(1994-2024)
30 años de la Consagración Constitucional de la Autonomía y Autarquía Universitaria en Argentina.

UNIVERSIDAD NACIONAL DEL LITORAL
LUX INDEFICIENS · 1919

FICH

ESTADÍSTICA
PROBABILIDAD Y ESTADÍSTICA

Notas de Clase #1 - Unidad 1 - *Probabilidad*

Ingenierías en: Recursos Hídricos, Ambiental y Agrimensura

Año 2024

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

ETIMOLOGÍA

La estadística (la forma femenina del término alemán Statistik, derivado a su vez del italiano statista, «hombre de Estado») es la disciplina que estudia la **variabilidad**, así como el proceso aleatorio que la genera siguiendo las leyes de la **probabilidad**. Cuando este conocimiento se aplica a las ciencias fácticas, el proceso de investigación requiere la recolección, organización, **análisis**, **interpretación** y **presentación** de los **datos**.

Fuente: Wikipedia

Nociones Iniciales

Experimento aleatorio:

Un experimento aleatorio (EA) es cualquier actividad, proceso o fenómeno que puede repetirse indefinidamente, respetando las mismas condiciones y producirá un resultado dentro de un conjunto de resultados conocido, aunque no se tenga la certeza de cuál de los resultados se obtendrá.

Ejemplo:

Supongamos que definimos como EA, el hecho de lanzar un dado de 6 caras. Si suponemos que el dado no está cargado, sabemos que el resultado será el que corresponda al número de alguna de las 6 caras, pero no podemos saber qué numero saldrá en cada lanzamiento. Otro ejemplo podría ser efectuar la medición de alguna magnitud de temperatura en un determinado tiempo y lugar, no sabemos qué valor arrojará pero será un número dentro de los reales.

Nociones Iniciales

Espacio Muestral:

El conjunto de resultados posibles de un experimento aleatorio se denomina Espacio Muestral (EM). Si bien no es posible saber qué resultado del EM se obtendrá como resultado, es seguro que el resultado que se obtenga obtenido será uno de los contenidos del EM. Cada resultado elemental posible dentro de este espacio se conoce como punto muestral. El EM se suele representar en forma simbólica con la letra E y gráficamente como un rectángulo opcionalmente puede contener una letra que lo referencia ya sea una E o una S.

Ejemplo:

Si se considera el ejemplo e EA de lanzar el dado de seis caras, el conjunto de resultados puede representarse simbólicamente de la siguiente forma:

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

Nociones Iniciales

- Experimento aleatorio
  - Espacio Muestral E

Nociones Iniciales

Evento o suceso:

Se denomina Evento o Suceso a un conjunto de uno o más puntos del espacio muestral. Un evento suele definirse mediante la asociación de una letra mayúscula, que representará la ocurrencia del mismo y podrá expresarse tanto simbólicamente como en forma coloquial.

Ejemplo:

Continuando con ejemplo anterior, podemos pensar en el evento que implica que se obtenga un número par con las siguientes definiciones:

P: "El número obtenido es par" (Coloquial)

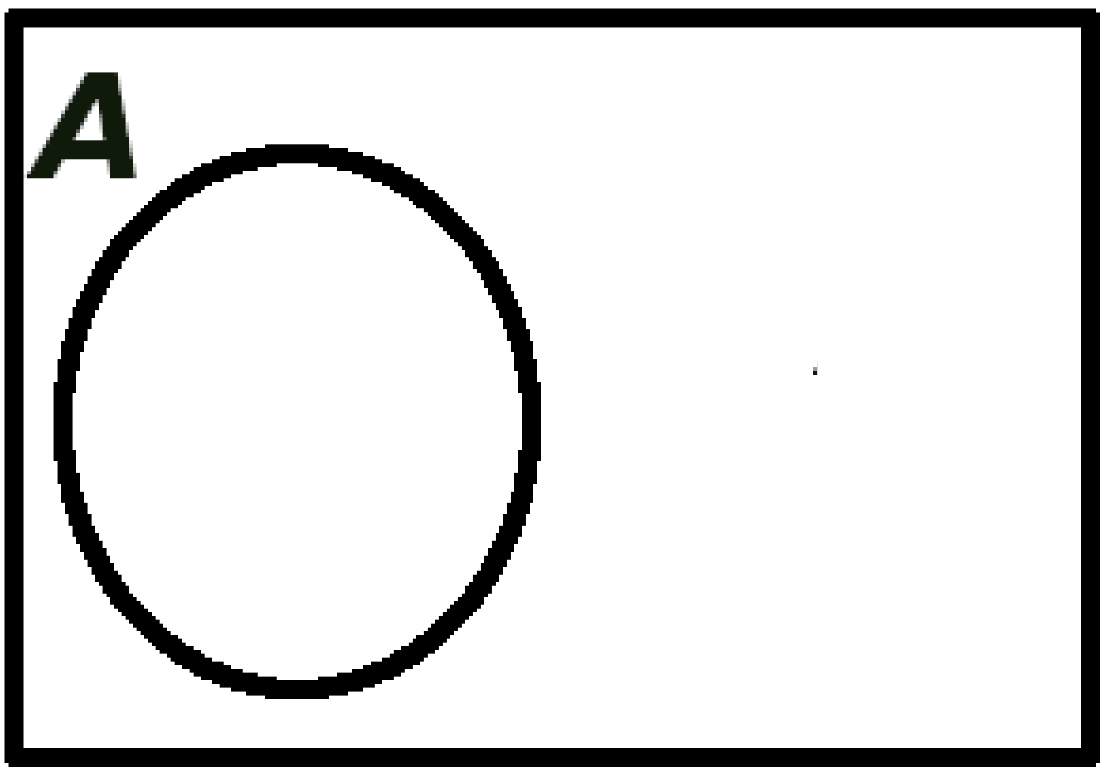
$P = \{2, 4, 6\}$ (Simbólica)

Nociones Iniciales

Representación gráfica

Para representar gráficamente uno o más eventos utilizaremos un diagrama de Venn. Imaginaremos que cada círculo representa un evento que contiene uno o más puntos elementales del EM.

Ejemplo:



Nociones Iniciales

- Eventos o Sucesos
  - Relaciones
    - Unión
    - Intersección
    - Complementarios
    - Excluyentes
    - Sistema Exhaustivo y Excluyente
    - Dependencia

UNIÓN E INTERSECCIÓN

La unión de dos eventos, representa la suma de los conjuntos de puntos en el espacio muestral que contiene cada evento. De esta forma, dados dos eventos denotados por A y B, el conjunto de puntos que representa la unión de dichos eventos, se puede denominar como *“la unión de A con B”* y contendrá todos los puntos del EM que se encuentren en A y/o en B. A su vez, se podrá expresar simbólicamente de la siguiente forma:

$$A \cup B$$

La interesección de dos eventos, representa el conjunto de los puntos del espacio muestral que se encuentran contenidos en ambos eventos. De esta forma, dados dos eventos denotados por A y B, el conjunto de puntos que representa la intersección de dichos eventos, se puede denominar la “intersección de A con B” y contendrá todos los puntos del EM que se encuentren en A y en B. A su vez, se podrá expresar simbólicamente de la siguiente forma:

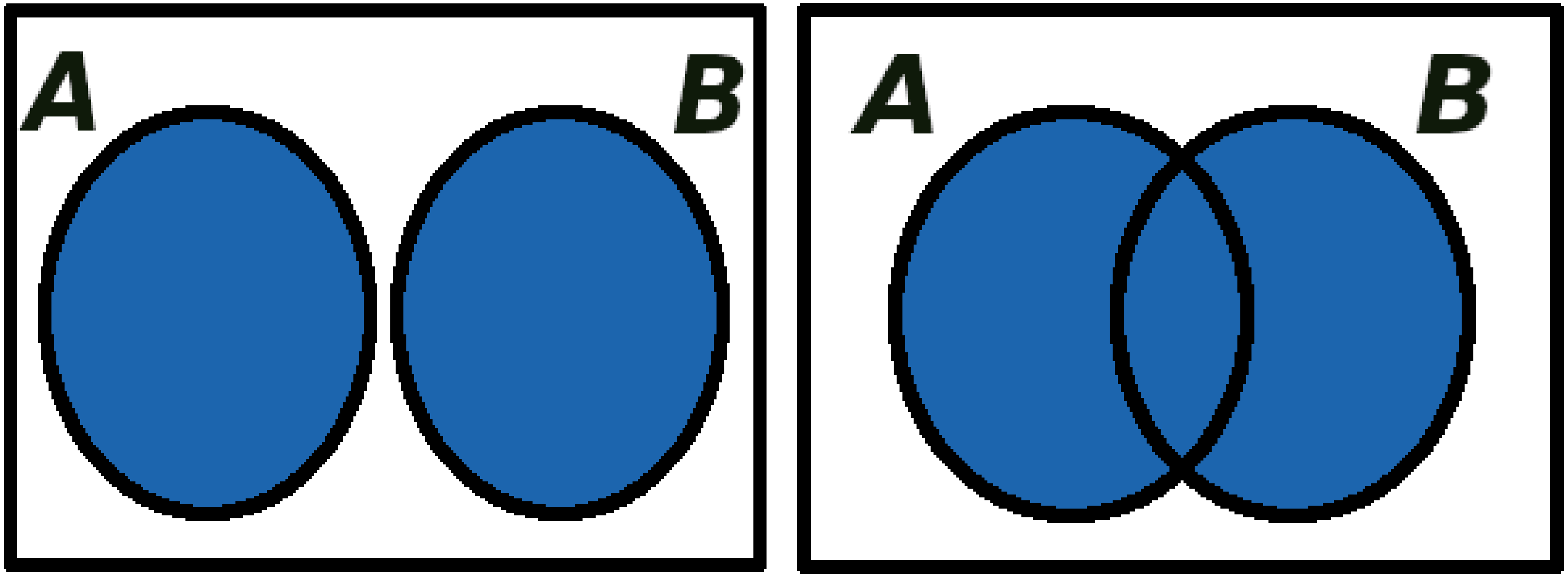
$$A \cap B$$

consultas: taulamet@unl.edu.ar

UNIÓN

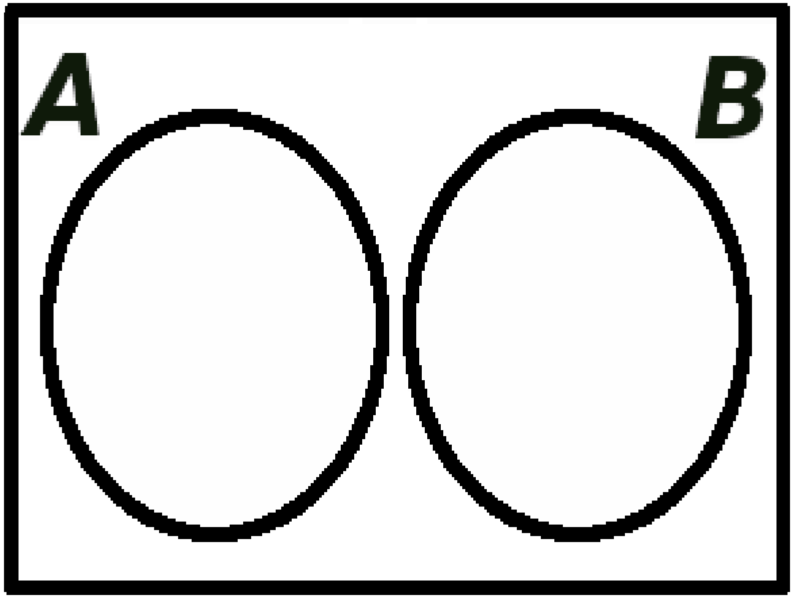
La representación gráfica de la unión, se puede realizar pintando los puntos incluidos en la misma, en el Diagrama de Venn.

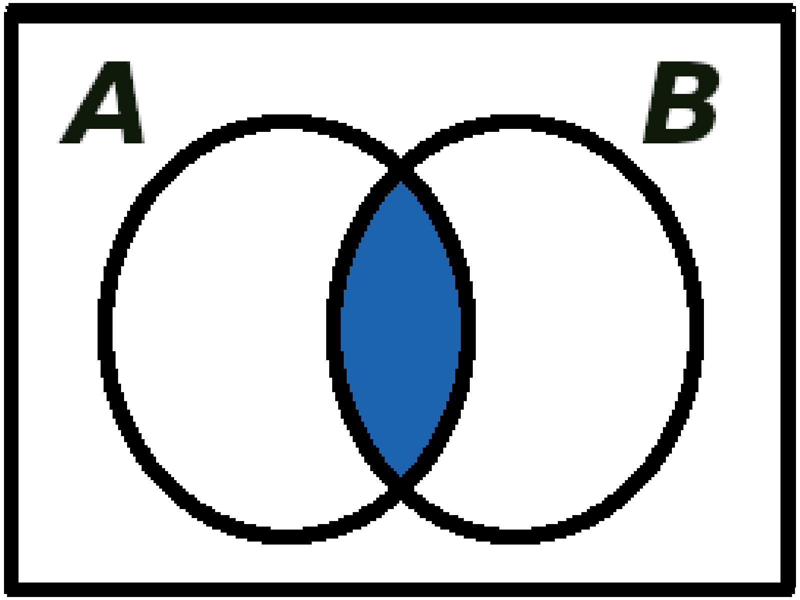
A continuación presentamos dos ejemplos, uno incluyendo dos eventos que no tienen puntos en común (Izquierda) y un ejemplo que incluye dos eventos que tienen puntos en común (Derecha).



INTERSECCIÓN

La representación gráfica de la intersección, se puede realizar pintando los puntos incluidos en la misma en el Diagrama de Venn. A continuación presentamos dos ejemplos, uno incluyendo dos eventos que no tienen puntos en común (Izquierda) y un ejemplo que incluye dos eventos que tienen puntos en común (Derecha). En el gráfico de la izquierda podemos observar que la intersección no representa ningún punto en el EM. Desarrollaremos ese caso a continuación.

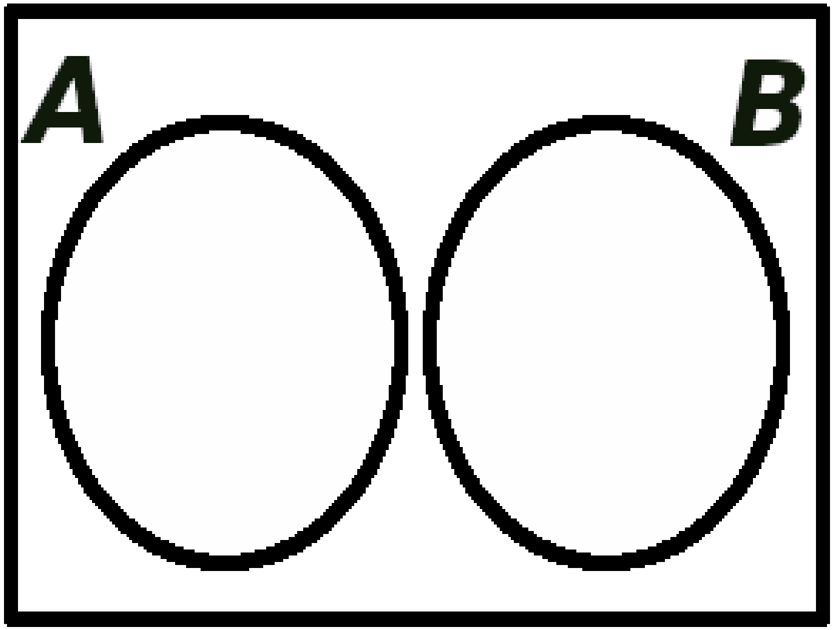




EXCLUYENTES

Cuando dos eventos no tienen puntos en común, la intersección de los mismos no incluye ningún punto en el espacio muestral, Cuando eso sucede, se dice que los eventos son Mutuamente Excluyentes (ME) y su intersección representa el conjunto vacío. A continuación les ofrecemos la representación gráfica y la expresión simbólica que se corresponden a dos eventos A, B que son mutuamente excluyentes.

$$A \cap B = \Phi$$



COMPLEMENTARIOS

Cuando un evento representa todos los puntos en el espacio muestral que no se encuentran incluidos algún evento en particular, se dice que dichos eventos son complementarios entre si. De esta forma, dado un evento A, complemento de A, representa todos los puntos del EM que no se encuentran en A y simboliza con una línea sobre el evento A. La unión de eventos complementarios configura todo el EM.

COMPLEMENTARIOS

Cuando un evento representa todos los puntos en el espacio muestral que no se encuentran incluidos algún evento en particular, se dice que dichos eventos son complementarios entre si. De esta forma, dado un evento A, complemento de A, representa todos los puntos del EM que no se encuentran en A y simboliza con una línea sobre el evento A.

Excluyencia de eventos complementarios

Si dos eventos son complementarios, entonces serán mutuamente excluyentes. Lo inverso no siempre puede asegurarse y dependerá de la unón de los mismos, tal como mencionaremos a continuación.

Complementariedad de de eventos ME

Si dos eventos son ME, serán además complementarios sí y solo sí son exhaustivos.

COMPLEMENTARIOS

Unión de eventos complementarios

Cuando dos eventos son complementarios, la unión de los mismos incluye a todos los puntos del EM. Esto puede expresarse simbólicamente de la siguiente forma:

$$A \cup \overline{A} = E$$

Exhaustividad

Cuando la unión de dos o más eventos representa todos los puntos del EM, se dice que dichos eventos conforman un sistema exhaustivo o son exhaustivos.

Exhaustividad de eventos complementarios

Por todo lo anterior, se sabe que si dos eventos son complementarios entre sí, serán además exhaustivos. A su vez si dos eventos son exhaustivos, podrán ser o no complementarios.

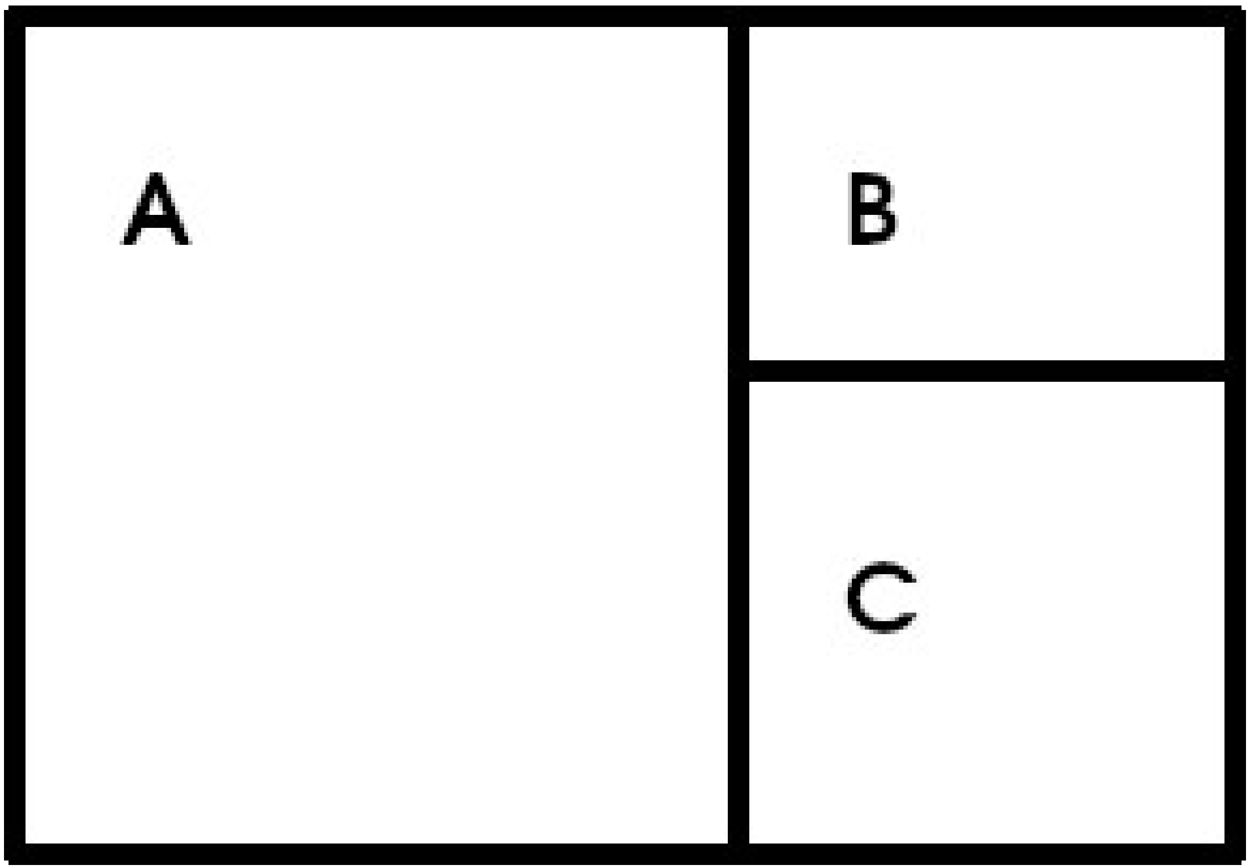
SISTEMA EXHAUSTIVO Y EXCLUYENTE

Si se combinan algunas nociones anteriores, podemos definir un sistema exhaustivo colectivo y excluyente de eventos, a un conjunto de eventos que cumple dos condiciones:

Los eventos son excluyentes entre sí, es decir que no poseen puntos en común entre sí.

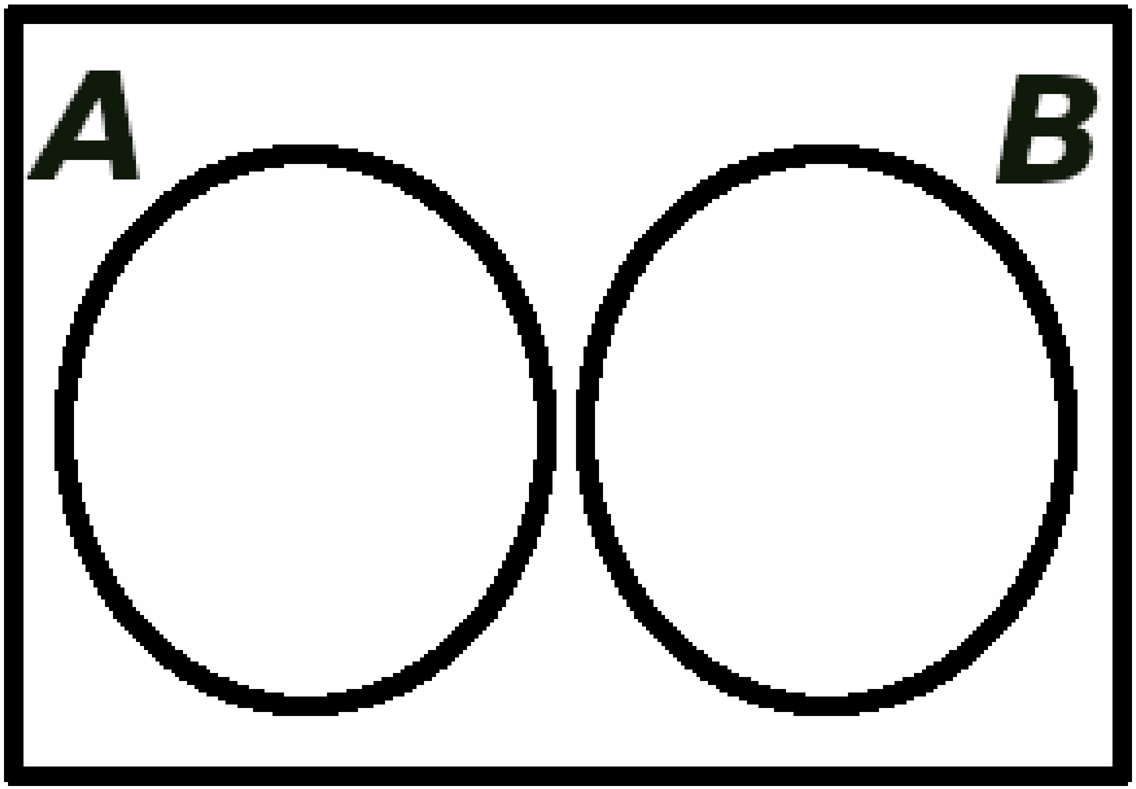
La unión de todos los eventos es exhaustiva, es decir que incluye a todos los puntos del EM.

