**Muestreo – Inferencia estadística**

**Muestreo – Inferencia estadística**

**1.- POBLACIÓN Y MUESTRA.**

Cuando una investigación va referida a un conjunto colección o colectivo de elementos, estadísticamente este colectivo se llama **población**.

**Tamaño** de la población es el número de elementos o unidades estadísticas que la componen. La población, por su tamaño, puede ser finita o infinita.

Cuando la población es muy grande no se suele hacer una observación exhaustiva, se estudia una parte de la misma llamada **muestra**, siendo imprescindible, para poder obtener conclusiones acerca de la población, que la muestra sea representativa.

***Muestra*** *es una parte de la población, debidamente elegida, que se somete a la observación científica en representación de la misma, con el propósito de obtener resultados válidos para el total de la población.*

Para que una muestra se considere válida debe cumplir que:

* Su tamaño sea proporcionado al tamaño de la población.
* No hay distorsión en la elección de los elementos de la muestra.
* Sea representativa.

Un estudio exhaustivo, cuyos datos se utilizan para multitud de trabajos e investigaciones es el Censo de Población, que se realiza en España cada diez años. Requiere un gran esfuerzo tanto económico como de medios y en él se recaba información de todos los habitantes del país. Sin embargo, para el conocimiento de algunas características de la población, se utilizan métodos alternativos que reducen el costo y el tiempo.

Los modelos reducidos de la población, constituidos por muestras, tienen como finalidad obtener resultados que puedan ser aplicables (extrapolables) a la población.

Los principales motivos que inducen a tomar muestras son:

* El coste económico y de tiempo.
* Que la población sea homogénea, pudiendo obtener buenos resultados a partir de cualquier muestra.
* Por la falta de personal preparado para llevar a cabo un buen estudio general.
* Por la necesidad de obtener unos datos de forma rápida.

El uso del muestreo presenta limitaciones, entre ellas:

* + El riesgo que supone la toma de una muestra que puede ser **no** representativa.
	+ Cuando es necesaria información de todos los elementos de la población.
	+ Cuando no se domina bien la técnica del muestreo.
	+ Cuando la población está formada por un número muy pequeño de elementos, ya que una ligera equivocación en la toma de la muestra puede originar grandes errores.

**2.- MUESTREO ELEATORIO SIMPLE Y SISTEMÁTICO.**

**Muestreo aleatorio simple.**

¿Cómo podemos conocer la estatura media de 240 alumnos de 1º de Bachillerato?

Podríamos utilizar una calculadora o un ordenador, y en casos en los que la muestra es grande, se puede obtener mediante un muestreo.

*El* ***muestreo*** *consiste en la elección de unidades estadísticas dentro de un conjunto que constituye la población.*

La muestra se puede elegir por distintos procedimientos. El principio que debe presidir la elección de una muestra es el principio aleatorio, es decir, que todos y cada uno de los elementos de la población tengan la misma probabilidad de ser elegidos y formar parte de la muestra. Se puede llevar a cabo mediante un sorteo riguroso y sacar una serie de unidades estadísticas (con o sin reemplazamiento) hasta completar la muestra fijada.

*El* ***muestreo aleatorio simple*** *consiste en seleccionar n elementos sin reemplazamiento de entre los N que componen la población, de tal modo que todas las muestras de tamaño n que se pueden formar, CN,n , tengan la misma probabilidad de ser elegidas.*

La probabilidad de elegir una muestra es: 

La probabilidad de que un elemento determinado de la población forme parte de la muestra viene dada por:

, ya que una vez fijado un elemento, con los *N* – 1 restantes se pueden hacer  combinaciones, que son los casos favorables.

En la práctica, la muestra se obtiene unidad a unidad. Para ello, se enumeran los elementos de la población desde 1 hasta *N* y se extraen a continuación *n* elementos al azar o bien se introducen en un bombo tantas bolas numeradas como elementos de la población; a continuación, se extraen las *n* bolas al azar que formarán la muestra.

Este procedimiento, aunque simple, requiere tener unos medios materiales: bombos, bolas suficientes, etc., por lo que a veces se utilizan en su lugar atrás alternativas como las tablas de números aleatorios, formadas por grupos de dígitos obtenidos al azar y ordenados por filas y columnas.

En la tabla se empieza por cualquier número y se continua hacia arriba, hacia abajo, a la derecha, a la izquierda o en diagonal.

