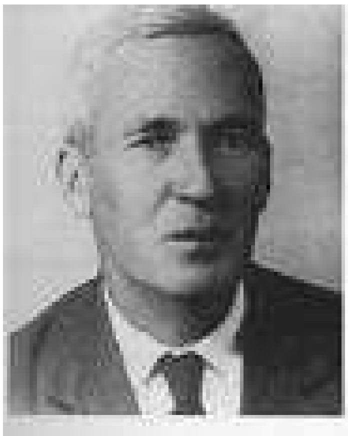


Laplace
1749-1827)

Kolmogorov
(1903-1987)

Fisher
(1890-1962)

BLOQUE III:

ESTADÍSTICA

TEMA1: SUCESOS ALEATORIOS

1.1.-EXPERIMENTO ALEATORIO

Se dice que un experimento es aleatorio si es imposible predecir el resultado cada vez que se realiza.

Ejemplos

A) Lanzar al aire una moneda y observar la cara superior
B) Lanzar un dado y observar la cara superior
C) Extraer una carta de la baraja
D) Lanzar un dado de quinielas

1.2.-ESPACIO MUESTRAL

El espacio muestral de un experimento aleatorio es el conjunto do todos los resultados posibles de dicho experimento . Se designa por $\mathbb{E}$.

Ejemplos

A) El espacio muestral asociado al experimento "lanzar una moneda" es
$\mathbb{E}$ ={Cara,cRuz}
B) Para el experimento "lanzar un dado" el espacio muestral es
$\mathbb{E}$ = {1,,2,3,4,5,6}
C) Para el experimento "lanzar un dado de quinielas" el espacio muestral es
$\mathbb{E}$ = {1,X,2}
D) Para el experimento "extraer una carta de la baraja" el espacio muestral consta do 40 elementos: cada una de las cartas de la baraja.

Consideraremos solamente experimentos cuyo espacio muestral $\mathbb{E}$ conste de un numero finito de elementos.

1.3.-SUCESOS

Dado un experimento aleatorio cuyo espacio muestral es $\mathbb{E}$, se llama suceso a cada uno de los subconjuntos de $\mathbb{E}$.

Ejemplos

En el experimento "lanzar un dado" son sucesos:
A={5} que podríamos llamar "salir un cinco"
B= {1,3,5} que podríamos llamar "salir impar"
C= {3,6} que podríamos llamar "salir múltiplo de 3"
D={4,5,6} que podríamos llamar "salir más de 3"

Al conjunto de todos los sucesos ,es decir, al conjunto P($\mathbb{E}$) (partes de $\mathbb{E}$) se le llama espacio de sucesos

Ejemplo

En el experimento "lanzar un dado de quinielas" el espacio de sucesos es
P($\mathbb{E}$) ={ $\varnothing$, {1} , {X} , {2} ,{1,x} , {1,2} , {X,2} , $\mathbb{E}$ }

Al realizar el experimento se dice que se ha verificado un suceso A si el resultado obtenido es uno de los que constituyen A

Ejemplo

Al lanzar un dado se verificará el suceso A={2,4,6} ="salir par " si el resultado obtenido es un 2 ó un 4 ó un 6.

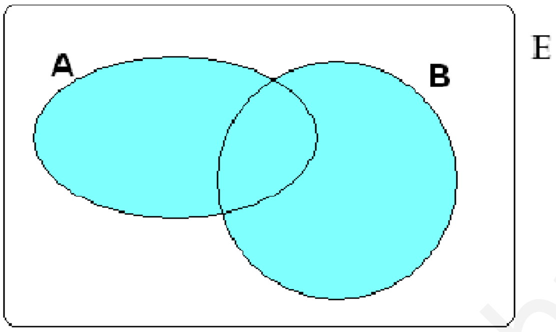
El suceso $\emptyset$ nunca se verifica: se le llama suceso imposible. El suceso E se verifica siempre: se le llama suceso seguro.
A los sucesos que constan de un solo elemento se les llama sucesos elementales.

1.4.- OPERACIONES CON SUCESOS

Los sucesos son conjuntos y como tales se pueden someter a las operaciones propias de los conjuntos, a saber:

1.4.1.- Unión de sucesos:

El suceso $A \cup B$ se llama suceso unión de A y B. El suceso $A \cup B$ se verifica cuando se verifica A ó se verifica B.

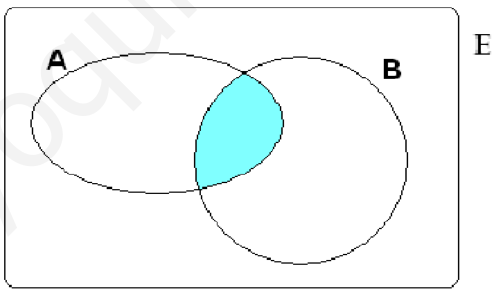


Ejemplo

Experimento:"lanzar un dado" .
Si A={2,4,6} y B={3,6} entonces $A \cup B$={2,3,4,6}

1.4.2.- Intersección de sucesos:

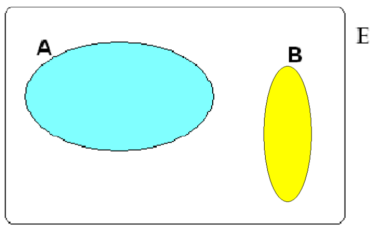
El suceso $A \cap B$ se llama suceso intersección de A y B. El suceso $A \cap B$ se verifica cuando se verifica A y se verifica B.



Ejemplo

Experimento:"lanzar un dado" .
Si A={2,4,6} y B={3,6} entonces $A \cap B$ ={6} .

Se dice que dos sucesos A y B son **incompatibles** si su intersección es el suceso imposible. Dos sucesos incompatibles no pueden verificarse simultáneamente.

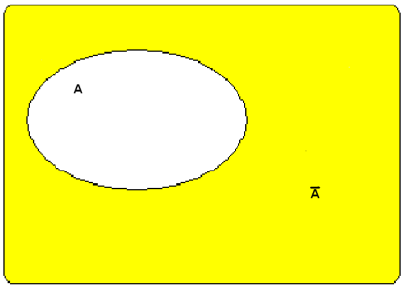


Ejemplo

En el experimento " lanzar un dado" son incompatibles los sucesos M={1,2} = "salir menos de 3" y N={5,6}="salir más de 4"

1.4.3.- Sucesos contrarios:

El suceso Ā constituido por los elementos de E que no pertenecen a A se llama suceso contrario de A. El suceso Ā se verifica cuando no se verifica A (y recíprocamente)



Ejemplo

Experimento:"lanzar un dado"

A={1,3,5} ="salir impar" ⇒ Ā ={2,4,6} = "salir par"

Es evidente que dos sucesos contrarios son incompatibles (aunque sucesos incompatibles no tienen por que ser contrarios)

1.5.- ÁLGEBRA DE SUCESOS

Así pues , con los sucesos se pueden realizar las operaciones de a unión e intersección. Ademas estas operaciones con sucesos gozan de las siguientes propiedades (que son características de una estructura llamada ALGEBRA DE BOOLE) :

1.-Conmutativa: $A\cup B=B\cup A$ $\qquad$ $A\cap B = B\cap A$
2.-Asociativa: $A\cup (B \cup C) = (A \cup B) \cup C$ $\qquad$ $A \cap (B \cap C) = (A \cap B) \cap C$
3.-Distributiva: $A\cup(B\cap C) = (A\cup B) \cap (A\cup C)$ $\qquad$ $A\cap (B \cup C) = (A\cap B) \cup (A\cap C)$
4.-Neutros: $A\cup\varnothing = A$ $\qquad$ $A\cap E = A$
5.-Complementación: para cada A existe un Ā tal que
$A\cup\bar{A} = E$ $\qquad$ $A\cap\bar{A} = \varnothing$

Por ello diremos que el espacio de sucesos con la unión y la intersección es un álgebra de Boole: el álgebra de sucesos

TEMA 2: PROBABILIDAD

2.1.-FRECUENCIA ABSOLUTA Y RELATIVA DE UN SUCESO

Consideremos un experimento aleatorio que repetimos **N** veces . Consideremos un suceso A de dicho experimento. Definimos

-Frecuencia absoluta del suceso A es el número de veces **n** que se ha verificado dicho suceso en los **N** ensayos del experimento

-Frecuencia relativa del suceso A : es el cociente entre su frecuencia absoluta y el número de veces que se ha repetido el experimento. Es decir:

$$f(A) = \frac{n}{N}$$

La frecuencia relativa tiene tres propiedades básicas:

1.- $0 \leq f(A) \leq 1$ para cualquier suceso A

2- $f(E) = 1$

3.- $f(A \cup B) = f(A)+f(B)$ si A y B son sucesos incompatibles

No es difícil demostrar estas propiedades. Hágase como ejercicio. Hay otras propiedades tales como $f(\emptyset)=0$, $f(A)+f(\bar{A})= 1$ que se pueden deducir de las tres anteriores.

2.2.-IDEA DE PROBABILIDAD

La persona que tira un dado podría afirmar que el resultado va a ser un número menor que 6, pero no puede estar totalmente seguro. Pero tiene más seguridad en esto que en que salga un 6 . Es decir, hay más seguridad de que unos sucesos se verifiquen que otros.

Surge, en consecuencia, la necesidad de medir de alguna manera la seguridad de que un determinado suceso se verifique al realizar el experimento. Esta medida es lo que se conoce como probabilidad.

La primera definición de probabilidad es debida a Laplace. Es la siguiente :

"La probabilidad de un suceso A es el cociente entre el número de casos favorables al suceso y el número de casos posibles si éstos son igualmente probables"

Probabilidad de Laplace: $$p(A) = \frac{\text{nº de casos favorables}}{\text{nº de casos posibles}}$$

Esta definición es fácil de aplicar. Por ejemplo consideremos el experimento "lanzar un dado" . Sea A el suceso "salir múltiplo de 3" .Está claro que los casos posibles son seis: 1 , 2 , 3 , 4 , 5 o 6 . Si el dado está bien construido estos seis casos son equiprobables. Los casos favorables al suceso A son dos: 3 y 6 . Por tanto P(A)= 2/6=1/3

Pero esta definición de probabilidad presenta inconvenientes. En primer lugar , desde el punto de vista formal es incorrecta ya que emplea el término a definir (*igualmente <u>probables</u>*) en la definición. Desde el punto de vista práctico solo es aplicable a los casos de equiprobabilidad , condición que por otra parte no es fácil de determinar.

Otro camino seguido para definir probabilidad es el que se basa en las frecuencias relativas teniendo en cuenta la siguiente ley:

Ley del azar: Al repetir un experimento un número elevado de veces la frecuencia relativa de un suceso tiende a estabilizarse alrededor de un cierto valor.

Es imposible demostrar que esta hipótesis es verdadera pero la experiencia muestra que se puede aceptar razonablemente. En efecto si por ejemplo lanzamos repetidamente una moneda vemos que la frecuencia relativa del suceso "salir cara" tiende a estabilizarse alrededor del valor 0,5 ,es decir, la cara tiende a salir la mitad de las veces.

De acuerdo con esta Ley del azar se puede dar la siguiente definición de probabilidad:

"La probabilidad de un suceso A es el número alrededor del cual se estabiliza su frecuencia relativa al repetir indefinidamente el experimento"

Esta definición no presenta los inconvenientes de la de Laplace, pero utiliza aspectos puramente experimentales que la hacen incómoda.

Para eliminar todas estas deficiencias, es Kolgomorov quien en 1933 introduce una definición axiomática de la probabilidad que vamos a estudiar en la siguiente pregunta.

2.3.-DEFINICIÓN DE PROBABILIDAD

La idea fundamental de esta definición es considerar la íntima relación entre frecuencia relativa y probabilidad. Para ello se toman como axiomas (principios indemostrables) los aspectos más esenciales de la frecuencia (vease 2.1) y a partir de ellos se obtienen, por deducción, los demás resultados Tenemos así la definición axiomática de la probabilidad:

Dado un experimento aleatorio cuyo espacio muestral es $\mathbb{E}$, se llama probabilidad a una aplicación

$$p: \mathbb{P}(\mathbb{E}) \longrightarrow \mathbf{R}$$

(*es decir , una regla por la que a cada suceso se le asigna un número - su probabilidad -) que verifica los siguientes axiomas*:

P1) $0 \leq p(A) \leq 1$ *para cualquier suceso A*

P2) $p(\mathbb{E}) = 1$

P3) $p(A \cup B) = p(A) + p(B)$ *para todo par do sucesos A y B incompatibles.*

Obsérvese que como axiomas se han tomado las idealizaciones de las tres propiedades básicas de las frecuencias relativas.

Ejemplo

Consideremos el experimento "lanzar una moneda" .El espacio muestral es $\mathbb{E}$= { C, R} Definimos la aplicación $p: \mathbb{P}(\mathbb{E}) \longrightarrow \mathbf{R}$ así :

$$p(\varnothing)=0 \ , \ p(\{C\})=1/2 \ , \ p(\{R\})=1/2, \ \ p(\mathbb{E})=1$$

Esta aplicación es una probabilidad.

También lo es la aplicación $p: \mathbb{P}(\mathbb{E}) \longrightarrow \mathbf{R}$ definida por:

$$p(\varnothing)=0 \ , \ p(\{C\})=2/3 \ , \ p(\{R\}) =1/3, \ \ p(\mathbb{E})=1.$$

Respondería esta probabilidad a un experimento en el que la moneda estaría tarada de tal manera que la cara tiende a salir el doble de veces que la cruz.