SISTEMA EXHAUSTIVO Y EXCLUYENTE



EVENTOS EXCLUYENTES

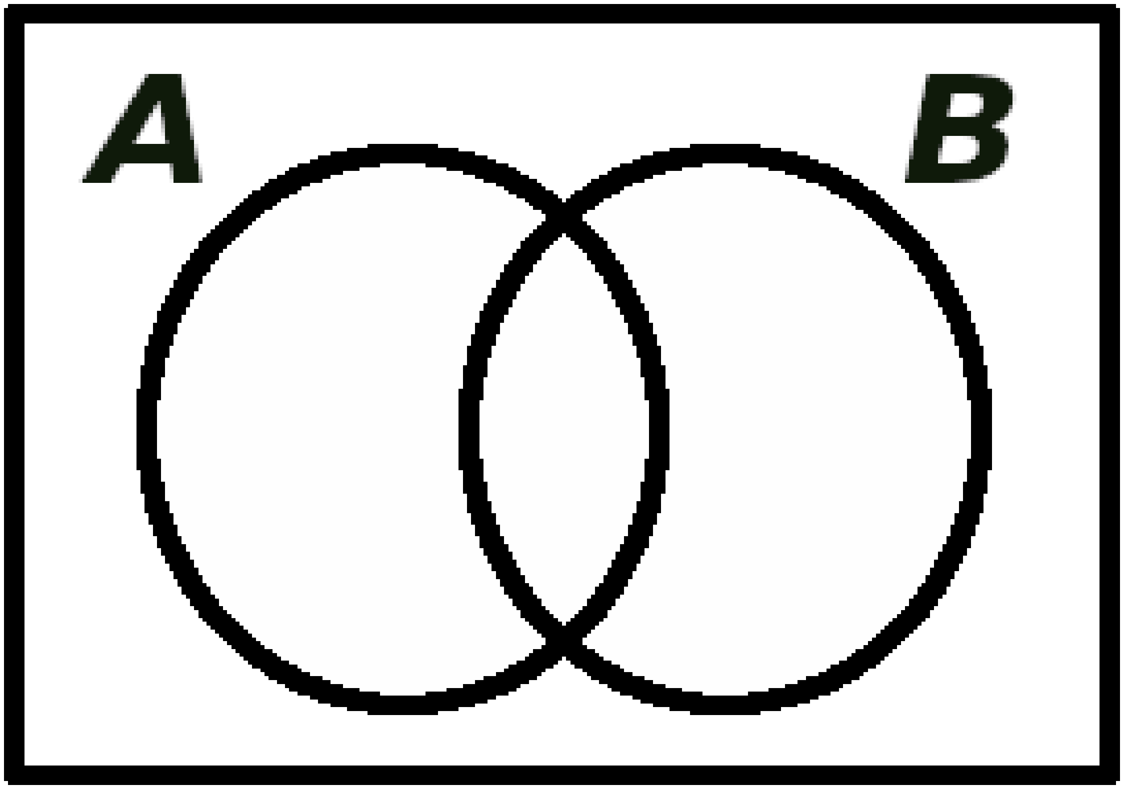
CASO: NO EXHAUSTIVO



$$A \cap B = \phi$$
$$A \cup B \neq E$$

EVENTOS NO EXCLUYENTES

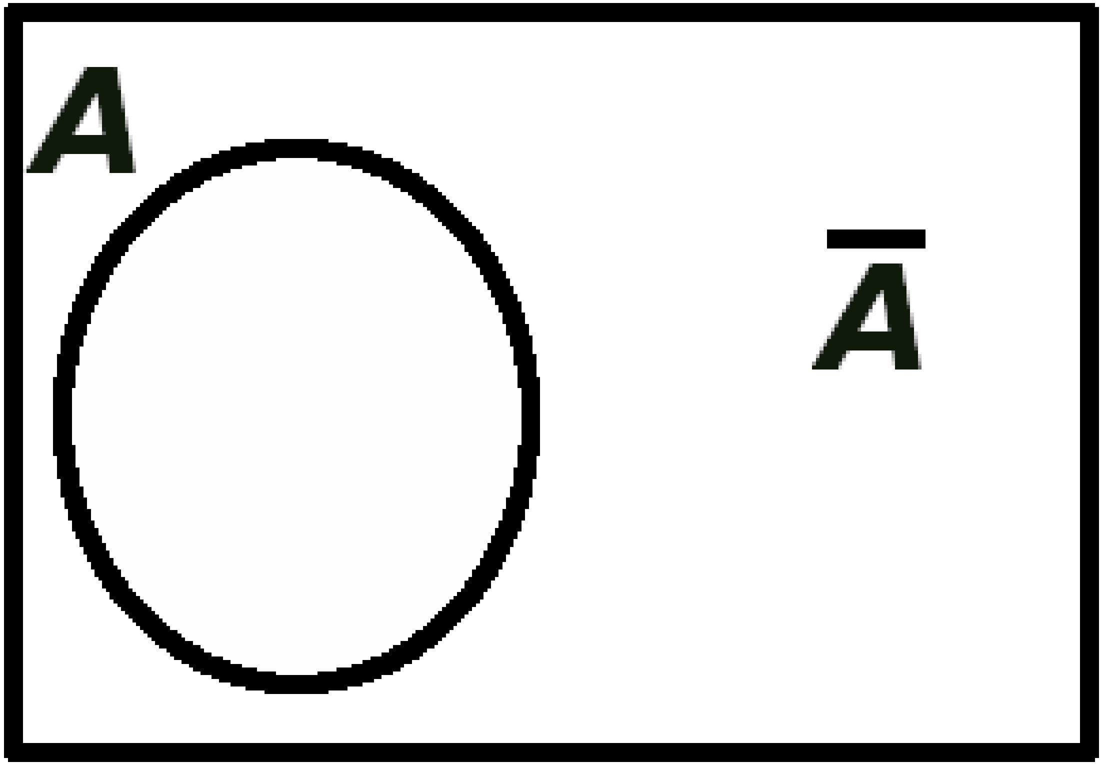
CASO: NO EXHAUSTIVO



$$A \cap B \neq \phi$$
$$A \cup B \neq E$$

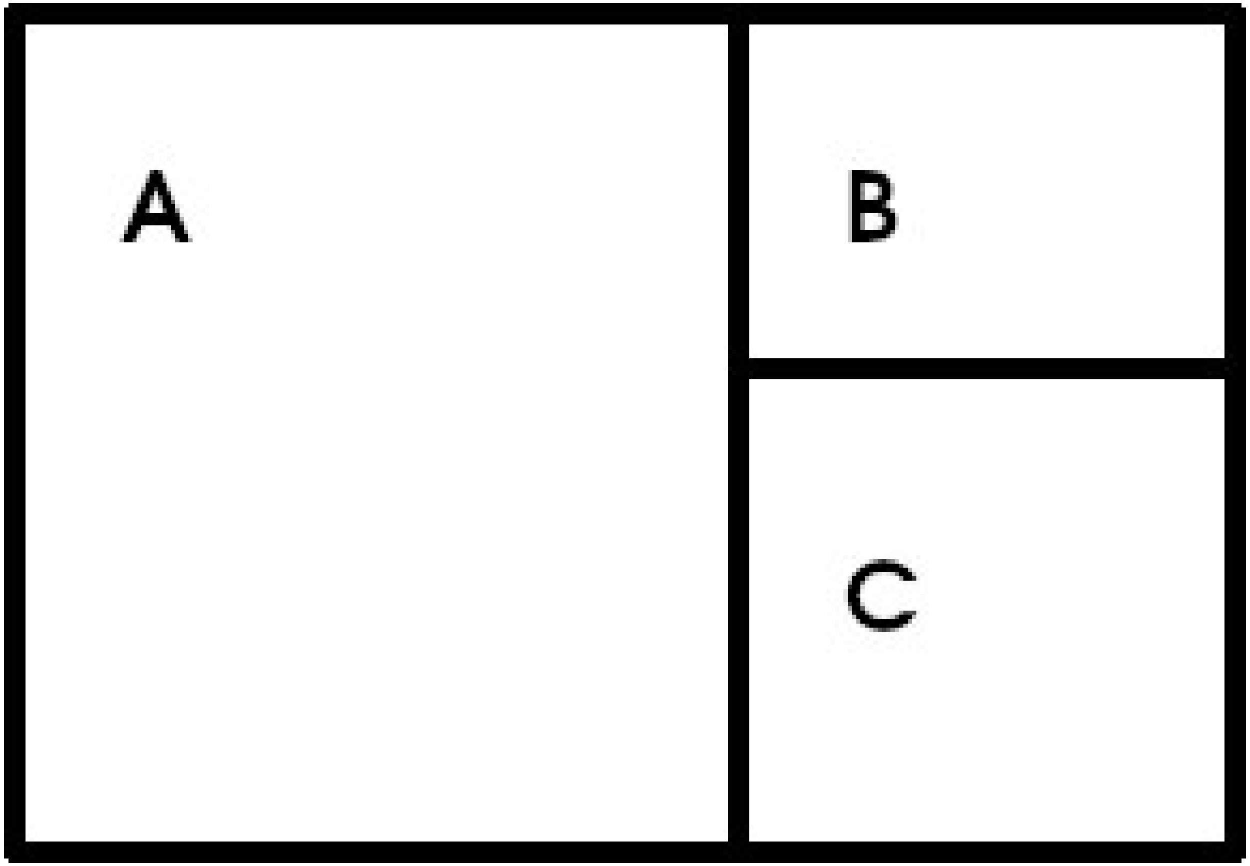
COMPLEMENTARIOS

A SU VEZ SON EXHAUSTIVOS



$$A \cap \overline{A} = \phi$$
$$A \cup \overline{A} = E$$

SISTEMA EXHAUSTIVO Y EXCLUYENTE



$$A \cap B = B \cap C = A \cap C = \phi$$
$$A \cup B \cup C = E$$

DEPENDENCIA

Se dice que dos eventos son dependientes, cuando la ocurrencia de uno afecta o influye en la ocurrencia del otro. Si por el contrario, la ocurrencia de un evento no influye en la ocurrencia del otro, se dice que son eventos independientes.

Dependencia de eventos ME

Si dos eventos son ME, la ocurrencia de uno implicaría la imposibilidad de ocurrencia del otro, ya que al no tener puntos en común, sería imposible que sucedan a la vez. De esta forma, dados dos eventos A, B que son ME entre sí, entonces A y B serán Dependientes. De manera similar, dos eventos complementarios, son además excluyentes, por lo que también serán dependientes.

Nociones Iniciales

- Probabilidad
  - Subjetiva
  - A priori
  - A posteriori o frecuencial
  - Contraria
  - Conjunta
  - Total

Nociones Iniciales

Probabilidad subjetiva

Las probabilidades subjetivas están basadas en las creencias de las personas que efectúan la estimación de probabilidad. De hecho, la probabilidad subjetiva se puede definir como la probabilidad asignada a un evento por parte de un individuo, basada en la evidencia que tenga disponible. (LEVIN, p.133)

Probabilidad de que mañana salga el sol:

Siendo d: cantidad de días en los que el sol ya ha salido.

A: “Mañana sadrá el sol”
P(A) = (d + 1) / (d + 2)

Si nos situásemos en un universo en el cual el sol ya hubiera d veces y otorgásemos diferentes valores a d, podríamos obtener las siguientes probabilidades:

$$d = 0 => P = 1/2 = 50\%$$
$$d = 1 => P = 2/3 \approx 66\%$$
$$d = 2 => P = 3/4 = 75\%$$
$$d = 3 => P = 4/5 = 80\%$$
$$d = 4 => P = 5/6 \approx 83\%$$
$$d = 5 => P = 6/7 \approx 86\%$$
$$d = 8 => P = 9/10 = 90\%$$
$$d = 98 => P = 99/100 = 99\%$$
$$d = 998 => P = 999/1000 \approx 99.9\%$$

Probabilidad de que mañana salga el sol:

Este planteo de Laplace nos permite pensar que en el caso inicial, tenemos la misma probabilidad de que el sol salga, que del suceso complementario, es decir, cuando d toma el valor 0, el evento A y su complementario son equiprobables.

A su vez, en la medida de que consideremos diferentes valores de d, cuanto mayor es el valor de d, más probable será que el sol salga y la probabilidad crecerá hasta acercarse aproximadamente al 99,99%.

PROBABILIDAD A PRIORI

La probabilidad “A priori”, también conocida como Probabilidad Clásica, se basa en el supuesto de que todos los puntos del EM tienen la misma probabilidad de ocurrir y en en ese contexto considera el cociente de los casos favorables sobre los casos posibles.

$$P(E) = \frac{n}{N}$$

n: Cantidad de casos favorables al evento E.

N: Cantidad de casos posibles en el espacio muestral.*

*Los N casos posibles son igualmente probables.

“A Priori”: “Previo a” – No se realizó el experimento

consultas: taulamet@unl.edu.ar

PROBABILIDAD A PRIORI

Retomando el EA de lanzar un dado de seis caras y considerando el evento que implica obtener un número par, tendremos que existen 6 casos posibles y 3 son favorables al evento planteado.

$$A: \text{“El número obtenido es par”}$$
$$P(A) = \frac{3}{6}$$
$$P(A) = 50\%$$

Para calcular esta probabilidad, no hace falta realizar el EA de lanzar el dado, puede calcularse antes realizar el lanzamiento del dado.

PROBABILIDAD A PRIORI

De manera análoga podríamos pensar un evento que represente la posibilidad de que el número sea el número 2:

$$B : \text{“El número obtenido es el 2”}$$
$$P(B) = \frac{1}{6}$$
$$P(B) \approx 17\%$$

De esta forma consideramos que de los seis casos posibles que existen, sólo uno de ellos es favorable a la posibilidad del número 2. De esta forma, la probabilidad de que salga el número dos es aproximadamente del 17% calculada mediante la probabilidad “A Priori”.

consultas: taulamet@unl.edu.ar

PROBABILIDAD A POSTERIORI

Este tipo de probabilidad, también llamada de Probabilidad Frecuencial, se calcula cuando se decide no suponer que todos los puntos del EM tienen la misma probabilidad de ocurrencia a priori, y propone repetir el evento n veces y luego ponderar la cantidad de veces que se obtuvo el resultado buscado.

De esta forma si quisiéramos saber cuál es la probabilida de que salga el número dos, podríamos lanzar el dado 100 veces, tomar nota de cuántas veces ha salido y calcular la probabilidad realizando el cociente de la cantidad de veces que ha salido dividido 100. La teoría plantea que a mayor cantidad de repeticiones, la probabilidad a posteriori se estabilizará y convergirá a una frecuencia relativa la probabilidad del suceso estudiado.

Podemos asociar este tipo de probabilidad al ejemplo presentado por Laplace sobre la probabilidad de que mañana salga el sol, que converge al 99,99% cuando d crece.

consultas: *taulamet@unl.edu.ar*

PROBABILIDAD A POSTERIORI

Este tipo de probabilidad se calcula cuando se decide no suponer que todos los puntos del EM tienen la misma probabilidad de ocurrencia a priori, y propone repetir el evento n veces y luego ponderar la cantidad de veces que se obtuvo el resultado buscado.

$$P(A) = \lim_{n \to \infty} \frac{n1}{n}$$

n1: Cantidad de veces que sucedió el evento A.

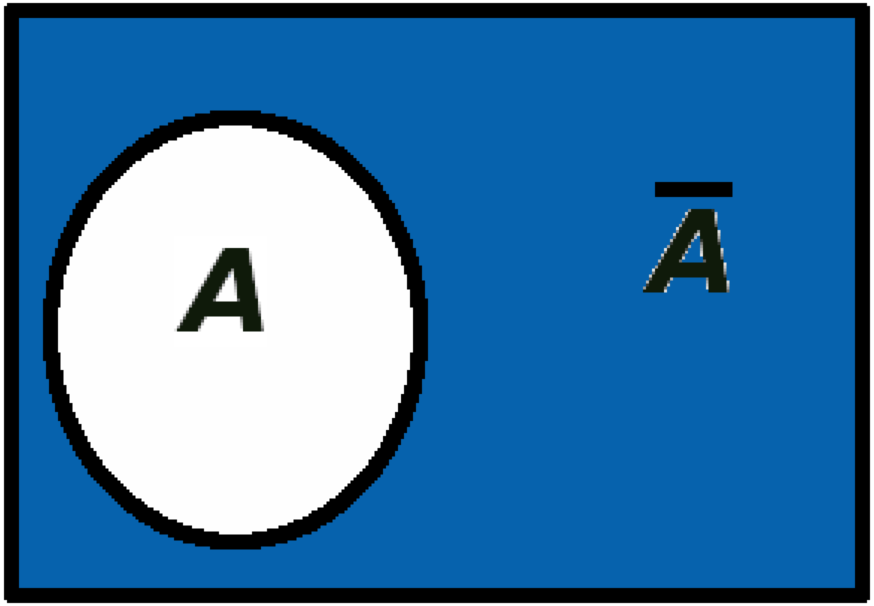
n: Cantidad de veces que se repitió el experimento.

n1: es una frecuencia absoluta (casos favorables)
n1/n: es una frecuencia relativa (casos favorables en relación
A la cantidad de repeticiones del experimento
"A Posteriori": "Posterior a"

PROBABILIDAD CONTRARIA

Considerando el EM incluye todos los resultados posibles, su probabilidad es del 100% (suceso seguro), la probabilidad contraria a un evento se puede calcular restando la probabilidad del EM a la probabilidad del evento. Es decir, dado un evento A, la probabilidad contraria de A será 1 menos la probabilidad de A.

$$P(\overline{A}) = 1 - P(A)$$



PROBABILIDAD CONJUNTA

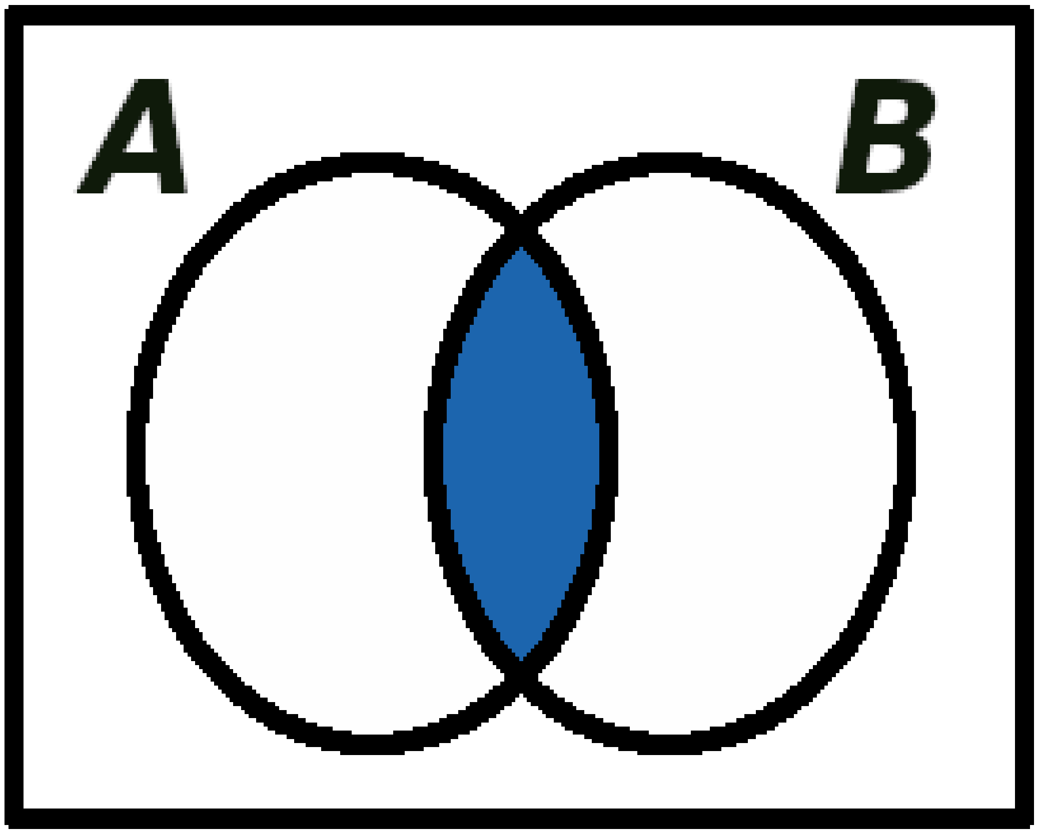
La probabilidad conjunta refiere a la intersección de dos o más eventos. De esta forma dados dos eventos A y B, la probabilidad conjunta de A y B, se relaciona con la probabilidad de que sucedan a la vez ambos eventos. Esta probabilidad refiere a que el resultado del EA que ocurra, se corresponda con un punto del EM que se encuentre incluido en el evento A y que a la vez se encuentre incluido en el evento B.

En eventos excluyentes
Si eventos son ME, los eventos no tienen puntos en común, por lo tanto la intersección es un suceso imposible y la probabilidad conjunta sera nula.

En condiciones de independencia estadística
Si dos eventos son independientes, la probabilidad conjunta de los mismos se puede calcular como el producto de las probabilidades de los eventos.

PROBABILIDAD CONJUNTA

CASO: EVENTOS NO EXCLUYENTES
INDEPENDENCIA O INDEPENDENCIA

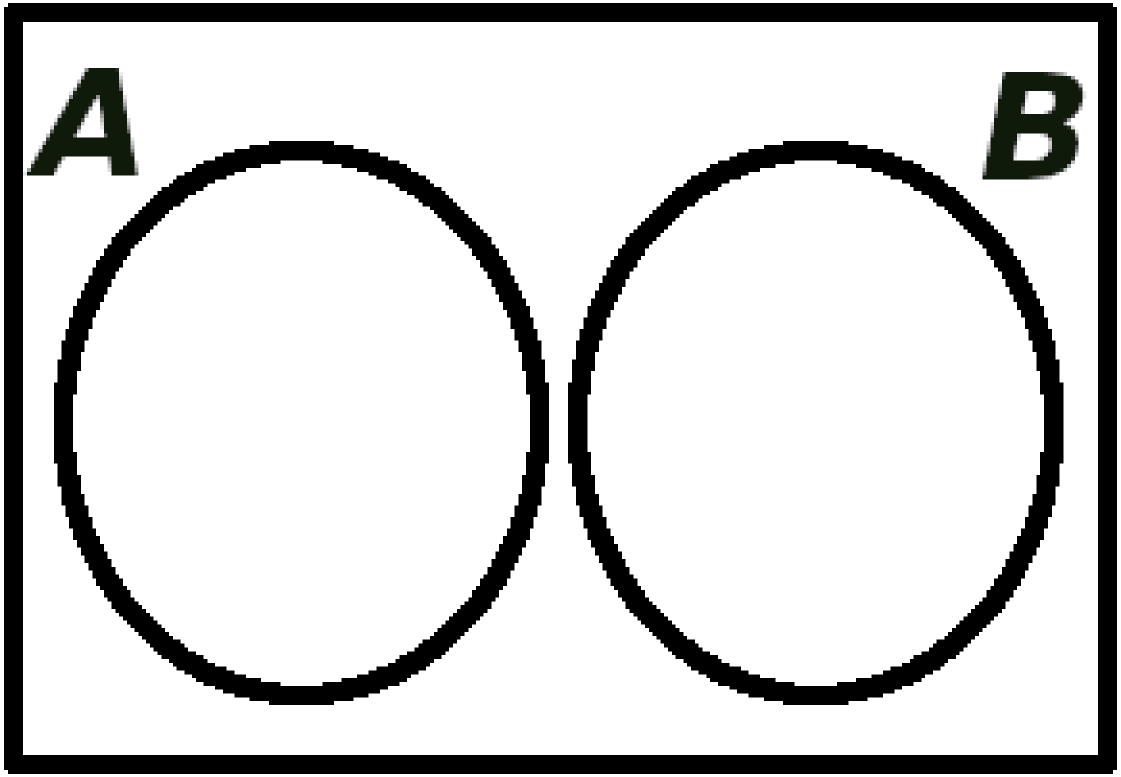


$$A, B : \text{Independientes} \iff P(A \cap B) = P(A) * P(B)$$
$$P(A \cap B) = P(A) * P(B) \implies A, B : \text{Independientes}$$
$$P(A \cap B) \neq P(A) * P(B) \implies A, B : \text{Dependientes}$$

consultas: taulamet@unl.edu.ar

PROBABILIDAD CONJUNTA

CASO: EVENTOS EXCLUYENTES
EXISTE DEPENDENCIA



$$A \cap B = \phi \implies P(A \cap B) = 0$$
$$P(A \cap B) = 0 \implies P(A \cap B) \neq P(A) * P(B)$$
$$A, B : \text{Excluyentes} \implies A, B : \text{Dependientes}$$

PROBABILIDAD TOTAL

La probabilidad total refiere a la unión de dos o más eventos. De esta forma dados dos eventos A y B, la probabilidad total de A y B, se relaciona con la probabilidad de la unión entre A y B.

Es importante considerar que cada uno de los puntos del espacio muestral incluidos en la probabilidad total, es decir en la probabilidad de la unión, deben ser considerados una sola vez.

Esto implica que cuando el conjunto de los eventos que deben considerarse son no excluyentes (es decir poseen puntos en común), éstos puntos, que representan la intersección, deberán contarse una sola vez. De esta forma, al considerar la probabilidad entre dos eventos A, B que son no excluyentes, si sumásemos P(A) + P(B), los puntos pertenecientes a ambos eventos, habrían sido contados dos veces. Es por ello que deberemos restar la probabilidad conjunta A y B.

PROBABILIDAD TOTAL

En base a lo anterior analizamos cada caso.