Así, si queremos seleccionar mediante las tablas de números aleatorios una muestra del listado numerado de alumnos de 1º de Bachillerato, elegimos al azar una de las diez columnas de números, por ejemplo la segunda, y como se trata de números entre 1 y 240 tomamos, por ejemplo, los tres primeros dígitos de los números de dicha columna, es decir: 362, 236, 012, …

**tabla de números aleatorios**

1324 3625 2188

2356 2369 2514

4589 9987 4788

2568 3652 4487

4478 1888 6987

9584 5454 6625

0236 0321 4569

1200 2012 2258

2569 3658 0633

2658 0021 0121

5554 6472 3625

6965 6912 0021

2020 3021 4785

0001 9090 3333

1698 0125 4478

1258 2016 6999

6987 1144 4114

0125 3641 2314

0154 7899 9658

6987 1111 7488

2365 2369 8554

7845 1467 2060

1987 2257 9825

7895 0639 3145

9887 1158 5566

 De aquí elegimos los menores de 240, sin que ningún número aparezca dos veces. Tenemos una muestra ordenada de 12 alumnos, formada por:

2, 12, 32, 63, 111, 114, 115, 146, 188, 201, 225 y 236.

**Muestreo aleatorio sistemático.**

Se empieza numerando todos los elementos de la población desde 1 hasta *N*. Para seleccionar los *n* elementos que constituyen la muestra es preciso obtener el **coeficiente** **de elevación** **. Después se elige al azar un número *i*, llamado **origen** comprendido entre 1 y *h* , que nos indica el punto de arranque de la selección.

La muestra está formada por los elementos: *i, i + h, i +* 2*h, ..., i +* (*n –* 1) *h*.

Este procedimiento exige, para que se pueda aplicar correctamente, que la población no presente ninguna ordenación por la variable objeto de estudio (sexo, estatura, peso, etc.) y, si la hay, previamente habrá que desordenarla.

Para elegir los 12 alumnos de 1º de Bachillerato a través del muestreo aleatorio sistemático obtenemos el coeficiente de elevación *h*: **

Elegimos en la tabla de números aleatorios un número al azar comprendido entre 1 y 20. Para ello seleccionamos una columna de números, por ejemplo la tercera, y en ella, por ejemplo los dos últimos dígitos, hasta encontrar un número *i* entre 1 y 20, que resulta ser el 14.

Los alumnos que componen la muestra serán: *i, i + h, i +* 2*h, ..., i +* 11 *h*, es decir, los alumnos que ocupan los lugares:

14, 34, 54, 74, 94, 114, 134, 154, 174, 194, 214, 224.

**3.- MUESTREO ELEATORIO ESTRATIFICADO.**

Supongamos que el listado de los 240 alumnos ha sido obtenido de 3 institutos, I1, I2, I3 proporcionando cada uno 72, 48 y 120 alumnos respectivamente. A su vez se encuentran distribuidos en 10 grupos.

Queremos elegir una muestra de 12 alumnos teniendo en cuenta el número de alumnos de cada instituto.

En este caso, la población de *N* elementos se ha dividido en tres subpoblaciones de elementos *N*1*, N*2*,* y *N*3.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | I1 | I2 | I3 | Total |
| Subpoblación | *N*1 | *N*2 | *N*3 | *N* |
| Muestra | *n*1 | *n*2 | *n*3 | *n* |

*Se llama muestreo aleatorio estratificado al procedimiento completo de seleccionar, en cada estrato o subpoblación, la muestra por muestreo aleatorio. La elección, en cada subpoblación, la podemos hacer bien por muestreo aleatorio simple o bien sistemático.*

El muestreo aleatorio estratificado suele llevarse a cabo en poblaciones no homogéneas, obteniéndose así una mayor precisión, y menor error, al tiempo que posibilita la utilización simultánea, en una misma muestra, de distintos tipos de muestreo.

La muestra total está formada por la suma de las muestras correspondientes a cada estrato.

● Cuando todas las muestras tienen el mismo tamaño en cada estrato, se dice que es un ***muestreo aleatorio estratificado constante* (*o de afijación igual*).**

Así, si hemos dividido la población en *L* subpoblaciones y la muestra es de tamaño *n*, en cada estrato tomamos el mismo número de unidades estadísticas o elementos:

*n*1 = *n*2 =… = *nL* = *n /L*

Así, en cada instituto o estrato tomaríamos: *n*1 = *n*2 = *n*3 = 12/3 = 4 alumnos, que resultaría poco representativo por la desigualdad en el número de alumnos de cada instituto.

● Cuando la selección de un número de elementos en cada estrato es proporcional a su tamaño, se trata de un ***muestreo aleatorio estratificado proporcional* (*o de afijación proporcional*).**

****

La muestra *n*i en cada estrato se toma de forma proporcional a su tamaño *N*i.

En nuestro caso:

 ****

Luego:



Ahora, en cada subpoblación se eligen los alumnos bien por muestreo estratificado simple o sistemático.

**Ej.** Un centro de Bachillerato tiene las especialidades de Tecnología, Ciencias Naturales y Ciencias Sociales con 200, 150 y 250 alumnos respectivamente. Si se quiere hacer una encuesta a 120 alumnos, ¿cuántos elegiríamos de cada especialidad en un muestreo aleatorio estratificado?