EJERCICIOS

2.1.-Un experimento consiste en extraer una bola de una urna que contiene una bola azul, 2 rojas y 3 verdes y observar el color de la bola extraida.Determinar el espacio muestral, el espacio de sucesos y definir una probabilidad

2.4.-PROPIEDADES DE LA PROBABILIDAD

A partir de los axiomas se demuestran las siguientes propiedades de la probabilidad:

1.-Probabilidad del suceso contrario: $p(\bar{A}) = 1 - p(A)$

2.-Probabilidad del suceso imposible: $p(\varnothing) = 0$.

3.-Probabilidad de la unión de sucesos incompatibles: Si A , B y C son tres sucesos incompatibles dos a dos entonces :

$$p(A \cup B \cup C) = p(A)+ p(B)+p(C)$$

4.-Probabilidad de la unión de sucesos compatibles:

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

5.-Si $A \subset B$ entonces $p(A) \leq p(B)$

6.-Regla de Laplace: Sean W_1 , W_2,....,W_n los **n** sucesos elementales de un experimento aleatorio, todos ellos equiprobables. Si A es la unión de **k** de estos sucesos elementales entonces:

$$p(A) = \frac{k}{n} = \frac{\text{casos favorables}}{\text{casos posibles}}$$

EJERCICIOS

2.2-Se tiran tres monedas diferentes al aire. Determinar el espacio muestral y la probabilidad de obtener dos caras

2.3.-Al extraer una carta de la baraja hallar las siguientes probabilidades: A) de obtener una copa B) de obtener un rey C) de obtener una figura D) de obtener una figura de oros E)de obtener un siete ó una espada F) de no obtener bastos G) de no obtener ni bastos ni figuras

2.4.-Al lanzar dos dados ¿Cual es la probabilidad de que la suma de puntos sea 7?

2.5.-Entre 4 tornillos la mitad son defectuosos. Si elegimos dos al azar ¿Cual es la probabilidad de que aparezca al menos uno defectuoso?¿Y si eligiéramos 3?

2.6.-Al extraer dos cartas simultáneamente de la baraja hallar las siguientes probabilidades: A) de obtener dos sotas B) de obtener dos copas C) de que una de las cartas sea figura D) de obtener una copa y un basto

2.7.-En un baile se reúnen diez matrimonios disfrazados. Al elegir dos personas hallar la probabilidad A) de que sean esposos B) de que sean dos mujeres C) de que sea una mujer y un hombre D) de que una por lo menos sea mujer

2.5.- PROBABILIDAD CONDICIONADA

La obtención de información adicional sobre el resultado de un experimento puede modificar la probabilidad de algún suceso.

Por ejemplo: Consideremos el experimento "lanzar un dado" .Consideremos los sucesos A= "salir impar" = {1, 3, 5 } y B="salir > 1"= {2,3,4,5,6} . Según la regla de Laplace p(A)=3/6 y p(B)=5/6

Si al lanzar el dado nos dicen que ha salido impar (es decir, que se ha verificado A) entonces la probabilidad de B queda condicionada por la verificación de A. En efecto, la probabilidad de B sabiendo que se ha verificado A es 2/3 ya que los casos posibles se han reducido a tres (de los cuales 2 son favorables a A).

Al suceso consistente en que se verifique B habiéndose verificado A se le llama **"suceso B condicionado a A"** . Se simboliza por B/A

Observemos en el ejemplo anterior que:

p(B/A)=2/3

p(A)=3/6

$A\cap B=\{3,5\}\Rightarrow p(A\cap B)=2/6$

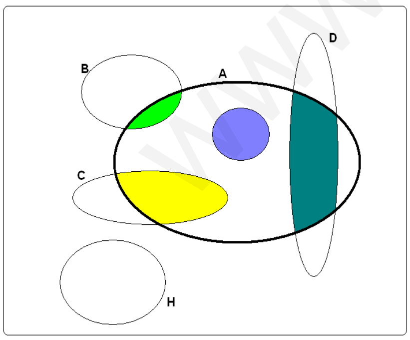
Vemos que se cumple :

$$\frac{p(A\cap B)}{p(A)}=\frac{2/6}{3/6}=\frac{2}{3}=p(B/A)$$

La anterior igualdad nos lleva a dar la siguiente definición de probabilidad condicionada:

Para cada suceso B se define la probabilidad condicionada p(B/A) ***como el*** *cociente:*

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



E

La probabilidad condicionada a A consiste en repartir la unidad de probabilidad de que disponemos dando a cada suceso según su "nivel de participación en A". Por ello los sucesos incompatibles con A tienen nula su probabilidad condicionada a A. Un suceso que contenga a A tiene su probabilidad condicionada a A igual a 1.

2.6.- PROBABILIDAD COMPUESTA

La noción de probabilidad condicionada se puede utilizar para calcular la probabilidad de la intersección a partir de la de uno de ellos y de la del otro condicionada a él (probabilidades que generalmente se puede calcular directamente). En efecto, sean A y B dos sucesos. Si de la fórmula de la probabilidad condicionada despejamos p(A∩B) , obtenemos:

$$p(A\cap B) = p(A)\cdot p(B/A)$$

Que se llama fórmula de la probabilidad compuesta. Se puede generalizar para un número cualquiera de sucesos. Así para tres sucesos A, B y C tendría la siguiente expresión:

$$p(A\cap B\cap C) = p(A)\cdot p(B/A)\cdot p(C/A\cap B)$$

EJERCICIOS

2.8.-Al extraer sucesivamente 2 cartas de la baraja ¿Cuál es la probabilidad de qué las dos sean copas?
2.9.-Al extraer sucesivamente tres cartas de la baraja ¿Cuál es la probabilidad de que sean 3 ases?
2.10.-En un grupo de 30 hombres y 32 mujeres el 40% de los hombres son fumadores y el 25% de las mujeres. Elegida una persona al azar ¿Cuál es la probabilidad de que sea una mujer no fumadora?

2.7.- SUCESOS INDEPENDIENTES: REGLA DEL PRODUCTO

Si p(B/A) > p(B) entonces el suceso A favorece la probabilidad de B y si p(B/A)< p(B) la perjudica. Si p(B/A)=p(B) entonces la probabilidad de B no se ve influenciada por el hecho de que se le condicione a A. Esto nos lleva a la siguiente definición:

Se dice que A y B son sucesos independientes si p(B/A)=p(B) (ó p(A/B)=p(A))

Si A y B son sucesos independientes entonces la fórmula de la probabilidad compuesta toma la siguiente forma:

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

Ejemplo

Consideremos el experimento "lanzar un dado" Sea A="salir par" (p(A)=1/2) . Sea B="salir múltiplo de 3" (p(B)=2/6)

Para el suceso B/A ("salir múltiplo de 3 si ha salido par") los casos posibles son 2,4 y 6, de los cuales sólo el 6 es favorable. Luego p(B/A)=1/3=p(B) .Por tanto los sucesos A y B son independientes. Obsérvese que esto es debido a que la proporción de múltiplos de 3 respecto a los resultados pares (6 frente a 2,4 y 6) es la misma que la de múltiplos de 3 respecto a la totalidad de los resultados (3 y 6 frente a 1,2,3,4,5 y 6)

La independencia de sucesos resulta especialmente clara en el caso de experimentos compuestos como el siguiente: Sea el experimento "lanzar una moneda dos veces" Sean los sucesos A="salir cara la primera vez" y B="salir cara la 2ª vez" .Estos sucesos son independientes pues el hecho de que salga cara la 1ª vez no modifica la probabilidad de que salga cara la 2ª. Como aplicación de la regla del producto tenemos que la probabilidad de que salga cara las dos veces es p(A∩B). P(A)·p(B)=(1/2)(1/2)=1/4

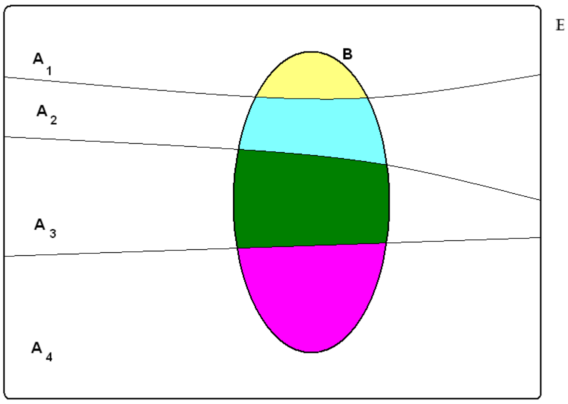
EJERCICIOS

2.11.-Al extraer una carta de la baraja encontrar dos sucesos independientes
2.12.-Se lanza un dado dos veces ¿Cual es la probabilidad de obtener par en la primera tirada e impar en la 2ª?
2.13-Se lanzan dos dados ¿Cual es la probabilidad de obtener par en un dado e impar en el otro?

2.8.- PROBABILIDAD TOTAL

Sean A_1 ,A_2 ,.......,A_n sucesos incompatibles dos a dos tales que su unión es $\mathbb{E}$. Para cualquier suceso B se verifica:

$$p(B) = p(A_1)\cdot p(B/A_1) + p(A_2)\cdot p(B/A_2) + \cdots\cdot + p(A_n)\cdot p(B/A_n)$$



Ejemplo

En una clase de 40 alumnos hay 10 que aprueban todas las asignaturas ,al 70% de los cuales les gustan las Matemáticas. Hay 15 que suspenden una ó dos asignaturas , al 50% de los cuales les gustan las Matemáticas. Finalmente hay 15 alumnos que suspenden más de dos asignaturas , al 20% de los cuales les gustan las Matemáticas. Se elige al azar un alumno de esta clase ¿Cual es la probabilidad de que le gusten las Matemáticas?

A_1="aprobar todo" $\Rightarrow p(A_1)=10/40$

A_2="suspender 1 ó 2" $\Rightarrow p(A_2)=15/40$

A_3="suspender más de 2" $\Rightarrow p(A_3)=15/40$

Si llamamos B="gustar las Matemáticas" tendremos que $p(B/A_1)= 70/100$, $p(B/A_2)= 50/100$ y $p(B/A_3)= 70/100$ y entonces según la probabilidad total:

$$p(B) = p(A_1)\cdot p(B/A_1) + p(A2)\cdot p(B/A_2) + p(A_3)\cdot p(B/A_3) =$$
$$= \frac{10}{40}\frac{70}{100} + \frac{15}{40}\frac{50}{100} + \frac{15}{40}\frac{20}{100} = 0.4375$$

2.8.- LA REGLA DE BAYES

La regla de Bayes sirve para calcular probabilidades a posteriori, es decir, para resolver los problemas del tipo: "Si un resultado B puede ser producido por varias causas incompatibles hallar la probabilidad de que ocurra debido a una de esas causas"

Consideremos el ejemplo de la pregunta anterior: Supongamos que sabemos que al alumno elegido le gustan las matemáticas.Podemos entonces preguntarnos si este alumno ha aprobado todas las asignaturas. A esto da respuesta el teorema de Bayes:

Sean A_1 ,A_2 ,.......,A_n sucesos incompatibles dos a dos tales que su unión es E .Se verifica:

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

Ejemplo

Vamos a responder a la pregunta planteada previamente. Habremos de calcular cual es la probabilidad de que el alumno elegido haya aprobado todas las asignaturas sabiendo que le gustan las Matemáticas. Esta probabilidad es $p(A_1/B)$ que aplicando la Regla de Bayes es:

$$P(A_i/B) = \frac{P(A_1)P(B/A_1)}{P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + P(A_3)P(B/A_3)} = \frac{\frac{10}{40}\frac{70}{100}}{\frac{10}{40}\frac{70}{100} + \frac{15}{40}\frac{50}{100} + \frac{15}{40}\frac{20}{100}} = 0.4$$

EJERCICIOS

2.14.- María y Laura idean el siguiente juego: cada una lanza un dado, si en los dos dados sale el mismo número, gana Laura; si la suma de ambos es 7 gana María; y en cualquier otro caso hay empate.

a) Calcular la probabilidad de que gane Laura.
b) Calcular la probabilidad de que gane María.

2.15-Dados dos sucesos aleatorios A y B, se sabe que: $p(\overline{B}) = 3/4$ y $p(A) = p(A/B) = 1/3$

a) Razonar si los sucesos A y B son independientes.
b)Calcular $P(A \cup B)$.

2.16.- Las probabilidades de aprobar los exámenes de Historia, Lengua e Inglés son, para un alumno determinado: 2/3, 4/5 y 3/5 respectivamente. Obtener las probabilidades de:
1. Suspender las tres asignaturas.
2. Suspender solo una de las tres.
3. Suspender Lengua si se sabe que solo suspendió una asignatura de las tres

2.17.-En una determinada asignatura hay matriculados 2500 alumnos. En Junio se presentaron 1800 de los que aprobaron 1015, mientras que en septiembre, de los 700 que se presentaron, suspendieron 270. Elegido al al azar un alumno matriculado en esa asignatura,
a) Calcular la probabilidad de que la haya aprobado.
b) Si ha aprobado la asignatura, cuál es la probabilidad de haberse presentado en Septiembre.

2.18.-En un centro de Secundaria, aprueban Biología 4 de cada 5 alumnos, las matemáticas las aprueban 2 de cada 3 alumnos y 3 de cada 5 alumnos aprueban la Lengua. Elegido al azar un alumno matriculado de esas asignaturas en ese centro, calcula la probabilidad de que:
1. Suspenda esas tres asignaturas.
2. Suspenda solo una de ellas.