En eventos excluyentes

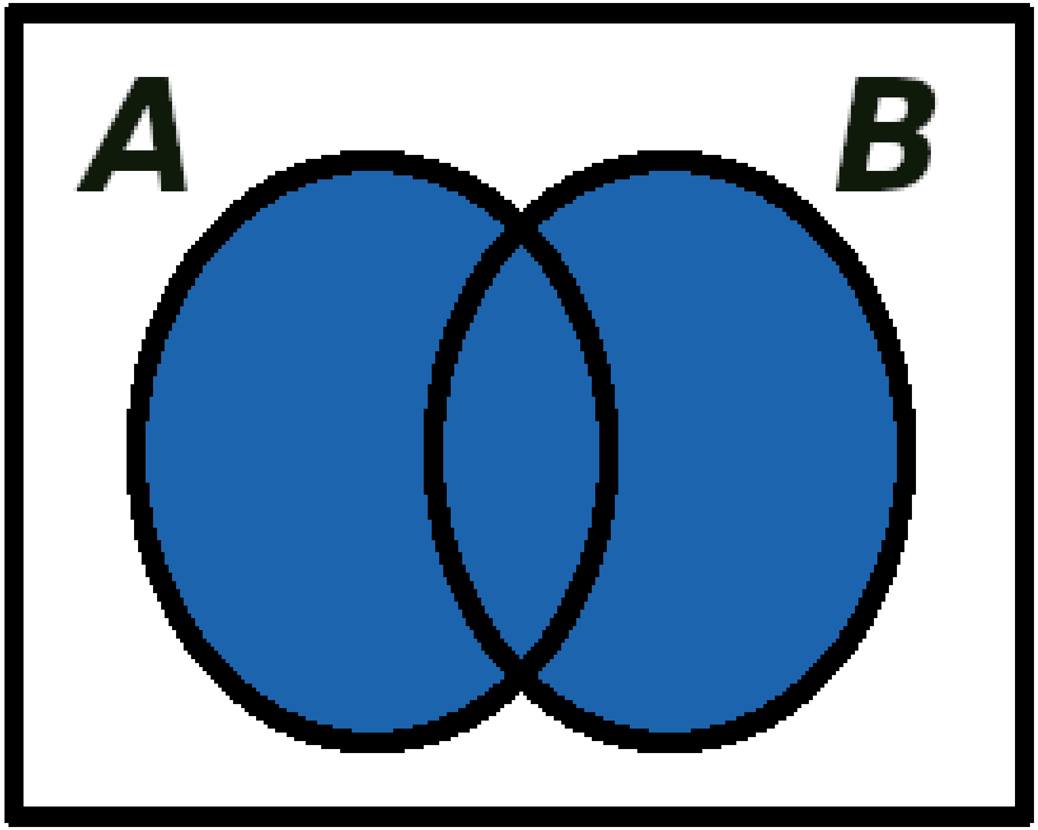
Si eventos son ME, los eventos no tienen puntos en común y la probabilidad total se obtiene sumando las probabilidades de los eventos. Todos los puntos han sido considerados una sola vez y por tanto no hace falta restar nada. Algo similar sucede cuando los eventos son complementarios, aunque en este caso la probabilidad total será del 100%, ya que refiere a todo el EM.

En eventos no excluyentes

Cuando dos eventos son no excluyentes, tienen puntos en común y la probabilidad total no puede calcularse simplemente sumando las probabilidades de los eventos, sino que debe considerarse aquellos puntos que pertenecen a la intersección entre los eventos una sola vez. De esta forma, dados dos eventos A y B, no excluyentes, la probabilidad total de A y B se obtendrá sumando las probabilidades de A y B y restando a dicha suma la probabilidad conjunta entre A y B.

PROBABILIDAD TOTAL

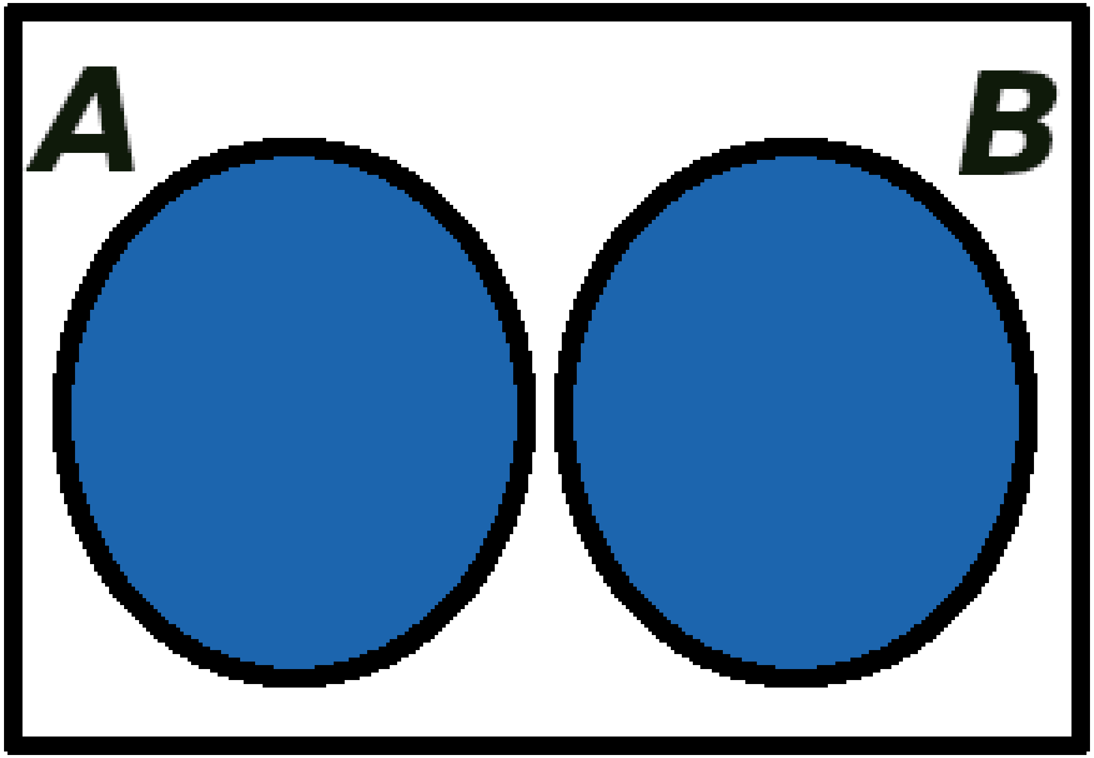
CASO: EVENTOS NO EXCLUYENTES



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

PROBABILIDAD TOTAL

CASO: EVENTOS EXCLUYENTES



$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$A, B \text{ Excluyentes: } \implies P(A \cap B) = 0$$
$$A, B \text{ Excluyentes: } \implies P(A \cup B) = P(A) + P(B)$$

consultas: taulamet@unl.edu.ar

ESTADÍSTICA
DIAPOSITIVAS DE EJEMPLO

Unidad 1 - *Probabilidad Total*

Ingenierías en Recursos Hídricos / Ambiental / Agrimensura

Año 2022

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

¿Cómo pensar un ejercicio?

- **Interpretar el enunciado**
  - Datos
  - Identificar Eventos y Espacio Muestral.
  - Interrogantes
- **Planteo**
- **Resolución**
- **Respuesta**

El 60% de las presas de hormigón construidas en el país tienen vertederos frontales. El 70% no están conectadas a la Red Interconectada Nacional (RIN), es decir, producen energía solamente para uso regional. El 10% de las represas tienen vertederos frontales y están conectadas a la RIN.

a) Si se seleccionara una represa al azar, ¿Cuál es la probabilidad de que la misma tenga vertederos frontales o este conectada a la RIN?
b) ¿El hecho de que una represa esté conectada a la RIN es independiende de si tiene o no vertederos frontales?

Datos

Definir Eventos o Sucesos

F: “La presa tiene vertedero frontal”

C: “La presa está conectada a la RIN”

60% = P(?)

70% = P(?)

10% = P(?)

Datos

Definir Eventos o Sucesos

F: “La presa tiene vertedero frontal”

C: “La presa está conectada a la RIN”

60% → $P(F) = 0{,}60$

70% = P(?) → $P(C) = 1 - 0{,}70 = 0{,}30$

10% = P(?) → $P(F \cap C) = 0{,}10$

Interrogante

¿Cuál es la probabilidad de que la misma tenga **vertederos frontales** o **esté conectada** a la RIN?

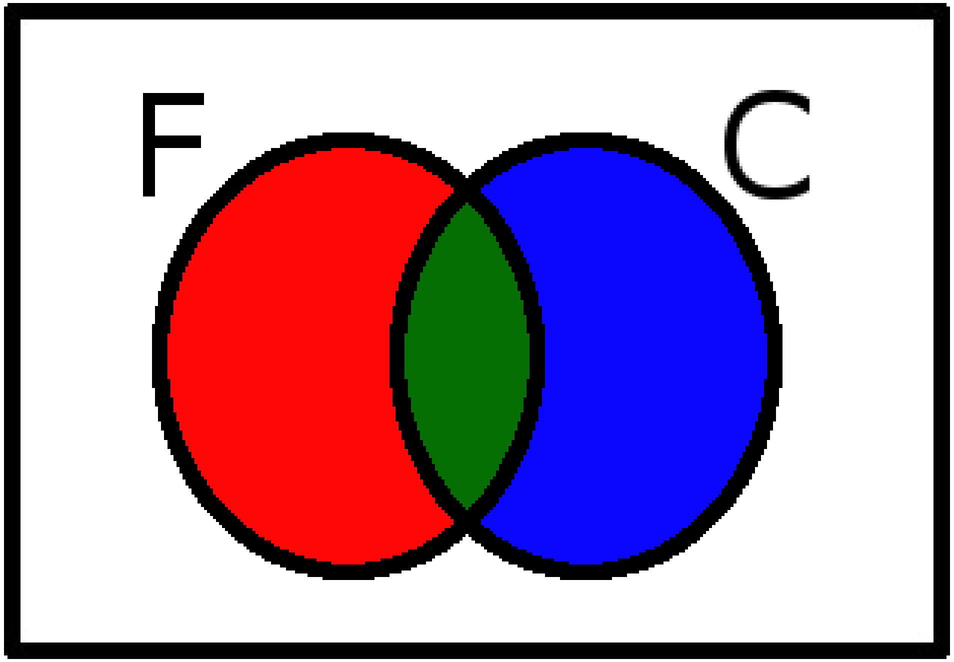
F o C → F U C

P(F U C) = ?

Probabilidad Total

Planteo: Probabilidad Total

$$P\,(F \cup C) = P(F) + P(C) - P(F \cap C)$$



Resolución

$$P(F) = 0,60$$
$$P(C) = 1 - 0,70 = 0,30$$
$$P(F \cap C) = 0,10$$
$$P(F \cup C) = P(F) + P(C) - P(F \cap C) =$$
$$= 0,60 + 0,30 - 0,10 = 0,80$$

Interrogante

¿El hecho de que una represa esté conectada a la RIN ***es independiente*** *de si tiene o no vertederos frontales?*

Resolución

$$F, C : \text{Independientes} \iff P(F \cap C) = P(F) * P(C)$$
$$P(F) = 0,60$$
$$P(C) = 1 - 0,70 = 0,30$$
$$\text{P(F) * P(C)} = 0{,}60*0{,}30 = 0{,}18$$
$$P(F \cap C) = 0,10 \neq 0,18$$

Respuestas

La probabilidad de que una presa tomada al azar tenga vertederos frontales o esté conectada a la RIN es del 80%.

El hecho de que una represa esté conectada a la RIN no es independiente de si tiene o no vertederos frontales.

(1994–2024)
30 años de la Consagración Constitucional de la Autonomía y Autarquía Universitaria en Argentina.

UNIVERSIDAD NACIONAL DEL LITORAL
LUX INDEFICIENS · 1919

FICH

ESTADÍSTICA

Notas de Clase #2 - *Teorema de Bayes*

Ingenierías en: Recursos Hídricos, Ambiental y Agrimensura

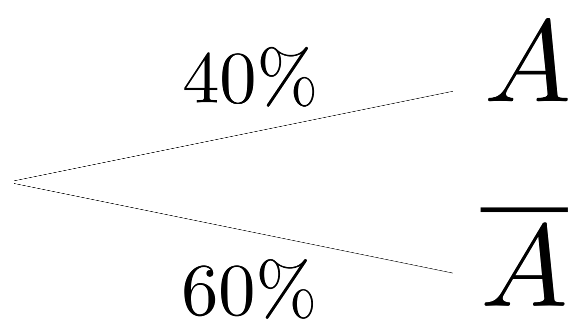
Año 2024

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

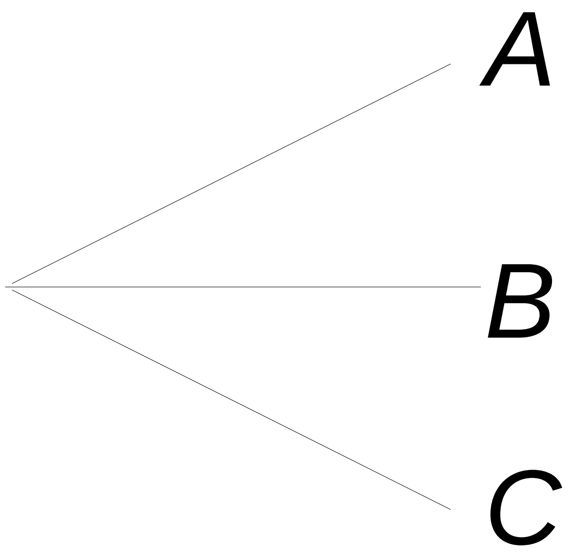
Árbol de probabilidad:

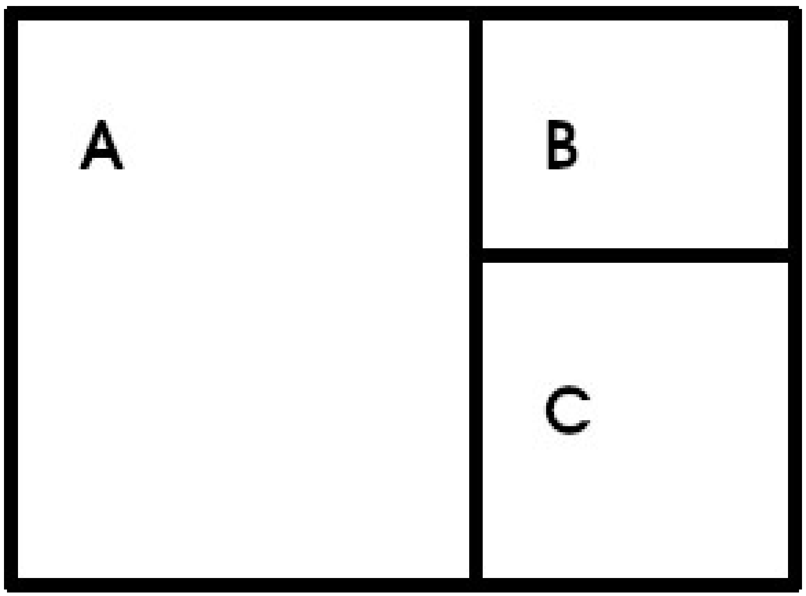
Se trata de un diagrama que permite representar el espacio muestral a partir de un conjunto de Eventos. El caso más sencillo podría ser el de dos eventos complementarios. De esta forma, si se define un evento A, que representa un conjunto de puntos en el espacio muestral, puede realizarse un diagrama de dos ramas, de tal forma que se asocie una rama a la ocurrencia del evento A y otra, a la ocurrencia del evento complementario a A. El diagrama permite incluir en la representación la probabilidad de ocurrencia de cada rama y la suma de las probabilidades de todas las ramas debe ser 1. Por ejemplo si el evento A, tiene una probabilidad de ocurrencia del 40%, la representación del EM podría realizarse mediante el siguiente diagrama de árbol de probabilidad:



Sistema Exhaustivo y Excluyente y Árbol:

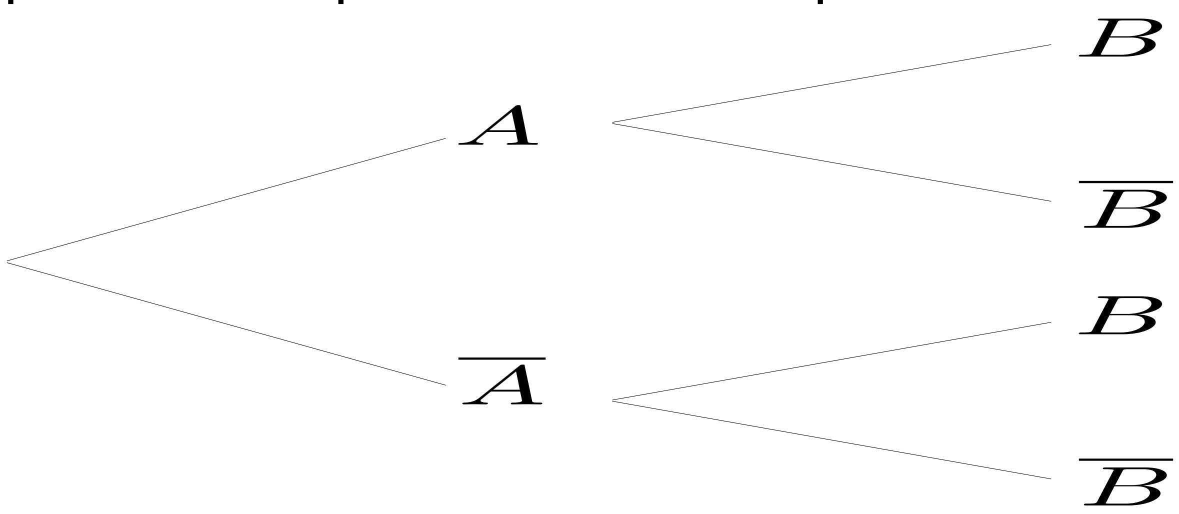
De manera similar al caso de los eventos complementarios, en el árbol de probabilidad se podrían representar cualquier conjunto de eventos que fueran excluyentes entre sí y que conformasen un sistema exhaustivo. Por ejemplo, dados 3 eventos A, B y C que conformen dicho sistema, se podrían generar los siguientes diagramas:





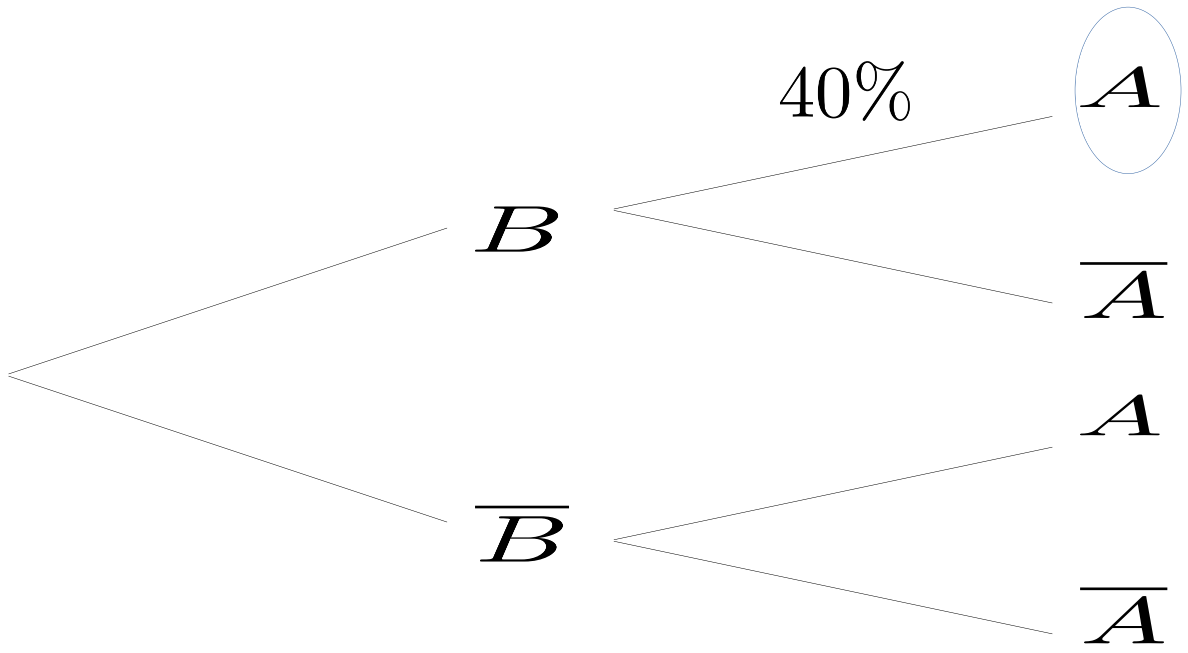
Árbol de dos niveles:

Se puede realizar un árbol que posea más de un nivel o generación de ramas, lo cual permitirá representar otro tipo de probabilidades. Por ejemplo, colocando en el primer nivel un evento A o su complementario y en un segundo nivel un evento B o su complementario. De esta forma, se pueden expresar 4 combinaciones de intersecciones. Tal como abordaremos a continuación, en el gráfico siguiente, no es lo mismo que suceda el evento B viniendo por el camino de la rama A, que viniendo por el camino complementario:



Probabilidad condicional:

Las probabilidades que se encuentran en el segundo nivel, se corresponden con probabilidades que hacen referencia determinados eventos, pero considerando el contexto de la rama del nivel precedente. Por ejemplo, el diagrama que sigue representa que la probabilidad de que suceda el evento A, en el contexto del evento B es del 40%. Esta probabilidad recibe el nombre de probabilidad condicional y se representa simbólicamente como P(A/B) = 40%.



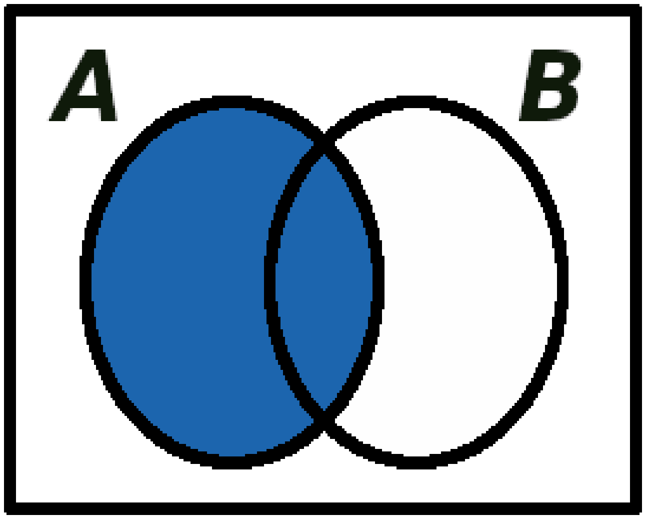
PROBABILIDAD CONDICIONAL

Desde el punto de vista conceptual, la probabilidad condicional de A dado B, no hace referencia a la probabilidad de que se obtenga como resultado alguno de los puntos del EM incluidos en A, sino más bien alguno de los puntos incluidos en A, dentro del universo de los puntos incluidos en el evento B. Es decir, en vez de considerar la probabilidad A en el contexto del espacio muestral, se considera la probabilidad de A en el contexto del evento B, que es un recorte del EM. Simbólicamente, la probabilidad condicional de A dado B, se define con la siguiente expresión:

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

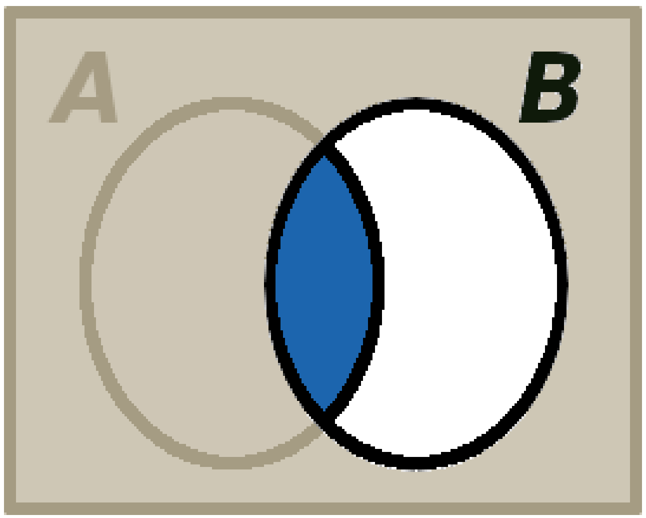
A continuación se presentan dos diagramas de Venn que señalan -en color azul-, al evento A, en dos contextos diferentes. En la imagen de la izquierda, se considera que no existe ningún condicionante, a la derecha, en el contexto del evento B, los únicos puntos de A que pueden considerarse, son aquellos que están además en incluidos en B.

Probabilidad de A en el espacio muestral.



$$P(A)$$

Probabilidad de A, en el contexto del evento B.



