuencia, su evaluación requiere el uso de técnicas de integración numéricas o de aproximación por tablas.

Sea X una variable aleatoria con distribución normal con parámetros μ y σ^2 . Entonces la variable aleatoria **Z** tiene **distribución normalizada o estándar**:

$$Z = \frac{X - \mu}{\sigma}, \text{ tal que } \Phi(x) = P(Z \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

Ejemplo 31:

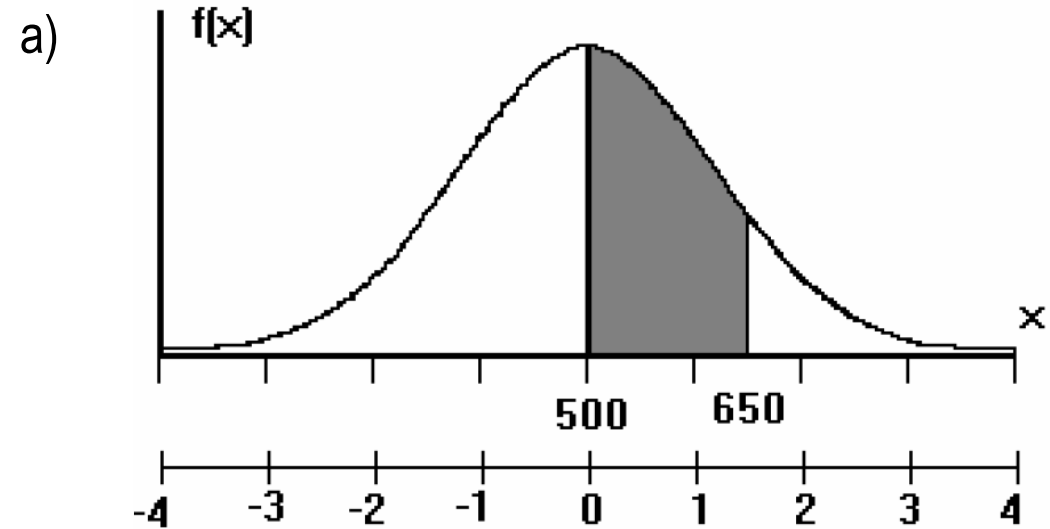
En una universidad **las edades** de los profesores es una **v.a. X** , que se distribuye **normalmente** con media de 50 años y desviación estándar de 5 años.

- a) ¿Cuál es la probabilidad que un profesor escogido aleatoriamente **no sea mayor de 45 años**?
- b) ¿Qué porcentaje de los profesores tiene **entre 50 y 52,5 años**?
- c) ¿Cuál es la probabilidad que un profesor tenga **entre 41 y 58 años**?
- d) El 20 % de los profesores tienen una edad **inferior a x años**, ¿Cuál es esa edad?

Ejemplo 32:

En un país, la matrícula de alumnos de escuelas primarias es una v.a. X con distribución Gaussiana, con parámetros $m = 500$ y desviación estándar $\sigma = 100$. ¿Cuál es la probabilidad de que el número de alumnos inscritos en una escuela elegida al azar sea?

- a) entre 500 y 650.
- b) entre 450 y 600.
- c) Máximo de 500.



Análisis Combinacional

- En muchos casos, el **número de observaciones o puntos de muestra** en un espacio de muestreo no es muy grande, por lo que la enumeración directa o el recuento de los puntos de muestra necesarios para obtener probabilidades no es difícil.
- Sin embargo, surgen problemas cuando el recuento directo es práctico. En tales casos se utiliza el **análisis combinatorio**, el cual es una forma sofisticada de contar.

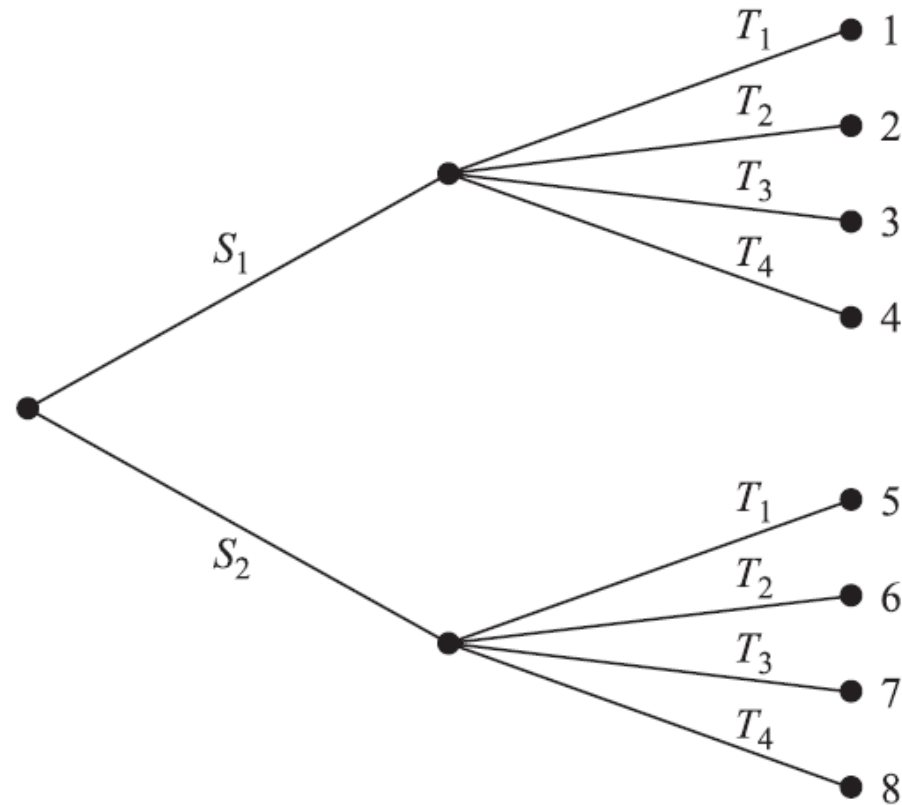
Tree Diagrams

If one thing can be accomplished in n_1 different ways, and after this a second thing can be accomplished in n_2 different ways, . .and finally a k th thing can be accomplished in n_k different ways, then all k things can be accomplished **in the specified order in $(n_1) (n_2) \dots (n_k)$ different ways.**

Ejemplo:

Si un alumno tiene 2 camisas y 4 corbatas, entonces se tienen $(2) \cdot (4) = 8$ formas de escoger una camisa y una corbata.

Un *diagrama de arbol* sirve para mostrar este principio:



Permutaciones

Un arreglo **ordenado** de **r** objetos distintos se denomina **permutación**.

El número de formas de **ordenar** (o de **escoger**) **n** objetos distintos tomados **r** a la vez, está designado por la notación: P_r^n ó ${}_n P_r$

Sin restitución ¿de cuantas formas r objetos pueden ser escogidos de entre un conjunto de n objetos?

- Para escoger **1** objeto: hay **n** formas de escoger.
- Para escoger **2** objetos. hay **n (n-1)** formas de escoger.
- Para escoger **r** objetos: hay **n(n-1)...(n-r+1)** formas
- La elección se realiza **sin reemplazo (sin restitución)**.
- El **orden** de los arreglos es importante: **abc** es una permutación diferente de **bca**.
- Por ejemplo, las “combinaciones” de un candado son “permutaciones”.