2.19.-El 20% de los habitantes de una determinada población son jubilados y otro 20% son estudiantes. La música clásica le gusta al 75% de los jubilados, al 50% de los estudiantes y al 20% del resto de la población. Calcula la probabilidad de que elegida al azar una persona a la que le gusta la música clásica sea jubilada.

2.20.-El 60% de las personas que visitaron un museo durante el mes de mayo eran españoles. De estos, el 40% eran menores de 20 años. En cambio, de los que no eran españoles, tenían menos de 20 años el 30%. Calcular:
a). La probabilidad de que un visitante elegido al azar tenga menos de 20 años.
b) Si se escoge un visitante al azar, !a probabilidad de que no sea español y tenga 20 años o más.

2.21.-Las máquinas A y B producen 50 y 250 piezas por hora, con un porcentaje de fallos del 1% y del 10% respectivamente. Tenemos mezcladas las piezas fabricadas en una hora y elegimos una pieza al azar. Calcular:
a)La probabilidad de que sea una pieza defectuosa fabricada en la máquina A
b)La probabilidad de que esté fabricada en la máquina A si sabemos que es defectuosa

2.22.- En el segundo curso de bachillerato de cierto instituto se han matriculado el doble mujeres que de varones. Sabiendo que un 25% de las mujeres fuman y que no lo hacen un 40% de los varones, determinar la probabilidad de que seleccionada al azar una persona en el segundo curso de bachillerato de ese instituto resulte ser una persona fumadora. Justificar la respuesta

2.23.-Una urna A contiene 5 bolas blancas y 4 negras, y otra urna B contiene 1 blanca y 2 negras. Se extrae una bola al azar de la urna A y se pone en la B. Después se extrae de la urna B una bola al azar.
a) Calcular la probabilidad de que la bola extraída de la urna B sea blanca.
b) Suponiendo que la bola extraída de la urna B ha sido blanca, calcular la probabilidad de que la bola extraída de la urna A también haya sido blanca

2.24.-Dos parejas de novios deciden ir al cine. Si se sientan al azar en cuatro butacas contiguas, ¿cuál es la probabilidad de que cada uno esté al lado de su pareja?

2.25.-Dos expertos, E1 y E2 realizan peritaciones para una cierta compañía de seguros. La probabilidad de que una peritación haya sido realizada por E1 es 0,55 y por E2 es 0.45. Si una peritación ha sido realizada por E1 la probabilidad de que de lugar al pago de una indemnización es 0.98 y si ha sido realizada por E2 la probabilidad de que dé lugar al pago de una indemnización es 0,90. Un siniestro ha supuesto a la compañía el pago de una indemnización. Hallar la probabilidad de que la peritación haya sido realizada por E2

2.26.-En una empresa se producen dos tipos de bombillas: halógenas y de bajo consumo, en una proporción de 3 a 4, respectivamente. La probabilidad de que una bombilla halógena sea defectuosa es 0,02 y de que una de bajo consumo sea defectuosa es 0,09. Se escoge al azar una bombilla y resulta defectuosa, ¿cuál es la probabilidad de que sea halógena?

2.27.-Un ordenador personal está contaminado por un virus y tiene cargados dos programas antivirus que actúan independientemente uno del otro. El programa p1 detecta la presencia del virus con una probabilidad de 0,9 y el programa p2 detecta el virus con una probabilidad de 0,8. ¿Cuál es la probabilidad de que el virus no sea detectado?

2.28.-En un colegio el 4% de los chicos y el 1 % de las chicas miden más de 175 cm de estatura. Además el 60% de los estudiantes son chicas. Si se selecciona al azar un estudiante y es más alto de 175 cm, ¿cuál es la probabilidad de que el estudiante sea chica?

2.29.-La probabilidad de que un estudiante universitario termine su carrera en los años establecidos por el plan de estudios es 3/5 y la de que su hermana finalice la suya sin perder ningún año es 2/3. Halla la probabilidad de que:
a)Ambos terminen sus estudios en los años establecidos
b) Solo el varón los termine en el plazo fijado
c) Al menos uno de los dos los termine en el tiempo establecido.

2.30.-En un grupo de personas, al 50% les han puesto alguna vez una multa de tráfico. Por otro lado, al 12,5% no les han puesto nunca una multa pero sí han sufrido alguna vez un accidente. Finalmente, al 60% de quienes nunca han tenido un accidente no les han puesto nunca una multa.
a) ¿Qué porcentaje no han tenido nunca un accidente ni les han puesto nunca una multa?
b) ¿Qué porcentaje no han tenido nunca un accidente?
c) Entre las personas que nunca han tenido una multa, ¿qué porcentaje no han tenido nunca un accidente?

2.31.-En la urna U1 hay 4 bolas blancas, numeradas de 1 a 4, y 2 bolas negras numeradas de 1 a 2. mientras que en la urna U2, hay 2 bolas blancas numeradas de 1 a 2 y 4 bolas negras numeradas de 1 a 4. Si se extraen al azar dos bolas, una de cada urna, hallar:
a) La probabilidad de que tengan el mismo número.
b) La probabilidad de que sean del mismo color.

2.32.-Para ir al trabajo un individuo toma el bus el 30% de las veces, o el metro (el 70% restante) y llega tarde el 40% de las veces que va en bus y e! 20% de las que va en metro. Cierto día llegó larde. ¿Cuál es la probabilidad de que tomara el bus?

2.33. -El 60% de los hombres son fumadores y el 30% de las mujeres .En una reunión hay 30 hombres y 20 mujeres. Si elegimos una persona al azar ¿Cual es la probabilidad de que sea fumadora?

2.34.-Para ir a su trabajo una persona sigue un itinerario A el 50% de los dias, un itinerario B el 40% y un itinerario C el 10%. Por el A suele llegar a tiempo el 95% de las veces, por el B el 80% y por el C el 30% . Un dia cualquiera ¿Cual es la probabilidad de que llegue a tiempo al trabajo?

2.35.-De una urna que contiene 10 bolas blancas y 20 negras se extraen dos bolas que se colocan en una segunda urna que ya contiene 8 blancas y seis negras y a continuación se extrae una bola de esta segunda urna ¿Cual es la probabilidad de que sea negra?

2.36.-Una urna contiene 4 bolas rojas, 2 negras y 3 amarillas. .Al sacar tres bolas al azar ¿Cual es la probabilidad de que al menos dos sean rojas?

2.37.-Pedro y Juan están en una reunión con otras 8 personas. Entre los presentes se reparten al azar números del 1 al 10 ¿Cuál es la probabilidad de que Pedro y Juan tengan números consecutivos?

2.38.-Se lanzan al aire 4 dados ¿Cual es la probabilidad de obtener más de 5 puntos en total?

2.39.-Al sacar tres cartas de una baraja hallar la probabilidad de que sean un as, un rey y una sota

2.40.-Una pieza de artillería dispone de 4 obuses para alcanzar un objetivo En cada disparo la probabilidad de alcanzarlo es 1/3 .¿Cual es la probabilidad de fallar los cuatro? ¿Y de fallar exactamente 2?

2.41.-Una bolsa contiene 20 bolas numeradas de 1 a 20 .Se extraen sucesivamente dos bolas. Se pide A)Probabilidad de que las dos bolas extraidas tengan número impar si se introduce la bola en la bolsa después de la primera extracción. B)Probabilidad de que la 1ª sea par y la 2ª impar , sin reemplazamiento

2.42.-Se distribuyen 4 bolas entre tres cajas sin dejar ninguna vacía.¿Cual es la probabilidad de que la primera caja tenga 2 bolas?

2.43.-El 30% de las piezas fabricadas por una máquina son defectuosas.¿Cual es la probabilidad de que al elegir 4 piezas al azar, haya al menos dos defectuosas?

2.44.-La probabilidad de que un alumno apruebe matemáticas es 0,6 , la de que apruebe lengua es 0,7 y de que apruebe las dos es 0,3 .Hallar A)La probabilidad de que apruebe al menos una. B)De que no apruebe ninguna. C)De que apruebe matemáticas y suspenda lengua

2.45.-Un dado está "cargado" de forma que al lanzarlo la probabilidad de obtener un número es proporcional a dicho número. Hallar la probabilidad de que al lanzarlo se obtenga un número par.

2.46.-Una enciclopedia consta de 8 tomos, uno de los cuales se titula "Matemáticas" .Calcular la probabilidad de que al elegir dos tomos al azar, uno de ellos sea el tomo "Matemáticas"

2.47.-Una urna contiene dos bolas , que pueden ser blancas , negras ó una blanca y una negra. Se añade una bola blanca a la bolsa y después se extrae una bola al azar.¿Cual es la probabilidad de que sea blanca?

2.48.-Un artículo es producido en 2 fábricas X e Y que producen el 60% y 40% respectivamente.La fábrica X produce un 1% de artículos defectuosos y la Y un 5% . Al comprar un articulo ¿Cuál es la probabilidad de que sea defectuoso ?

2.49. -Un alumno sabe 8 de las 10 preguntas que entran para un examen. Si en el examen le ponen 3 preguntas ¿Que probabilidad tiene de aprobar si para ello tiene que A) contestar bien las tres preguntas B) contestar por lo menos dos bien

2.50. -¿Cual es la probabilidad de que al extraer cuatro cartas de la baraja , 2 sean ases?

2.51. -Se lanza una moneda cuatro veces consecutivas ¿Cual es la probabilidad de que se obtenga por lo menos una cara? ¿Y de que salgan todas caras ó todas cruces?

2.52.- A un alumno lo lleva en coche a la facultad un amigo el 80% de los días. Cuando lo lleva el amigo llega tarde el 20% de las veces. Cuando el amigo no lo lleva llega temprano a clase el 10% de las veces. Determinar

a)La probabilidad de que llegue pronto a clase y le haya llevado el amigo
b)La probabilidad de que llegue tarde a clase
c)Ha llegado pronto a clase ¿ Cuál es la probabilidad de que no le haya llevado el amigo?

2.53.-Juan es el responsable de un aula de informática y no se puede confiar en él pues la probabilidad de que olvide hacer el mantenimiento de un ordenador en ausencia del jefe es 2/3. Si Juan le hace el mantenimiento a un ordenador, éste tiene la misma probabilidad de estropearse que de funcionar correctamente, pero si no le hace el mantenimiento solo hay una probabilidad de 0.25 de funcionar correctamente. . ¿Cuál es la probabilidad de que un ordenador funcione correctamente a la vuelta del jefe?. A su regreso, el jefe se encuentra un ordenador averiado ¿Cuál es la probabilidad de que Juan no le hiciera el mantenimiento?

TEMA 3: POBLACIÓN Y MUESTRA

3.1.-POBLACIÓN Y MUESTRA

POBLACIÓN O UNIVERSO: es el conjunto de todos los individuos que tienen en común un carácter del que queremos hacer un estudio.

MUESTRA: es un subconjunto extraido de la población, cuyo estudio va a servir para deducir las características de toda la población. Al número de elementos de la muestra se le llama tamaño muestral.

3.2.-MUESTREO

¿Por qué se recurre a las muestras?

-Población excesivamente numerosa
-Población muy difìcil o imposible de controlar
-Proceso de medición destructivo
-Necesidad de conocer rápidamente ciertos datos de la población

¿Cómo deben ser las muestras?

Hay dos aspectos a los que se debe prestar atención:

Su tamaño:

Si es demasiado pequeño las conclusiones acerca de la población pueden ser incorrectas. Por el contrario una muestra muy grande puede resultar muy costosa y difícil de manejar

Su forma de elección:

Es lo que se llama el muestreo. Ha de realizarse de forma que sea representativa de la población. Para ello es indispensable que sea aleatorio, es decir, que todos sus elementos se elijan al azar de modo que todos los individuos de la población tengan a priori la misma probabilidad de ser elegidos

¿Cuáles son los principales tipos de muestreo?