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

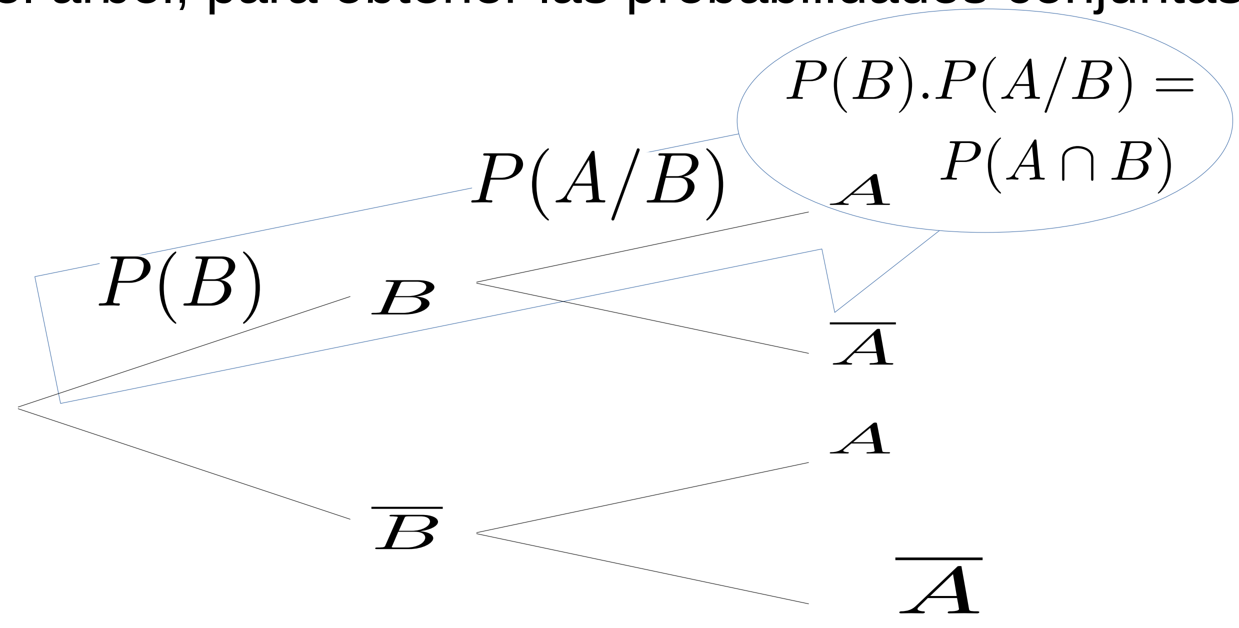
A partir de la fórmula de probabilidad la condicional siguiente:

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

Se puede despejar la probabilidad conjunta, dando origen a la expresión siguiente:

$$P(B).P(A/B) = P(A \cap B)$$

De manera equivalente a lo que puede obtenerse analíticamente, se puede considerar el producto de las probabilidades de cada rama del árbol, para obtener las probabilidades conjuntas:



De manera análoga, pueden agregarse en el árbol todas las probabilidades complementarias y el producto de las probabilidades de cada rama permitirá obtener las probabilidades conjuntas para cada combinación:

$P(B)$ — B

$P(A/B)$ — A — $P(B).P(A/B) = P(B \cap A)$

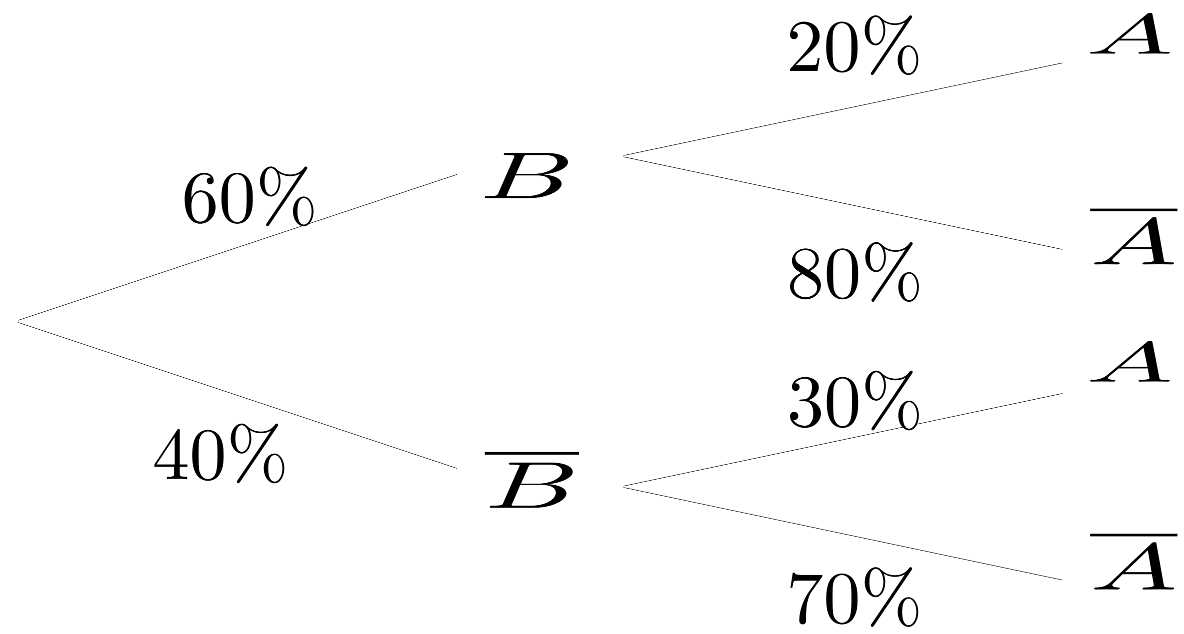
$P(\overline{A}/B)$ — $\overline{A}$ — $P(B).P(\overline{A}/B) = P(B \cap \overline{A})$

$P(\overline{B})$ — $\overline{B}$

$P(A/\overline{B})$ — A — $P(\overline{B}).P(A/\overline{B}) = P(\overline{B} \cap A)$

$P(\overline{A}/\overline{B})$ — $\overline{A}$ — $P(\overline{B}).P(\overline{A}/\overline{B}) = P(\overline{B} \cap \overline{A})$

A partir de lo dicho anteriormente, ante un diagrama de árbol como el que sigue,

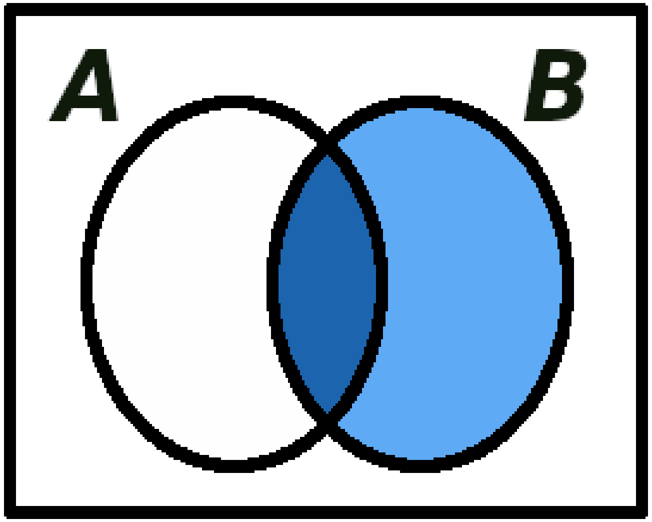


se puede deducir que la probabilidad conjunta de B y A, se puede calcular haciendo 0,60 x 0,20 = 0,12. Por otra parte si se analizan las probabilidades de A en los diferentes contextos, resulta evidente que la ocurrencia del evento B, influye reduciendo la probabilidad de que suceda A. De esta forma se puede concluir que existe dependencia entre A y B.

Si A y B son eventos no excluyentes, la fórmula de la probabilidad condicional se puede pensar como un cociente de casos favorables sobre posibles. Los casos posibles, serán naturalmente aquellos pertenecientes al evento que determina el contexto, es decir P(B). Lo casos favorables serán todos aquellos puntos que se encuentran en A en el contexto actual, es decir aquellos puntos que se encuentran en la intersección entre A y B.

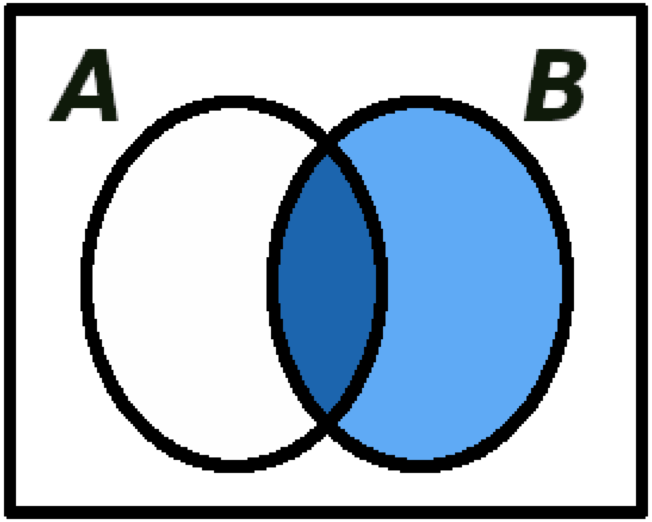
$$P(A/B) = P(A \cap B)/P(B)$$

CASO: EVENTOS NO EXCLUYENTES



Si A y B son eventos no excluyentes, pero además independientes, la fórmula de la probabilidad condicional se puede expresar reemplazando la probabilidad conjunta del numerador por el producto de las probabilidades de los eventos simples. De esta forma, como el evento B no influye en el evento A porque son independientes, la probabilidad del evento A en el contexto de B toma el mismo valor que en el espacio muestral, es decir P(A).

CASO: EVENTOS NO EXCLUYENTES E INDEPENDIENTES



$$P(A/B) = P(A \cap B)/P(B)$$
$$A, B : \text{Independientes} \implies P(A \cap B) = P(A) * P(B)$$
$$\implies P(A/B) = \frac{P(A) * P(B)}{P(B)} = P(A)$$

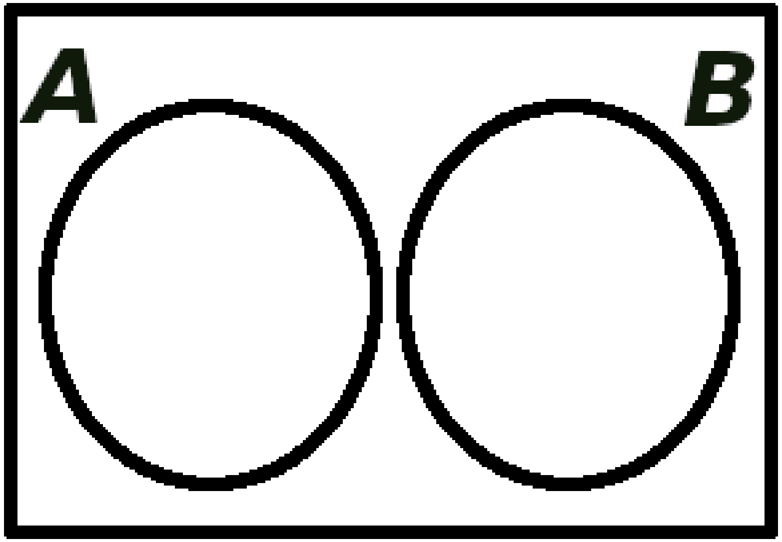
consultas: taulamet@unl.edu.ar

PROBABILIDAD CONDICIONAL

Si A y B son eventos excluyentes, la probabilidad conjunta es cero por lo tanto la probabilidad condicional también se anulará.

$$P(A/B) = P(A \cap B)/P(B)$$
$$A, B : \text{Excluyentes} \implies P(A \cap B) = 0$$
$$P(A/B) = 0/P(B) = 0$$

CASO: EVENTOS EXCLUYENTES

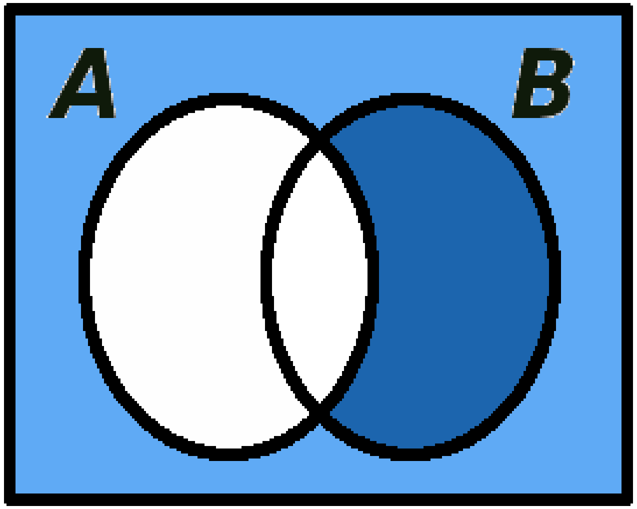


consultas: taulamet@unl.edu.ar

PROBABILIDAD CONTRARIA

Las probabilidades de los eventos complementarios pueden calcularse de la misma forma que en el espacio muestral siempre que se respete el mismo condicionante. De esta forma, en un mismo contexto, el complemento de un evento posee una probabilidad igual a 1 menos la probabilidad del evento complementado.

CASO: CONDICIONAL



$$P(\overline{A}/B) = 1 - P(A/B)$$

Tal como se ha establecido, de la definición de la probabilidad condicional de A dado B, se puede despejar la probabilidad conjunta, dando origen a la expresión siguiente:

$$P(A \cap B) = P(A/B).P(B)$$

A su vez, como la intersección de A y B es el mismo conjunto de puntos en el espacio muestral que la intersección de B y A, puede plantearse la siguiente igualdad:

$$P(B \cap A) = P(B/A).P(A)$$

Con lo cual se tiene además que:

$$P(A/B).P(B) = P(B/A).P(A)$$

Como cierre de todo lo abordado en probabilidad, se presenta a continuación la generalización propuesta en el Teorema de Bayes.

TEOREMA DE BAYES

Dados un conjunto de eventos A_1, A_2, …, A_n, que conforman un sistema exhaustivo y excluyente de eventos -que se denominan hipótesis-, la probabilidad de ocurrencia de alguno de estos eventos, conociendo la ocurrencia de cualquier evento B perteneciente al espacio muestral podrá calcularse con la siguiente expresión:

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

A su vez, probabilidad total:

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

Y por probabilidad condicional:

$$P(A_i \cap B) = P(B/A_i).P(A_i)$$

Por lo que:

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

consultas: taulamet@unl.edu.ar

Proponemos pensar el teorema de Bayes, como una forma de integrar, relacionar y resumir todo lo que hemos visto en la presente unidad.

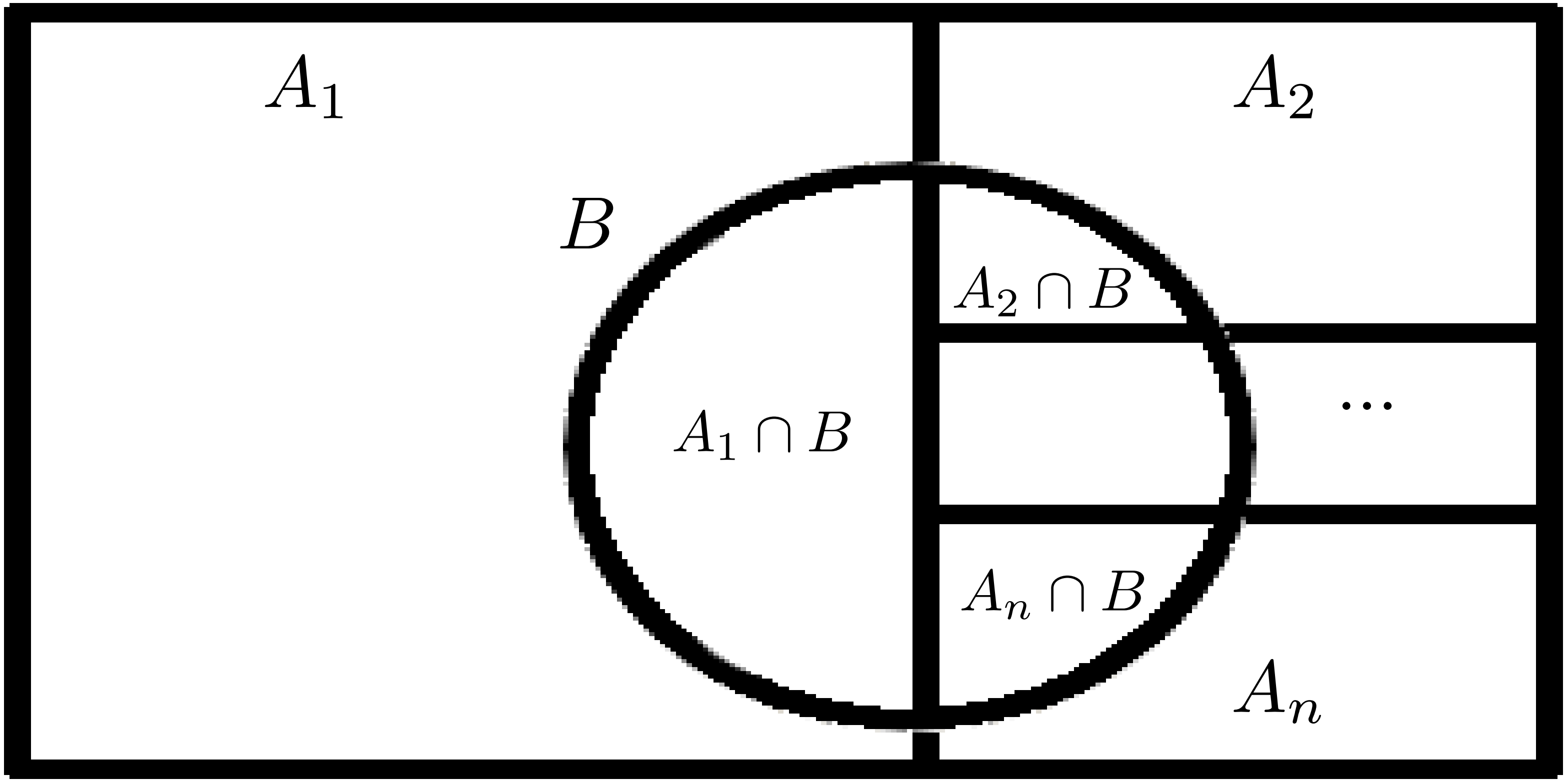
A partir de las nociones iniciales podemos pensar en la partición del espacio muestral en un conjunto hipótesis que forman un sistema exhaustivo y excluyente de eventos.

A partir de las probabilidades de Laplace, pensar el cociente de lo favorable a lo que nos interesa estudiar sobre lo posible.

Considerando la probabilidad condicional, proponer el cociente de la probabilidad conjunta sobre la probabilidad del evento que condiciona.

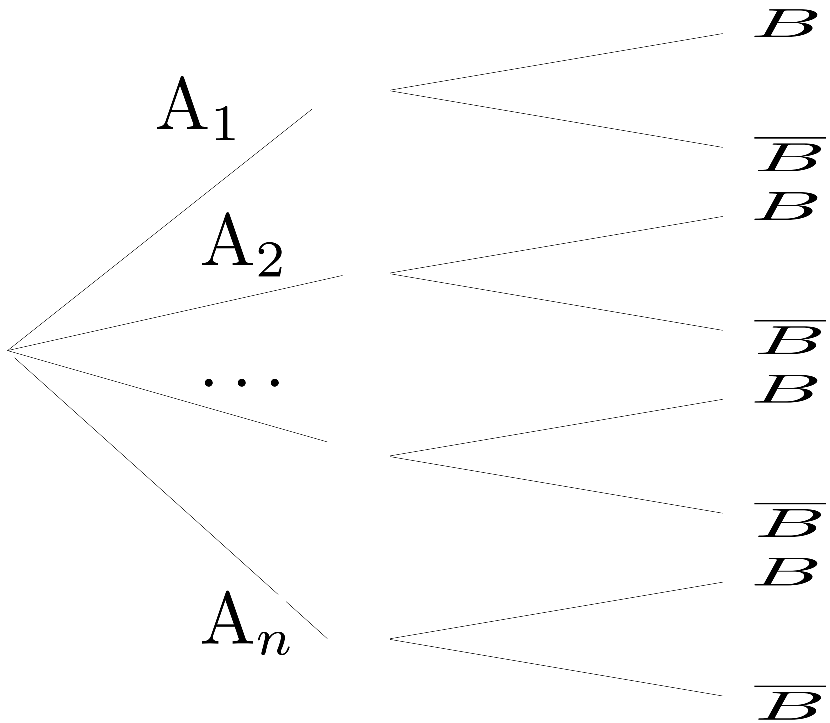
A partir de la combinación de todo lo anterior, podremos pensar la probabilidad condicional de cualquiera de las hipótesis dada la ocurrencia de evento de interés, como la probabilidad conjunta del evento en el contexto de la ocurrencia del mismo.

Presentamos a continuación algunos diagramas, de Venn y de Árbol que ilustran el Teorema de Bayes.



Si el espacio muestral es particionado en un conjunto de hipótesis A_1, A_2,..., A_n, que son eventos que conforman un sistema exhaustivo y excluyente, cualquier evento B de dicho EM, podrá conformarse como la unión de las intersecciones de cada hipótesis con dicho evento.

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



$$\mathrm{P}(\mathrm{A}_1 \cap B) = P(B/A_1).P(A_1)$$

$$\mathrm{P}(\mathrm{A}_2 \cap B) = P(B/A_2).P(A_2)$$

$$\mathrm{P}(\mathrm{A}_i \cap B) = P(B/A_i).P(A_i)$$

$$\mathrm{P}(\mathrm{A}_n \cap B) = P(B/A_n).P(A_n)$$

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

consultas: taulamet@unl.edu.ar

¿Cómo pensar un ejercicio?

- **Interpretar el enunciado**
  - Datos
  - Identificar Eventos y Espacio Muestral.
  - Interrogantes
- **Planteo**
- **Resolución**
- **Respuesta**

2021 ~ Año de homenaje
al Premio Nobel de Medicina
Dr. César Milstein

ESTADÍSTICA
DIAPOSITIVAS DE EJEMPLO

Unidad 1 – *Teorema de Bayes*

Ingenierías en Recursos Hídricos / Ambiental / Agrimensura

Año 2021

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

Tres empresas se presentan a una licitación para la construcción de un puente.

Las probabilidades de que E1, E2 y E3 obtengan el contrato son 0.5, 0.3, y 0.2 respectivamente.

Se sabe además que si E1 es seleccionada, subcontratará a A como ejecutor de obra con probabilidad 0.8; si lo obtiene E2 la seleccionará con probabilidad 0.4; y si lo obtiene E3 lo hará con probabilidad 0.1.

Si se sabe que se ha subcontratado a la empresa A, ¿Cuál es la probabilidad de que haya obtenido el contrato E2?

Datos: Eventos y Probabilidades

E_1: “E1 obtiene el contrato”

E_2: “E2 obtiene el contrato”

E_3: “E3 obtiene el contrato”

$P(E_1) = 50\%$

$P(E_2) = 30\%$

$P(E_3) = 20\%$

Datos: Eventos y Probabilidades

“Si E_1 obtiene el contrato, subcontratará a A con probabilidad 0,8.”

A: “A es subcontratada”

$P(A/E_1) = 80\%$

$P(A/E_2) = 40\%$

$P(A/E_3) = 10\%$

¿Pregunta?

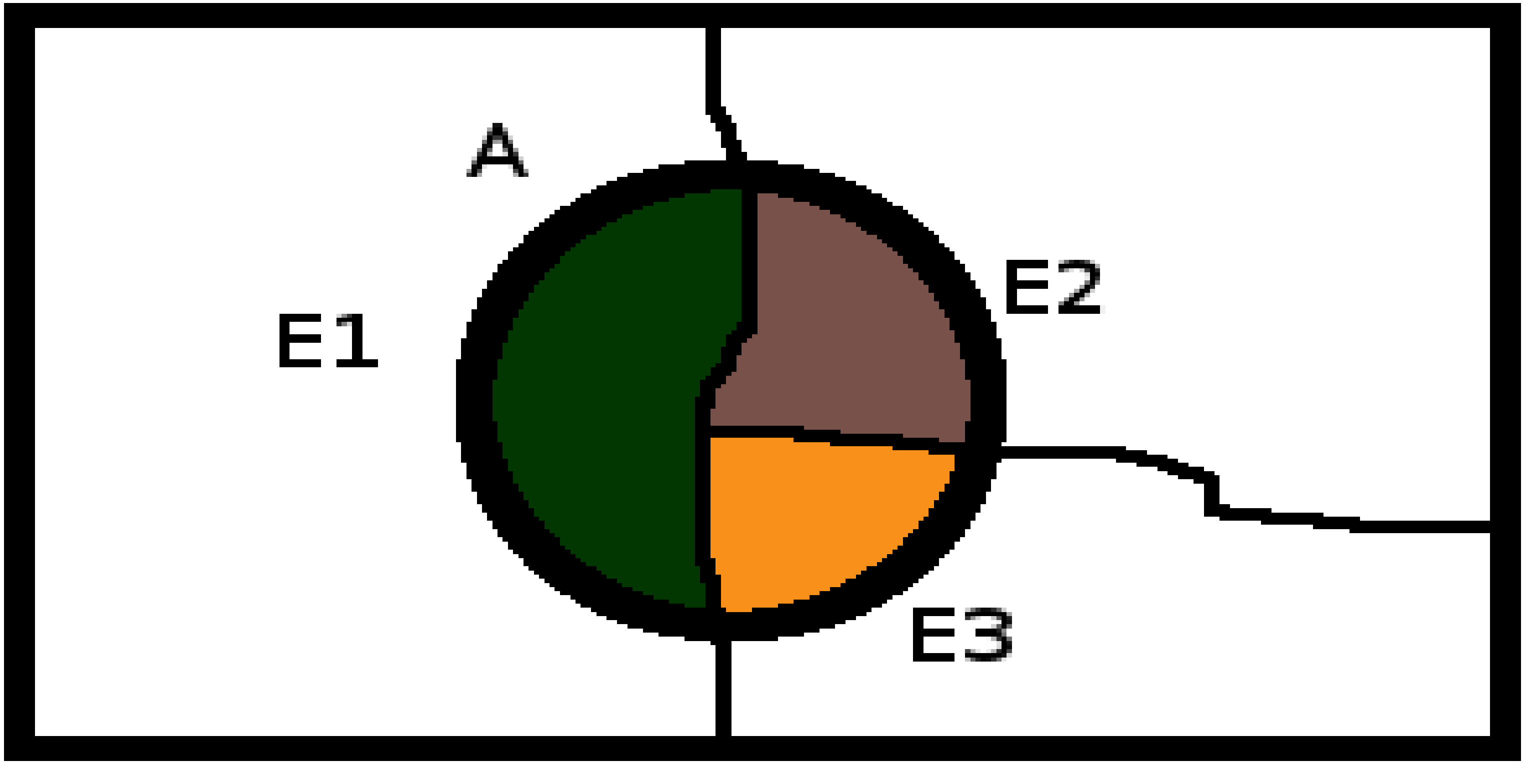
Si se sabe que se ha subcontratado a la empresa A, ¿Cuál es la probabilidad de que haya obtenido el contrato E2?