Entonces, el número de permutaciones de **n** objetos tomando **r** a la vez:

$$P_r^n = n(n-1)(n-2) \cdots (n-r+1)$$

$$P_r^n = n(n-1)(n-2) \cdots (n-r+1) \frac{(n-r)!}{(n-r)!} = \frac{n!}{(n-r)!}$$



Ejemplo 33:

¿Cuántas **permutaciones** se pueden obtener con las letras A, B, C, D, E, F, G, al tomar arreglos de

- a) De 2 letras
- b) De 3 letras

Ejemplo 34:

En un sorteo, los nombres de 3 empleados se obtienen seleccionando aleatoriamente y **sin restitución**, de una urna que contiene los nombres de 30 empleados. La persona cuyo nombre sea sacado primero recibe \$A y aquellos cuyos nombres se saquen en segundo y tercero recibirán \$B y \$C, respectivamente. Determina:

a) n y r

b) El número de arreglos ordenados que se pueden obtener en el sorteo.

Ejemplo 35:

Se tiene a 10 personas para ser acomodados en una fila con cuatro asientos.

Determina

a) n y r

b) De cuantas formas se podrían asignar los lugares

Ejemplo 36:

Un estudiante tiene un conjunto de libros, los cuales son todos diferentes:

4 de matemáticas, 6 de Física y 2 de Química.

Si se desea colocarlos todos en UNA SOLA fila de un librero, de cuantas maneras diferentes se podrían acomodar tal que:

- a) Los libros de cada tema permanezcan juntos.
- b) Solo los de matemáticas queden juntos.

Combinaciones

El número de *combinaciones* de n objetos tomados r a la vez es el número de subconjuntos, cada uno de tamaño r , que se pueden formar a partir de los n objetos. Este número estará denotado por

$$C_r^n \quad \text{ó} \quad \binom{n}{r}$$

$$\binom{n}{r} = C_r^n = \frac{P_r^n}{r!} = \frac{n!}{r!(n-r)!}$$

- La elección es **sin restitución**
- **El orden para escoger a los r objetos no es importante.**
- El orden del contenido de los subconjuntos no es importante: **abc** es igual a **bca**.
- Dos combinaciones **son diferentes solo si los contenidos son diferentes.**

Ejemplo 37:

Determina el número de formas de integrar una comisión de **5** personas de entre un grupo de **20** ejecutivos.

Ejemplo 38:

De entre un grupo de 5 matemáticos y 7 físicos, se debe integrar una comisión formada por 2 matemáticos y 3 físicos. De cuantas formas se podría hacer lo anterior si:

- a) Se toma en cuenta a las 12 personas,
- b) Un físico en particular DEBE ser incluido en la comisión.
- c) 2 matemáticos en particular NO DEBEN formar parte del comité.

LA DISTRIBUCION BINOMIAL

Esta distribución fue elaborada por Jacobo Bernoulli y es aplicable a un gran número de problemas como:

- Juegos de azar
- Control de calidad de un producto
- En educación
- En las finanzas.

La distribución binomial posee las siguientes propiedades:

- 1.- El espacio muestral contiene **n ensayos idénticos**..
- 2.- Cada observación se puede clasificar en una de dos categorías conocidas como **éxito E** o **fracaso E'**, las cuales son mutuamente excluyentes es decir $E \cap E' = 0$.
- 3.- Las probabilidades de éxito es **p** y de fracaso es **q** = 1 – p.
- 4.- El resultado de cualquier observación es **independiente** del resultado de cualquier otra observación.

Ejemplos:

- Lanzamiento de una moneda.
- Recepción de “1” o “0” en un flujo de bits.

PROCEDIMIENTO: Realizar o repetir **n** veces el experimento y determinar la probabilidad con la que el evento de interés **E** ocurre **r** veces.

La probabilidad de que el evento **E ocurra r veces** en cualquier orden en **n ensayos independientes**, está dada por la fórmula binomial:

$$P(r, n, p) = \binom{n}{r} p^r q^{n-r}$$

donde:

p = Probabilidad de éxito.

q = Probabilidad de fracaso

r = Número de ocurrencias del evento **E**

n = Número de ensayos efectuados

MEDIDAS DE TENDENCIA CENTRAL Y DE DISPERSIÓN:

m = np

Var = npq

Ejemplo 40:

Una canción popular la ha escuchado el 90% de la población. De un grupo de 12 alumnos, determina

- a) n , p y q .
- b) El parámetro r y la probabilidad de que la hayan visto 4 de ellos
- c) La probabilidad de que la hayan visto un máximo de 2 de ellos.

Ejemplo 41:

Si se lanza 4 veces una moneda honesta, sea el número resultante de “águilas” la variable A.

Determina la probabilidad del “número de águilas que caen.”

datos:

$n = 4$ ensayos.

$P=0.5$, probabilidad de éxito en un ensayo.

$q = 1 - p = 0.5$

$x = 0, 1, 2, 3, 4$

$S = \{\text{lanzar 4 veces la moneda}\}$

$A = \{\text{número de águilas que caen}\}$

a) $r=0$, cero águilas

b) $r=1$, un águila

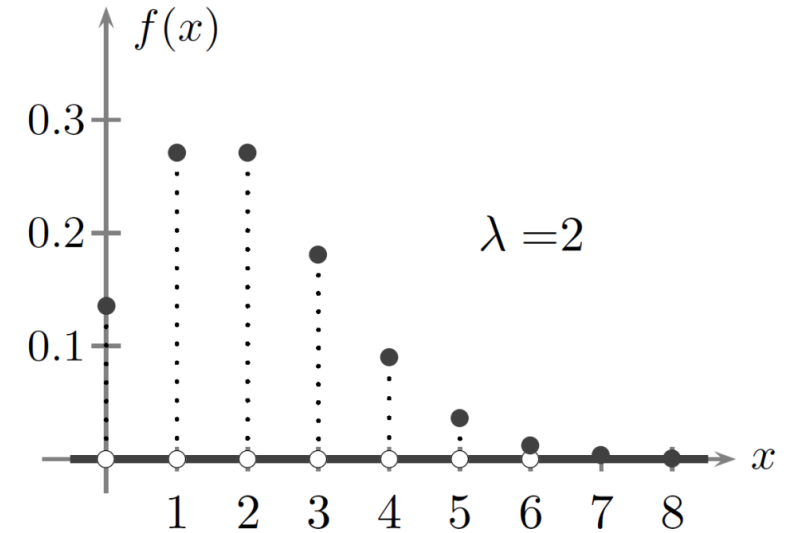
c) $r=3$, tres águilas

LA DISTRIBUCION DE POISSON

Se emplea para calcular las probabilidades asociadas a una variable aleatoria discreta respecto a un intervalo continuo de tiempo o espacio (m, m², kg, litros, etc). Debe existir **independencia** en la ocurrencia de eventos.

Algunos de los problemas que presentan distribución de Poisson son:

- Autos que utilizan el 2do piso entre las 7 y 9am..
- Número de llamadas realizadas en CU por hora en un día cualquiera.
- Número de defectos por m² en una tela.
- Número de defectos de un proceso de producción en un lote.
- Número de negocios cerrados por semana.
- Personas que se forma en un cajero automático entre 10 y 11am de un día cualquiera.