Muestreo aleatorio simple:

Es el más sencillo. Se numeran los elementos de la población y se seleccionan al azar los n elementos que debe contener la muestra

Muestreo aleatorio estratificado:

Si la población puede dividirse en estratos , se elige la muestra fijando de antemano el número de individuos de cada estrato. Cuando estos números son proporcionales a los tamaños de los estratos se dice que el muestreo es estratificado con reparto proporcional

3.3.-PARÁMETROS Y ESTADÍSTICOS

Parámetros de una población ó parámetros:

Media μ**:** indica el “centro” de la población
Varianza σ^2 : indica en que grado está dispersa la población en torno a la media
Desviación típica σ

Estadísticos muestrales ó estadísticos:

Media $\overline{X}$: indica el “centro” de la muestra
Varianza S^2 : indica el grado de dispersión de la muestra respecto a la media
Cuasivarianza $S_{n-1}^2 = \dfrac{n}{n-1} S^2$
Desviación típica S

3.4.-VARIABLE ALEATORIA

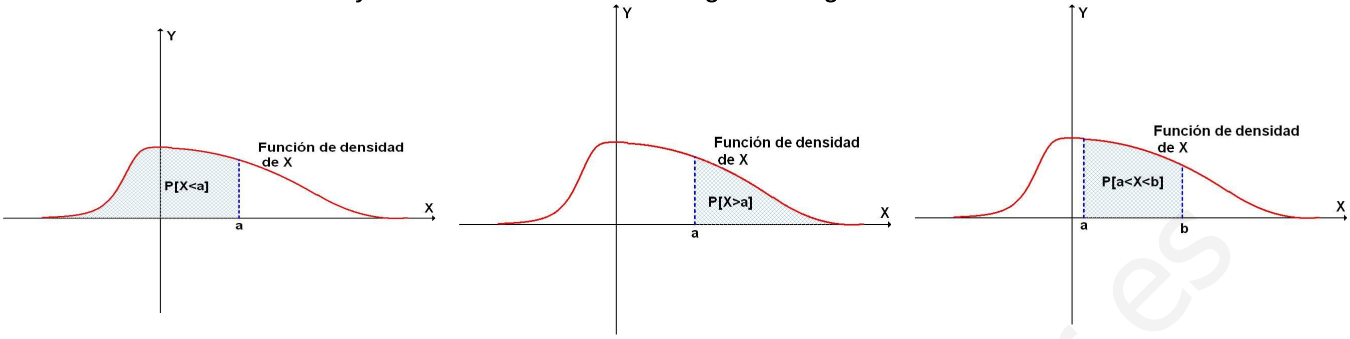
Una variable aleatoria es una variable X que puede tomar cualquiera de los valores numéricos asociados al carácter que estudiamos en la población. Puede ser :

-Continua: si toma cualquier valor en un intervalo

- Discreta: si solo toma valores enteros

3.5.-VARIABLE ALEATORIA NORMAL

Función de densidad de una variable aleatoria continua: es una función a partir de cuya gráfica se calculan las probabilidades (como área por debajo de la gráfica) de que la variable aleatoria tome determinados valores tal y como se muestra en la siguiente figura:

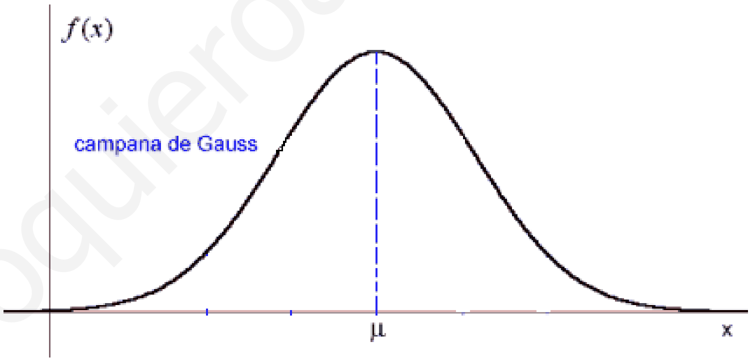


Variable aleatoria normal

Una variable aleatoria X continua se dice que es normal (o que tiene una distribución normal) de media μ y desviación típica σ , si su función de densidad es la curva normal (o campana de Gauss) que describimos a continuación:

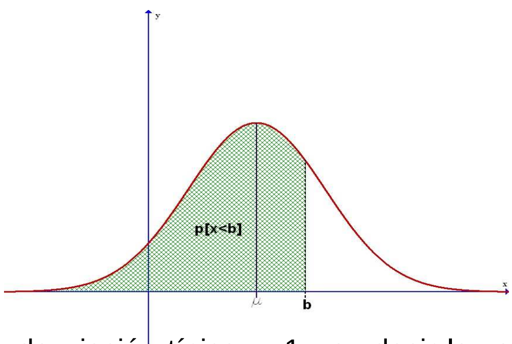
Curva normal o campana de Gauss:

Es la gráfica de la función $f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ que vemos a continuación:



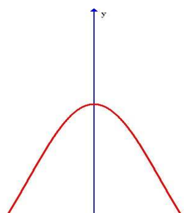
Esta curva es simétrica respecto de la recta vertical x=μ y el área que limita con el eje OX mide 1 u^2 (como en toda función de densidad)

Que X es normal de media μ y desviación típica σ se expresa escribiendo $X\in N(\mu\,;\,\sigma)$. En este caso la probabilidad **p[X<b]** es lo que mide el area limitada por la curva normal con OX hasta x=b , tal y como se muestra en la siguiente figura



Variable normal tipificada:

Se llama así a la variable normal de media μ=0 y desviación típica σ=1 , es decir la variable N(0; 1) Se designa con la letra Z .Su curva normal es la siguiente:

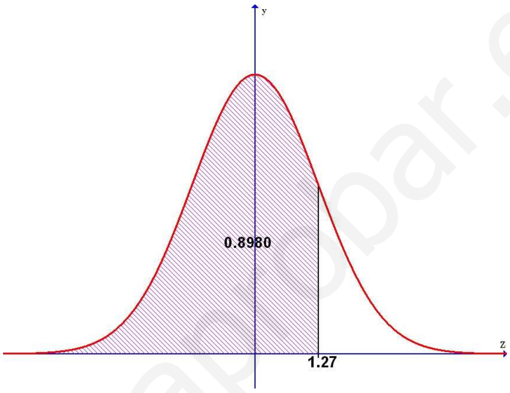


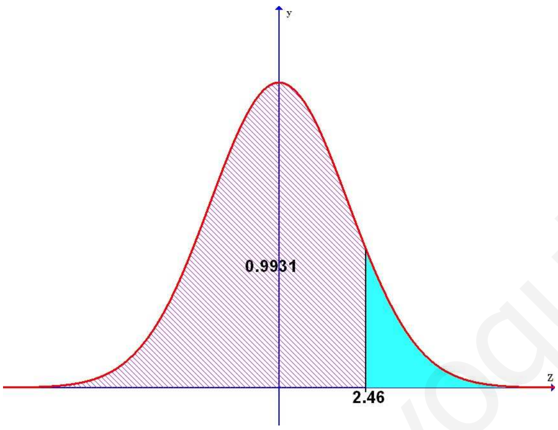
que es simétrica respecto del eje OY. Las áreas por debajo de esta curva están tabuladas (ver página 110)

3.6.-CALCULO DE AREAS BAJO LA CURVA NORMAL TIPIFICADA

Caso 1

P[Z<1,27] = (tabla)= 0,8980



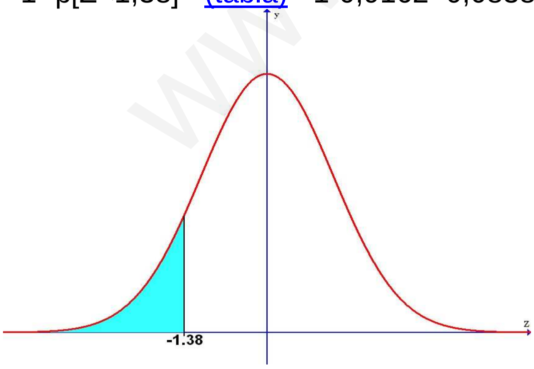


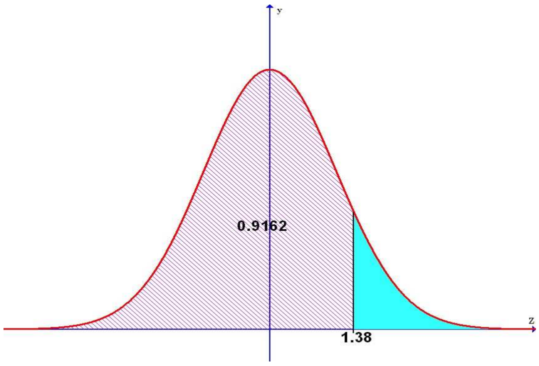
Caso 2

P[Z>2,46] = 1- p[Z<2,46] = (tabla)= 1-0,9931=0,0069

Caso 3

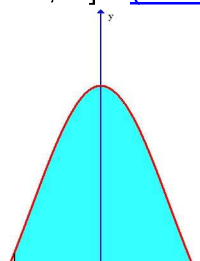
P[Z<-1,38] = p[Z>1,38] =
=1- p[Z<1,38]= (tabla)= 1-0,9162=0,0838

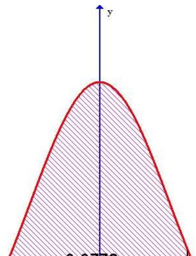




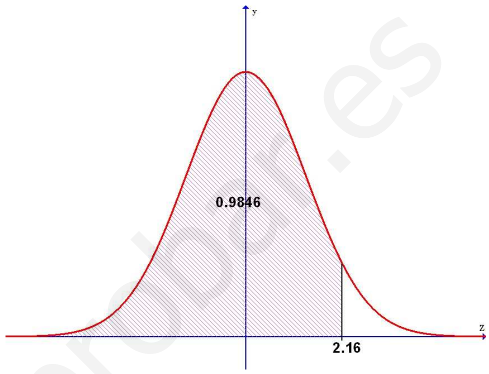
Caso 4

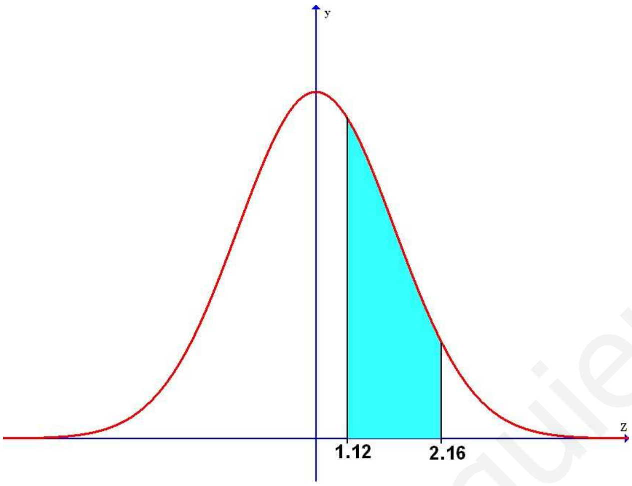
P[Z>-2,01] = p[Z<2,01] = (tabla)= 0,9778

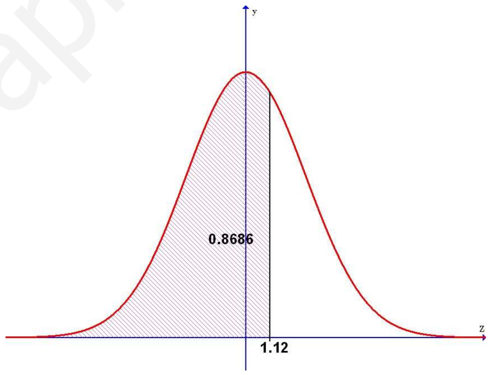




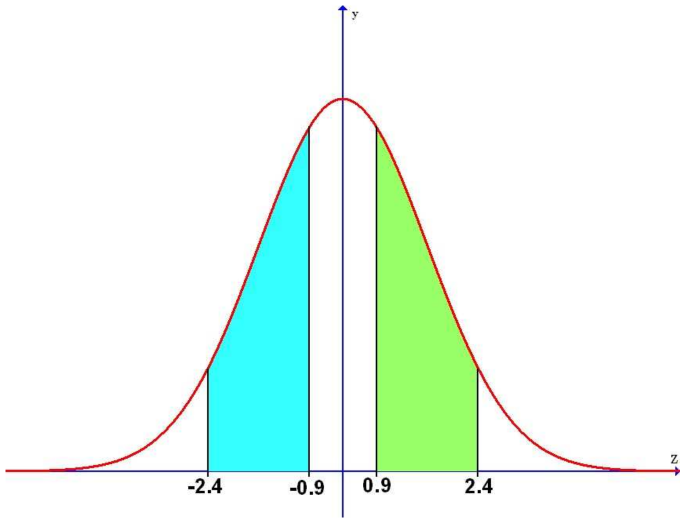
Caso 5
P[1,12<Z<2,16] = p[Z<2,16] - p[Z<1,12]
= [(tabla)]= 0,9846- 0,8686 =0,1160





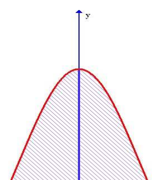


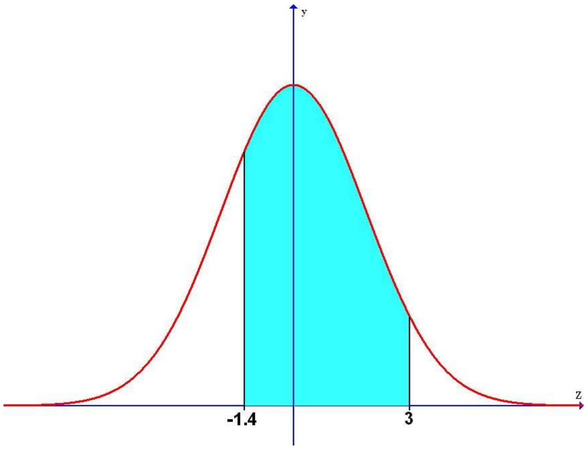
Caso 6
P[-2,4<Z<-0,9] = P[0,9<Z<2,4] =
(caso anterior) P[Z<2,4]- P[Z<0,9] =
(tabla)= 0,9918- 0,8159 =0,1759

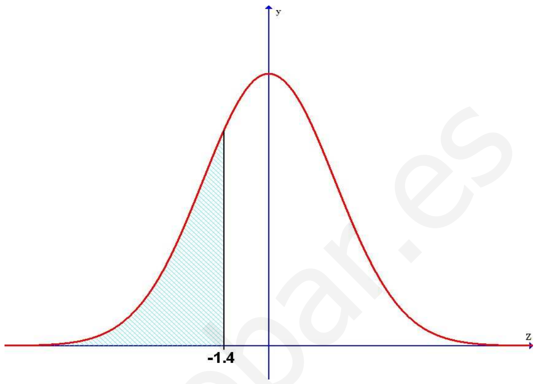


Caso 7

P[-1,4<Z<3] = P[Z<3] - P[Z<-1,4] = (tabla) 0,9987- {1-P[Z<1,4]} (caso3) = 0,9987- {1- (tabla) 0,9192} =0,9987+0,9192 -1=0,9179