¿Pregunta?

Si se sabe que se ha subcontratado a la empresa A, ¿Cuál es la probabilidad de que haya obtenido el contrato E2?

$$P(E_2/A) = ?$$

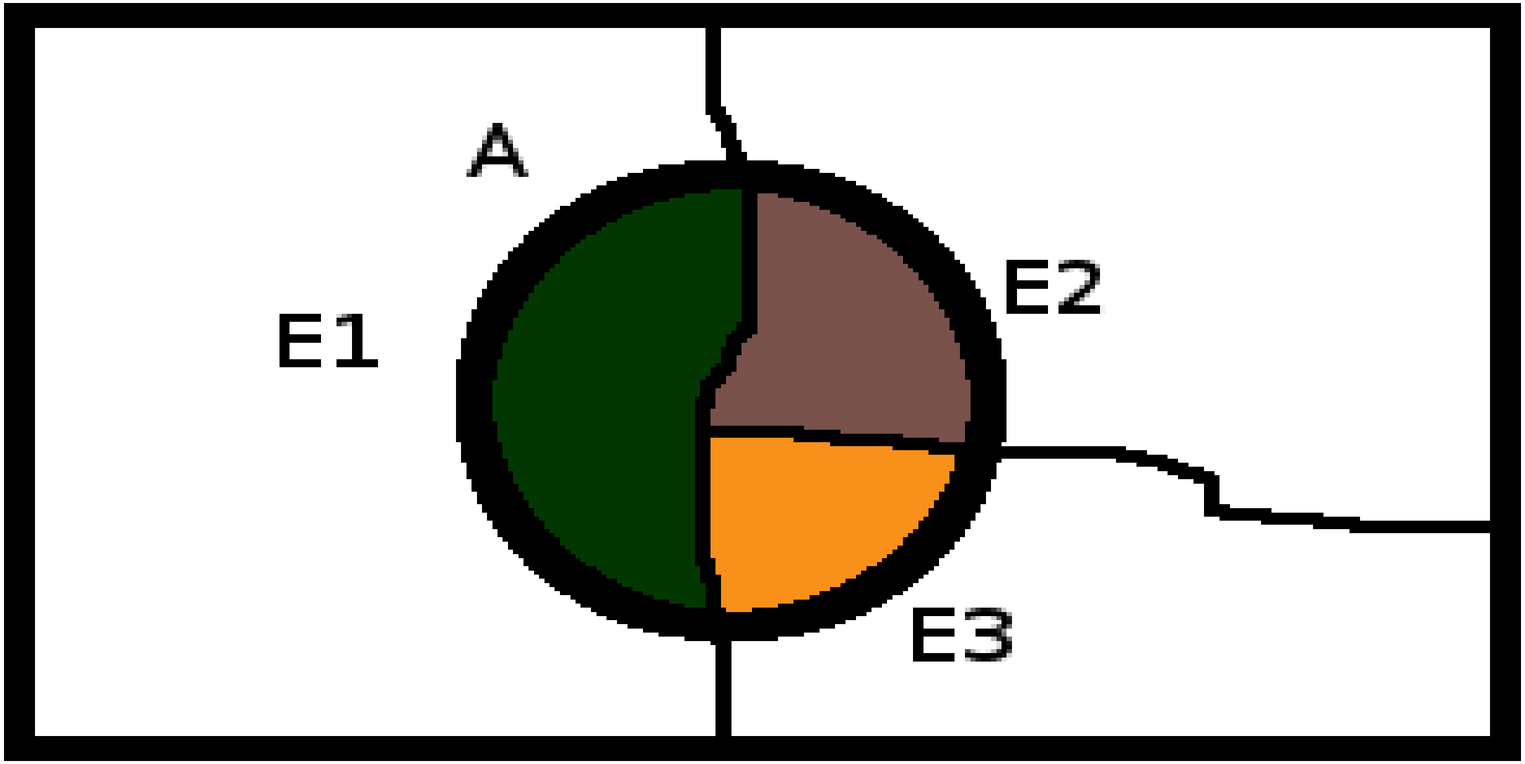
Teorema de Bayes



¿Pregunta?

$$P(E_2/A) = P(E_2 \cap A) / P(A)$$

$$P(A) = P(A \cap E_1) + P(A \cap E_2) + P(A \cap E_3)$$



¿Pregunta?

$$P(E_2/A) = P(E_2 \cap A) \;/\; P(A)$$

$$P(A) = P(A \cap E_1) + P(A \cap E_2) + P(A \cap E_3)$$

$$\begin{aligned} P(A) = \; & P(A/E_1) * P(E_1) + \\ & P(A/E_2) * P(E_2) + \\ & P(A/E_3) * P(E_3) \;\rightarrow \end{aligned}$$

$$\begin{aligned} P(A) = \; & 0{,}8*0{,}5 + 0{,}4*0{,}3 + 0{,}1*0{,}2 \\ & 0{,}40 + 0{,}12 + 0{,}02 = 54\% \end{aligned}$$

$$P(E_2/A) = 0{,}12 \;/\; 0{,}54 = 22{,}22\%$$

Respuesta

Si se sabe que se ha subcontratado a la empresa A, la probabilidad de que la empresa que la haya subcontratado sea *E2* es del 22%.

(1994–2024)
30 años de la Consagración Constitucional de la Autonomía y Autarquía Universitaria en Argentina.

UNIVERSIDAD NACIONAL DEL LITORAL
LUX INDEFICIENS · 1919

FICH

ESTADÍSTICA
DIAPOSITIVAS DE EJEMPLO

Unidad 1 - *Probabilidad*

Ingenierías en Recursos Hídricos / Ambiental / Agrimensura

Año 2024

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

¿Cómo pensar un ejercicio?

- **Interpretar el enunciado**
  - Datos
  - Identificar Eventos y Espacio Muestral.
  - Interrogantes
- **Planteo**
- **Resolución**
- **Respuesta**

Un test detecta la presencia de cierto tipo de bacterias en el agua con probabilidad 0.9, en caso de haberlas. Si no las hay, detecta la ausencia con probabilidad de 0.8. Sabiendo que la probabilidad de que una muestra de agua contenga bacterias de ese tipo es 0.2, calcular las probabilidades siguientes:

a) Realmente haya presencia de bacterias cuando el test ha dado resultado positivo.

b) Realmente haya presencia de bacterias cuando el test ha dado resultado negativo.

c) Haya bacterias y además el test dé positivo.

d) O haya bacterias, o el test dé positivo.

Datos

Definir Eventos o Sucesos

H: “Hay bacterias en el agua”

T: “El test da positivo”

Un test detecta la presencia de cierto tipo de bacterias en el agua con probabilidad 0.9, en caso de haberlas. *Si no las hay, detecta la ausencia con probabilidad de 0.8. Sabiendo que la probabilidad de que una muestra de agua contenga bacterias de ese tipo es 0.2, calcular la probabilidad de...*

H: "Hay bacterias en el agua"
T: "El test da positivo"

$$90\% = P(?)$$

Un test detecta la presencia de cierto tipo de bacterias en el agua con probabilidad 0.9, en caso de haberlas. *Si no las hay, detecta la ausencia con probabilidad de 0.8. Sabiendo que la probabilidad de que una muestra de agua contenga bacterias de ese tipo es 0.2, calcular la probabilidad de...*

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$90\% = P(T/H)$$

consultas: taulamet@unl.edu.ar

Un test detecta la presencia de cierto tipo de bacterias en el agua con probabilidad 0.9, en caso de haberlas. ***Si no las hay, detecta la ausencia con probabilidad de 0.8.*** *Sabiendo que la probabilidad de que una muestra de agua contenga bacterias de ese tipo es 0.2, calcular la probabilidad de que:*

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$80\% = P(?)$$

Un test detecta la presencia de cierto tipo de bacterias en el agua con probabilidad 0.9, en caso de haberlas. ***Si no las hay, detecta la ausencia con probabilidad de 0.8.*** *Sabiendo que la probabilidad de que una muestra de agua contenga bacterias de ese tipo es 0.2, calcular la probabilidad de que:*

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$80\% = P(\overline{T}/\overline{H})$$

Un test detecta la presencia de cierto tipo de bacterias en el agua con probabilidad 0.9, en caso de haberlas. Si no las hay, detecta la ausencia con probabilidad de 0.8. Sabiendo que ***la probabilidad de que una muestra de agua contenga bacterias de ese tipo es 0.2,*** *calcular la probabilidad de que:*

H: "Hay bacterias en el agua"
T: "El test da positivo"

$$20\% = P(?)$$

Un test detecta la presencia de cierto tipo de bacterias en el agua con probabilidad 0.9, en caso de haberlas. Si no las hay, detecta la ausencia con probabilidad de 0.8. Sabiendo que ***la probabilidad de que una muestra de agua contenga bacterias de ese tipo es 0.2,*** *calcular la probabilidad de que:*

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$20\% = P(H)$$

Datos

Definir Eventos o Sucesos

H: “Hay bacterias en el agua”

T: “El test da positivo”

$$P(T/H) = 90\%$$

$$P(\overline{T}/\overline{H}) = 80\%$$

$$P(H) = 20\%$$

Interrogantes

H: "Hay bacterias en el agua"
T: "El test da positivo"

$$P(T/H) = 90\%$$
$$P(\overline{T}/\overline{H}) = 80\%$$
$$P(H) = 20\%$$

a) Realmente haya presencia de bacterias cuando el test ha dado resultado positivo.

b) Realmente haya presencia de bacterias cuando el test ha dado resultado negativo.

c) Haya bacterias y además el test dé positivo.

d) O haya bacterias, o el test dé positivo.

consultas: taulamet@unl.edu.ar

Planteo: P. Condicional

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$P(T/H) = 90\%$$
$$P(\overline{T}/\overline{H}) = 80\%$$
$$P(H) = 20\%$$

a) Realmente haya presencia de bacterias cuando el test ha dado resultado positivo.

$$P(H/T) = ?$$

Planteo: P. Condicional

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$P(T/H) = 90\%$$
$$P(\overline{T}/\overline{H}) = 80\%$$
$$P(H) = 20\%$$

a) Realmente haya presencia de bacterias cuando el test ha dado resultado positivo.

$$P(H/T) = \frac{P(H \cap T)}{P(T)}$$

Entendiendo T

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$P(T/H) = 90\%$$
$$P(\overline{T}/\overline{H}) = 80\%$$
$$P(H) = 20\%$$

a) Realmente haya presencia de bacterias cuando el test ha dado resultado positivo.

$$P(H/T) = \frac{P(H \cap T)}{P(T \cap H \cup T \cap \bar{H})}$$

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$P(T/H) = 90\%$$
$$P(\overline{T}/\overline{H}) = 80\%$$
$$P(H) = 20\%$$

$$P(H/T) = \frac{P(H \cap T)}{P(T \cap H) + P(T \cap \bar{H})}$$

Aplicando Prob. Condicional

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$P(T/H) = 90\%$$
$$P(\overline{T}/\overline{H}) = 80\%$$
$$P(H) = 20\%$$

$$P(H/T) = \frac{P(H \cap T)}{P(T \cap H) + P(T \cap \bar{H})}$$

$$\frac{P(T/H) * P(H)}{P(T/H) * P(H) + P(T/\bar{H}) * P(\bar{H})}$$

Resultado

H: “Hay bacterias en el agua”
T: “El test da positivo”

$$P(T/H) = 90\%$$
$$P(\overline{T}/\overline{H}) = 80\%$$
$$P(H) = 20\%$$

$$P(H/T) =$$

$$\frac{P(T/H)*P(H)}{P(T/H)*P(H)+P(T/\bar{H})*P(\bar{H})}$$

$$\frac{90\%*20\%}{90\%*20\%+(1-80\%)*(1-20\%)} \approx 53\%$$

a) Realmente haya presencia de bacterias cuando el test ha dado resultado positivo.

$$P(H/T) = \frac{P(H \cap T)}{P(T \cap H \cup T \cap \bar{H})}$$

$$\frac{90\% * 20\%}{90\% * 20\% + (1 - 80\%) * (1 - 20\%)} \approx 53\%$$

consultas: taulamet@unl.edu.ar

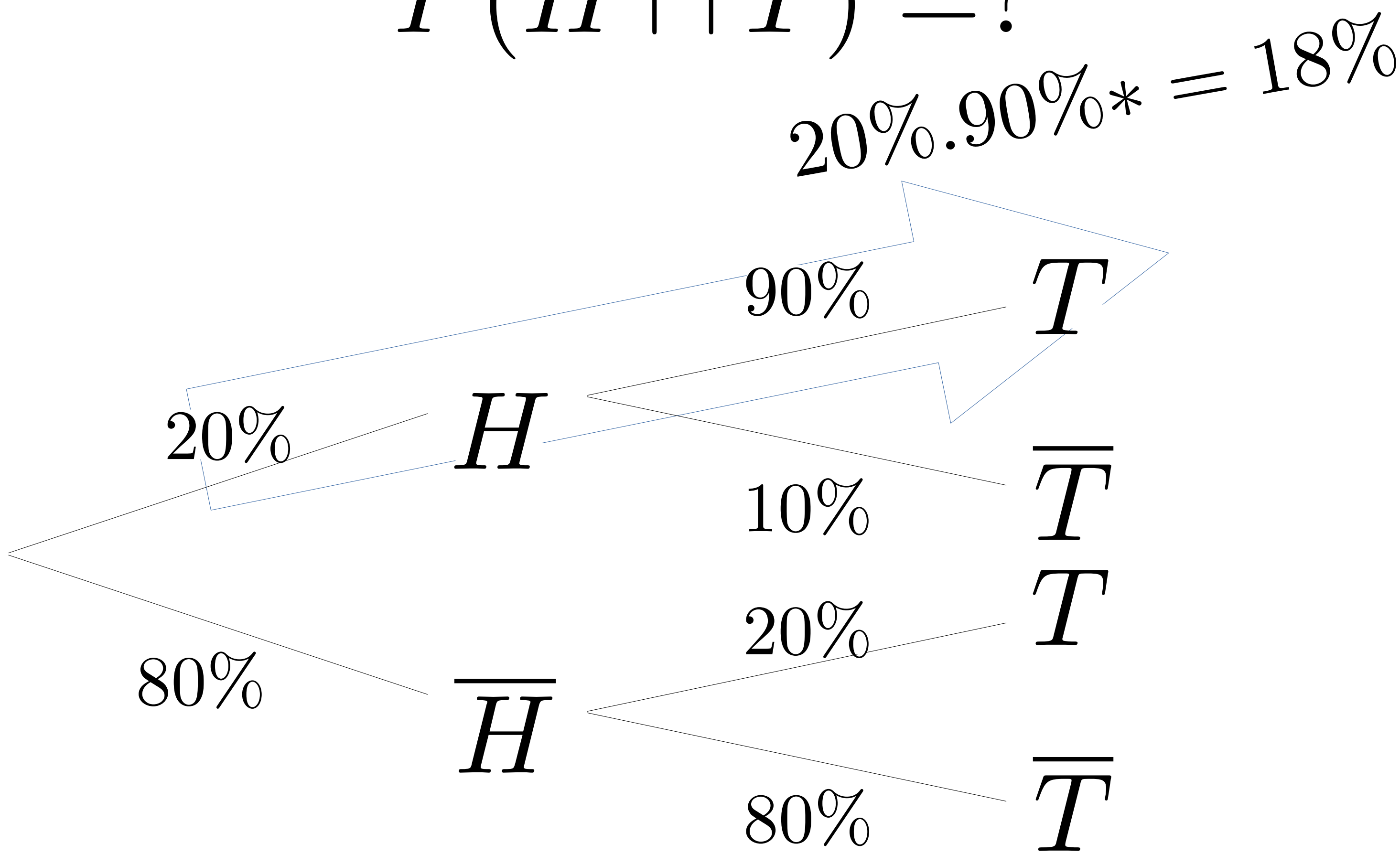
b) Realmente haya presencia de bacterias cuando el test ha dado resultado negativo.

$$P(H/\bar{T}) = \frac{P(H \cap \bar{T})}{P(H \cap \bar{T} \cup \bar{H} \cap \bar{T})}$$

$$\frac{20\% * 10\%}{20\% * 10\% + 80\% * 80\%} \approx 3\%$$

c) Haya bacterias y además el test dé positivo.

$$P(H \cap T) = ?$$



d) O haya bacterias, o el test dé positivo.

$P(H \cup T) = ?$

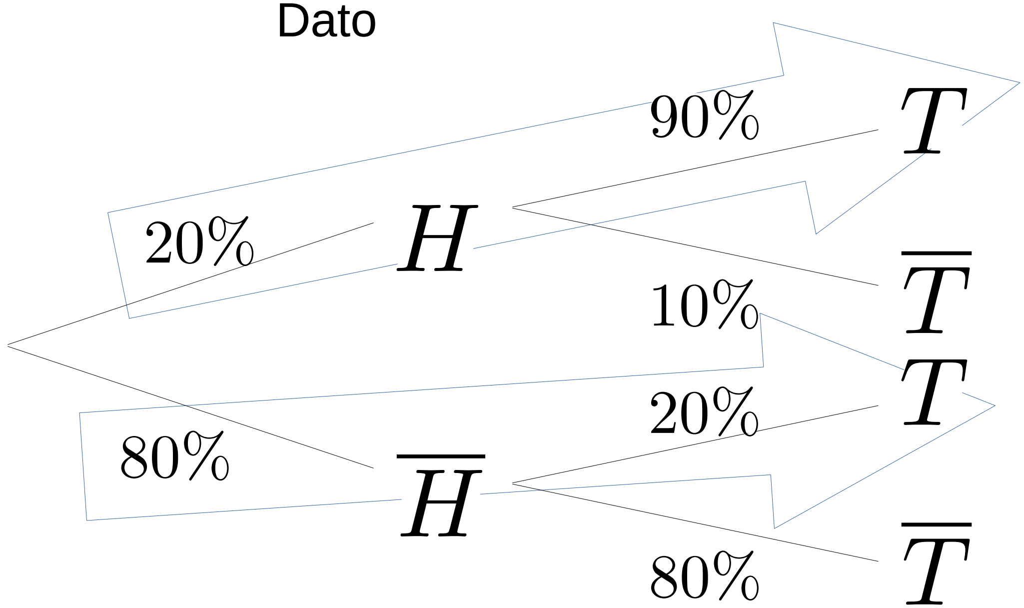
Por P. Total:

$P(H) + P(T) - P(H \cap T)$

$P(T) = P(T \cap H) + P(T \cap \overline{H}) = 18\% + 16\%) = 34\%$

$P(H \cup T) = 20\% + 34\% - 18\% = 36\%$

Dato



$\mathrm{P}(\mathrm{T} \cap H)$

$20\%.90\%* = 18\%$

$P(T \cap \overline{H})$

$80\%.20\%* = 16\%$

Respuestas

a) La probabilidad de que haya presencia de bacterias cuando el test ha dado resultado positivo es 53%.

b) La probabilidad de que haya presencia de bacterias cuando el test ha dado resultado negativo es 3%. (Falso negativo).

c) La probabilidad de que haya bacterias y además el test dé positivo es 18% y es una probabilidad conjunta.

d) La probabilidad de que haya bacterias, o el test dé positivo, es una probabilidad total que vale 36%.

(1994–2024)
30 años de la Consagración Constitucional de la Autonomía y Autarquía Universitaria en Argentina.

ESTADÍSTICA

Notas de Clase #3 - *Variables Aleatorias*

Ingenierías en: Recursos Hídricos, Ambiental y Agrimensura

Año 2024

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

VARIABLES ALEATORIAS

- Definición
- Clasificación
- Distribución de probabilidades
  - Funciones f(x) y F(x)
    - Los símbolos y los nombres
  - Momentos y propiedades
  - Funciones marginales
  - Independencia y condicionalidad

DEFINICIÓN

Una variable aleatoria es una función que asocia un número real con cada elemento del espacio muestral.

(WALPOLE et al, 2012, p. 81)

CLASIFICACIÓN

	UNI	BIDI
DISCRETA		
CONTINUA		

¿Cómo pensar un ejercicio?

- **Interpretar el enunciado**
  - Pasar en limpio los Datos
  - Identificar Eventos y Espacio Muestral.
  - **Definir y clasificar Variables Aleatorias**
  - Interrogantes
- **Planteo**
- **Resolución**
- **Respuesta**

¡Nuevo!

VARIABLES ALEATORIAS

Distribución de probabilidades

Ejercicio 1

La distribución de probabilidad de X: número de errores en un canal de transmisión es la siguiente:

x	0	1	2	3	4
f(x)	0.41	0.37	0.16	0.05	0.01

VARIABLES ALEATORIAS

x	0	1	2	3	4
f(x)	0.41	0.37	0.16	0.05	0.01

- *Condición de cierre* $\sum_{\forall i} f(x_i) = 1$
- *Función acumulativa*

$$F(x) = P(X \leq x) = \sum_{\forall i / x_i \leq x} f(x_i)$$

- $P(a < x \leq b) = F(b) - F(a)$

consultas: taulamet@unl.edu.ar

VARIABLES ALEATORIAS

Distribución de probabilidades

Ejercicio 1

Ejercicio 3

Se conoce que la duración (en días) de un disco rígido es una variable aleatoria que se distribuye según la siguiente función:

$$f(x) = \begin{cases} \frac{1000}{x^2} & \text{si } x > 1000 \\ 0 & \text{en otro rango} \end{cases}$$

LOS SÍMBOLOS Y NOMBRES

- *Clasificación*
  - *Discreta*
    - *F. Masa / F. Cuantía* $f(x)$
    - *F. de Distribución* $\mathrm{F(x)}$
  - *Continua*
    - *F. de Densidad* $f(x)$
    - *F. de Distribución* $\mathrm{F(x)}$
- *Unidimensional* $f(x), F(x)$
- *Bidimensional*

$$f(x,y), F(x,y)$$

$$f_x(x), f_y(y), F_x(x), F_y(y)$$

LOS SÍMBOLOS Y NOMBRES

DISCRETAS

$f(x)$ cuantía

Probabilidad

$$0 < f(x) < 1$$

$$F(x) \sum$$

CONTINUAS

$f(x)$ densidad

~~Probabilidad~~

$$0 < f(x)$$

$$F(x) \int$$

consultas: taulamet@unl.edu.ar

Función f(x): caso discreto

Condición de cierre

1. $f(x) \geq 0$,
2. $\sum_{x} f(x) = 1$,
3. $P(X = x) = f(x)$.

(WALPOLE et al, 2012, p. 84)

Función f(x): caso continuo

1. $f(x) \geq 0$, para toda $x \in R$.

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

Condición de cierre

3. $P(a < X < b) = \int_a^b f(x)\, dx$.