1. Constante.
2. Proporcional.

Sol:

1. *n*1 = *n*2 = *n*3 = 120/3 = 40 alumnos por especialidad.
2. ****



**4.- DISTRIBUCIÓN MUESTRAL DE LAS MEDIAS.**

Supongamos que deseamos determinar el rendimiento medio de los estudiantes de segundo de Bachillerato. Como el número de estudiantes es muy elevado, podemos considerar una muestra y obtener su rendimiento medio .

Si elegimos otra muestra distinta, pero del mismo tamaño que la anterior, y hallamos su rendimiento medio, posiblemente no obtengamos el mismo resultado. Y así para cada muestra elegida.

Consideremos una población de tamaño *N* de la que tomamos muestras de tamaño *n*. En estas condiciones, teniendo en cuenta la teoría combinatoria, podríamos elegir muestras distintas (X1, X2, ...) y, para cada una de ellas, obtendríamos medias .

Si llamamos  a la variable tal que asigna a cada muestra su media muestral, es decir:

**

se puede demostrar que:

* **La media de la variable aleatoria** , , es igual a la media de la población:



En el caso de los rendimientos de los estudiantes, podemos afirmar que la media de los rendimientos medios de cada una de las muestras es el rendimiento medio de todos los alumnos de segundo de Bachillerato.

* **La desviación típica de una variable aleatoria** , , es igual al cociente entre la desviación típica de la población y .



De aquí se deduce que los rendimientos medios obtenidos estarán más dispersos si el tamaño de la muestra es pequeño y serán más uniformes cuando el tamaño de la muestra sea grande.

* La distribución de la variable aleatoria , se llamada **distribución muestral de medias,** sigue una distribución normal N (,) cuya representación gráfica es:



Este hecho es cierto siempre que consideremos un muestreo aleatorio simple y que la población considerada siga también una distribución normal. Si la población no sigue una distribución normal, pero *n* > 30, la distribución muestral de medias se aproxima a una normal, y esta aproximación será mejor cuanto mayor sea *n*.

**Ej.** Consideremos una población formada por cuatro estudiantes y las notas que obtuvieron en el último examen: 8, 9, 5 y 6.

* La media de esta distribución es:



* Y su desviación típica:



Si ahora consideramos todas las muestras de tamaño 2 y las notas medias de estas muestras, obtenemos una nueva variable :

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Muestra | 8-8 | 8-9 | 8-5 | 8-6 | 9-8 | 9-9 | 9-5 | 9-6 | 5-8 | 5-9 | 5-5 | 5-6 | 6-8 | 6-9 | 6-5 | 6-6 |
| Nota media | 8 | 8.5 | 6.5 | 7 | 8.5 | 9 | 7 | 7.5 | 6.5 | 7 | 5 | 5.5 | 7 | 7.5 | 5.5 | 6 |

Esta variable,, cumple lo siguiente:

* Su media es igual a la media poblacional:



* Su desviación típica es igual a la desviación típica poblacional dividida entre :



**Ej.** Si tenemos una población con media 250 y desviación típica 50 y tomamos una muestra de tamaño 49, como sabemos que la variable aleatoria  de las medias muestrales se aproxima a una normal  podemos obtener que la probabilidad de que la media de esta muestra se encuentre entre los valores 249 y 251 es:



Sin embargo, si consideramos una muestra de tamaño 400, entonces  se aproximará a una normal *N* (250, 2.5) y, por tanto, la probabilidad será, en este caso:



Esto es debido a que los valores de las medias muestrales están más concentrados en torno a la media, pues la desviación típica es menor.

**Ej.** Una variable X hace corresponder los valores 1, 2, 3, 4 y 5 a una población formada por cinco individuos.

1. Construye una tabla con todas las muestras de tamaño 2.
2. Halla la media y la desviación típica de X.
3. Halla la media y la desviación típica de , cuya distribución es la de las medias muestrales.
4. Comprueba que  y que 

**Ej.** Se sabe que la edad de los españoles al empezar el ya desaparecido servicio militar sigue una distribución *N*(20.8, 2).

1. Si se escoge una muestra de 100 jóvenes, ¿qué probabilidad hay de que su edad media no supere los 21 años?
2. Si aumentamos el tamaño de la muestra y seleccionamos 1000 jóvenes, ¿cuánto valdrá ahora la probabilidad de que la edad media de estos jóvenes no sea superior a 21 años?

**5.- DISTRIBUCIÓN MUESTRAL DE LAS PROPORCIONES.**

Supongamos una variable aleatoria que sólo puede tomar dos valores, si o no, éxito o fracaso, A o B...En este caso, la población objeto de estudio sigue una distribución binomial y lo que tratamos de estimar es qué proporción *p* de esta población tiene uno de estos dos valores.

Si en lugar de toda la población consideramos una serie de muestras de tamaño *n*, obtenemos, para cada una de estas muestras, unas proporciones *p*1, *p*2, ... de individuos de la población que poseen esta característica.