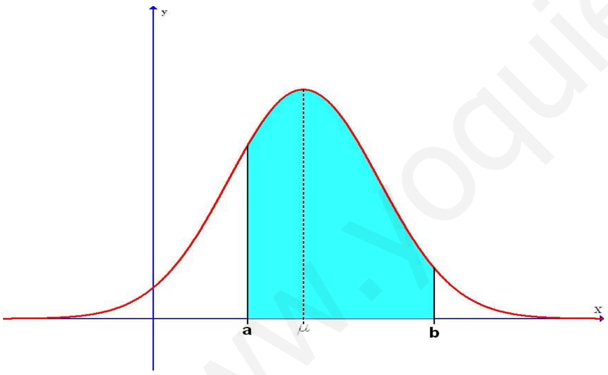
3.7.-CALCULO DE AREAS BAJO UNA CURVA NORMAL N(μ ; σ)

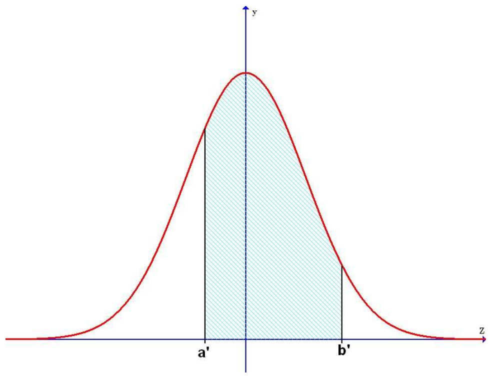
Si X es una variable con una distribución N(μ ; σ) entonces

$$\mathbf{Z = \frac{X-\mu}{\sigma}}$$

tiene una distribución N(0 ;1) , es decir, es la variable normal tipificada

Entonces: **P[a<X<b] = P[a´<Z< b´]**





teniendo en cuenta que:

$$X = a \quad \Rightarrow \quad Z = \frac{a-\mu}{\sigma} = a´$$

$$X = b \quad \Rightarrow \quad Z = \frac{b-\mu}{\sigma} = b´$$

EJERCICIOS

3.1.-Si X sigue una distribución N(2,1 ; 1,6)

a) hallar las siguientes probabilidades p[X>1,3] , p[1<X<4]

b) Si p[2,3<X<k]=0,3212 ¿Cuánto vale k?

3.2.-Calcula k para que p[-k<Z<k] valga a) 0.9 ,b) 0.95 , c) 0.99

3.8.-VARIABLE ALEATORIA BINOMIAL

Consideremos el siguiente experimento aleatorio,llamado “Proceso de Bernouilli” consistente en realizar N pruebas (ó ensayos) de un experimento con las siguientes características:

a)En cada prueba solo se consideran dos resultados posibles: uno que llamaremos éxito y su contrario que llamaremos fracaso

b)Los N ensayos son independientes, es decir, que el resultado de un ensayo no influye en el resultado del siguiente. En consecuenciala probabilidad del éxito y del fracaso se mantienen constantes en todo el proceso. Llamaremos **p** y **q**=1-p a sus respectivas probabilidades.

Considerenos la variable aleatoria X = número de éxitos obtenidos en los N ensayos .Esta variable se llama **variable binomial** y los valores que puede tomar son : 0,1,2,……,N.
La probabilidad de obtener r éxitos , es decir, la probabilidad de que X=r viene dada por:

$$P[X=r]=\binom{N}{r}p^r q^{N-r}$$

Ejemplo

Lanzamos un dado 4 veces ¿Cuál es la probabilidad de que salga un 6 tres veces?. Consideramos un éxito obtener un 6 y un fracaso no sacar un 6.
Probabilidad del éxito: p= 1/6
Probabilidad del fracaso: q=1- (1/6)= 5/6
Variable binomial X= nº de éxitos

Nos piden la probabilidad de obtener 4 éxitos : $P[X=4]=\binom{4}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^1 \cong 0.015$

Aproximación de la binomial mediante la normal

Cuando N es un número grande resulta engorroso el cálculo de la expresión $P[X=r]=\binom{N}{r}p^r q^{N-r}$. Una propiedad importante de la distribución normal es que puede utilizarse para el cálculo de probabilidades de una binomial. En efecto si Np≥5 y Nq≥5 entonces la distribución binomial es de forma muy aproximada igual due una distribución $\mathbf{N}(Np;\sqrt{Npq})$.
Para utilizar esta aproximación se ha de tener en cuenta que el suceso X=r de la binomial se considera equivalente al suceso $r-\frac{1}{2}\le X\le r+\frac{1}{2}$ de la normal

Ejemplo

Supongamos que una máquina produce un 25% de piezas defectuosas. Si se toma una muestra de 100 piezas ¿Cuál es la probabilidad de que contenga 6 defectuosas? ¿Y de que contenga un máximo de 6?
Estamos ante una variable binomial con N=100 , p=0.25 y q=0.75 y las probabilidades buscadas son:

$$P[X=6]=\binom{100}{6}0.25^6\cdot 0.75^{94} \quad y \quad P[X\le 6]=P[X=0]+P[X=1]+.....+P[X=6]=......$$

Como Np=25 y Nq=75 resulta que la binomial es prácticamente igual a la normal $\mathbf{N}(Np;\sqrt{Npq})=(25;4.3)$ con lo que para hallar las anteriores probabilidades haremos

$P[X_B=6]=P[5.5\le X_N\le 6.5]$ (tipificar y usar las tablas)

$P[X_B\le 6]=P[X_N\le 6.5]$ (tipificar y usar las tablas)

EJERCICIOS

3.3.-Se lanzan un dado 6 veces ¿Cual es la probabilidad de obtener 4 veces un múltiplo de 3?

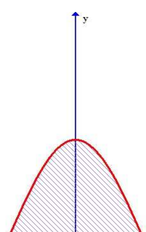
3.4.-Una pieza de artillería dispone de 10 obuses para alcanzar un objetivo En cada disparo la probabilidad de alcanzarlo es 3/4 .¿Cual es la probabilidad de fallar los 10?

3.5.-El 30% de las piezas fabricadas por una máquina son defectuosas.¿Cual es la probabilidad de que en una caja que coniene 100 piezas haya a lo sumo 6 defectuosas?

3.6.-Se extrae una carta de la baraja 12 veces devolviéndola de cada vez ¿Cuál es la probabilidad de que por lo menos la mitad de las veces salga una figura?

3.7.-Se lanza una moneda 10000 veces consecutivas ¿Cual es la probabilidad de que se obtenga cara entre 4980 y 5050 veces?

TABLA DE DISTRIBUCIÓN NORMAL TIPIFICADA N(0,1)



K=

	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
0,0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0,1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0,2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0,3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0,4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0,5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0,6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0,7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0,8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079	0.8106	0.8133
0,9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1,0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1,1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1,2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1,3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1,4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1,5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1,6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1,7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1,8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1,9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2,0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2,1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2,2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2,3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2,4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2,5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2,6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2,7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2,8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2,9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3,0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

TEMA 4: DISTRIBUCIÓN DE PROBABILIDAD DE LA MEDIA MUESTRAL TEOREMA CENTRAL DEL LÍMITE

4.1.-DISTRIBUCIÓN DE LAS MEDIAS MUESTRALES

Consideremos una variable aleatoria **X**. Cada muestra de tamaño **n** que podemos tomar de la población nos proporciona un valor para la media . Estos son los valores que puede tomar la variable aleatoria $\overline{X}$ (media muestral). Acerca de su distribución consideraremos dos resultados:

Si la variable **X** es normal con media **μ** y desviación típica **σ** entonces la variable media muestral $\overline{X}$ también es normal con la misma media **μ** y con desviación típica $\frac{\sigma}{\sqrt{n}}$.

Es decir: $X \in N(\mu;\sigma) \Rightarrow \overline{X} \in N\left(\mu;\frac{\sigma}{\sqrt{n}}\right)$

El otro resultado se conoce como teorema central del límite

4.2.-TEOREMA CENTRAL DEL LÍMITE

Si la variable **X** sigue una distribución de probabilidad cualquiera (normal o no) con media **μ** y desviación típica **σ** y se toman muestras de tamaño **n>30 ,** entonces la variable media muestral $\overline{X}$ es aproximadamente normal con la misma media **μ** y con desviación típica $\frac{\sigma}{\sqrt{n}}$.

Es decir : Si **X** tiene media **μ** y desviación típica **σ** y **n>30** entonces

$$\overline{X} \in N\left(\mu;\frac{\sigma}{\sqrt{n}}\right)$$

4.3.-DISTRIBUCIÓN DE LA PROPORCIÓN

Además de la media, existen otros parámetros de uso frecuente. Tal es el caso de la proporción de individuos de una población que presentan una determinada característica.
Consideremos una población en la que la proporción de individuos que presentan una determinada característica es **p (** y **q=1-p** la proporción de los que no la presentan) . En cada muestra de tamaño **n** habrá una proporción $\hat{p}$ de individuos con esa característica . Si n>30 (ó np y nq=n(1-p) son ambas ≥5) entonces $\hat{p}$ es normal de media **p** y desviación típica $\sqrt{\frac{pq}{n}}$

Es decir : Si n>30 (ó np y nq=n(1-p) son ≥5) entonces $\hat{p} \in N(p;\sqrt{\frac{pq}{n}})$

EJERCICIOS

4.1.-El peso X , en Kgs, del alumnado de un instituto sigue una distribución N(58;8) . Los alumnos están distribuidos en 20 aulas. Se toma una muestra eligiendo al azar un alumno de cada aula. Hallar la probabilidad de que la media de los pesos de esos alumnos esté comprendida entre 60 y 65 Kgs

4.2.- Las bolsas de azúcar envasadas por una cierta máquina tiene una media de 500 grs con una desviación típica de 35 grs. Las bolsas se empaquetan en cajas de 100 unidades. Hallar la probabilidad de que la media de los pesos de las bolsas de una caja sea menor que 495 grs.

4.3.-En un centro de enseñanza se ha determinado que el 20% de los alumnos no tienen hremanos. En una excursión participan 50 alumnos de ese centro ¿Cuál es la probabilidad de más del 30% de esos alumnos no tengan hermanos?

TEMA 5: ESTIMACIÓN DE PARÁMETROS

5.1.-ESTIMACIÓN PUNTUAL

Cuando desconocemos un parámetro de la población , tomamos una muestra de tamaño n y calculamos el estadístico correspondiente que será una estimación puntual del parámetro de la población.

Así la media muestral $\overline{\mathbf{X}}$ sirve para estimar la media poblacional **μ.** La desviación típica muestral **S** sirve para estimar la desviación típica poblacional **σ.**

Pero la estimación puntual sirve de poco si no conocemos el grado de aproximación entre el estadístico y el parámetro que estima.

Por este motivo se procede a la estimación por intervalos

5.2.-ESTIMACIÓN POR INTERVALOS

A partir de una muestra de tamaño n podemos estimar el valor de un parámetro de la población del siguiente modo:

1.-Dando un intervalo dentro del cual confiamos que esté el parámetro . Se llama intervalo de confianza

2.-Hallando la probabilidad de que tal cosa ocurra. A dicha probabilidad de le llama nivel de confianza

El nivel de confianza se designa por 1-**α** (a **α** se le llama nivel de significación)

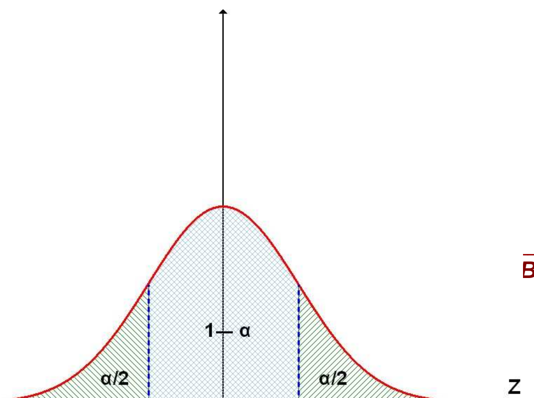
En este curso solo haremos estimaciones de la media poblacional y de la proporción. Conseguiremos mayor eficacia en la estimación aumentando el tamaño de la muestra. Esta mayor eficacia se manifiesta de dos formas:

a) Disminuyendo el intervalo de confianza (obviamente cuanto más pequeño sea el intervalo más precisos estamos siendo)
b) Aumentando el nivel de confianza (obviamente a mayor nivel de confianza más seguridad en la estimación)

5.3.-INTERVALO DE CONFIANZA PARA LA MEDIA

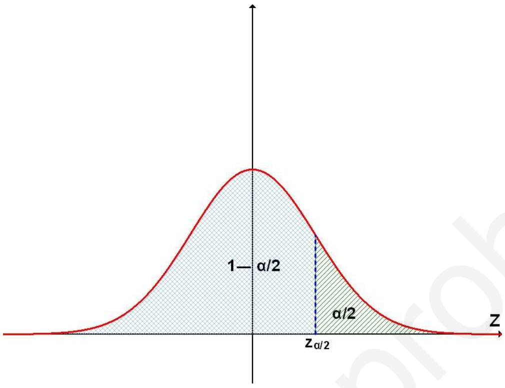
Se desea estimar la media **μ** de una población cuya desviación típica **σ** es conocida . Para ello se recurre a una muestra de tamaño **n** en la cual se obtiene una media muestral $\overline{x}$.