(WALPOLE et al, 2012, p. 89)

Función F(x): caso discreto

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t),$$

$$\text{para} \quad -\infty < x < \infty.$$

(WALPOLE et al, 2012, p. 85)

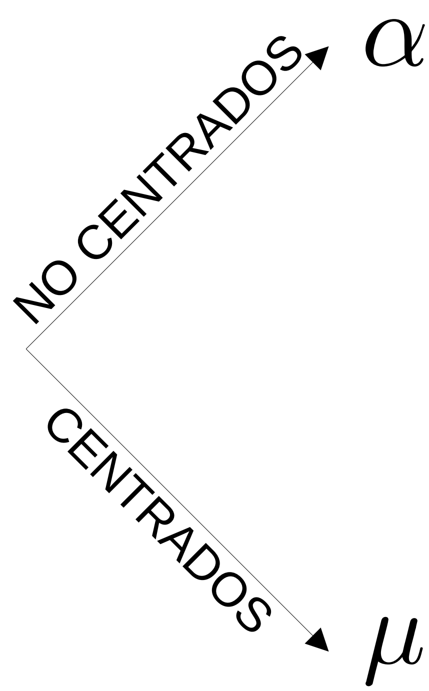
Función F(x): caso continuo

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt,$$

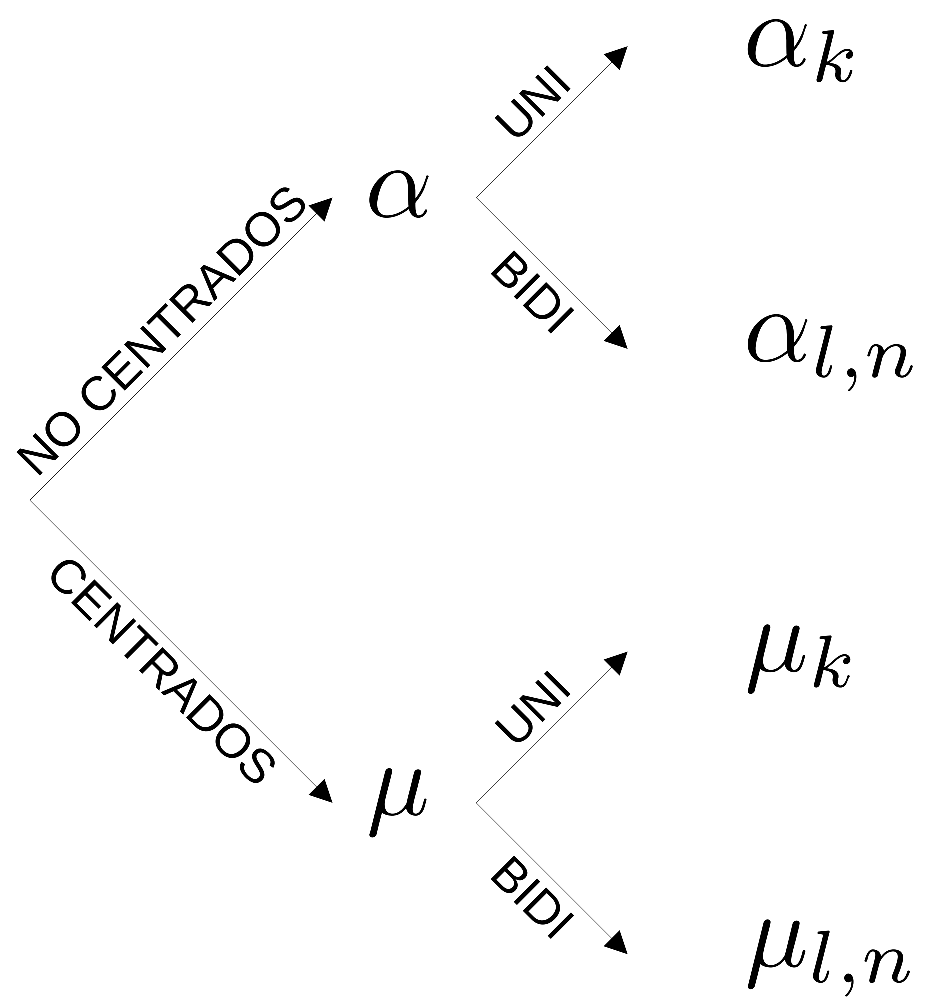
$$\text{para } -\infty < x < \infty.$$

(WALPOLE et al, 2012, p. 90)

MOMENTOS

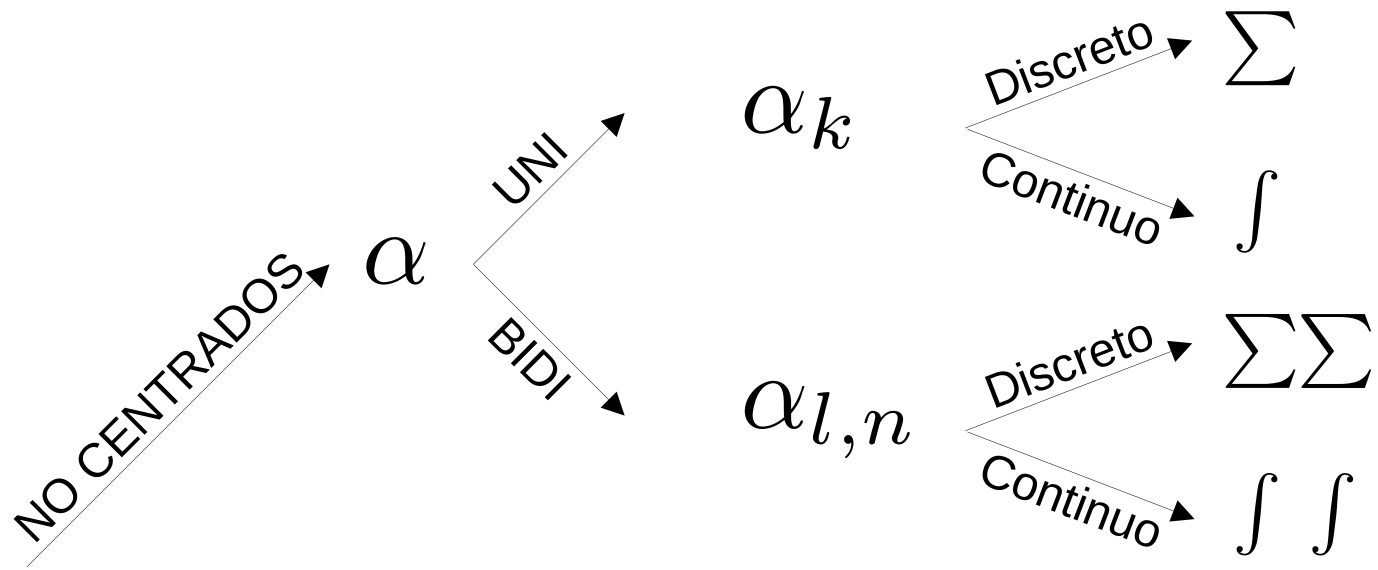


MOMENTOS



consultas: taulamet@unl.edu.ar

MOMENTOS


NO CENTRADOS
α
UNI
BIDI
α_k
$\alpha_{l,n}$
Discreto
Continuo
$\sum$
$\int$
Discreto
Continuo
$\sum\sum$
$\int\int$

MOMENTOS NO CENTRADOS

(No Centrados o Centrados con respecto al origen)

$$Orden\ k$$

$$\alpha_k \xrightarrow{\text{Discreto}} \alpha_k = \sum_{\forall i} x_i^k . f(x_i)$$

$$\alpha_k \xrightarrow{\text{Continuo}} \alpha_k = \int_{-\infty}^{+\infty} x^k . f(x) dx$$

consultas: taulamet@unl.edu.ar

MOMENTOS NO CENTRADOS

(No Centrados o Centrados con respecto al origen)

$Orden\ 1$

$$\alpha_1 \begin{cases} \xrightarrow{\text{Discreto}} & \alpha_1 = \sum_{\forall i} x_i^1 . f(x_i) \\ \xrightarrow{\text{Continuo}} & \alpha_1 = \int_{-\infty}^{+\infty} x^1 . f(x) dx \end{cases}$$

MOMENTOS NO CENTRADOS

(No Centrados o Centrados con respecto al origen)

$Orden\ 1$

$$\begin{tikzcd} & \alpha_1 = \sum_{\forall i} x_i . f(x_i) \\ \alpha_k \arrow[ur, "\text{Discreto}"] \arrow[dr, "\text{Continuo}"'] & \\ & \alpha_1 = \int_{-\infty}^{+\infty} x.f(x)dx \end{tikzcd}$$

consultas: taulamet@unl.edu.ar

Propiedades - Condición de cierre

2. $\sum_{x} f(x) = 1,$ 2. $\int_{-\infty}^{\infty} f(x)\,dx = 1.$

Condición de cierre

α_0

Discreto

$$\alpha_0 = \sum_{\forall i} x_i^0 . f(x_i)$$

Continuo

$$\alpha_0 = \int_{-\infty}^{+\infty} x^0 . f(x) dx$$

VARIABLES ALEATORIAS

$$P(a < x < b) = ?$$

VARIABLES ALEATORIAS

$$P(a < x < b) = ?$$

VARIABLES ALEATORIAS

$$P(a < x < b) = ?$$

CASO DISCRETO

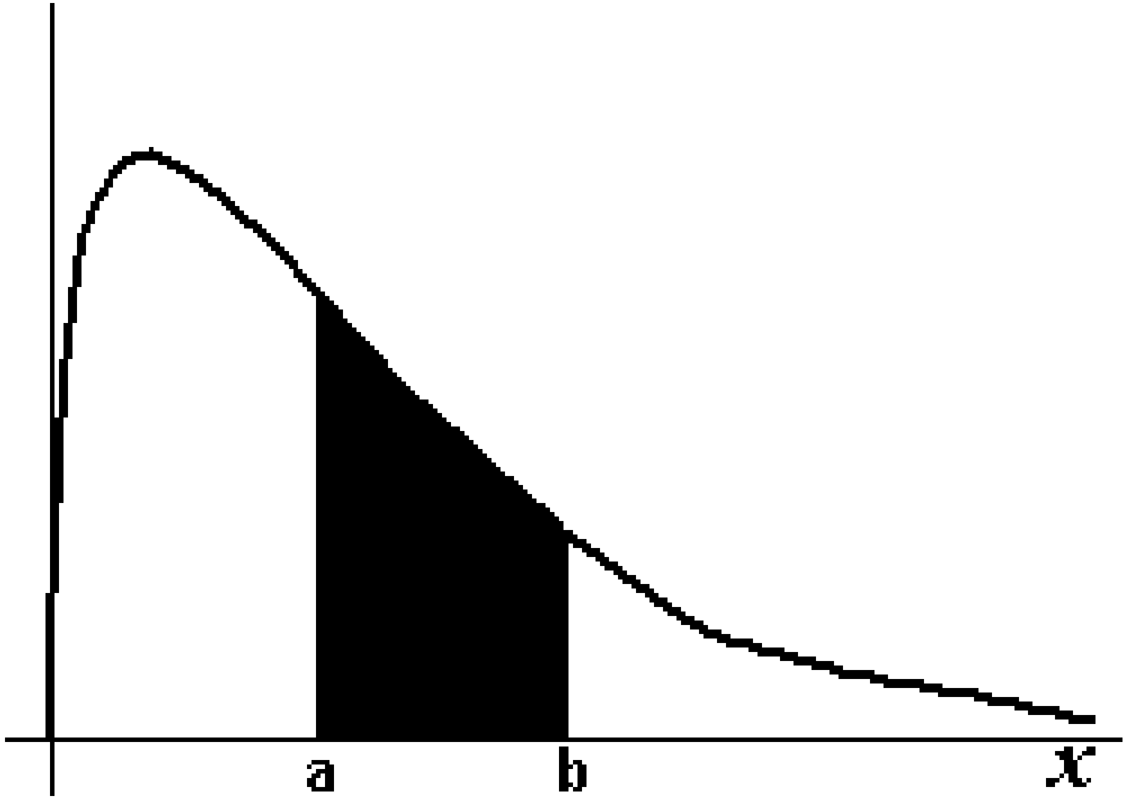
x	0	1	2	3	4
f(x)	0.41	0.37	0.16	0.05	0.01

$$P(a < x \leq b) = F(b) - F(a)$$

VARIABLES ALEATORIAS

CASO CONTINUO

$$P(a < x < b) = \int_a^b f(x)dx$$



VARIABLES ALEATORIAS

CASO CONTINUO

$$P(a < x < b) = \int_a^b f(x)dx$$

$$F(x) = \int_{-\infty}^x f(z)dz$$

$$P(a < x < b) = F(b) - F(a)$$

FUNCIONES MARGINALES

Un teodolito se clasifica de acuerdo al número de defectos de fabricación y a la fábrica que lo produce. Sean X e Y las variables aleatorias que representan respectivamente el número de defectos por unidad, con valores posibles 0, 1, 2 y 3; y el número de fábrica con valores posibles 1 y 2. La siguiente tabla muestra la distribución de probabilidades conjuntas.

		x			
		0	1	2	3
y	1	1/8	1/16	3/16	1/8
	2	1/16	1/16	1/8	1/4

a) Determinar la distribución de probabilidad marginal de X y de Y.

FUNCIONES MARGINALES

Un teodolito se clasifica de acuerdo al número de defectos de fabricación y a la fábrica que lo produce. Sean X e Y las variables aleatorias que representan respectivamente el número de defectos por unidad, con valores posibles 0, 1, 2 y 3; y el número de fábrica con valores posibles

Las **distribuciones marginales** sólo de X y sólo de Y son

$$g(x) = \sum_y f(x, y) \quad \text{y} \quad h(y) = \sum_x f(x, y)$$

para el caso discreto, y

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \quad \text{y} \quad h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

para el caso continuo.

Ejemplo: Función Acumulativa

Con objeto de establecer un plan de producción, una empresa que administra una represa hidroeléctrica ha estimado que la demanda semanal de energía es una variable aleatoria X (expresada en millones de unidades), cuya función de probabilidad viene dada por:

$$F(x) = \begin{cases} 1 - e^{-(x/3)^2} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

¿Podemos obtener la probabilidad de que la demanda semanal sea mayor a 2? ¿Cómo?

Planteo: F(x) y Prob. Contraria

Dado que la función de distribución nos brinda la posibilidad de calcular la probabilidad acumulada hasta un valor determinad y aprovechando el concepto de probabilidad contraria hacemos:

$$P(X > 2) = 1 - P(X \leq 2)$$

$$P(X > 2) = 1 - F(2)$$

Resolviendo...

$$F(x) = \begin{cases} 1 - e^{-(x/3)^2} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

$$P(X > 2) = 1 - F(2) =$$

$$1 - (1 - e^{-(2/3)^2}) = \frac{1}{e^{\frac{4}{9}}} = 0,6411$$

La probabilidad de que la demanda semanal sea superior a 2 es del 64%.

Segunda Parte: V.A. Bidim

.

Si se sabe que una nueva variable aleatoria Y también es necesaria para el estudio de manera que la función conjunta de ambas variables es:

$$f(x,y) = \begin{cases} \frac{1}{9} x e^{-(x/3)^2 - y/2} & \text{si } x > 0, y > 0 \\ 0 & \text{en otro caso.} \end{cases}$$

¿Se puede considerar que ambas variables son independientes?

Calcular $P(Y < 1/X > 2)$ e interpretar su significado.

consultas: taulamet@unl.edu.ar

Independencia

¿Son X e Y Independientes?

DATOS:

$$F(x) = \begin{cases} 1 - e^{-(x/3)^2} & \text{si } x < 0 \\ 0 & \text{si } x \geq 0 \end{cases}$$

$$f(x,y) = \begin{cases} \frac{1}{9} x e^{-(x/3)^2 - y/2} & \text{si } x > 0, y > 0 \\ 0 & \text{en otro caso.} \end{cases}$$

Independencia

CASO DISCRETO

¿Son X e Y Independientes?

$$f(x_i, y_j) = f(x_i) * f(y_j)$$
$$\forall x_i, y_j$$

Independencia

CASO CONTINUO

¿Son X e Y Independientes?

$$f(x,y) = f(x) * f(y)$$

Hallamos *f(x)*

$$f(x) = \int f(x,y)dy =$$

$$\int \frac{1}{9} x e^{-(x/3)^2 - y/2} dy =$$

$$\boxed{f(x) = \frac{2}{9} x e^{\frac{-x^2}{9}}}$$

También puede obtenerse f(x) a partir de la derivada de F(x)

Hallamos *f(y)*

$$f(y) = \int f(x,y)dx =$$

$$\int \frac{1}{9} x e^{-(x/3)^2 - y/2} dx =$$

$$\boxed{f(y) = \frac{e^{\frac{-y}{2}}}{2}}$$

Hallamos *f(x)* * *f(y)*

$$f(x) = \frac{2}{9}xe^{\frac{-x^2}{9}}$$

$$f(y) = \frac{e^{\frac{-y}{2}}}{2}$$

$$f(x) \cdot f(y) = \frac{2}{9}xe^{\frac{-x^2}{9}} \cdot \frac{e^{\frac{-y}{2}}}{2}$$

$$\boxed{f(x) \cdot f(y) = \frac{1}{9}xe^{-(x/3)^2 - y/2}}$$

Independencia

$$\boxed{f(x,y) = \frac{1}{9} x e^{-(x/3)^2 - y/2}}$$

$$\boxed{f(x) \cdot f(y) = \frac{1}{9} x e^{-\frac{x^2}{9} - y/2}}$$

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

X e Y son Independientes

Cálculo de Probabilidad

$$P(Y < 1/X > 2) = ?$$

Cálculo de Probabilidad

Recordando Planteo Condicional

$$P(Y < 1 / X > 2) =$$

$$\frac{P(Y < 1 \cap X > 2)}{P(X > 2)} =$$

Cálculo de Probabilidad

$$\frac{P(Y < 1 \cap X > 2)}{P(X > 2)} =$$

$$\frac{\int_0^1 \int_2^\infty \frac{1}{9} x e^{-(x/3)^2 - y/2} dx dy}{\int_0^\infty \int_2^\infty \frac{1}{9} x e^{-(x/3)^2 - y/2} dx dy} =$$

$$\frac{0,25}{0,64} \approx 39\%$$

(1994–2024)
30 años de la Consagración Constitucional de la Autonomía y Autarquía Universitaria en Argentina.

ESTADÍSTICA

Notas de Clase #4 - *Características*

Ingenierías en: Recursos Hídricos, Ambiental y Agrimensura

Año 2024

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

PROGRAMA ANALÍTICO

UNIDAD III : CARACTERÍSTICAS
3.1 - **Características y Momentos** de una distribución de probabilidades.

3.1.1 - Medidas de la Tendencia Central.
3.1.2 - Momentos - Medidas de Variabilidad.
3.1.3 - Medidas de Asimetría.
3.1.4 - Medidas de Kurtosis.

CARACTERÍSTICAS

"Existen una serie de números que resumen las <u>características</u> dominantes del comportamiento de una variable aleatoria."

(Vanlesberg, 2021, Apuntes de Teoría)

CLASIFICACIÓN

- TENDENCIA CENTRAL
- VARIABILIDAD
- FORMA
  - ASIMETRÍA
  - CURTOSIS

Recapitulando...

DISCRETAS	CONTINUAS
$f(x)$ cuantía	$f(x)$ densidad
Probabilidad	~~Probabilidad~~
$0 < f(x) < 1$	$0 < f(x)$
$F(x) \sum$	$F(x) \int$

Recapitulando...

	UNI	BIDI
DISCRETA		
CONTINUA		

CARACTERÍSTICAS

- TENDENCIA CENTRAL
  - PROMEDIOS
  - UBICACIÓN

CARACTERÍSTICAS

- TENDENCIA CENTRAL
  - PROMEDIOS
  - UBICAC

Promedios

a- Esperanza matemática
b- Media geométrica
c- Media armónica

Medidas de ubicación

a- Mediana
b- Modo
c- Cuantiles

- TENDENCIA CENTRAL
  - PROMEDIOS
    - ESPERANZA MATEMÁTICA

$$\begin{tikzcd} & \sum_{\forall i} x_i . f(x_i) \\ E(X) \arrow[ur, "\text{Caso Discreto}"] \arrow[dr, "\text{Caso Continuo}"'] & \\ & \int_{-\infty}^{+\infty} x . f(x) dx \end{tikzcd}$$

ESTADÍSTICA
DIAPOSITIVAS DE EJEMPLO

Unidad 3 – *Características*

Ingenierías en Recursos Hídricos / Ambiental / Agrimensura

Año 2022

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

Para rellenar una zona baja aledaña a un río, se utilizan dos camiones: A y B. La distribución de la carga diaria (en tn) transportada por cada camión se puede adaptar a la siguiente función:

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

¿Se podrá decir cuál es el valor medio de carga de cada camión?¿Alcanza con sólo este valor para caracterizar a las cargas? ¿Qué más se podría calcular?

Antes que nada...

¿Es una función de densidad válida?:

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

- ¿Definida para todo posible x?
- ¿No negativa para todo posible x?
- ¿Cumple la condición de cierre?

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

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

$$\int_{\infty}^{11} 0dx + \int_{11}^{15} \frac{x}{52} dx + \int_{15}^{\infty} 0dx =$$

$$0 + \frac{x^2}{52 * 2}\Bigg|_{11}^{15} dx + 0 = 1$$

Lo último que se pierde...

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx =$$

$$\int_{-\infty}^{11} x0dx + \int_{11}^{15} \frac{x^2}{52} dx + \int_{15}^{\infty} x0dx =$$

$$0 + \frac{x^3}{52 * 3}\Bigg|_{11}^{15} dx + 0 => E(X) \approx 13.10$$

consultas: taulamet@unl.edu.ar

Propiedades de la Esperanza

Destacamos las siguientes

$$E(X) \begin{array}{l} \xrightarrow{\text{Caso Discreto}} \displaystyle\sum_{\forall i} x_i . f(x_i) \\ \xrightarrow{\text{Caso Continuo}} \displaystyle\int_{-\infty}^{+\infty} x.f(x)dx \end{array}$$

$$E(c) = c$$

$$E(cX) = cE(X)$$

$$E(a + bX) = a + bE(X)$$

consultas: taulamet@unl.edu.ar

Propiedades de la Esperanza

Ejercicio 2

La producción de trigo en una determinada región es una variable aleatoria X (en miles de toneladas) cuya distribución está caracterizada por la siguiente función:

$$f(x) = \begin{cases} \frac{3}{22}(x+3)(2-x) & \text{para } 0 \leq x \leq 2 \\ 0 & \text{en otro caso} \end{cases}$$

El beneficio por cada mil toneladas producidas se obtiene como función de la cantidad producida resulta $B = 5000X - 1000$. ¿Cuál es el beneficio esperado?

$$E(a + bX) = a + bE(X)$$

CARACTERÍSTICAS

- TENDENCIA CENTRAL
  - PROMEDIOS
    - ESPERANZA
  - UBICACIÓN
    - MEDIANA

Centro de masa de la distribución

Mitad de la probabilidad acumulada

- TENDENCIA CENTRAL
  - PROMEDIOS
  - UBICACIÓN
    - MEDIANA

$$Mediana \xrightarrow{\text{Caso Discreto}} \sum_{x}^{Mediana} f(x) \geq 0.5$$

$$Mediana \xrightarrow{\text{Caso Continuo}} \int_{-\infty}^{Mediana} f(x)dx = 0.5$$

**Suponiendo que el primer valor de X sea 1*

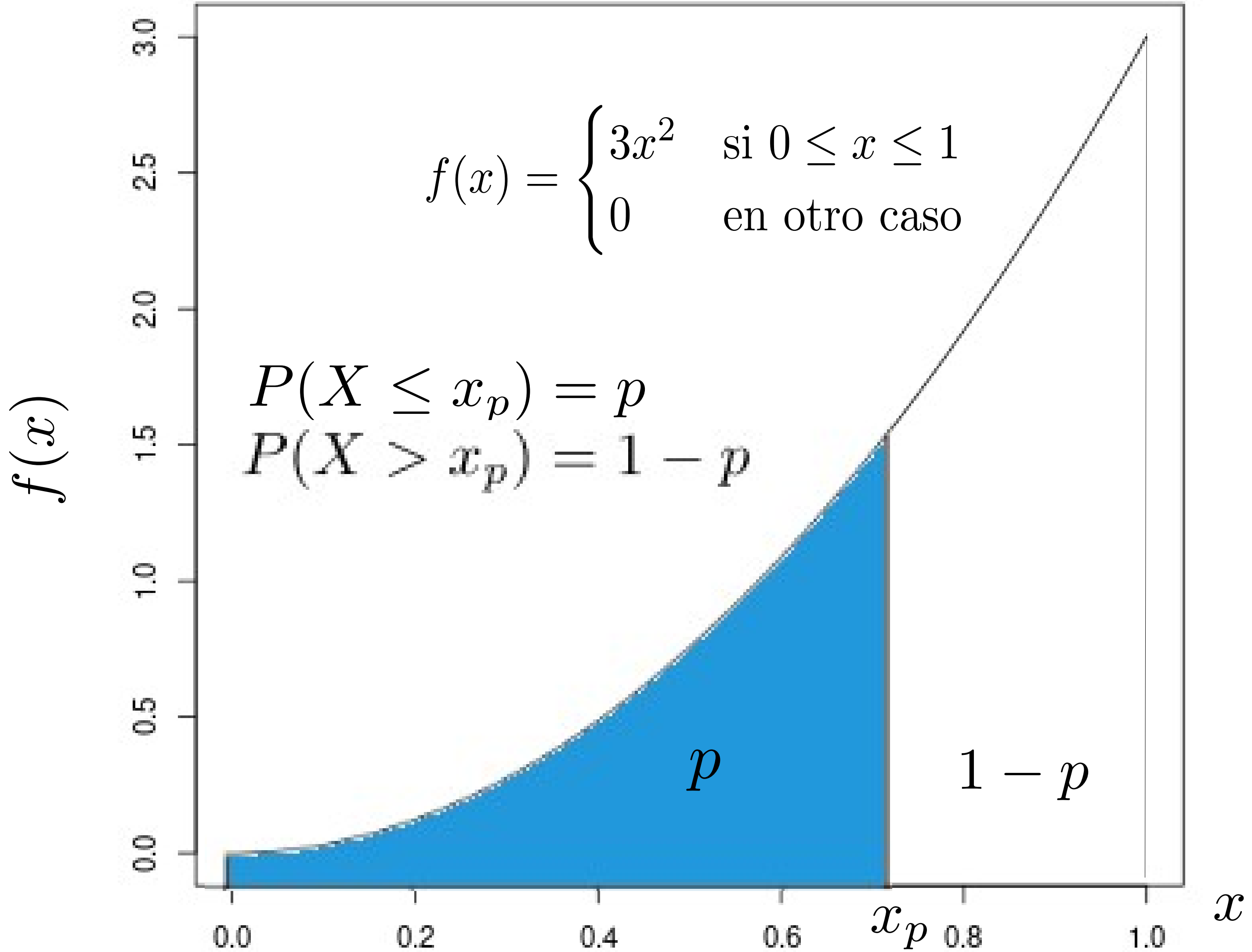
- MEDIDAS DE UBICACIÓN
  - CUANTILES

Se denomina ***cuantil de orden p*** *(siendo p un número perteneciente al intervalo [0,1]) al valor de la variable Xp que cumple con al menos una de las siguientes condiciones:*

$$P(X \leq x_p) = p$$
$$P(X > x_p) = 1 - p$$

consultas: taulamet@unl.edu.ar

Ejemplo: Cuantil de orden *p*



➢ MEDIDAS DE UBICACIÓN

➢ CUANTILES

Dividen la distribución en partes con igual probabilidad y los veremos en cuatro grupos:*

$$
\begin{aligned}
&MEDIANA\\
&CUARTILES\ Q_i\\
&DECILES\ D_i\\
&PORCENTILES\ P_i
\end{aligned}
$$

*Pueden existir otros, como los Terciles o los Quintiles.