Si  es la variable aleatoria que toma los valores *p*1, *p*2, ..., se puede demostrar que:

* La media de la variable aleatoria , , es igual a la proporción poblacional *p*.



* La desviación típica de la variable aleatoria , , es aproximadamente igual a la raíz cuadrada positiva del cociente entre *p* · (1 – *p*) y *n*.



* La distribución de la variable aleatoria , llamada **distribución muestral de las proporciones**, se aproxima a una normal 

Esta aproximación es tanto mejor cuanto mayor sea *n* y más próximo sea *p* a 0,5. Así si *n·p* y *n·*(1 – *p*) son mayores que 5, podemos considerar que la aproximación es buena. En caso contrario aumentaríamos el tamaño de la muestra.

**Ej.** Supongamos que el porcentaje de familias españolas con un solo hijo es del 20%. Consideremos una muestra de 1000 familias, y queremos saber cuál es la probabilidad de que, al menos el 21% de estas familias, tenga un solo hijo.

En este caso *n* = 1000 y *p* = 0,2; por tanto, la variable aleatoria que mide las proporciones muestrales  se aproxima a la distribución normal:



Se cumple que 

Luego 

**Ej.** En una población de seis individuos, tres poseen una determinada característica y tres no.

1. ¿Qué proporción *p* de individuos posee esta característica?
2. Considera todas las posibles muestras de tamaño dos y halla en cada caso la proporción *p*i. ¿Cuál es la media de estas proporciones?¿Y la desviación típica?

**Ej.** Se sabe que la eficacia de una determinada vacuna contra una enfermedad es del 88%. Si cogemos una muestra de 100 individuos expuestos a esta enfermedad, ¿cuál es la probabilidad de que al menos 85 de ellos no la contraigan?

**6.- DISTRIBUCIÓN MUESTRAL DE LA DIFERENCIA DE MEDIAS.**

Supongamosque queremos comparar la efectividad de dos métodos de enseñanza en estudiantes de Bachillerato. Podríamos considerar dos grupos de estudiantes; al primer grupo le aplicaríamos un método de enseñanza y al segundo, otro método. Se trataría entonces de hacer inferencia sobre la diferencia en el aprovechamiento medio de ambos métodos de enseñanza.

Tan importante como estimar la media de una población es comparar las medias de dos poblaciones; para ello, supongamos que la primera población tiene media **μ1** y desviación típica **σ1,** y la segunda población tiene media **μ2** y desviación típica **σ2**. Tomamos muestras de tamaño **n1** para la primera población y **n2** para la segunda, y para cada una de estas parejas de muestras calculamos la diferencia de sus medias 

Si llamamos a la variable aleatoria que toma los valores , se puede demostrar que:

* La media de la variable aleatoria , , es igual a la diferencia de las medias de las dos poblaciones.



* La desviación típica de , , viene dada, aproximadamente, por la siguiente expresión:



* A medida que *n*1 y *n*2 crecen, la distribución de la variable aleatoria , llamada **distribución muestral de la diferencia de medias**, se aproxima a una distribución normal 

Si σ1  y σ2 son desconocidas, podemos considerar como estimación de estos parámetros las desviaciones típicas muestrales para muestras grandes.

**Ej.** Las pilas del tipo A tienen un tiempo medio de vida de 1000 horas, con una desviación típica de 100 horas, mientras que las del tipo B tienen un tiempo medio de vida de 900 horas con una desviación típica de 200 horas.

Tomamos una muestra de 400 pilas del tipo A y 100 del tipo B. ¿Qué probabilidad hay de que el tiempo medio de vida de las pilas del tipo A supere al menos en 50 horas el tiempo medio de vida de las pilas del tipo B?

La variable aleatoria  sigue una distribución normal: ; por tanto:



**Ej.** Supongamos que la altura media de los jugadores de baloncesto es de 1,95 metros, con una desviación típica de 0,2 metros, mientras que la altura media de los jugadores de fútbol es de 1,70 metros, con una desviación típica de 0,25 metros.

Si elegimos una muestra de 25 jugadores de baloncesto y 36 futbolistas, ¿cuál es la probabilidad de que la altura media de los jugadores de baloncesto supere a la de los futbolistas al menos en 0,3 metros?

**7.- ESTIMACIÓN PUNTUAL Y POR INTERVALOS DE CONFIANZA.**

Si consideramos una población cuya distribución es conocida, pero desconocemos alguno de sus parámetros, podemos estimar el valor de este parámetro a partir de una muestra representativa por medio de un proceso denominado **estimación paramétrica.**

* **La estimación puntual** es un procedimiento mediante el cual se obtiene un único valor para el parámetro desconocido.
* **La estimación por intervalos de confianza** es un procedimiento por el que obtenemos un intervalo que contiene al parámetro desconocido con un cierto nivel de confianza fijado de antemano.