Sea un proceso de Poisson con parámetro λ , observado en un intervalo S de tiempo. K es el número de eventos que se analizan, siendo X la v. a. de Poisson. La función de distribución es

$$p(X \leq K) = F(K; \lambda s) = e^{-\lambda s} \sum_{k=0}^K \frac{(\lambda s)^k}{k!} u(x-k)$$

Tal que

$$p(X = K) = [(\lambda s)^K e^{-\lambda s}] / K!$$

Siendo $E[X] = Var[X] = \lambda$

¿Cómo se interpreta $E[X] = \lambda$?

Ejemp. 42: El conmutador de una aseguradora a nivel nacional recibe en promedio 5 llamadas por hora, reportando autos siniestrados.

¿Cuál es la probabilidad de que en una hora seleccionada aleatoriamente se reciba?

- a) Ninguna llamada
- b) Exactamente 3 llamadas
- c) No más de 3 llamadas.

¿Cómo se interpreta $E[X] = \lambda$?

$$p(X = K) = \frac{(\lambda s)^K e^{-\lambda s}}{K!}$$

$$p(X \leq K) = F(K; \lambda s) = e^{-\lambda s} \sum_{k=0}^K \frac{(\lambda s)^k}{k!} u(x - k)$$

DISTRIBUCION EXPONENCIAL

$$X \sim \exp(\lambda),$$

Una variable aleatoria **continua** X tiene una distribución exponencial con parámetro $\lambda > 0$, cuando su función de densidad es

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

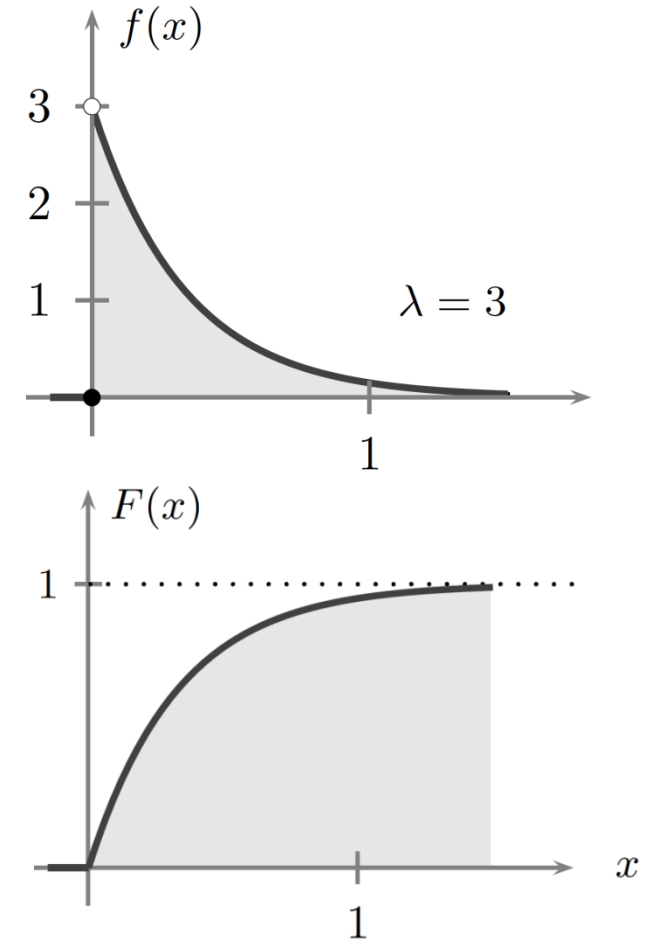
Tal que

$$E(X) = 1/\lambda$$

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

IMPORTANTE:

Si una variable aleatoria sigue una distribución de Poisson discreta, **el tiempo entre ocurrencia de eventos** tiene una **distribución exponencial**



¿Diferencia entre los valores esperados de las distribuciones Poisson y exponencial?

Ejemplo 43. El número de accesos a una página de red social sigue una distribución de Poisson. A lo largo del día los usuarios se conectan un gran número de veces. El tiempo promedio que los usuarios permanecen conectados es de **5 minutos**. Calcula la probabilidad de que un usuario cualquiera permanezca conectado a la página por un tiempo de:

- a) menos de un minuto.
- b) mas de una hora (60 minutos).

Nota: Realiza los cálculos en minutos

$\lambda = ?$

4.- Variables aleatorias múltiples

Un vector aleatorio de dimensión dos es un vector de la forma (X, Y) en donde cada coordenada es una variable aleatoria. De manera semejante se pueden tener vectores aleatorios multidimensionales (X_1, \dots, X_n) . Un vector aleatorio es discreto, o continuo, si las todas las variables aleatorias que lo conforman lo son.

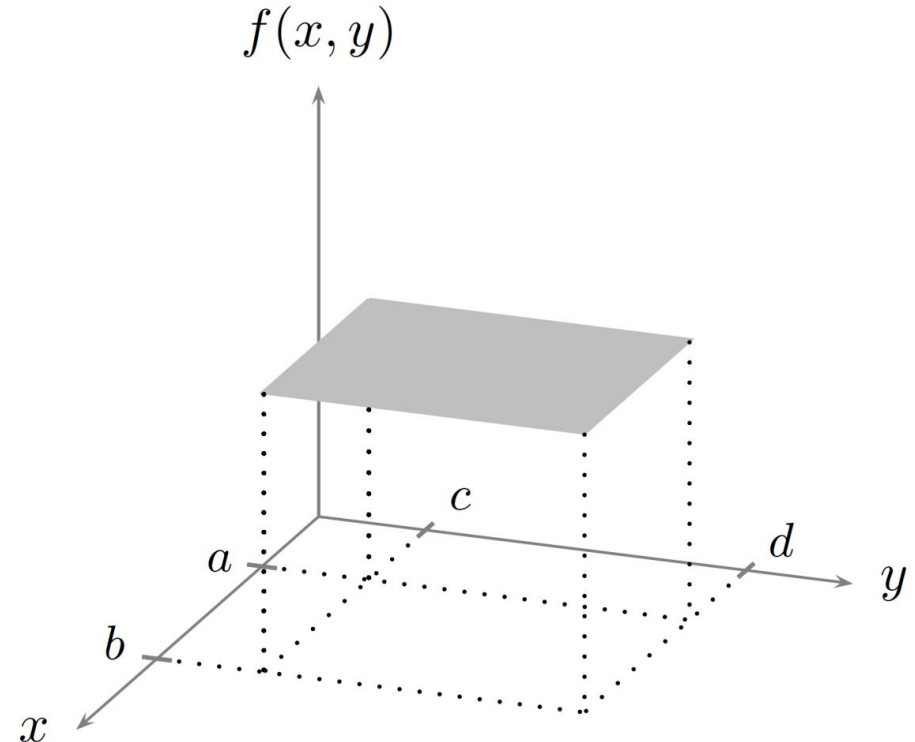
$$P(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du$$

a) $f(x, y) \geq 0$.

b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$

Ejemp.