Fijamos un nivel de confianza 1-**α** (los niveles más usuales son 0.90=90% , 0.95=95% , 0.99=99%). Por $z_{\alpha/2}$ designamos el valor de la variable normal tipificada Z que cumple que $P[-z_{\alpha/2} \leq Z \leq z_{\alpha/2}] = 1-\alpha$ tal y como se indica en la siguiente figura:



EJERCICIOS

5.1.- Utiliza la siguiente figura



Para comprobar que los valores de $z_{\alpha/2}$ para los niveles de confianza más usuales son los siguientes:

$1-\alpha$	$z_{\alpha/2}$
0.90=90%	1.645
0.95=95%	1.960
0.99=99%	2.575

Si la población de partida es normal (o si el tamaño de la muestra es **n>30**), entonces

$$\overline{X} \in N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$$

y el intervalo de confianza de **μ** con un nivel de confianza del 1-**α** es

$$I = \left(\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\ ,\ \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

Este es el intervalo que contiene la media poblacional **μ** para el el (1-**α**)·100% de las muestras

El valor $E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ se llama **<u>error máximo admisible</u>**. Evidentemente depende de :

-El tamaño muestral **n**: de manera que cuanto mayor es el tamaño muestral , más pequeño es el error E y el intervalo de confianza es más pequeño (mejor estimación)

-El nivel de confianza 1-**α** : de manera que cuanto mayor sea el nivel de confianza , es decir cuanta mayor seguridad deseemos, mayor será $z_{\alpha/2}$ (y en consecuencia el error) por lo que el intervalo aumentará.

Observación:

Si la desviación típica σ de la población es desconocida , hay que estimarla a partir de la muestra. La forma más correcta de hacerlo es mediante el estadístico

$$s_{n-1} = \sqrt{\frac{n}{n-1}}\,s$$

(es decir, la raíz de la cuasivarianza) . Sin embargo para valores relativamente grandes de **n** puede utilizarse la desviación típica de la muestra **s.**

5.3.1.-Determinación del tamaño de la muestra

A veces el nivel de confianza y el error máximo admisible los fijamos de antemano y entonces tenemos que hallar el tamaño de la muestra para conseguir esas condiciones. Conseguiremos el valor mínimo del tamaño de la muestra despejando **n** en la expresión de E:

$$z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le E \Rightarrow \sqrt{n} \ge \frac{z_{\alpha/2}\cdot\sigma}{E} \Rightarrow n \ge \left(\frac{z_{\alpha/2}\cdot\sigma}{E}\right)^2$$

5.3.2-Determinación del nivel de confianza

A veces se desea hallar el intervalo de confianza para la media poblacional con un error E determinado y para ello se toma una muestra de tamaño **n.** En este caso interesa saber el nivel de confianza con el que se realiza la estimación. Se tiene que ::

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow z_{\alpha/2} = \frac{E\sqrt{n}}{\sigma}$$

Conocido $z_{\alpha/2}$, mediante las tablas obtenemos el valor de $\alpha/2$ que nos da el nivel de confianza $1-\alpha$

5.4.INTERVALO DE CONFIANZA PARA LA PROPORCION

Se desea estimar la proporción **p** de individuos que presentan una determinada caraterística. Para ello se recurre a una muestra de tamaño **n** de la cual se obtiene una proporción muestral $\hat{p}$ (siendo $\hat{q}$=1-$\hat{p}$)

Si n>30 (ó n$\hat{p}$≥5 y n(1-$\hat{p}$)≥5) entonces el intervalo de confianza de **p** con un nivel de confianza del 1-α es

$$I = \left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\ ,\ \overline{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$$

donde el error máximo admisible es en este caso $E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$

5.4.1.-Determinación del tamaño de la muestra

A veces el nivel de confianza y el error máximo admisible los fijamos de antemano y entonces tenemos que hallar el tamaño de la muestra para conseguir esas condiciones. Al Al igual que ocurria con la media conseguiremos el valor mínimo del tamaño de la muestra despejando **n** en la expresión de **E**

5.4.2-Determinación del nivel de confianza

A veces se desea hallar el intervalo de confianza para la proporción poblacional con un error **E** determinado y para ello se toma una muestra de tamaño **n.** En este caso interesa saber el nivel de confianza con el que se realiza la estimación. Para ello en la expresi´´on de **E** despejaremos $z_{\alpha/2}$ cuyo valor nos permite calcular 1-α

EJERCICIOS

5.2.-Un fabricante de pilas alcalinas sabe que el tiempo de duración, en horas, de las pilas que fabrica sigue una distribución Normal de media desconocida y varianza 3600. Con una muestra de su producción, elegida al azar, y un nivel de confianza del 95% ha obtenido para la media el intervalo de confianza (372.6; 392.2).
a)Calcular el valor que obtuvo para la media de la muestra y el tamaño muestral utilizado.

b) ¿Cuál sería el error de su estimación, si hubiese utilizado unl muestra de tamaño 225 y un nivel de confianza del 86,9% ?

5.3.-El peso de los paquetes enviados por una determinada empresa de transportes se distribuye según una ley Normal, con una desviación típica de 0,9 kg. En un estudio realizado con una muestra aleatoria de 9 paquetes, se obtuvieron los siguientes pesos en kilos;
9.5 10 8.5 10.5 12,5 10,5 12,5 13 12
a)Hallar un intervalo de confianza, al 99%, para el peso medio de los paquetes enviados por esa empresa.
b) Calcular el tamaño mínimo que debería tener una muestra, en el caso de admitir un error máximo de 0,3 kg, con un nivel de confianza del 90%.

5.4.-Las alturas, expresadas en centímetros, de los estudiantes de segundo de Bachillerato se distribuye normalmente con una desviación típica de 20 cm. En un colectivo de 500 estudiantes de segundo de Bachillerato se ha obtenido una media de 160 cm.
Calcula, con una probabilidad del 98%, entre qué valores estará la media de la altura de la población total de estudiantes de segundo de Bachillerato.

5.5.-Se quiere estimar la media de la nómina mensual que reciben los directivos de las compañías multinacionales que operan en Europa.
a) Si la varianza de la nómina en la población es de 1000 €^{2}. ¿Cuál es la varianza de la media muestral cuando el tamaño de la muestra es de 100?
b)Si en las condiciones del apartado anterior obtenemos un intervalo de confianza (4000,4016). ¿Qué nivel de confianza tiene la estimación?

5.6.-El peso medio de una muestra aleatoria de 100 naranjas de una determinada variedad variedad es de 272 grs. Se sabe que la desviación típica poblacional es de 20 grs. A un nivel de significación de 0.05 hallar el intervalo de confianza para la media poblacional

5.7.-En un pais se sabe que la altura de la población se distribuye normalmente con una desviación típica de 10 cms.
a)Si la media poblacional fuera de 170 cms, calcular la probabilidad que la media muestral de una muestra de 64 personas difiera menos de un centímetro de la media de la población.
b) ¿Cuál es el tamaño muestral que se debe tomar para estimar la media de la altura de la población con un error menor de 2 cms y con un nivel de confianza del 95%?
c) Y si en el apartado anterior aumentamos el nivel de confianza al 99 % ¿qué tamaño muestral se necesitará?

5.8.-Para determinar el porcentaje de alumnos que tienen acceso a Internet en su domicilio se ha encuestado a 60 alumnos resultando que 42 sí lo tienen. Determinar el intervalo de confianza para la proporción de alumnos con acceso a Internet en su domicilio a un nivel de confianza del 95%

5.9.-La estatura de los miembros de una población se distribuye según la ley normal de media desconocida y desviación típica 9 cm. Con el fin de estimar la media se toma una muestra de 9 individuos de la población, obteniéndose para ellos una media aritmética igual a 170 cm.
a) Calcula el intervalo de confianza al nivel del 95% para la estatura media de la población
b) Calcula el tamaño muestral necesario para estimar la media de la población con una precisión de ±5 cm y un nivel de confianza del 99%.

5.10.-El precio de ciertos electrodomésticos puede considerarse una variable aleatoria con distribución normal de desviación típica 100 euros Los precios en euros correspondientes a una muestra de 9 de estos electrodomésticos son
255 85 120 290 80 80 275 290 135
a) Construir un intervalo de confianza al 98% para la media población
b) Hallar el tamaño mínimo que debe tener la muestra, para que con el nivel de confianza del 99% el error de estimación del precio medio no supere los 50 euros.

5.11.-Un estudio realizado por el departamento de Lengua de un Instituto estima que la proporción de alumnos que durante un curso leen al menos 6 libros de lectura es del 0.45.
a)Por la dirección de la biblioteca escolar se decide realizar una encuesta para determinar dicho porcentaje con un error no superior a 0.05 a un nivel de confianza del 95% ¿Cuál será el tamaño mínimo de la muestra?
b)Si se decide tomar una muestra de tamaño 300 ¿cuál es el nivel de confianza de que el error no supera 0.05?

5.12.-Se quiere conocer la permanencia media de pacientes en un hospital, con el fin de estudiar una posible ampliación del mismo. Se tienen datos referidos a la estancia, expresada en días, de 800 pacientes, obteniéndose los siguientes resultados: $\overline{X} = 81$ días; .S = 9 días. Se pide obtener un intervalo de confianza del 95% para la estancia media.

5.13.-El salario medio correspondiente a una muestra de 900 personas de una población dada es de 725 euros. Se sabe que los salarios de esa población siguen una distribución normal con desviación típica de 84 euros. ¿Se puede afirmar que el salario medio de dicha población es de 700 euros con un nivel de confianza del 95%?

5.14.-Una máquina de refrescos está ajustada de manera que la cantidad de líquido despachada se distribuye en forma normal con una desviación típica de 0.15 dl. Encontrar un intervalo de confianza del 97% para la media de todos los refrescos si una muestra aleatoria de 36 refrescos tiene un contenido promedio de 2.25dl

5.15.-De una encuesta realizada a 1200 hogares de una ciudad se ha obtenido que la proporción de las que utilizan microondas es del 0.30.
a) Calcular el intervalo de confianza para la proporción de hogares que utilizan el microondas en dicha ciudad a un nivel de confianza del 92%
b) Para una estimación con un error menor que 0.04 a un nivel de confianza del 95% ¿de que tamanño deberíamos elegir la muestra?

TEMA 6: CONTRASTE DE HIPÓTESIS

6.1.-INTRODUCCIÓN. DEFINICIONES

En el tema anterior estudiamos como obtener, a partir de una muestra, un intervalo que con un nivel de confianza determinado contiene a un parámetro desconocido de la población.

En esta tema abordamos la teoría de la decisión estadística que consiste en utilizar la información obtenida de una muestra para tomar una decisión sobre la aceptación de una hipótesis relativa a la población.

Definiciones:

✧Se llama **hipótesis nula**, denotada por H_0 , la hipótesis sobre la población que queremos contrastar. Será aceptada o rechazada después de obtener información de la muestra.

✧La hipótesis contraria a la hipótesis nula se llama **hipótesis alternativa** y se denota por H_1. Obviamente si se acepta H_0 se rechaza H_1 y viceversa

✧Un contraste de hipótesis es un procedimiento para, a partir de una muestra significativa, decidir si una hipótesis se acepta o se rechaza. Los contrastes de hipótesis pueden ser de dos tipos:

- ✷ Contraste bilateral: Si la hipótesis nula adopta la forma de igualdad
- ✷ Contraste unilateral: Si la hipótesis nula adopta la forma de desigualdad

✧Estadístico de contraste es el estadístico muestral utilizado en el contraste

✧Nivel de significación es un valor probabilístico , muy pequeño, que se fija previamente a la aplicación del contraste. Se simboliza por α **.** Es la probabilidad de que la hipótesis nula se rechace siendo cierta

✧Se llama **región crítica** al conjunto de valores del estadístico de contraste que nos llevan a rechazar la hipótesis nula y **región de aceptación** el conjunto de valores que nos lleva a aceptarla.