**Ej.** Si decimos que la estimación del porcentaje de votos de un determinado partido es del 48%, estaremos realizando una estimación puntual.

Si decimos que la estimación del porcentaje de votos se encuentra entre el 46% y el 50%, con un nivel de confianza del 95%, estaremos considerando una estimación por intervalos de confianza.

Para realizar una estimación puntual de un parámetro tendremos que utilizar un estadístico denominado *estimador puntual*.

Así, por ejemplo, si quisiéramos estimar la media poblacional, podríamos tomar una muestra, hallar su media y concluir que esta coincide con el parámetro buscado. Pero, ¿cómo saber si este método de estimación puntual es bueno?

Se puede demostrar que el procedimiento empleado en la estimación puntual es bueno si, al considerar no una única estimación obtenida de una única muestra, sino todas las posibles estimaciones obtenidas de distintas muestras, estas forman una variable aleatoria que sigue una distribución tal que:

* Su media se concentra alrededor del valor del parámetro que se está estimando, es decir, es un estimador *insesgado*.
* Su desviación típica es lo más pequeña posible, es decir, es un estimador *eficiente.*

****

Estimador insesgado Estimador sesgado





Estimador eficiente Estimador no eficiente

Si el procedimiento de estimación empleado cumple estas dos condiciones, tendremos una probabilidad alta de que el parámetro estimado a partir de una muestra difiera poco del valor real.

Esta probabilidad será mayor cuanto más eficiente sea el estimador.

Algunos buenos estimadores utilizados para estimar puntualmente parámetros poblacionales son:

* La media muestral, que es un estimador insesgado de la media poblacional.
* La proporción muestral, que es un estimador insesgado de la proporción poblacional.

Estos dos estimadores tienen en común que son tanto más eficaces cuanto mayor sea el tamaño de las muestras consideradas.

**Ej.** Si para evaluar el tanto por ciento de españoles entre 18 y 25 años que fuman, elegimos una muestra de 5000 personas y obtenemos que el 41% son fumadores, podemos estimar que el 41% de los españoles entre 18 y 25 años fuman, pues sabemos que la proporción muestral es un buen estimador de la proporción poblacional.

Más completa que la estimación puntual es la estimación por **intervalos de confianza**, ya que en vez de proporcionarnos un único valor como estimación del parámetro desconocido, nos da todo un intervalo donde puede encontrarse dicho valor.

Este procedimiento de estimación nos permite calcular dos valores entre los que esperamos que esté el parámetro buscado con un cierto nivel de confianza, que llamaremos 1 – α, donde α es el *nivel de riesgo* fijado de antemano; y los límites inferior y superior del intervalo de confianza se llaman *límite de confianza inferior* y *superior,* respectivamente.

*Llamamos* ***intervalo de confianza*** *para un parámetro λ, con un nivel de confianza* 1 – α, *con* 0 < α < 1, *a un intervalo real* (*a, b*), *tal que la probabilidad de que el parámetro λ pertenezca a dicho intervalo es* 1 – α, *es decir*:



Supongamos que tenemos una población de N individuos y que consideramos todas sus muestras de tamaño *n.* Para cada una de estas muestras hallamos su intervalo de confianza, con un determinado nivel de confianza 1 – α; entonces comprobaremos que, al menos el 100 · (1 – α) % de los intervalos de confianza hallados, contienen el parámetro poblacional que deseamos estimar.

**Ej.** Si tenemos una población de 60 individuos, en la que tomamos muestras de tamaño 5, tendremos 605 = 777 600 000 muestras distintas.

Si para cada una de estas muestras construimos el intervalo de confianza de un parámetro desconocido con un 0,95 de nivel de confianza, lo que podemos afirmar es que, de todos los intervalos hallados, al menos el 95% de ellos, es decir, 0,95 · 777 600 000 = 738 720 000 intervalos contienen el parámetro que queremos estimar.

Antes de hallar el intervalo de confianza de un parámetro desconocido, con un nivel de confianza 1 – α, tendremos que calcular un valor llamado **valor crítico, **, tal que si consideramos la curva de la distribución normal estándar N(0, 1), este valor deja a su derecha un área igual a α/2.

**Ej.** Si 1 – α = 0,95, entonces  = 0,025. Luego si **** deja a su derecha un área igual a 0,025, a su izquierda dejará un área igual a 1 – 0,025 = 0,975.

Buscando en las tablas de la N(0, 1), tenemos que **=** 1,96.

**Ej.** Normalmente se usan como niveles de confianza los valores 0,95; 0,9545; 0,99 y 0,9973. Encuentra para cada uno de ellos su valor crítico.

**8.- INTERVALO DE CONFIANZA PARA LA MEDIA POBLACIONAL.**

Para calcular el intervalo de confianza de un determinado parámetro comenzaremos estudiando el caso de una población con media, μ, desconocida y que deseamos estimar, y desviación típica, σ, conocida.