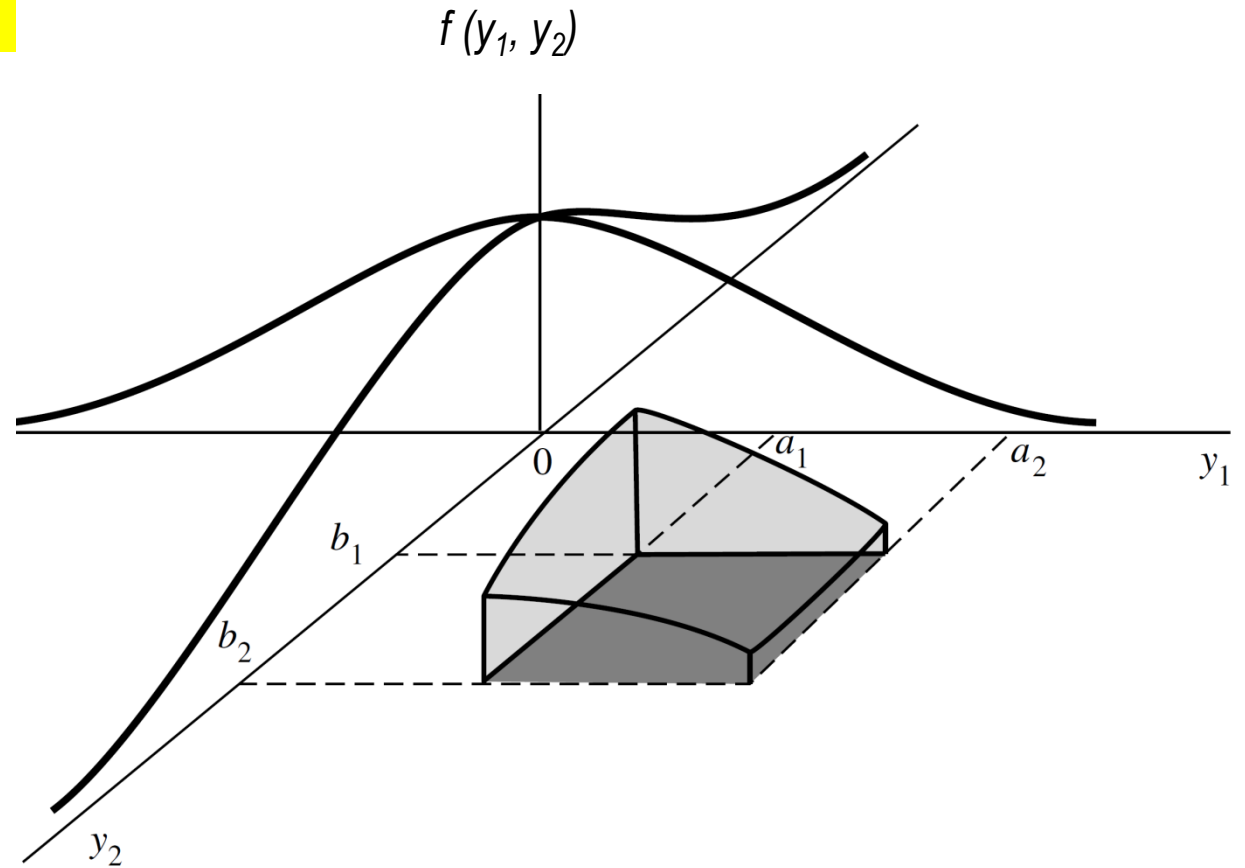
$$f(x, y) = \begin{cases} \frac{1}{(b-a)(d-c)} & \text{si } a < x < b, c < y < d, \\ 0 & \text{otro caso,} \end{cases}$$



Función de densidad conjunta bivalente $f(y_1, y_2)$

$P(a_1 \leq Y_1 \leq a_2, b_1 \leq Y_2 \leq b_2)$ es el volumen sombreado y es igual a

$$\int_{b_1}^{b_2} \int_{a_1}^{a_2} f(y_1, y_2) dy_1 dy_2.$$



$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$$

Caso continuo:

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

Caso discreto:

$$f(x) = \sum_y f(x, y)$$

En general:

$$f(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_2 \cdots dx_n$$

INDEPENDENCIA DE VARIABLES ALEATORIAS.

Sean X_1, \dots, X_n variables aleatorias con distribución conjunta $F(x_1, \dots, x_n)$, y con respectivas distribuciones marginales $F(x_1), \dots, F(x_n)$.

Se dice que estas variables son independientes si para toda x_1, \dots, x_n se cumple:

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_n}(x_n).$$

Puede definirse la independencia en términos de la función de densidad:

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f(x_1) f(x_2) \cdots f(x_n).$$

Para calcular el valor esperado marginal:

a) Cuando se tiene la función de densidad marginal $E(Y_1) = \int_{-\infty}^{\infty} y_1 f_1(y_1) dy_1$

b) Cuando se tiene la función de densidad conjunta $E(Y_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 f(y_1, y_2) dy_2 dy_1$
 $= \int_{-\infty}^{\infty} y_1 \left[\int_{-\infty}^{\infty} f(y_1, y_2) dy_2 \right] dy_1$

Ejemp. 50: Sea Y_1 y Y_2 que tienen una densidad conjunta dada por

$$f(y_1, y_2) = \begin{cases} 2y_1, & 0 \leq y_1 \leq 1, 0 \leq y_2 \leq 1 \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

Determina $E[Y_1]$

Sea $g(Y_1, Y_2, \dots, Y_k)$ una función de las variables aleatorias Y_1, Y_2, \dots, Y_k que tienen función de probabilidad conjunta $p(y_1, y_2, \dots, y_k)$. **El valor esperado** de $g(Y_1, Y_2, \dots, Y_k)$ es

Caso continuo

$$E[g(Y_1, Y_2, \dots, Y_k)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y_1, y_2, \dots, y_k) f(y_1, y_2, \dots, y_k) dy_1 dy_2 \cdots dy_k$$

Caso discreto

$$E[g(Y_1, Y_2, \dots, Y_k)] = \sum_{\text{toda } y_k} \cdots \sum_{\text{toda } y_2} \sum_{\text{toda } y_1} g(y_1, y_2, \dots, y_k) p(y_1, y_2, \dots, y_k)$$

Sean Y_1 y Y_2 variables aleatorias **independientes** y sean $g(Y_1)$ y $h(Y_2)$ funciones de Y_1 y Y_2 , respectivamente. Entonces

$$E[g(Y_1)h(Y_2)] = E[g(Y_1)]E[h(Y_2)]$$

Ejemp. 51: Considera que Y_1 y Y_2 tienen una densidad conjunta dada por

$$f(y_1, y_2) = \begin{cases} 2y_1, & 0 \leq y_1 \leq 1, 0 \leq y_2 \leq 1 \\ 0, & \text{en cualquier otro punto} \end{cases}$$

Determina $E[Y_1 Y_2]$.

Covarianza de dos variables aleatorias

$$\text{Cov}(Y_1, Y_2) = E[(Y_1 - \mu_1)(Y_2 - \mu_2)]$$

Cuanto mayor sea la covarianza, *mayor será la dependencia lineal* entre Y_1 y Y_2 .