CUANTILES

Existen algunas equivalencias

$$MEDIANA = \mathrm{Q}_2 = D_5 = P_{50}$$

$$CUARTILES\ Q_i : Q_1 = P_{25}; Q_2 = P_{50}; Q_3 = P_{75};$$

$$DECILES\ D_i : D_1 = P_{10}; D_2 = P_{20}; ...; D_9 = P_{90};$$

$$PORCENTILES\ P_i : Todos\ los\ anteriores.$$

**Pueden existir otros, como los Quintiles.*

consultas: taulamet@unl.edu.ar

$$Mediana \xrightarrow{\text{Caso Discreto}} \sum_{i=1*}^{Mediana} f(x_i) = 0.5$$

$$Mediana \xrightarrow{\text{Caso Continuo}} \int_{-\infty}^{Mediana} f(x)dx = 0.5$$

$$Percentil\ P_i \xrightarrow{\text{Caso Discreto}} \sum_{j=1*}^{P_i} f(x_j) = i$$

$$Percentil\ P_i \xrightarrow{\text{Caso Continuo}} \int_{-\infty}^{P_i} f(x)dx = i$$

**Suponiendo que el primer valor de X sea 1*

➢ MEDIDAS DE UBICACIÓN

➢ MODO

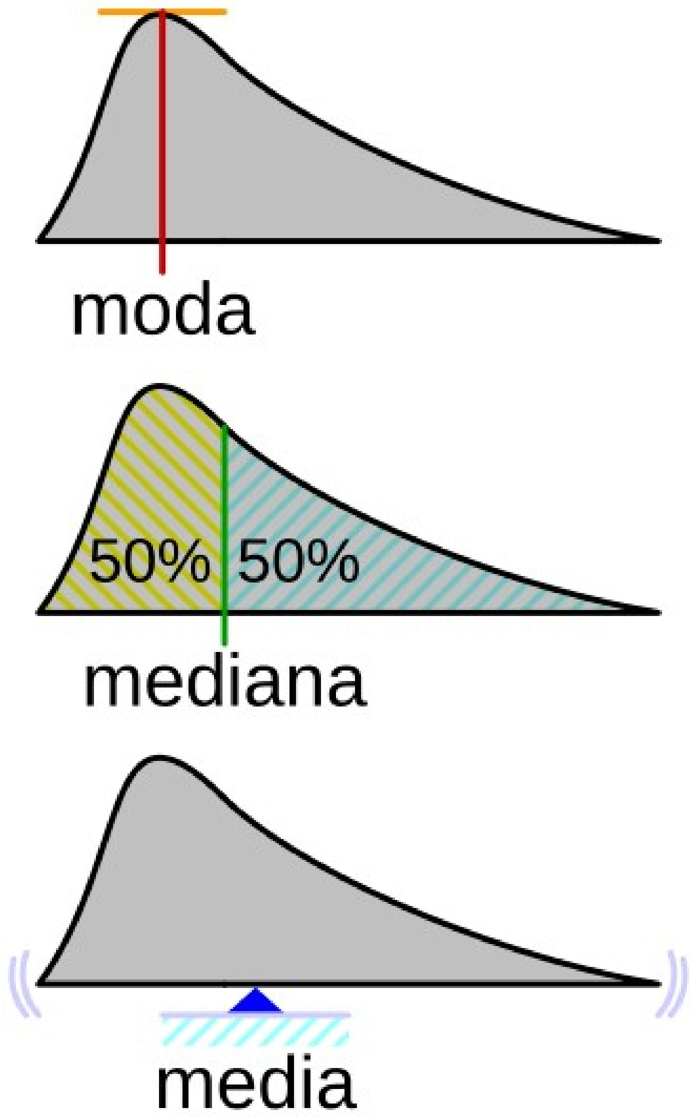
*El **modo o la moda** Es el valor de la variable aleatoria que se corresponde con el máximo de la función de probabilidad. En el caso discreto será el valor asociado a una mayor probabilidad de ocurrencia y en el caso continuo podrá investigarse a partir de las derivadas de la función de densidad y eventualmente observando los límites de cada tramo de la función.*

Caso Discreto

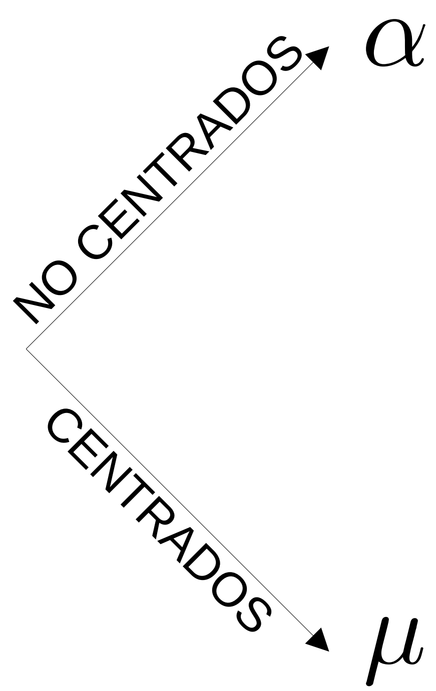
$$M_x : f(M_x) = max_x f(x)$$

Caso Continuo

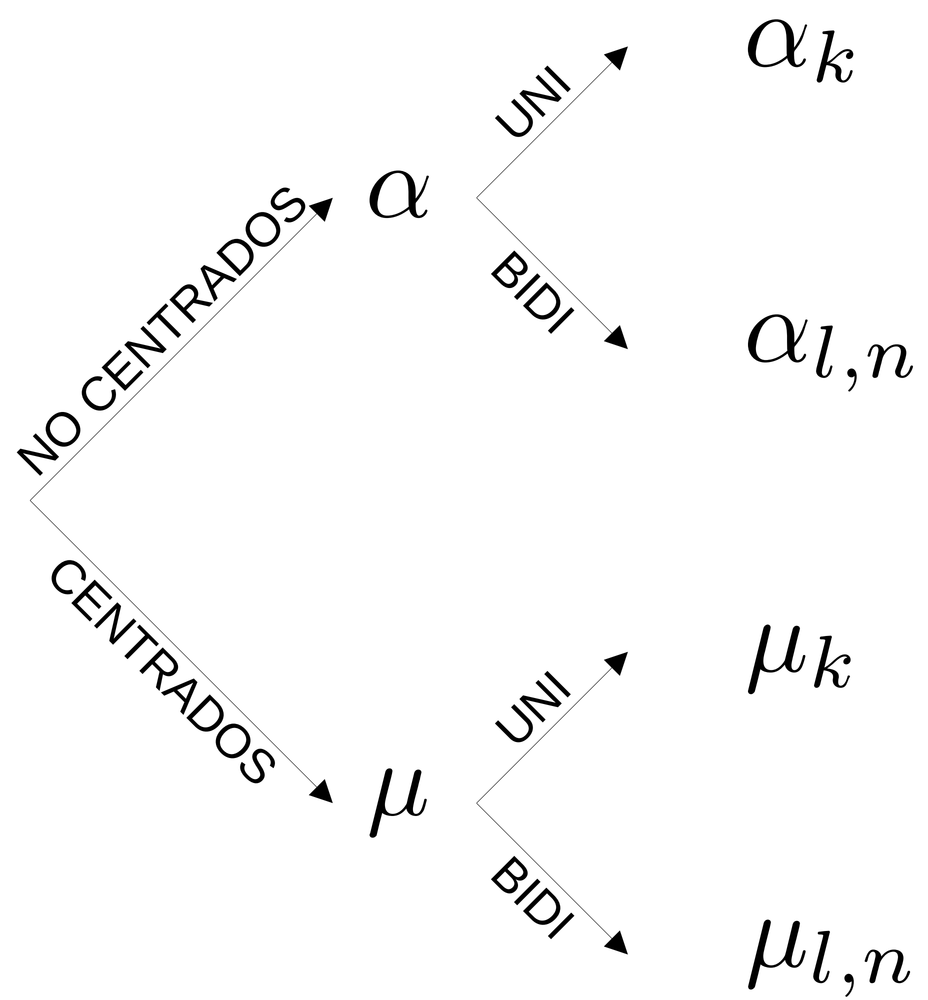
$$\frac{df(x)}{dx} = 0 \ y \ \frac{d^2 f(x)}{dx} \leq 0$$


moda
50%
50%
mediana
media

MOMENTOS

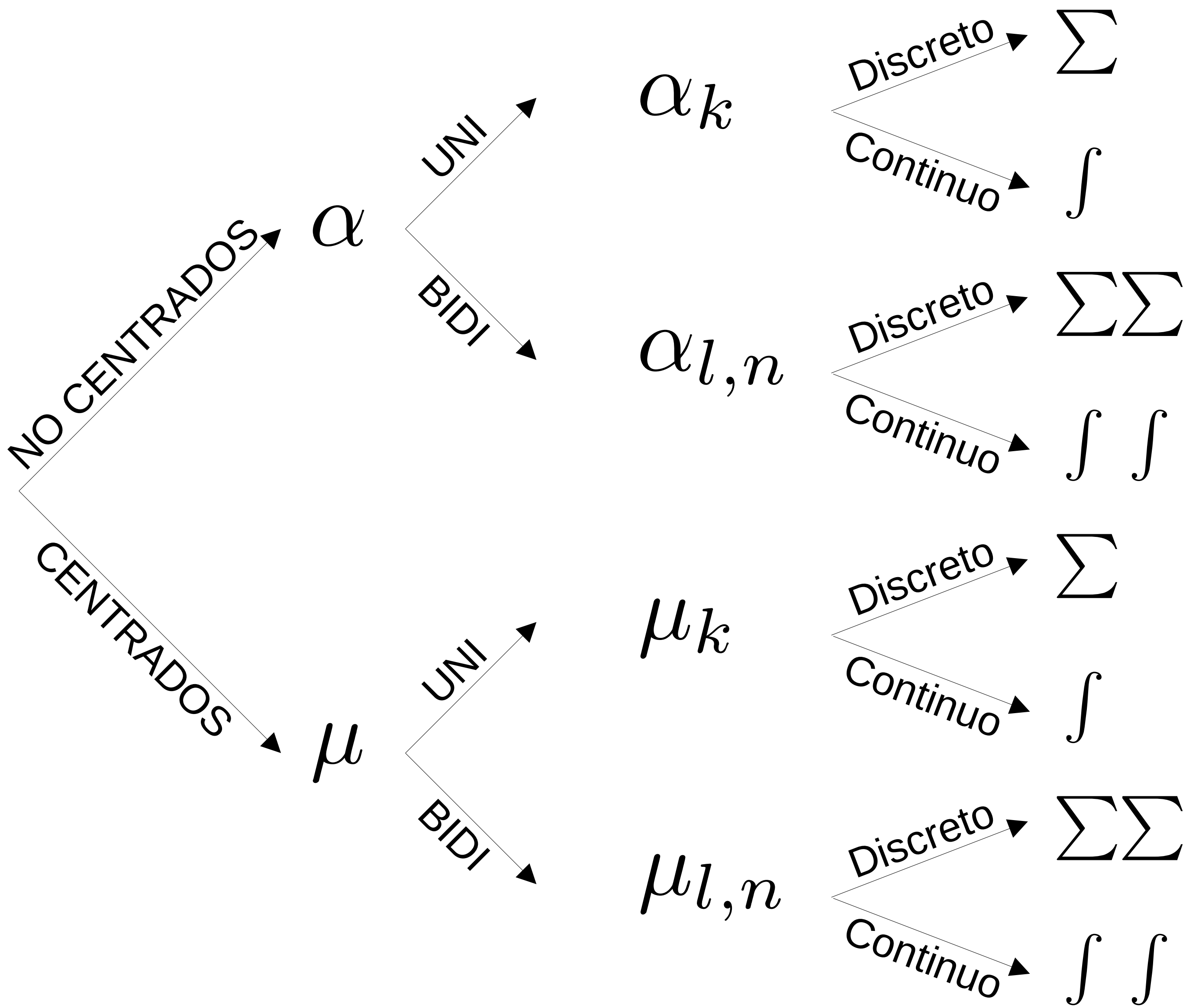


MOMENTOS



consultas: taulamet@unl.edu.ar

MOMENTOS



NO CENTRADOS
CENTRADOS
α
μ
UNI
BIDI
UNI
BIDI
α_k
$\alpha_{l,n}$
μ_k
$\mu_{l,n}$
Discreto
Continuo
Discreto
Continuo
Discreto
Continuo
Discreto
Continuo
$\sum$
$\int$
$\sum\sum$
$\int\int$
$\sum$
$\int$
$\sum\sum$
$\int\int$

MOMENTOS

No Centrados o Centrados con respecto al origen

$Orden\ k$

$$\alpha_k = E(X^k)$$

Discreto

$$\alpha_k = \sum_{\forall i} x_i^k . f(x_i)$$

Continuo

$$\alpha_k = \int_{-\infty}^{+\infty} x^k . f(x) dx$$

consultas: taulamet@unl.edu.ar

MOMENTOS

No Centrados o Centrados con respecto al origen

$$Orden\ 1$$

$$\alpha_1 = E(X^1) \begin{cases} \xrightarrow{\text{Discreto}} \alpha_1 = \sum_{\forall i} x_i^1 . f(x_i) \\ \xrightarrow{\text{Continuo}} \alpha_1 = \int_{-\infty}^{+\infty} x^1 . f(x) dx \end{cases}$$

MOMENTOS

No Centrados o Centrados con respecto al origen

$Orden\ 1$

$$\alpha_1 = E(X)$$

Discreto

$$\alpha_1 = \sum_{\forall i} x_i . f(x_i)$$

Continuo

$$\alpha_1 = \int_{-\infty}^{+\infty} x . f(x) dx$$

MOMENTOS

No Centrados o Centrados con respecto al origen

$$Orden\ 2$$

$$\alpha_2 = E(X^2)$$

Discreto

$$\alpha_2 = \sum_{\forall i} x_i^2 . f(x_i)$$

Continuo

$$\alpha_2 = \int_{-\infty}^{+\infty} x^2 . f(x) dx$$

MOMENTOS

Centrados (con respecto al valor medio)

$$\mu_k = E[(X - E[X])^k]$$

$$\mu_k = E[(X - \mu)^k]$$

$$Orden\ k$$

$$\mu_k \begin{cases} \xrightarrow{\text{Discreto}} & \mu_k = \sum_{\forall i} (x_i - \mu)^k . f(x_i) \\ \xrightarrow{\text{Continuo}} & \mu_k = \int_{-\infty}^{+\infty} (x - \mu)^k . f(x) dx \end{cases}$$

VARIABILIDAD

- RANGO (Máx - Mín)
- VARIANZA
- DESVÍO
- COEF. DE VARIABILIDAD

VARIANZA

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

$$Var(X) = \sigma^2 = E[(X - \mu)^2]$$

Discreto

$$Var[X] = \sum_{\forall i} (x_i - E[X])^2 f(x_i)$$

Continuo

$$Var[X] = \int_{-\infty}^{+\infty} (x - E[X])^2 f(x)dx$$

VARIANZA

Forma resumida

$$
\begin{aligned}
\operatorname{Var}(X) &= \mathrm{E}[(X-\mu)^2] \\
&= \mathrm{E}[(X^2 - 2X\mu + \mu^2)] \\
&= \mathrm{E}[X^2] - 2\mu\,\mathrm{E}[X] + \mu^2 \\
&= \mathrm{E}[X^2] - 2\mu^2 + \mu^2 \\
&= \mathrm{E}[X^2] - \mu^2 \\
&= \mathrm{E}[X^2] - \mathrm{E}^2[X]
\end{aligned}
$$

Para rellenar una zona baja aledaña a un río, se utilizan dos camiones: A y B. La distribución de la carga diaria (en tn) transportada por cada camión se puede adaptar a la siguiente función:

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

¿Se podrá decir cuál es el valor medio de carga de cada camión?**¿Alcanza con sólo este valor para caracterizar a las cargas? ¿Qué más se podría calcular?**

consultas: taulamet@unl.edu.ar

¿Variabilidad?

$$D(X) = \sqrt{V(X)}; V(X) = E(X^2) - E(X)^2$$

$$E(X^2) = \int_{11}^{15} x^2 \cdot \frac{x}{52} dx = \int_{11}^{15} \frac{x^3}{52} dx =$$

$$\left.\frac{x^4}{52 * 4}\right|_{11}^{15} dx = 173) => V(X) = 173 - 13.10^2 =$$

$$V(X) \approx 1.39 => D(X) \approx 1.18$$

CARACTERÍSTICAS Y MOMENTOS

$$E(X) = \mu = \alpha_1$$

$$V(X) = \sigma^2 = \mu_2$$
$$V(X) = \sigma^2 = \alpha_2 - \alpha_1^2$$

$$\alpha_0 = E(1) = 1$$

$$\mu_0 = E(1) = 1$$

$$\mu_1 = E(x - \mu)) = 0$$

Propiedades de la Varianza

$$Var(C) = 0$$

$$E[C - E(C)]^2 = E[C - C]^2 = E(0) = 0$$

Propiedades de la Varianza

$$Var(CX) = C^2 Var(X)$$

$$E[CX - E(CX)]^2 = E[CX - CE(X)]^2 =$$

$$= E\{[C.(X - E(X))]^2\} = E\{C^2.[X - E(X)]^2\} =$$

$$= C^2 E[X - E(X)]^2 = C^2 Var(X)$$

Propiedades de la Varianza

Discreto

$$Var[X] = \sum_{\forall i} (x_i - E[X])^2 f(x_i)$$

Continuo

$$Var[X] = \int_{-\infty}^{+\infty} (x - E[X])^2 f(x)dx$$

$$Var(C) = 0$$

$$Var(CX) = C^2 Var(X)$$

$$Var(a + bX) = Var(a) + Var(bx) = 0 + b^2 Var(X)$$

consultas: taulamet@unl.edu.ar

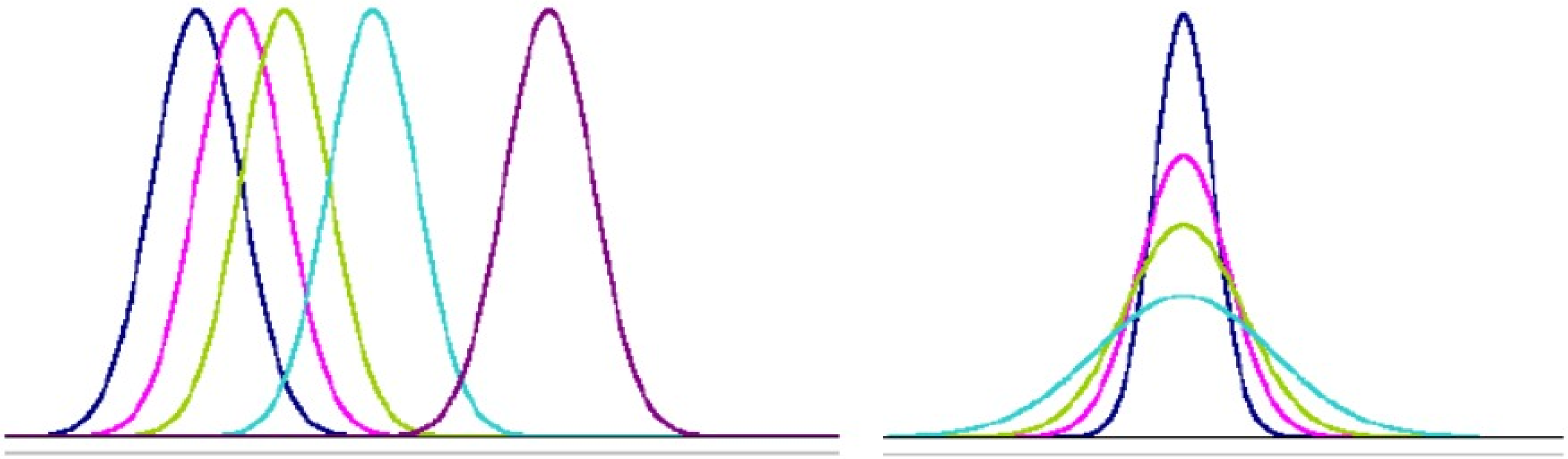
DESVÍO ESTÁNDAR y CV

$$\sigma_x = +\sqrt{Var(X)}$$

$$\sigma_x = +\sqrt{\alpha_2 - \alpha_1^2}$$

$$C_v = \frac{\sigma(X)}{E(X)}$$

VARIACIONES DE E(X) Y V(X)



Ejercicio 3

El tiempo requerido para transportar un cargamento por via fluvial entre dos puertos sigue una distribución de probabilidades con una media de 90 hs. y desvío de 3 hs. El capitán de un barco pretende llegar a un puerto entre 80 y 100 hs. después de haber salido del primer puerto. ¿Puede con esta información determinar la probabilidad de que realmente esto suceda?