Sabemos que la distribución muestral de las medias se aproxima a una distribución normal y, por tanto, la distribución tipificada  seguirá una distribución normal estándar N(0, 1).

Si observamos la figura anterior, resulta claro que:



Sustituyendo:

 

En realidad, si es una media muestral, como es un valor de , sabemos que, al menos en el (1 – α)% de los casos, se cumple que:



Despejando: 



Luego el intervalo de la media con un nivel de confianza del (1 – α)% es:



Si σ es desconocida y *n* suficientemente grande  podemos sustituir σ por la desviación típica de la muestra, obteniendo un intervalo aproximado.

Los valores son los límites de confianza inferior y superior.

**Ej.** Sabemos que la desviación típica de una población es σ = 2. Considerada una muestra de 100 individuos obtenemos que = 8, ¿cuál es el intervalo de confianza para la media poblacional a un nivel de confianza 1 – α = 0,95.

Sabemos que si 1 – α = 0,95 entonces **=** 1,96, y el intervalo de confianza es:



**Ej.** A fin de estimar el número medio de hijos de las familias españolas, consideramos una muestra de 1000 personas y encontramos que esta muestra tiene una media de 2,1 hijos y una desviación típica de 0,5.

Si deseamos hallar el intervalo de confianza del número medio de hijos por familia, con un nivel de confianza del 0,99, primero tendremos que hallar .

Como 1 – α = 0,99 , entonces si  deja a su derecha un área igual a 0,005, dejará a su izquierda un área igual a 1- 0,005 = 0,995, y buscando en las tablas de

la N(0, 1), tendremos que = 2,58.

Como el tamaño de la muestra es suficientemente grande, podemos sustituir la desviación típica poblacional por la muestral; por tanto, el intervalo de confianza será:



**9.- INTERVALO DE CONFIANZA PARA LA PROPORCIÓN Y PARA LA DIFERENCIA DE MEDIAS.**

Al estudiar la distribución muestral de las proporciones, vimos que esta se aproxima a una normal 

Para hallar el intervalo de confianza aproximado de la proporción poblacional realizaremos el mismo razonamiento que en caso del intervalo de confianza de la media poblacional. Así, el intervalo buscado será:



Donde es la proporción muestral, es decir, el número de veces que ocurre un suceso dividido por el tamaño de la muestra.

Del mismo modo, si tenemos dos poblaciones con medias μ1 y μ2 desconocidas y desviaciones típicas σ1 y σ2 conocidas, podemos hallar el intervalo de confianza para la diferencia de las medias procediendo de igual forma que en los dos casos anteriores; para tenemos en cuenta que la distribución muestral de la diferencia de medias sigue la distribución normal:



El intervalo buscado será pues:



Donde  es la media de la primera muestra e , la de la segunda.

Si σ1 y σ2  son desconocidas, se pueden estimar a partir de las muestras. Los intervalos que se obtienen es este caso serán aproximados.

**Ej.** El número de piezas fabricadas por una máquina A en cinco días ha sido: 50, 48, 53, 60 y 37; y, en esos mismos días, otra máquina B ha fabricado: 40, 51, 62, 55, y 64 piezas.

Si sabemos que la desviación típica de la máquina A es de 8,3 y la de la máquina B, de 9,6, determina el intervalo de confianza para la diferencia de medias con un nivel de confianza del 95%.

1 – α = 0,95  = 1.96





El intervalo buscado será: 

**Ej.** De una muestra de 100 individuos que probaron dos modelos de automóviles, 60 manifestaron que preferían el modelo A frente al modelo B. Determina un intervalo de confianza del porcentaje de individuos que prefiere el modelo A al modelo B, si consideramos un nivel de confianza del 0,98.

**10.- TAMAÑO DE LA MUESTRA.**

Sabemos que ha medida que aumenta el tamaño de la muestra obtenemos estimaciones más fiables. Pero, este aumento encarece el estudio a realizar.

Cuando estimamos un intervalo de confianza estamos cometiendo un error máximo que es igual al radio del intervalo de confianza, con un nivel de riesgo α, previamente establecido. El objetivo es hallar el tamaño mínimo de la muestra para un nivel de confianza y un error máximo determinados.

* En el caso de la media poblacional, el intervalo de confianza obtenido para un nivel 1–α es: . El error máximo que cometemos en esta estimación vendrá dado por el radio de este intervalo, es decir, E=

Si conocemos o predeterminamos el error máximo E que queremos cometer y el nivel de confianza, podemos calcular el tamaño de la muestra *n*, despejando de la expresión anterior.

Así, en este caso, *n* es el menor entero mayor que  

*n* = 

* Razonando de la misma forma, en el caso de la estimación de la proporción poblacional, obtenemos que *n* es el menor entero mayor que , donde debe ser previamente estimado.

*n* = 

Si no disponemos de información sobre , podemos suponer que = 0.5, ya que este valor corresponde al caso de máxima varianza.