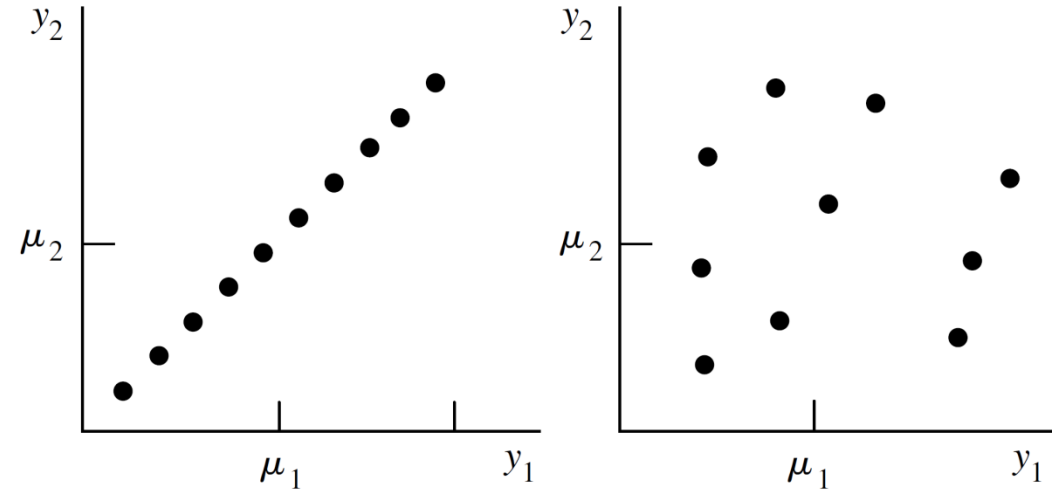
- Covarianza positiva indica que Y_1 aumenta cuando Y_2 aumenta;
- Covarianza negativa indica que Y_1 disminuye cuando Y_2 aumenta.
- Un valor cero de covarianza indica que las variables no están correlacionadas y que no hay dependencia lineal entre Y_1 y Y_2 .

Coefficiente de correlacion $\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}, -1 \leq \rho \leq 1$

Si Y_1 y Y_2 son variables aleatorias **independientes**, entonces $\text{Cov}(Y_1, Y_2) = 0$

Las **variables aleatorias independientes** no están correlacionadas linealmente

¿Dependencia entre variables aleatorias?



Dos v.a. X e Y son consideradas conjuntamente Gaussianas si su función de densidad conjunta es:

$$f_{X, Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left[-\frac{\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho_{XY}\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}{2(1-\rho_{XY}^2)}\right]$$

Las funciones marginales también deben ser Gaussianas:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X, Y}(x, y) dy = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X, Y}(x, y) dx = \frac{1}{\sqrt{2\pi\sigma_Y^2}} \exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right)$$

La **correlación** entre dos variables aleatorias es $R_{X, Y} = E[XY] = \iint xyf_{X, Y}(x, y) dx dy$

$$\text{Cov}(X, Y) = R_{X, Y} - \mu_X\mu_Y$$

Dos variables aleatorias con correlación cero se dice que son **ortogonales**.

La función de densidad Gaussian de un vector de N v. a., , con un vector de media μ y matriz de covarianza C_{XX} es

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N \det(\mathbf{C}_{XX})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_X)^T \mathbf{C}_{XX}^{-1}(\mathbf{x} - \boldsymbol{\mu}_X)\right)$$

Tal que $\boldsymbol{\mu}_X = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ and $\mathbf{C}_{XX} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$

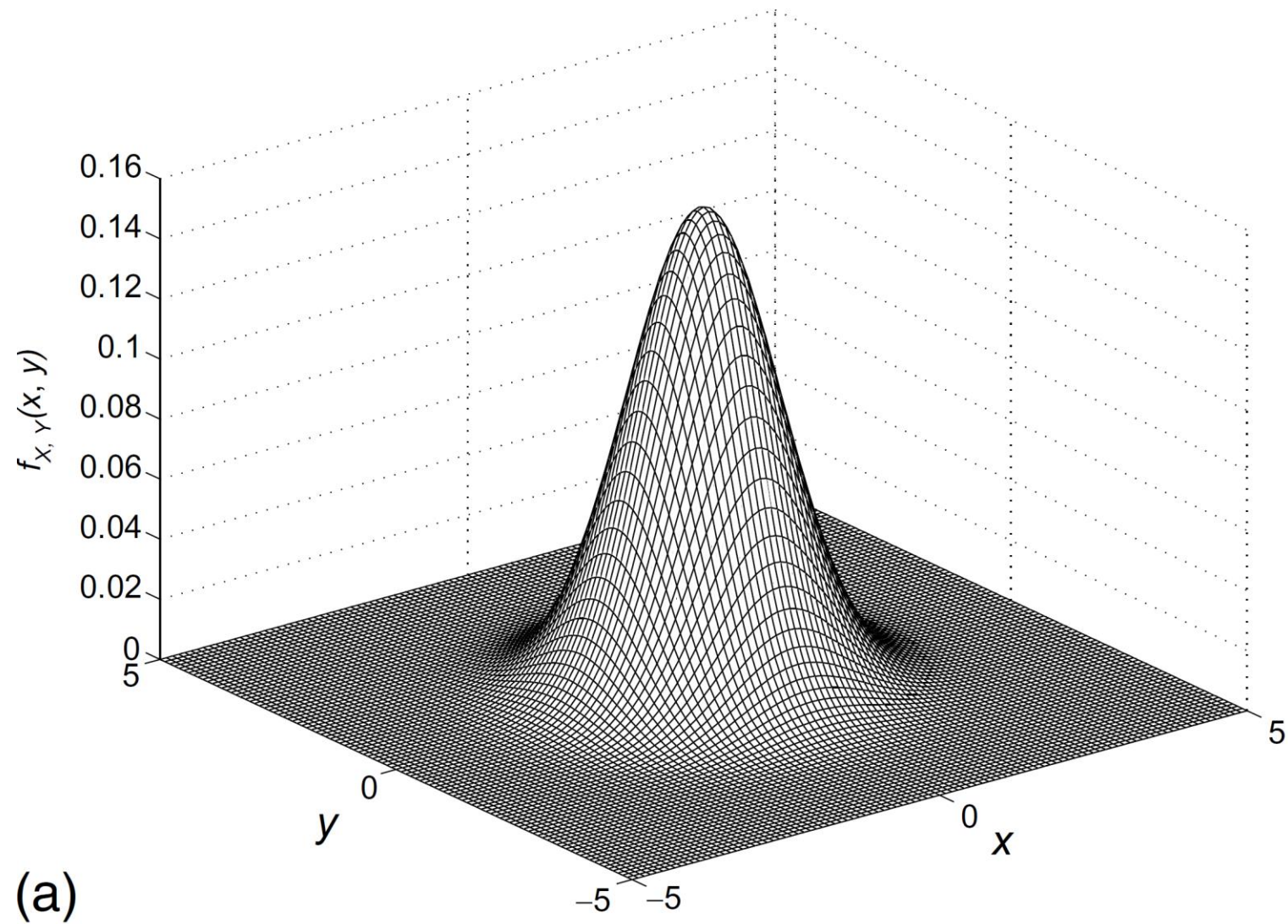
Cuando las variables
no están correlacionadas:

$$\mathbf{C}_{XX} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_N^2 \end{bmatrix}$$