6.2.-CONTRASTE DE HIPÓTESIS

Los 5 pasos a seguir en un contraste de hipótesis son los siguientes:

A.- Formular las hipótesis nula y alternativa

B.- Elegir un nivel de significación

C.- Determinación de la región crítica y de la región de aceptación

D.- A partir de la muestra (de tamaño mayor que 30 si la población no es normal) determinar el valor del estadístico de contraste

E.-Aceptar la hipótesis nula si el valor del estadístico de contraste pertenece a la región de aceptación o rechazarla si pertenece a la región crítica

6.3.-CONTRASTE DE HIPÓTESIS PARA LA MEDIA

6.3.1.-Contraste bilateral

Partimos de una población con una media μ desconocida y una desviación típica σ_0 conocida

A.- Hipótesis nula H_0 : $\mu=\mu_0$

Hipótesis alternativa : H_1 : $\mu\neq\mu_0$

B.- Elegimos un nivel de significación α

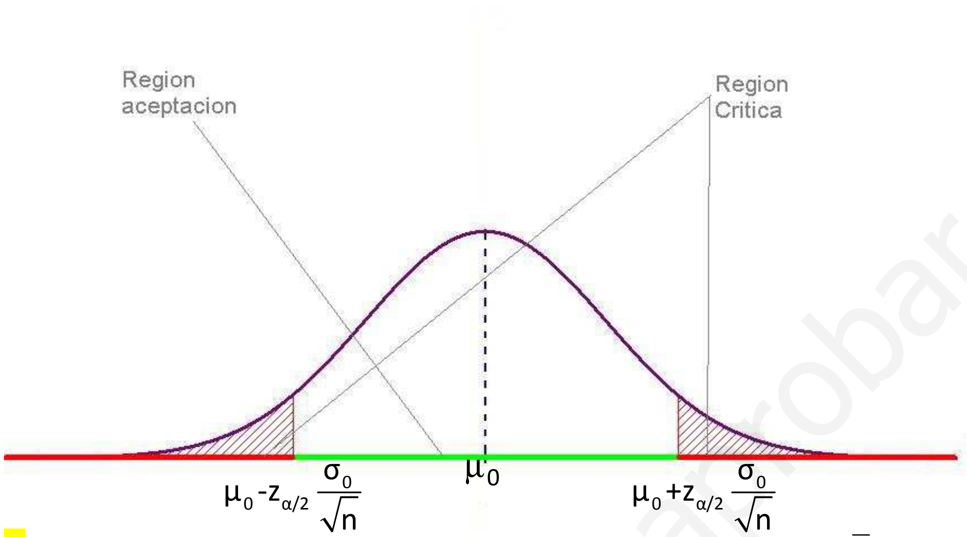
C.-Determinación de la región crítica y de la región de aceptación

Región de aceptación: $\left(\mu_0 - z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}, \mu_0 + z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right)$

$$\left(-\infty, \mu_0 - z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right) \cup \left(\mu_0 + z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}, +\infty\right)$$

Región crítica :

En el grafico siguiente representamos las dos regiones



D.-Determinamos el valor del estadístico de contraste (la media $\overline{X}$ de la muestra)

E.- Se acepta (si $\overline{X}$ pertenece a la región de aceptación) o rechaza la Hipótesis nula

Nota: Si la desviación típica no es conocida utilizaremos en su lugar la desviación típica **S** de la muestra

Ejemplo 1

Las pilas de una determinada marca tienen una duración de media desconocida con una desviación típica de 80 horas. Se toma una muestra de 100 pilas y se comprobó que su duración media era de 1670 horas. ¿Se puede afirmar , con un nivel de significación del 5% , que la duración media de las pilas de esa marca es de 1700 horas?

A.- Hipótesis nula : μ**=1700**

Hipóteis alternativa: μ**≠1700**

B.- Nivel de significación α**=0.05 .** Por lo tanto $z_{\alpha/2}$=1.96

C.- Región de aceptación:

$$\left(\mu_0 - z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}, \mu_0 + z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right) = (1700\text{-}1.96\cdot 80/10,\ 1700\text{-}1.96\cdot 80/10) = (1684.32, 1715{,}68)$$

D.-Media muestral : 1670

E.-Como la media muestral no pertenece a la región de aceptación se rechaza la hipótesis de que la duración media de las pilas de la marca es de 1700 horas

6.3.2.-Contraste unilateral

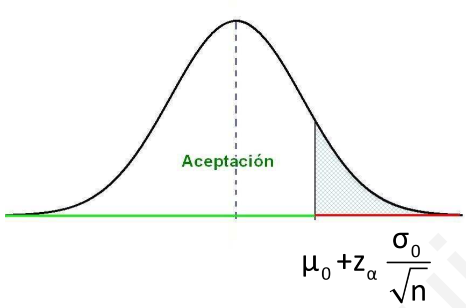
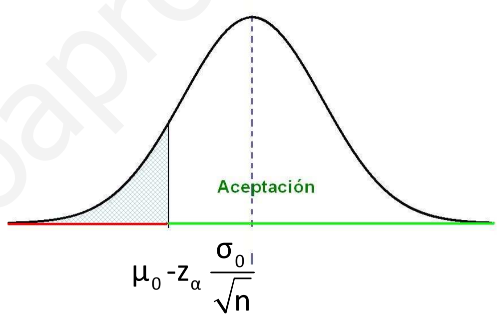
Partimos de una población con una media μ desconocida y una desviación típica σ_0 conocida

A.- Hipótesis nula H_0 : $\mu \le \mu_0$ (ó $\mu \ge \mu_0$)

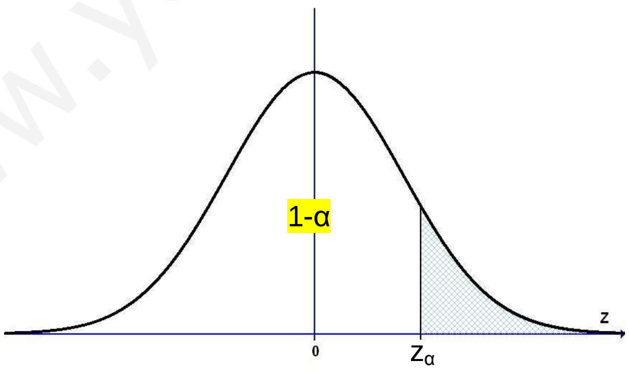
Hipótesis alternativa : H_1 : $\mu > \mu_0$ (ó $\mu < \mu_0$)

B.- Elegimos un nivel de significación α

C.-Determinación de la región crítica y de la región de aceptación

$H_0 : \mu \le \mu_0$	$H_0 : \mu \ge \mu_0$
Región de aceptación: $\left(-\infty, \mu_0 + z_\alpha \frac{\sigma_0}{\sqrt{n}}\right)$ Región crítica : $\left(\mu_0 + z_\alpha \frac{\sigma_0}{\sqrt{n}}, +\infty\right)$	Región de aceptación: $\left(\mu_0 - z_\alpha \frac{\sigma_0}{\sqrt{n}}, +\infty\right)$ Región crítica : $\left(-\infty, \mu_0 - z_\alpha \frac{\sigma_0}{\sqrt{n}}\right)$
	

donde z_α es el valor que verifica $P[Z< z_\alpha]=1-\alpha$ (ver la siguiente figura)



D.-Determinamos el valor del estadístico de contraste (la media $\overline{X}$ de la muestra)

E.- Se acepta o rechaza la Hipótesis nula

Nota: Si la desviación típica no es conocida utilizaremos en su lugar la desviación típica **S** de la muestra

Ejemplo 2

Para la realización de una prueba se considera que el tiempo apropiado es de 45 minutos. Se mide el tiempo empleado por 50 personas elegidas al azar obteniéndose un tiempo medio de 45.2 minutos con una desviación típica de de 4 minutos. ¿Se puede considerar , con un nivel de significación del 10% , que los 45 minutos es un tiempo suficiente para realizar la prueba?

A.- Hipótesis nula : $\mu \leq 45$

Hipóteis alternativa: $\mu > 45$

B.- Nivel de significación $\alpha = 0.10$. Por lo tanto $z_{\alpha} = 1.28$

C.- Región de aceptación:

$$\left(-\infty, \mu_0 + z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}\right) = (-\infty\ ,\ 45+1.28 \cdot 4/7.07) = (-\infty, 45.72)$$

D.-Media muestral : 45.2

E.-Como la media muestral pertenece a la región de aceptación se acepta la hipótesis de que 45 minutos es un tiempo adecuado para la realización de la prueba

---

6.4.-CONTRASTE DE HIPÓTESIS PARA LA DIFERENCIA DE MEDIAS

6.4.1.-Contraste bilateral

Partimos de dos poblaciones : una con media μ_1 desconocida y una desviación típica σ_1 conocida ; otra media μ_2 desconocida y una desviación típica σ_2 conocida. De la primera extraemos una muestra de tamaño n_1 con una media $\overline{X}_1$ y de la segunda otra de tamaño n_2 con una media $\overline{X}_2$. Se trata de contrastar si las dos poblaciones tienen la misma media.

A.- Hipótesis nula H_0 : $\mu = \mu_1 - \mu_2 = 0$ (es decir $\mu_1 = \mu_2$)

Hipótesis alternativa : H_1 : $\mu = \mu_1 - \mu_2 \neq 0$ (es decir $\mu_1 \neq \mu_2$)

B.- Elegimos un nivel de significación α

C.-Determinación de la región de aceptación: $\left(0 - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\ ,\ 0 + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$

D.-Determinamos el valor del estadístico de contraste (la diferencia de medias $\overline{X} = \overline{X}_1 - \overline{X}_2$)

E.- Se acepta o rechaza la Hipótesis nula

6.4.2.-Contraste unilateral

A.- Hipótesis nula H_0 : $\mu = \mu_1 - \mu_2 \leq 0$ (es decir $\mu_1 \leq \mu_2$) ; ó $\mu = \mu_1 - \mu_2 \geq 0$ (es decir $\mu_1 \geq \mu_2$)

Hipótesis alternativa : H_1 : $\mu = \mu_1 - \mu_2 > 0$ (es decir $\mu_1 > \mu_2$) ; ó $\mu = \mu_1 - \mu_2 < 0$ (es decir $\mu_1 < \mu_2$)

B.- Elegimos un nivel de significación α

C.-Determinación de la región de aceptación:

$$\left(-\infty, 0 + z_{\alpha}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \quad ; \quad ó \quad \left(0 - z_{\alpha}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\ ,+\infty\right)$$

D.-Determinamos el valor del estadístico de contraste (la diferencia de medias $\bar{X}=\bar{X}_1-\bar{X}_2$)

E.- Se acepta o rechaza la Hipótesis nula

Ejemplo 3

Para comprobar que entre el alumnado de un centro escolar las alumnas dedican más tiempo a la lectura que los alumnos, el departamento de matemáticas realiza una encuesta a una muestra de 50 alumnas y una muestra de 40 alumnos. Para las alumnas se obtiene que por término medio se dedican 10.1 horas semanales con una desviación de 1.4 horas y para los alumnos 10.3 horas con una desviación de 1.8 horas. ¿Se puede considerar , con un nivel de significación del 5% , que las alumnas dedican más tiempo a la lectura que los alumnos del centro?

A.- Hipótesis nula : $\mu=\mu_1-\mu_2>0$ (es decir $\mu_1>\mu_2$)

Hipóteis alternativa: $\mu=\mu_1-\mu_2\leq 0$

B.- Nivel de significación $\alpha=0.05$ **.** Por lo tanto $z_\alpha=1.645$

C.- Región de aceptación:

$$\left(0-z_\alpha\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\ ,+\infty\right)=\left(0-1.645\sqrt{\frac{1.4^2}{50}+\frac{1.8^2}{40}}\ ,+\infty\right)=(-0.57,+\infty)$$

D.-Diferencia de medias: $\bar{X}=\bar{X}_1-\bar{X}_2=10.1-10.3=-0.2$

E.-Como el valor de $\bar{X}$ pertenece a la región de aceptación se acepta la hipótesis de que las alumnas dedican más tiempo a la lectura que los alumnos.

Nota: Lo que realmente concluimos e que no se puede rechazar la hipótesis debido a que estamos aceptando como improbables solo el 5% de los resultados posibles. (Se puede comprobar que si aumentamos el nivel de significación al 20% , es decir aceptaríamos como improbables el 20% de los resultados, la hipótesis ya no resulta aceptable)

6.5.-CONTRASTE DE HIPÓTESIS PARA LA PROPORCIÓN

6.5.1.-Contraste bilateral

Partimos de una población con una proporción desconocida **p** de individuos que poseen una cierta característica (y una proporción **q=1-p** que no la tienen)

A.- Hipótesis nula H_0 : $p=p_0$

Hipótesis alternativa : H_1 : $p\neq p_0$

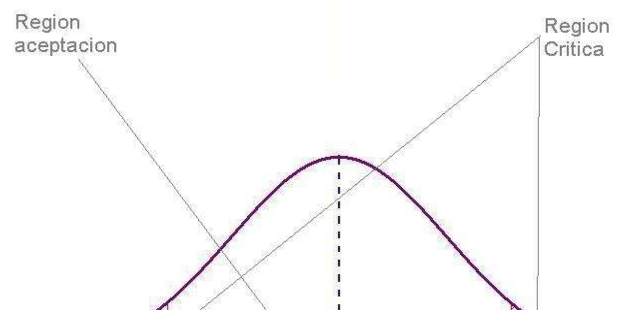
B.- Elegimos un nivel de significación α

C.-Determinación de la región crítica y de la región de aceptación

Región de aceptación: $\left(p_0-z_{\alpha/2}\sqrt{\frac{p_0q_0}{n}}\ ,\ p_0+z_{\alpha/2}\sqrt{\frac{p_0q_0}{n}}\right)$

Región crítica : $\left(-\infty,\ p_0-z_{\alpha/2}\sqrt{\frac{p_0q_0}{n}}\right)\cup\left(p_0+z_{\alpha/2}\sqrt{\frac{p_0q_0}{n}}\ ,+\infty\right)$

En el grafico siguiente representamos las dos regiones



$p_0 - z_{\alpha/2}\sqrt{\frac{p_0 q_0}{n}}$ $\quad p_0 \quad$ $p_0 + z_{\alpha/2}\sqrt{\frac{p_0 q_0}{n}}$

D.-Determinamos el valor del estadístico de contraste (la proporción $\hat{p}$ de la muestra)

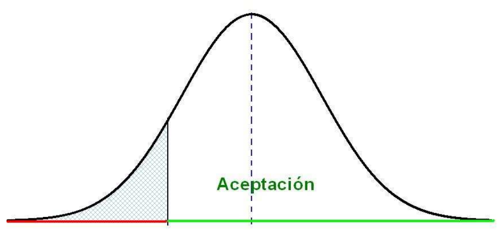
E.- Se acepta o rechaza la Hipótesis nula