**Ej.** Supongamos que queremos estimar la producción media de leche al día de un determinado tipo de vacas con un error menor que 0.5 litros y un nivel de confianza del 0.95. Si de estudios anteriores sabemos que la desviación típica es de 1.5 litros, ¿qué tamaño de muestra debemos tomar?

Como se trata de estimar una media, sabemos que el intervalo de confianza es:

. El error máximo será: E =

Tenemos que E = 0.5 σ = 1.5 

Por tanto: 0.5 = 

Luego el tamaño mínimo de la muestra será de 35 individuos.

**Ej.** Sabemos que el peso de los recién nacidos sigue una distribución N(3.3, 0.5). Con el fin de actualizar estos datos seleccionamos una nueva muestra. ¿Qué tamaño de muestra debemos elegir para cometer un error máximo de 0.1 kilogramos con un nivel de confianza del 95%?

E = 0.1 σ = 0.5   *n* = = 

El tamaño mínimo de la muestra será de 97 recién nacidos.

**11.- EJERCICIOS PROPUESTOS.**

**1.-** Una empresa dispone de 300 impresoras matriciales, 1500 de inyección de tinta y 200 láser. Para estudiar la eficacia de las mismas decide extraer una muestra de 40 impresoras mediante un muestreo estratificado.

¿Cuál es el tamaño correspondiente a cada estrato?

**2.-** Una empresa de mensajería dispone de 600 automóviles monovolumen, 1200 furgonetas y 300 camiones. Para estudiar la eficacia de los vehículos en el reparto de mercancías decide extraer una muestra de 60 vehículos mediante un muestreo estratificado proporcional. ¿Cuál es el tamaño correspondiente a cada estrato?

**3.-** En una población de cuatro individuos se ha medido una característica X, obteniendo como valores 1, 8, 10 y 12.

 **a)** Halla todas las posibles muestras de tamaño 2.

 **b)** Calculala distribución **** de las medias muestrales.

**c)** Averigua la media poblacional y la de la distribución ****. ¿Qué relación existe entre estos dos valores?

**d)** Calcula de desviación típica poblacional y la de ****.

**4.-** Si las notas de Matemáticas en las pruebas de acceso a la Universidad siguen una distribución normal N(5, 2) y elegimos al azar una muestra de 100 estudiantes.

**a)** ¿Qué probabilidad hay de que la nota media en Matemáticas de estos 100 alumnos esté entre 4,5 y 5?

**b)** Si la muestra hubiese sido de 1000 estudiantes, ¿qué probabilidad tendríamos de que la nota media estuviera entre 4,5 y 5?

**c)** ¿Por qué es mayor el segundo resultado?

**5.-** Si el 60% de los licenciados de una facultad encuentran trabajo el primer año después de acabar la carrera y seleccionamos al azar 25 de estos estudiantes, ¿qué probabilidad hay de que al menos 15 de ellos encuentren trabajo el primer año?

**6.-** Una empresa decide hacer un estudio sobre el interés de los padres por la introducción de la lengua extranjera en primero de Primaria. Encuestados 900 padres, obtenemos que el 60% está a favor. ¿Cuál es el intervalo de confianza para los padres que están a favor de esta medida, si consideramos un nivel de confianza del 0,95?

**7.-** Tenemos una población de 20 individuos y consideramos las 1140 muestras de tamaño 3. Para cada una de estas muestras hallamos el intervalo de confianza para la media poblacional con nivel 0,95.

 **a)** ¿Pueden existir 114 intervalos que no contienen la media poblacional?

 **b)** ¿Pueden existir 1090 intervalos que contienen la media poblacional?

**8.-** Una variable X sigue una distribución normal de media desconocida y desviación típica 12. Consideramos una muestra de 120 individuos y obtenemos una media muestral igual a 20.

**a)** Halla los intervalos de confianza con niveles del 90% y del 95% y compara los resultados.

**b)** Si queremos tener un error máximo menor que 2 en cada uno de los intervalos anteriores, ¿se necesitará aumentar el tamaño de la muestra? En caso afirmativo, ¿podrías decir cuál debería ser ese aumento?

**9.-** Se sabe que de cada 100 nacimientos, 52 son niñas. ¿Qué probabilidad hay de que en una muestra de 400 recién nacidos haya más hombres que mujeres?

**10.-** Supongamos que el 70% de los españoles ha seguido por televisión la retransmisión del debate sobre los presupuestos del año 2003.

**a)** Si seleccionamos al azar una muestra de 400 españoles, ¿qué probabilidad hay de que, al menos 300, hayan seguido este debate por televisión?

**b)** Si aumentamos la muestra a 1000 españoles, ¿cuál será ahora la probabilidad de que el mismo porcentaje de españoles haya seguido el debate por televisión?