Cuando las variables no están correlacionadas:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N \sigma_1^2 \sigma_2^2 \dots \sigma_N^2}} \exp\left(-\frac{1}{2} \sum_{n=1}^N \left(\frac{x_n - \mu_n}{\sigma_n}\right)^2\right) = \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{(x_n - \mu_n)^2}{2\sigma_n^2}\right)$$

En el caso Gaussiano: **no correlación** \longleftrightarrow **independencia**



Función conjunta Gussiana
 $m_x=m_y=0$, y $\sigma^2=1$

Teorema del límite central

Let X_1, X_2, \dots, X_n be a sequence of mutually independent and identically distributed random variables each of which has a finite mean μ_X and variance σ_X^2 . Let

$$S_n = X_1 + X_2 + \dots + X_n$$

Los parámetros de S_n son:

$$\begin{aligned}\bar{S}_n &= E[S_n] = n\mu_X \\ \sigma_{S_n}^2 &= n\sigma_X^2\end{aligned}$$

Converting S_n to a standard normal random variable (zero mean and variance = 1) we obtain:

$$Z_n = \frac{S_n - \bar{S}_n}{\sigma_{S_n}} = \frac{S_n - n\mu_X}{\sqrt{n\sigma_X^2}} = \frac{S_n - n\mu_X}{\sigma_X\sqrt{n}}$$

Then the central limit theorem states that if $F_{Z_n}(z)$ is the CDF of Z_n , then

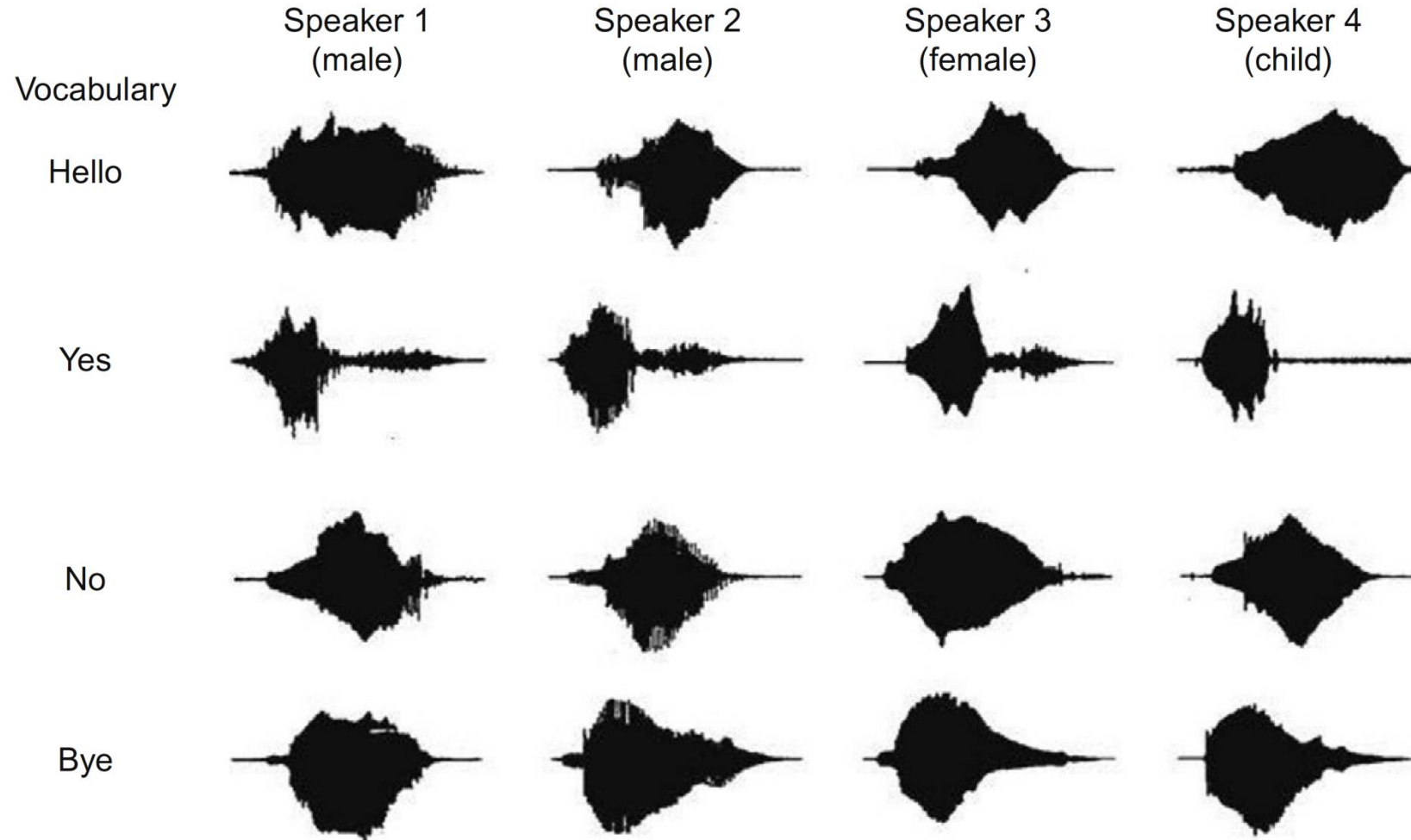
$$\lim_{n \rightarrow \infty} F_{Z_n}(z) = \lim_{n \rightarrow \infty} P[Z_n \leq z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du = \Phi(z)$$

This means that $\lim_{n \rightarrow \infty} Z_n = N(0; 1)$

The central limit theorem states that for large n the distribution of S_n is approximately **normal, regardless** of the form of the distribution of the X_k .

5.- Procesos estocásticos

Templates



Variations in speech templates for different speakers

- The concept of a random process allows us to study systems involving signals which are not entirely predictable. These random signals play fundamental roles in the fields of communications, signal processing, and control systems, and many other engineering disciplines.
- Examples of random processes include the population growth, the failure of a piece of equipment, the price of a network technology over time, and the number of cellular phone calls that arrive at a radio base station.
- The concept of random processes enlarges the random variable concept to include time. Thus, instead of thinking of a random variable X that maps an event $s \in S$, where S is the **sample space**, to some number $x(s)$, we think of how the random variable maps the event to different numbers **at different times**.
- This implies **that instead of the number $x(s)$ we deal with $X(t, s)$** , where $t \in T$ and T is usually a set of times.