6.5.2.-Contraste unilateral

A.- Hipótesis nula H_0 : $\mathbf{p \leq p_0}$ (ó $\mathbf{p \geq p_0}$)
Hipótesis alternativa : H_1 : $\mathbf{p > p_0}$ (ó $\mathbf{p < p_0}$)

B.- Elegimos un nivel de significación α

C.-Determinación de la región crítica y de la región de aceptación

H_0 : $\mathbf{p \leq p_0}$	H_0 : $\mathbf{p \geq p_0}$
Región de aceptación $\left(-\infty, p_0 + z_\alpha\sqrt{\frac{p_0 q_0}{n}}\right)$ Región crítica: $\left(p_0 + z_\alpha\sqrt{\frac{p_0 q_0}{n}}, +\infty\right)$	Región de aceptación: $\left(p_0 - z_\alpha\sqrt{\frac{p_0 q_0}{n}}, +\infty\right)$ Región crítica $\left(-\infty, p_0 - z_\alpha\sqrt{\frac{p_0 q_0}{n}}\right)$
 $p_0 + z_\alpha\sqrt{\frac{p_0 q_0}{n}}$	 $p_0 - z_\alpha\sqrt{\frac{p_0 q_0}{n}}$

D.-Determinamos el valor del estadístico de contraste (la proporción $\hat{p}$ de la muestra)

E.- Se acepta o rechaza la Hipótesis nula

6.6.-ERRORES EN LOS CONTRASTES DE HIPÓTESIS

Al aplicar un contraste de hipótesis podemos cometer dos tipos de error:

Error tipo I : Se comete cuando se rechaza la hipóteis nula siendo realmente verdadera. La probabilidad de cometer un error de este tipo es **α**

Error tipo II : Se comete cuando se acepta la hipótesis nula siendo realmente falsa. La probabilidad de cometer este tipo de error depende del tamaño de la muestra. Cuanto mayor es dicho tamaño más pequeña es la probabilidad de cometer este tipo de error

EJERCICIOS

6.1.- En los años 60, la estatura media de los españoles varones que hacían el servicio militar era de 170 cm,. En la actualidad, se ha realizado un muestreo a 36 adultos varones dando una media de 172 cm. Con una desviación típica S=2 cm
Se pide:
a) ¿Podemos afirmar, con una confianza del 95%, que esa diferencia es debida al azar?
b) ¿Qué se puede decir si esa media se ha calculado utilizando una muestra de 900 jóvenes?

6.2.- En una comunidad autónoma se estudia el número medio de hijos por mujer a partir de los datos disponibles en cada municipio. Se supone que este número sigue una distribución normal con desviación típica igual a 0,08. El valor medio de estos datos para 36 municipios resulta ser igual a 1,17 hijos por mujer.
Se desea contrastar, con un nivel de significación de 0,01, si el número medio de hijos por mujer en la comunidad es de 1,25.

6.3.- Según los datos de un censo de 1970, el analfabetismo en cierto país alcanzaba el 40% de la población. Una reciente encuesta realizada sobre una muestra aleatoria de 800 personas, arroja el dato de que 300 de ellas son analfabetas.
¿Puede considerarse, con un nivel de confianza del 95%, que se ha reducido el nivel de analfabetismo?

6.4.- Según un estudio realizado por una empresa hotelera durante el año 1992, la distribución del tiempo de estancia de cada viajero fue normal con una media de 3,7 días y una desviación típica de 1,1 días. A lo largo del año 2000 se analizó el tiempo de estancia de 49 viajeros elegidos al azar, obteniéndose una media de 3,5 días. ¿Podemos afirmar que esta diferencia es debida al azar con una confianza del 88%? Con el mismo nivel de confianza, ¿cambiaría la respuesta si esta media de 3,5 días se hubiera obtenido al analizar el tiempo de estancia de 100 viajeros elegidos al azar?

6.5.- Se afirma que la proporción de personas que contratan un determinado servicio telefónico es, como mínimo, del 23%. Sin embargo, la compañía telefónica sospecha que actualmente dicha proporción ha variado. Para comprobarlo hace una encuesta a 500 clientes potenciales entre los que solo 98 piensan contratar dicho servicio.
Con un nivel de significación del 5%, determinar si es aceptable la afirmación inicial.

6.6.- A partir de los datos recogidos sobre una muestra aleatoria de 121 pequeñas y medianas empresas de una región se ha calculado, para el año 2000, un beneficio medio de 89 millones de euros con una varianza de 30,25 euros^2.
Contestar justificando las respuestas:
a) ¿Podríamos rechazar (con un nivel de significación del 0,001) la afirmación de que los beneficios medios en la pequeña y mediana empresa de dicha región son de 90 millones de euros?

6.7.- En las últimas elecciones, celebradas hace un año, el 52 por ciento de los votantes de una ciudad estaban a favor del alcalde. Una encuesta, realizada recientemente, indica que, de 350 ciudadanos elegidos al azar, 198 están a favor del alcalde:
a) ¿Se puede afirmar, con un nivel de confianza del 90%, que el alcalde gana popularidad?
b) ¿Se obtiene la misma respuesta que en el apartado anterior si el nivel de confianza es igual a 0,99?

6.8.- Una fábrica de ordenadores produce dos modelos diferentes: uno portátil y otro de sobremesa. Tras unos estudios realizados, se concluye que el número de errores del sistema del modelo portátil se distribuye normalmente con varianza 5; y que en el modelo de sobremesa, el número de errores sigue una distribución normal con varianza 7.
Elegidos al azar 50 ordenadores de cada moodelo, se observa un número medio de errores de 15 y 10, respectivamente.
¿Podemos afirmar, a un nivel de confianza del 95%, que el número medio de errores, en ambos modelos, es el mismo?

b) ¿Qué podemos decir respecto al 1% de nivel de significación?

6.9.- Se sabe que la edad (en años) de los aspirantes a un puesto de trabajo en un determinado organismo oficial es una variable normal con desviación típica igual a 5. Se observa una muestra de 125 personas que se presentan a una prueba para optar a un puesto de trabajo en el citado organismo, obteniéndose una edad media igual a 22,3 años:
a) ¿Se puede afirmar, con un nivel de significación del 5%, que es igual a 21 la edad media de los que optan a un puesto de trabajo en el organismo oficial?
b) ¿Se puede afirmar, si el nivel de significación es del 1 %, que dicha edad media es menor o igual que 22?

6.10.- Un establecimiento vende paquetes de carbón para barbacoa de peso teórico 10 kg. Se supone que el peso de los paquetes sigue una distribución normal con desviación típica 1 kg. Para contrastar la citada hipótesis, frente a que el peso teórico sea distinto de 10 kg, se escogen al azar 4 paquetes que pesan en kilogramos, respectivamente: 8, 10, 9, 8
Se desea que la probabilidad de aceptar la hipótesis nula, cuando esta es cierta, sea 0,95. Se pide:
a) La región crítica del contraste.
b) ¿Se debe rechazar la hipótesis nula?

6.11.- Se supone que el peso de las sandías de cierta variedad sigue una distribución normal con desviación tipíca de 1 kg. Se toma una muestra aleatoria de 100 sandías y se observa que el peso medio es de 6 kg.
a) Calcúlese un intervalo de confianza al 95% para el peso medio de esa variedad de sandía.
b) ¿Puede aceptarse la hipótesis de que el verdadero peso medio de las sandías es de 5 kg, frente a que sea diferente, con un nivel de significación de 0,05?

6.12.- En un estudio diseñado para demostrar si existe una diferencia entre las alturas medias de mujeres adultas nacidas en dos países diferentes, las muestras arrojan los siguientes resultados:
n = 120; X= 162,7; (Sx = 2,50)
m =150; Y= 61, 8; (Sy = 2, 62)
Utilizar un nivel de significación de 0,05 para contrastar la hipótesis nula de que las medias de la población correspondientes son iguales contra la hipótesis alternativa de que no son iguales.

6.13.- En el año 1990 el 25% de los partos fueron de madres de más de 30 años. Este año se ha tomado una muestra de 120 partos de los cuales 34 fueron de madres de más de 30 años.
a) Con una significación del 10%, ¿se puede aceptar que la proporción de partos de madres de más de 30 años sigue siendo como mucho del 25%, frente a que ha aumentado?

6.14.- Una empresa de automóviles está estudiando las mejoras que ha incluido en la nueva generación de su gama de utilitarios. Hasta ahora, los kilómetros que uno de estos automóviles podía recorrer -con un uso normal- sin que fueran necesarias reparaciones importantes seguía una Normal con media 220 (en miles de kilómetros) y desviación típica 15 (en miles de kilómetros). Las mejoras parecen haber surtido efecto, puesto que con 100 automóviles de la nueva generación se ha obtenido una media de 225 (en miles de kilómetros) sin ningún tipo de problema grave. Suponiendo que la desviación típica se ha mantenido:
a) Plantea un test para contrastar la hipótesis de que las mejoras no han surtido efecto o incluso que han empeorado la situación, frente a que sí han surtido efecto, como parecen indicar los datos. Si se concluyera que la media sigue igual o incluso bajó, y sin embargo esta conclusión fuera falsa, ¿cómo se llama el error cometido?
b) Con un nivel de significación del 1 %, ¿a qué conclusión se llega?

6.15.- Se quiere estimar la media de la nómina mensual que reciben los directivos de las compañías multinacionales que operan en Europa.
a) Si la varianza de la nómina en la población es de 1000 €^{2}. ¿Cuál es la varianza de la media muestral cuando el tamaño de la muestra es de 100?
b) Si en las condiciones del apartado anterior, la media muestral es de 4008 €. ¿Se rechazaría, con un nivel de confianza del 0,95, la hipótesis de que la nómina media es de 4000 €?

6.16.- A principios de año, un estudio en cierta ciudad indicaba que un 15% de los conductores utilizaban el móvil con el vehículo en marcha. Con el fin de investigar la efectividad de las campañas que se han realizado desde en- tonces para reducir estos hábitos, recientemente se ha hecho una encuesta a 120 conductores y 12 hacían un uso indebido del móvil.
a) Plantea un test para contrastar que las campañas no han cumplido su objetivo frente a que sí lo han hecho, como parecen indicar los datos. ¿A qué conclusión se llega con un nivel de significación de14%?

6.17.-Se quiere contrastar si el nivel de colesterol en sangre de un grupo de enfermos es mayor que el de una población que se ha tomado como referencia, y que es de 160 u. .Se sabe que la desviación típica de la cantidad de colesterol en sangre es de 20 u.

Para ello se toma una muestra de 50 enfermos resultando una media muestral de 165 u. ¿Se rechazará la hipótesis con un nivel de significación de 0.025?

6.18.- Se desea determinar el número de familias que les gustaría disponer de televisión digital en sus domicilios. Para ello se considera una muestra de 400 familias y se observa el número de los que les gustaría, obteniendo un resultado de 188. En estas condiciones:
¿Se puede afirmar que la proporción de personas que les gustaría disponer de televisión digital es 0,5 a un nivel de significación de 0,02?
¿Se puede afirmar que la proporción de personas que les gustaría disponer de televisión digital es menor que 0,5 a un nivel de significación de 0,02?

6.19.-Una empresa dedicada a la fabricación de luminosos publicitarios anuncia que , como máximo, hay un 1% de luminosos defectuosos. Se selecciona una muestra de 100 rótulos y se observa que aparecen 3 defectuosos. Con un nivel de significación del 5% ¿Podemos aceptar la hipótesis del fabricante?¿Y con un nivel de confianza del 99%?

6.20.- Antes de la puesta en marcha del carné por puntos, la velocidad en cierta carretera seguía una Normal de media 80 kilómetros por hora y desviación típica 10. Pasado unos meses de la introducción de dicha medida, sobre 40 vehículos observados a diferentes horas del día se obtuvo una media de 75 kilómetros por hora. Si la velocidad sigue siendo una Normal con la misma desviación típica:
Plantea un test para contrastar la hipótesis de que con dicha medida la situación sigue igual, frente a que, como parece, ha mejorado. ¿A qué conclusión se llega para un nivel de significación de15%?

6.21.- Según los datos de cierta comunidad autónoma relativos al impuesto sobre la renta, en el pasado ejercicio fiscal la contribución media fue de 4000 euros. En una muestra de 500 declaraciones del año en curso, elegidas al azar, la contribución media ha sido de 4120 euros, con una desviación típica de 1200 euros. ¿Puede decirse, con un 95% de confianza, que ha variado la aportación media de los contribuyentes? ¿Y con un 99% de confianza?

6.22.- Un banco decide contratar guardias de seguridad para sus nuevas sucursales. Para ello contacta con dos empresas de seguridad; para estar seguros de elegir la opción correcta deciden aplicarles un test a los aspirantes de ambas empresas. Se sabe que las desviaciones típicas de los guardias de la primera empresa es 8 y la desviación típica de los de la segunda 7. De la primera empresa se seleccionaron 60 aspirantes y tras aplicarles el test obtuvieron una puntuación media de 70 puntos; de la segunda empresa se seleccionaron 50 aspirantes y la puntuación media fue de 73 puntos. A la vista de estos resultados
a) ¿podemos afirmar que existe alguna diferencia entre las puntuaciones medias de los aspirantes de las dos empresas de seguridad para un nivel de significación de 0,05?
b) ¿podemos afirmar que la puntuación media de la 1ª empresa es peor que la de la segunda para un nivel de significación del 0,10?