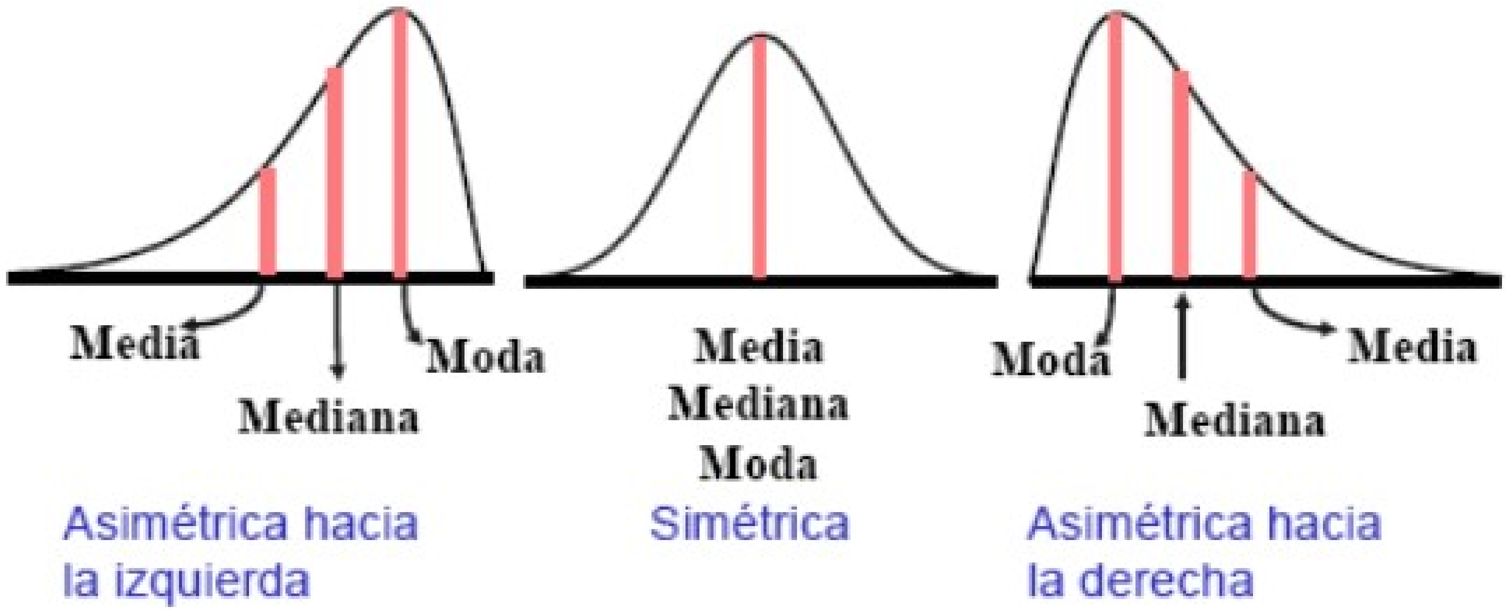
consultas: taulamet@unl.edu.ar

TEOREMA DE TCHEVYCHEFF

$$P(|X - \mu| \geq h\,\sigma) \leq \frac{1}{h^2}$$

$$P(|X - \mu| < h\,\sigma) \geq 1 - \frac{1}{h^2}$$

ASIMETRÍA


Media
Mediana
Moda
Media
Mediana
Moda
Moda
Mediana
Media
Asimétrica hacia la izquierda
Simétrica
Asimétrica hacia la derecha

ASIMETRÍA

$$As = \frac{E(X) - Modo}{\sigma(X)}$$

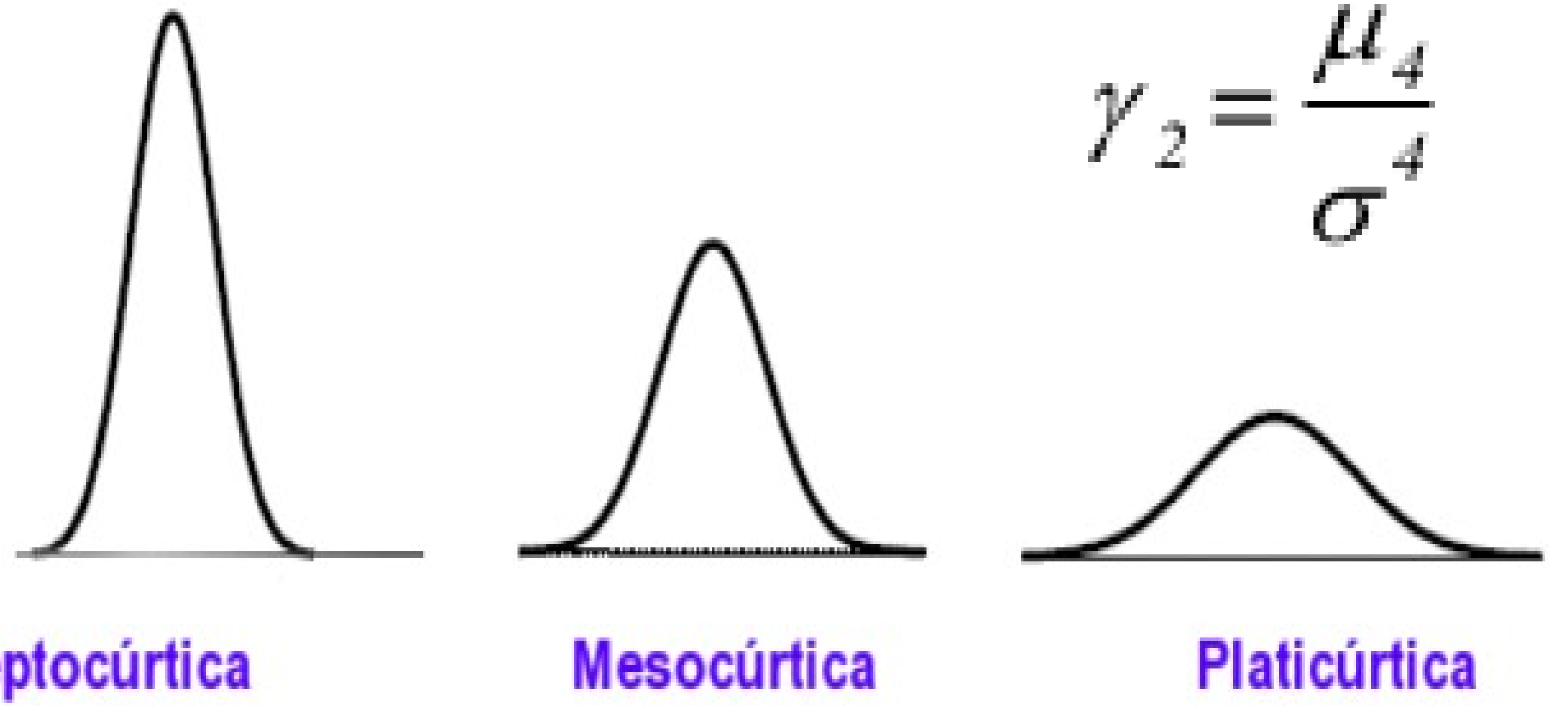
$$As = \frac{3\,[\,E(X) - Mediana\,]}{\sigma(X)}$$

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \qquad \gamma_1 = \frac{\alpha_3 - 3\alpha_2\alpha_1 + 2\alpha_1^3}{(\alpha_2 - \alpha_1^2)^{3/2}}$$

consultas: taulamet@unl.edu.ar

CURTOSIS

$$\gamma_2 = \frac{\mu_4}{\sigma^4}$$



$$\mu_4 = \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4$$

$$\gamma_2 - 3$$

consultas: taulamet@unl.edu.ar

CARACTERÍSTICAS DE V.A. DISTRIBUIDAS CONJUNTAMENTE

- MOMENTOS BIDIMENSIONALES
- MARGINALES
- DEPENDENCIA
- X + Y
- X - Y

MOMENTOS BIDIMENSIONALES

No Centrados o Centrados con respecto al origen

$$\alpha_{l,n} = E(X^l, Y^n)$$

$$\alpha_{l,n} \begin{cases} \xrightarrow{\text{Discreto}} \displaystyle\sum_{\forall i}\sum_{\forall j} x_i^l . y_j^n . f(x_i, y_j) \\ \xrightarrow{\text{Continuo}} \alpha_{l,n} = \displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^l . y^n . f(x,y) dxdy \end{cases}$$

$Orden\ l{,}n$

consultas: taulamet@unl.edu.ar

ESPERANZAS MARGINALES

Calculadas a partir de momentos no centrados

$$E(X) = \alpha_{1,0}$$

$$\alpha_{1,0} \begin{cases} \xrightarrow{\text{Discreto}} \displaystyle\sum_{\forall i} \sum_{\forall j} x_i^1 . y_j^0 . f(x_i, y_j) \\ \xrightarrow{\text{Continuo}} \displaystyle\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^1 . y^0 . f(x, y) dy dx \end{cases}$$

ESPERANZAS MARGINALES

Calculadas a partir de momentos no centrados

$$E(Y) = \alpha_{0,1}$$

$\alpha_{0,1}$

Discreto: $$\sum_{\forall i} \sum_{\forall j} x_i^0 . y_j^1 . f(x_i, y_j)$$

Continuo: $$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^0 . y^1 . f(x, y) dx dy$$

consultas: taulamet@unl.edu.ar

MOMENTOS BIDIMENSIONALES

Centrados con respecto los valores medios

$$\mu_{l,n} = E[X - E(X)]^l . E[Y - E(Y)]^n)$$

$$\mu_{l,n} = E[X - \mu_X]^l . E[Y - \mu_Y]^n)$$

Discreto

$$\sum_{\forall i} \sum_{\forall j} (x_i - \mu_X)^l . (y_j - \mu_Y)^n . f(x_i, y_j)$$

Continuo

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)^l . (y - \mu_Y)^n . f(x, y) dx dy$$

$Orden\ l,n$

consultas: taulamet@unl.edu.ar

VARIANZAS MARGINALES

$$V(X) = \sigma_X^2 = \mu_{2,0} = \alpha_{2,0} - \alpha_{1,0}^2$$

$$V(Y) = \sigma_Y^2 = \mu_{0,2} = \alpha_{0,2} - \alpha_{0,1}^2$$

COVARIANZA

$$COV(X,Y) = \mu_{1,1} = E[X - \mu_X]^1.E[Y - \mu_Y]^1)$$

$$COV(X,Y) = \mu_{1,1} = \alpha_{1,1} - \alpha_{1,0}.\alpha_{0,1}$$

INDEPENDENCIA

$$X, Y\, Indep$$

$$\mathrm{E(X,Y) = E(X)} \cdot E(Y)$$

INDEPENDENCIA

$$\sum_{\forall x_i}\sum_{\forall y_j}[x_i - E(X)].[y_j - E(Y)].f(x).f(y)=$$

$$=\sum_{\forall x_i}[x_i - E(X)].f(x_i)\sum_{\forall y_j}[y_j - E(Y)].f(y_j)=$$

$$=E[x_i - E(X)].E[y_j - E(Y)]=\mu_{1,0}.\mu_{0,1}$$

$$COV(X,Y)=\mu_{1,1}=\alpha_{1,1}-\alpha_{1,0}.\alpha_{0,1}$$
$$COV(X,Y)=\mu_{1,1}=\alpha_{1,0}.\alpha_{0,1}-\alpha_{1,0}.\alpha_{0,1}$$
$$COV(X,Y)=0$$

COEFICIENTE DE CORRELACIÓN

$$\rho_{x,y} = \frac{cov_{x,y}}{\sigma_x \cdot \sigma_y} = \frac{\mu_{1,1}}{\sqrt{\mu_{2,0} \cdot \mu_{0,2}}}$$

$$-1 \leq \rho \leq 1$$

SUMA DE VARIABLES

$$E(X \pm Y) = E(X) \pm E(Y)$$

$$V(X \pm Y) = V(X) + V(Y) \pm 2.COV(X, Y)$$

Si se supone que las cargas transportadas por los dos camiones son independientes; **¿Se podría decir cuál es el valor medio de carga total transportada por los camiones?** ¿Alcanza con sólo este valor para caracterizar a la carga total? ¿Qué más se podría calcular?

consultas: taulamet@unl.edu.ar

Pensando en 2 camiones

Suponiendo que X mide la carga del camión A e Y la carga del camión B, pasamos en limpio las funciones de densidad de cada variable:

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

$$f(y) = \begin{cases} \frac{y}{52} & \text{si } 11 < y < 15 \\ 0 & \text{e.o.c} \end{cases}$$

Carga Total Media

Luego si definimos la carga total T como la suma de las cargas de los camiones y aplicamos la propiedad de la esperanza de una suma:

$$T = X + Y$$

$$E(T) = E(X + Y) = E(X) + E(Y)$$

$$=> E(T) = 13.10 + 13.10 = 26.20$$

Desvío Total Medio

Análogamente, por la propiedad de la varianza de una suma, necesitaremos ***calcular la covarianza:***

$$V(T) = V(X+Y) =$$

$$V(X) + V(Y) + 2COV(X,Y)$$

$$COV(X,Y) = E(X,Y) - E(X) * E(Y)$$

Covarianza

$$COV(X,Y) = E(X,Y) - E(X) \cdot E(Y)$$

$$E(X,Y) = \int_{11}^{15}\int_{11}^{15} xyf(x,y) \cdot dx \cdot dy =$$

$$\int_{11}^{15}\int_{11}^{15} xy\frac{x}{52}\frac{y}{52} \cdot dx \cdot dy =$$

$$\int_{11}^{15}\int_{11}^{15} \frac{x^2y^2}{2704} \cdot dx \cdot dy \approx 171.68$$

Covarianza: Interpretación

$$COV(X,Y) = E(X,Y) - E(X) \cdot E(Y)$$

$$=> 171.68 - (13.10^2) \approx 0$$

$$=> COV(X,Y) = 0$$

Recordemos que si la covarianza es cero, podemos interpretar que no existe relación lineal entre las dos variables, lo cual es coherente con lo supuesto en el enunciado del problema.

Desvío Total Medio

$$V(T) = V(X+Y) =$$

$$V(X) + V(Y) + 2COV(X,Y)$$

$$=> V(T) = 1.3 + 1.3 = 2.6$$

$$=> D(T) \approx 1.61$$

(1994–2024)

30 años de la Consagración Constitucional de la Autonomía y Autarquía Universitaria en Argentina.

UNIVERSIDAD NACIONAL DEL LITORAL
LUX INDEFICIENS · 1919

FICH

ESTADÍSTICA

Notas de Clase #5 - *Modelos*

Ingenierías en: Recursos Hídricos, Ambiental y Agrimensura

Año 2024

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

PLANTEO

- NOCIONES DE PROBABILIDAD
- IDENTIFICAR EVENTOS
- RELACIONES
- IDENTIFICAR Y DEFINIR VARIABLES
- CLASIFICARLAS
- CARACTERIZARLAS
- ¿SIGUE ALGÚN MODELO PROBABILÍSTICO?

¿CÓMO ABORDAR LA GUÍA 4A?

- MODELOS DISCRETOS
  - BERNOULLI
  - BINOMIAL
  - GEOMÉTRICO
  - POISSON
  - HIPERGEOMÉTRICO

TABLAS & PROGRAMAS

consultas: taulamet@unl.edu.ar

GUÍA 4A Y GUÍA 4B

- MODELOS DISCRETOS
  - BERNOULLI
  - BINOMIAL
  - GEOMÉTRICO
  - POISSON
  - HIPERGEOMÉTRICO

- MODELOS CONTINUOS
  - EXPONENCIAL
  - ~~GAMA~~
  - NORMAL
  - ~~LOGNORMAL~~
  - ~~RELACIONADOS AL NORMAL~~
  - DE VALORES EXTREMOS

consultas: taulamet@unl.edu.ar

Un reciente estudio revela que el 40% de los adultos están a favor de un control de contaminación del aire y acústica en una ciudad. Si se seleccionaran 5 adultos aleatoriamente:

a) Determine la probabilidad de que ninguno esté a favor del citado control.
b) ¿Cuál es la probabilidad de que como máximo 3 estén a favor del control?
c) ¿Cuál el valor esperado de adultos seleccionados que están a favor del control?

Ensayo de Bernoulli

Éxito o Fracaso

Experimento: Se le consulta a un adulto su opinión sobre el control ambiental:

A: El adulto está a favor

~A: El adulto está en contra

Datos

"...el 40% de los adultos están a favor de un control ambiental..."

En este caso: Éxito → A Favor

E: El adulto está a favor del control

- 5 Adultos
- P(E) = 0,40

Interrogante a)

Si se seleccionan 5 adultos al azar:

¿Determine la probabilidad de que **ninguno** esté a favor del citado control?

Si X: “Número de adultos a favor”

$$P(X=0) = ?$$

Planteo a)

X: Número de adultos a favor

X $\rightarrow$ V.A Discreta

X $\rightarrow$ Unidimensional

X ~ ?

Planteo a)

X: Número de de adultos a favor (V.A. Discreta)

X $\rightarrow$ V.A Discreta

X $\rightarrow$ Unidimensional

X ~ Binomial(n=5, p=0,40)

$$P(X=0) = \binom{n}{x} \mathrm{p}^{x}(1-p)^{n-x} = \binom{5}{0} 0,40^{0} * 0,60^{5}$$

$$P(X=0) = 0,078 \quad = binomdist(0;5;0,40;1)$$

Interrogante b)

Si se seleccionan 5 adultos al azar:

¿Determine la probabilidad de que **como máximo 3** estén a favor del citado control?

Si X: “Número de adultos a favor”

$$P(X \leq 3) = ?$$

Planteo b)

X: Número de de adultos a favor (V.A. Discreta)

X → V.A Discreta

X → Unidimensional

X ~ Binomial(n=5, p=0,40)

$$P(X \leq 3) = \sum_{i=0}^{3} \binom{n}{x_i} \; \mathrm{p}_i^x (1-p)^{n-x_i}$$

$$P(X \leq 3) = 0,91 \qquad = binomdist(3;5;0,40;1)$$

Interrogante c)

Si se seleccionan 5 adultos al azar:

¿Cuál es el **valor esperado** de adultos a favor del control?

Si X: “Número de adultos a favor”

$$E(X) = ?$$

Planteo c)

Si se seleccionan 5 adultos al azar:

¿Cuál es el **valor esperado** de adultos a favor del control?

Si X: “Número de adultos a favor”

Si X ~Binom

$E(X) = n*p = 5*0,40 = 2$

Respuestas

a) La probabilidad de que ninguno esté a favor del citado control es 0,078.

b) La probabilidad de que como máximo 3 estén a favor del control es 0,91.

c) El valor esperado de adultos seleccionados que están a favor del control es 2.

GUÍA 4A Y GUÍA 4B

- MODELOS DISCRETOS
  - BERNOULLI
  - BINOMIAL
  - GEOMÉTRICO
  - POISSON
  - HIPERGEOMÉTRICO

- MODELOS CONTINUOS
  - EXPONENCIAL
  - GAMMA
  - NORMAL
  - LOGNORMAL
  - ~~RELACIONADOS AL NORMAL~~
  - DE VALORES EXTREMOS

MODELO EXPONENCIAL

Este modelo surge al considerar el ***tiempo transcurrido hasta*** la primera ocurrencia de un evento que pueda ser considerado como un proceso de Poisson.

MODELO EXPONENCIAL

Este modelo surge al considerar el ***tiempo transcurrido hasta*** la primera ocurrencia de un evento que pueda ser considerado como un proceso de *Poisson*.

Recordemos:

El modelo *Poisson* considera la **cantidad de veces** que sucede un evento en un período de tiempo.

RELACIÓN ENTRE MODELOS EXPONENCIAL - POISSON

Consideremos un evento *A* de tipo *Poisson y sea:*

$$\text{X: "Cantidad de veces que sucede A en un tiempo K"}$$

Tal como se ha visto,

$$\text{X} \sim Poisson \qquad E(X) = \lambda_k$$

Luego, la probabilidad de que no se produzcan eventos A en el intervalo de tiempo K se puede obtener como:

$$P(X=0) = \frac{e^{-\lambda} * \lambda^0}{0!} = e^{-\lambda}$$

consultas: taulamet@unl.edu.ar

RELACIÓN ENTRE MODELOS EXPONENCIAL - POISSON

Consideremos ahora la variable *T* definida como:

$$\text{T: ”Tiempo hasta el primer evento A”}$$

Resulta interesante que pensar la probabilidad de que *T* supere al tiempo *K*, es equivalente a pensar la probabilidad de que no se produzca ningún evento A en el período de tiempo *K*, en símbolos:

$$P(X=0) = P(T>K)$$

Con lo cual:

$$P(T>K) = e^{-\lambda}$$

RELACIÓN ENTRE MODELOS EXPONENCIAL - POISSON

Consideremos ahora la variable *T* definida como:

$$\text{T:“Tiempo hasta el primer evento A”}$$

De igual forma, si consideramos ahora una fracción *t* del período de tiempo estudiado (siendo *t* positivo):

$$P(T > t); t > 0$$

Tendremos entonces:

$$P(X = 0) = e^{-\lambda t}$$
$$P(T > t) = e^{-\lambda t}$$

MODELO EXPONENCIAL

Resumiendo, sea X ~ exponencial:

$$P(T > t) = e^{-\lambda t}$$

Luego, por probabilidad contraria:

$$F(t) = \begin{cases} 0 & \text{para } t \leq 0 \\ 1 - e^{-\lambda t} & \text{para } t > 0 \end{cases}$$

derivando,

$$f(t) = \begin{cases} 0 & \text{para } t \leq 0 \\ \lambda e^{-\lambda t} & \text{para } t > 0 \end{cases}$$

$$\mathrm{E(T)} = 1 \,/\, \lambda \qquad \mathrm{V(T)} = 1/\, \lambda^2$$

TABLA RESUMEN

Biblioteca 2024 | 02 - Primer Parcial

Modificar Eliminar Visible

Secretaría de Educación

Biblioteca correspondiente a los temas del primer parcial del 2° Cuatrimestre de 2024.

[PDF] Tabla Resumen de Modelos

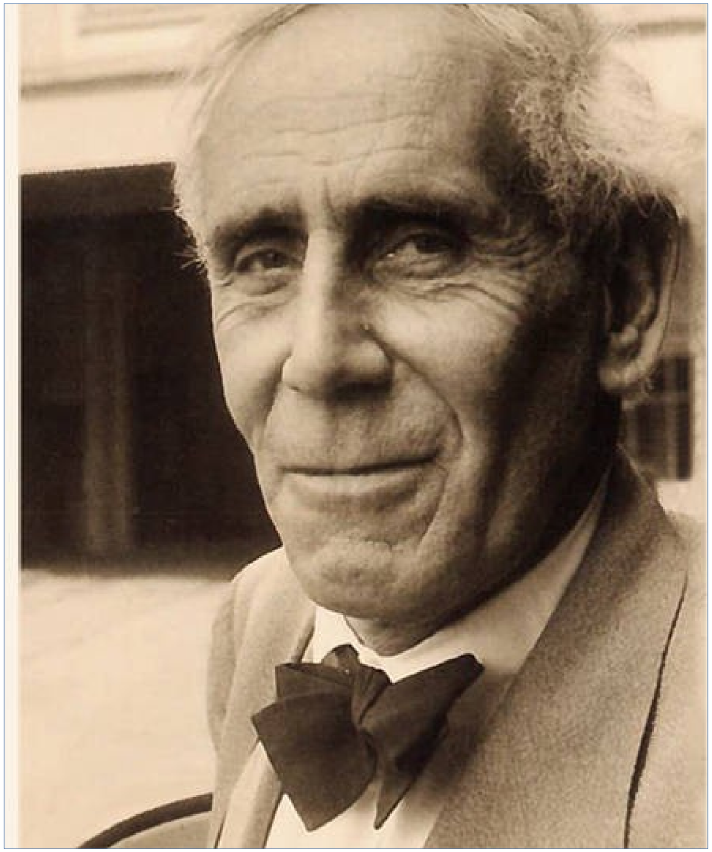
consultas: taulamet@unl.edu.ar

Normal	$f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$
Normal estándar	$f(z)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$
Normal	$F(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x}e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}dt$
Normal estándar	$F(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\displaystyle\int_{-\infty}^{z}e^{-\frac{1}{2}t^2}dt$

$E(x)=\mu$	$Var(x)=\sigma^2$	μ σ
$E(z)=0$	$Var(z)=1$	$z=\dfrac{x-\mu}{\sigma}$ $\mu=0$ $\sigma=1$

Situación Real Aplicada

Se cuenta con registros de lluvias máximas en la estación meteorológica Chajarí. Suponiendo que se obtuvo para esa región un valor medio de 200 mm con un desvío de 50 mm, obtener los parámetros del modelo adecuado para ajustarle a este tipo de variable y calcular la probabilidad de que ocurra una lluvia máxima superior a 150 mm en esa zona.



(1891 Munich, 1966 Nueva York)

Fue matemático, activista político y pacifista.

Era judío y huyó de Alemania en 1933, año en que se le privó de la nacionalidad alemana como parte de la Primera Lista de Expatriación del Reich Alemán.

"Es imposible que lo improbable nunca ocurra."

Emil Julius Gumbel

Ya que estamos trabajando con valores máximos sabemos que podemos utilizar la distribución de Gumbel.

X: "Lluvia máxima en la estación Chajarí"

X ~ Gumbel(μ,β) $$F(x) = \exp(-e^{-(x-\mu)/\beta})$$

Se conoce la esperanza y el desvío, y se deben calcular μ y β para poder hallar la probabilidad que se pide:

$$E(X) = 200 \quad D(X) = 50 \quad V(X) = 2500$$

$$V(X) = \frac{(\beta\pi)^2}{6} \Longrightarrow \beta = \sqrt{\frac{2500*6}{\pi}}$$

$$\Longrightarrow \beta \approx 38.98$$

$$\mathrm{E}(X) = \mu + \gamma\beta \Longrightarrow \mu = 200 - 0.57 * 38.98$$

$$\Longrightarrow \mu \approx 177,05$$

Siendo: $\pi \approx 3.1415 \ \gamma \approx 0.5772$

Calculando...

$$\mu \approx 177,05 \quad \beta \approx 38.98$$

$$F(x) = \exp(-e^{-(x-\mu)/\beta})$$

$$P(X > 150) = 1 - F(150)$$

$$1 - e^{-e^{(150-177,51)/38.98}} \approx 87\%$$

En GNUmeric =r.pgumbel(150;177,51;38.98;0)

Calculando...

$$\mu \approx 177,05 \quad \beta \approx 38.98$$

$$F(x) = \exp(-e^{-(x-\mu)/\beta})$$

$$P(X > 150) = 1 - F(150)$$

$$1 - e^{-e^{(150-177,51)/38.98}} \approx 87\%$$

En GNUmeric =r.pgumbel(150;177,51;38.98;0)