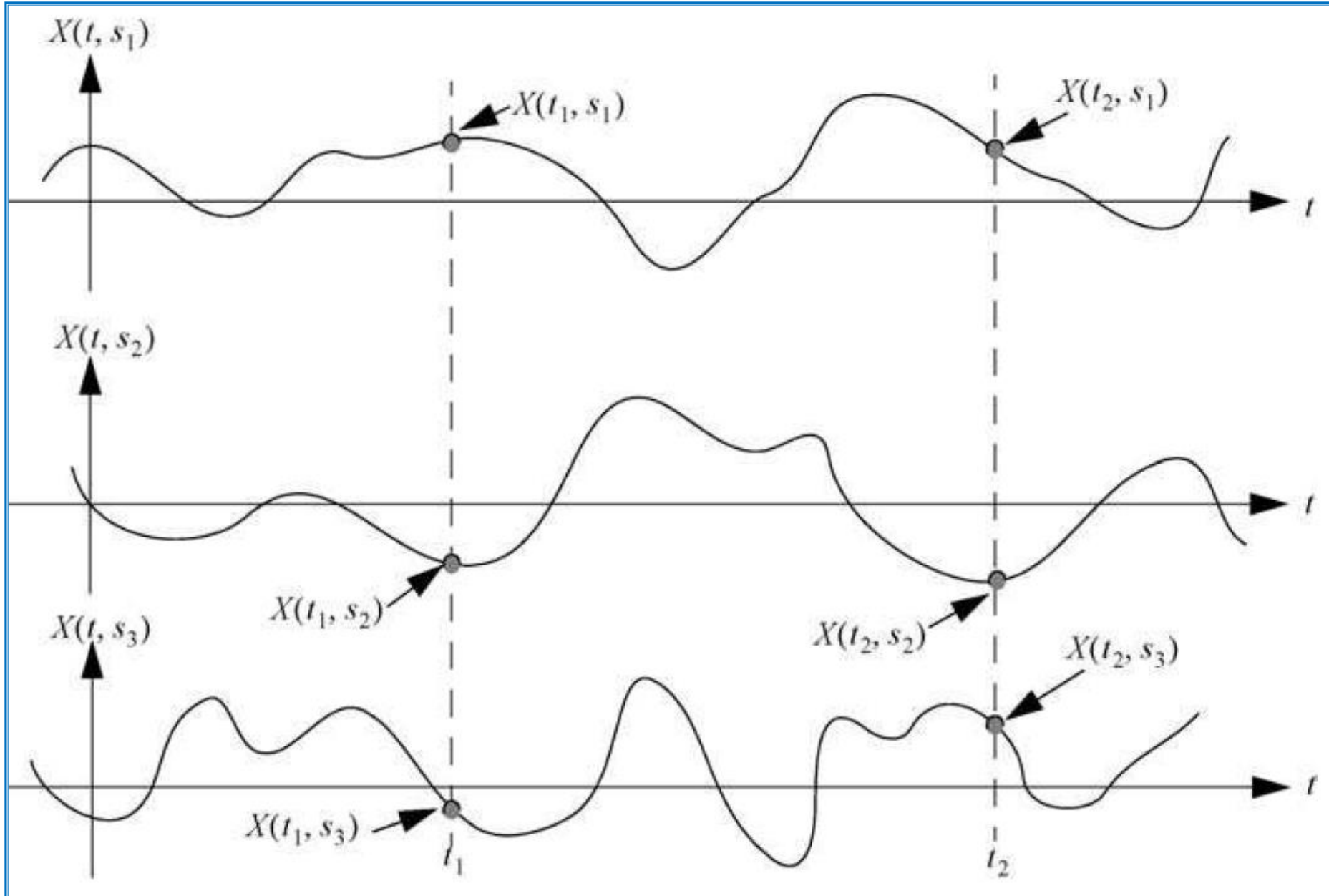
Definition: A random process is a function of the elements of a sample space, S , as well as another independent variable, t . Given an experiment, E , with sample space, S , the random process, $X(t)$, maps each possible outcome, $\zeta \in S$, to a function of t , $x(t, \zeta)$ as specified by some rule.

If we fix t , we have a function $X(s)$ that depends only on s and thus is a random variable. Thus, a random process becomes a random variable when time is fixed at some particular value. With many values of t we obtain a collection of random variables. Thus, we can define a random process as a family of random variables $\{X(t, s) \mid t \in T, s \in S\}$ defined over a given probability space and indexed by the time parameter t .

Consider a communication system example. Assume we have a set of possible messages that can be transmitted over a channel. The set of possible messages then constitutes our sample space. For each message M generated by our source, we transmit an associated waveform $X(t, s)$ over the channel. The channel is not perfect; it selectively adds a noise waveform $N(t, s)$ to the original waveform so that what is seen at the receiver is a random signal $R(t, s)$ that is the sum of the transmitted waveform and the noise waveform. That is,

$$R(t, s) = X(t, s) + N(t, s)$$

$$R(t, s) = X(t, s) + N(t, s)$$



Because the noise waveform is probabilistically selected by the channel, different noise waveforms can be associated not only with the same transmitted waveform but also with different transmitted waveforms.

Caracterización de un proceso aleatorio

Por simplicidad en la notación $X(t, s)$ se denota como $X(t)$.

Un proceso aleatorio p.a. queda completamente descrito por su función de distribución, tal que el valor del p.a. $X(t)$ en el tiempo t_i es $X(t_i)$ es una variable aleatoria.

Proceso aleatorio de **primer orden**:

$$\left\{ \begin{array}{l} F_X(x_1, t_1) = F_X(x_1) = P[X(t_1) \leq x_1] \\ F_X(x_2, t_2) = F_X(x_2) = P[X(t_2) \leq x_2] \\ \dots \\ F_X(x_n, t_n) = F_X(x_n) = P[X(t_n) \leq x_n] \end{array} \right. \quad \text{Donde } 0 < t_1 < t_2 < \dots < t_n.$$

Proceso aleatorio de **n orden**:

$$\left\{ \begin{array}{l} F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P[X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n] \\ f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n}{\partial x_1 \partial x_2 \dots \partial x_n} F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \end{array} \right.$$

$\forall n$

The **mean function** of a random process is simply the expected value of the process. For continuous time processes, this is written as

$$\mu_X(t) = E[X(t)] = \int x f_X(x;t) dx$$

Ejemp: Consider a sinusoidal random process where $X(t) = A \sin(\omega_0 t)$ and A is a uniform random variable over $[-1, 1]$. In this case,

$$\mu_X(t) = E[X(t)] = E[A \sin(\omega_0 t)]$$

$$= E[A] \sin(\omega_0 t)$$

$$= \frac{1}{2} \sin(\omega_0 t)$$

Autocorrelación entre dos muestras de datos, observadas en instantes diferentes:

$$R_{XX}(t_1, t_2) = E[X(t_1)X(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X_1, X_2}(x_1, x_2; t_1, t_2) dx_1 dx_2$$

$$R_{XX}(t, t) = E[X^2(t)]$$

- The autocorrelation function describes the relationship between two samples of a random process.
- This correlation will depend on when the samples are taken; thus, the autocorrelation function is, in general, a function of two time variables.
- Quite often we are interested in how the correlation between two samples depends on how far apart the samples are spaced.

The **crosscorrelation function** essentially measures how similar two different processes (or *signals*) are when one of them is shifted in time relative to the other

$$R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)]$$

Basically the **autocorrelation function** defines **how much a signal is similar to a time-shifted version of itself**.

To explicitly draw out this relationship, define a time difference variable, $\tau = t_2 - t_1$, and the autocorrelation function can then be expressed as

$$R_{XX}(t, t + \tau) = E[X(t)X(t + \tau)]$$

$X(t)$ y $Y(t)$ son **procesos ortogonales** si $R_{XY}(t, s) = 0, \forall t$ y s .

The **autocovariance function** of the random process $X(t)$ is another quantitative measure of the statistical coupling between $X(t_1)$ and $X(t_2)$.

The autocovariance function, C_{XX} , of a continuous time random process, $X(t)$, is defined as the covariance of $X(t_1)$, and $X(t_2)$:

$$C_{XX}(t_1, t_2) = \text{Cov}(X(t_1), X(t_2)) = E[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))]$$

Al emplear la autocorrelación:

$$C_{XX}(t_1, t_2) = R_{XX}(t_1, t_2) - \mu_X(t_1)\mu_X(t_2)$$

Ejemp. The autocovariance function is helpful when studying random processes which can be represented as the sum of a **deterministic signal**, $s(t)$, plus a statistically independent **zero-mean** noise process, $N(t)$.

If $X(t) = s(t) + N(t)$, then the **autocorrelation function** of $X(t)$ is

$$R_{XX}(t_1, t_2) = E[(s(t_1) + N(t_1))(s(t_2) + N(t_2))] = s(t_1)s(t_2) + R_{NN}(t_1, t_2)$$

the **autocovariance function** is

$$C_{XX}(t_1, t_2) = R_{XX}(t_1, t_2) - s(t_1)s(t_2) = R_{NN}(t_1, t_2) = C_{NN}(t_1, t_2)$$

Therefore, the autocovariance function allows us to isolate the noise which is the source of randomness in the process.

La **independencia entre $X(t_1)$ y $X(t_2)$** ¿cómo afecta **la autocovarianza** ?

$$C_{XX}(t_1, t_2) = R_{XX}(t_1, t_2) - \mu_X(t_1)\mu_X(t_2)$$

Si $X(t_1)$ y $X(t_2)$ son independientes, ¿la covarianza es?

$$C_{XX}(t_1, t_2) = 0$$

Por lo que: $0 = R_{XX}(t_1, t_2) - \mu_X(t_1)\mu_X(t_2)$

∴

- $R_{XX}(t_1, t_2) = \mu_X(t_1)\mu_X(t_2)$
- $C_{XX}(t_1, t_2) = 0$ indica que no hay acoplamiento entre $X(t_1)$ y $X(t_2)$ y que por lo mismo, no están correlacionadas.

Ejemp. 52: Sea el proceso aleatorio $X(t) = K \cos wt$, con $t \geq 0$, donde w es una constante y K está uniformemente distribuida entre 0 y 2. Determina:

- (a) $E[X(t)]$
- (b) La función de autocorrelación de $X(t)$
- (c) La función de autocovarianza de $X(t)$

Ejemp. 53: Sean los procesos aleatorios $Y(t)$, $X(t)$ y $N(t)$. Dado que $Y(t) = X(t) + N(t)$, donde $N(t)$ es un ruido aleatorio y existe independencia entre $X(t)$ y $N(t)$.

Determina la función de correlación cruzada entre $X(t)$ y $Y(t)$.

Stationary and Ergodic Random Processes

The **mean function** and the **autocorrelation** (or autocovariance) function can provide information about the **temporal structure** of a random process.

Definition: A continuous time random process $X(t)$ is **strict sense stationary (SSS)** if the statistics of the process are **invariant to a shift** in the time origin. Specifically, for **any arbitrary time shift** τ and any integer $n \geq 1$

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau)$$

The PDF does not depend on t but is a function of τ .

Definition: A random process is **wide sense stationary (WSS)** if the **mean function** and **autocorrelation function** are **invariant to a time shift**. In particular, this implies that

$$\begin{aligned}\mu_X(t) &= \mu_X = \text{constant} \\ R_{XX}(t, t + \tau) &= R_{XX}(\tau) \quad (\text{function only of } \tau)\end{aligned}$$

Many practical problems that we encounter require that we deal with **only the mean and autocorrelation function** of a random process. Solutions to these problems **are simplified** if these quantities do not depend on absolute time.

All strict sense stationary random processes are also WSS, provided that the mean and autocorrelation function exist.

- The converse is not true. A WSS process does not necessarily need to be stationary in the strict sense.
- A process which is not WSS is called *non-stationary*.

Ergodicidad

In order to calculate the mean or autocorrelation function of a random process, it is necessary to perform an ensemble average.

- In many cases, this may not be possible as we may not be able to observe all realizations (or a large number of realizations) of a random process.
- In fact, quite often we may be able to observe only a single realization. This would occur in situations where the conditions of an experiment cannot be duplicated and therefore the experiment is not repeatable.
- Is it possible to calculate the mean and/or autocorrelation function from a single realization of a random process?

Suppose a WSS random process $X(t)$ has a mean μ_x . We are able to observe one realization of the random process, $x(t)$, and wish to try to determine μ_x from this realization. One obvious approach would be to calculate the time average of the realization:

If the two averages are the same, then we say that the random process is **ergodic in the mean**.

$$\langle x(t) \rangle = \lim_{t_0 \rightarrow \infty} \frac{1}{2t_0} \int_{-t_0}^{t_0} x(t) dt$$

Given a single realization, $x(t)$, form the time-average autocorrelation function:

$$\mathfrak{R}_{xx}(\tau) = \langle x(t)x(t+\tau) \rangle = \lim_{t_0 \rightarrow \infty} \frac{1}{2t_0} \int_{-t_0}^{t_0} x(t)x(t+\tau) dt$$

If $\mathfrak{R}_{xx}(\tau) = R_{XX}(\tau)$ for any realization, then the random process is said to be **ergodic in the autocorrelation**

Definition: A WSS random process is *ergodic* if ensemble averages involving the process can be calculated using time averages of any realization of the process. Two limited forms of ergodicity are:

- Ergodic in the mean - $\langle x(t) \rangle = E[X(t)]$,
- Ergodic in the autocorrelation - $\langle x(t)x(t + \tau) \rangle = E[X(t)X(t + \tau)]$

UNAMSAT



REFERENCIAS

- *Probability, random signals and statistics: a textgraph with integrated software for electrical & computer engineers*, Li, X. Rong, CRC Press LLC
- *Fundamentals of applied probability and random processes*, Ibe, Oliver C. , Second edition, Academic Press.
- *PROBABILITY, RANDOM VARIABLES, AND RANDOM SIGNAL PRINCIPLES*, PEYTON PEEBLES, Jr., Irwin/McGraw-HILL
- *Probability, statistics, and random processes for electrical engineering*, Alberto Leon-Garcia, 3rd ed., Pearson Prentice Hall
- *Probability and Random Processes With Applications to Signal Processing and Communications*, 2d Ed, Scott L. Miller, Elsevier Inc.
- *Schaum's Outline of Theory and Problems of PROBABILITY, RANDOM VARIABLES, AND RANDOM PROCESSES*, Hsu, Hwei P., McGraw-Hill Companies, Inc.
- *Probabilidad básica para ingenieros*, Luis Rodriguez Ojeda, Guayaquil, Ecuador.
- *Probability and Stochastic Processes*, Roy D. Yates, David J. Goodman, 2Ed, John Wiley & Sons, Inc,