**11.-** En un municipio de 100 000 habitantes realizamos una encuesta para determinar el porcentaje de población que no tiene decidido su voto para las próximas elecciones. Si de 1000 encuestados, 150 indican que no saben qué votarán, establece el intervalo en el que se encuentra el porcentaje de dudosos con un nivel de confianza del 95,45%

**12.-** Una determinada fábrica de legumbres lanza la mercado envases de 1 kg. Para comprobar que, efectivamente, el peso medio de estos paquetes es de 1 kg, se coge una muestra de 100 paquetes y se obtiene que el peso medio es de 0,963 kg. con una desviación típica de 0,012 kg. ¿Puedes afirmar que el peso medio es realmente de 1 kg. con un nivel de confianza del 99%?

**13.-** Queremos determinar el porcentaje de estudiantes que necesitan gafas. De un estudio realizado hace 3 años sabemos que el 65% de ellos usaban gafas.

**a)** ¿Qué tamaño de la muestra debemos coger para cometer un error máximo del 5% con un nivel de riesgo del 5%?

**b)** Si no tenemos información previa, ¿qué tamaño de muestra debemos tomar?

**14.-** Se sabe que la estatura de los individuos de una población es una variable aleatoria que sigue una distribución normal con desviación típica 6 cm.

Se toma una muestra aleatoria de 225 individuos que da una media de 176 cm.

1. Obtenga un intervalo, con un 99% de confianza, para la media de la estatura de la población.
2. Calcula el mínimo tamaño de muestra que se ha de tomar para estimar la estatura media de los individuos de la población con un error inferior a 1 cm y con un nivel de confianza del 95%.

**15.-** Se sabe que los estudiantes de una provincia duermen un número de horas diarias que se distribuye según una ley normal de media μ y desviación típica σ = 2 horas.

1. A partir de una muestra de 64 alumnos se ha obtenido el siguiente intervalo de confianza (7,26; 8,14) para la media de la población. Determina el nivel de confianza con que se ha construido dicho intervalo.
2. Determina el tamaño muestral mínimo necesario para que el error que se cometa al estimar la media de la población por un intervalo de confianza sea, como máximo, de 0,75 horas, con un nivel de confianza del 98%.

**16.-** En una universidad se toma al azar una muestra de 100 alumnos y se encuentra que han aprobado todas las asignaturas 62. e pide hallar:

1. Con un nivel de confianza del 95%, un intervalo para estimar el porcentaje de alumnos que aprueban todas las asignaturas.
2. A la vista del resultado anterior se pretende repetir la experiencia para conseguir una cota de error de 0,03, con el mismo nivel de confianza del 95%. ¿Cuántos individuos ha de tener la muestra?

**17.-** En una población de estudiantes de bachillerato se quiere estimar la proporción de estudiantes que tiene posibilidad de conectarse a Internet desde su domicilio. Se selecciona al azar una muestra de 300 estudiantes de dicha población y a partir de la información obtenida con ellos, se determina el intervalo de confianza (0,22; 0,28) para dicha proporción con una confianza del 99%. Teniendo en cuenta esta información, contestar justificando las respuestas:

1. ¿Qué estimación puntual daríamos a la proporción de estudiantes de esa población que pueden conectarse a Internet desde su domicilio?
2. ¿Qué número mínimo de estudiantes tendríamos que seleccionar al azar con objeto de conseguir, con una confianza del 99%, un error máximo en la estimación de dicha proporción menor que 0,05?

**18.-** En una muestra aleatoria de 400 personas de una población hay 80 que tienen teléfono móvil. Calcular el intervalo de confianza aproximado para la proporción poblacional para un nivel de confianza del 95%.

**19.-** Un laboratorio farmacéutico afirma que el número de horas que un medicamento de fabricación propia tarda en curar una determinada enfermedad sigue una variable normal con desviación típica igual a 8. Se toma una muestra de 100 enfermos a los que se les suministra el medicamento y se observa que la media de horas que tardan en curarse es igual a 32.

1. Encontrar un intervalo de confianza, con un nivel de confianza del 99% para la media del número de horas que tarda en curar el medicamento.
2. Si el nivel de significación es igual a 0,05, ¿cuál es el tamaño de la muestra que habría que considerar para estimar el valor de la media con un error menor de 3 horas?

**20.-** En una muestra de 600 personas de una ciudad se observa que 30 son inmigrantes.

1. Determinar un intervalo de confianza de nivel 0,95 para el porcentaje de inmigrantes en la ciudad.
2. Se quiere estimar el porcentaje de inmigrantes con un error máximo de 0,02, ¿cuál es el tamaño de la muestra que habría que considerar si se da un nivel de significación del 1%?

**21.-** La duración de las llamadas de teléfono, en una oficina comercial, sigue una distribución normal con desviación típica 10 segundos. Se hace una encuesta entre 50 llamadas y la media de duración obtenida es esa muestra es 35 segundos. Calcular un intervalo de confianza al 99% para la duración media de las llamadas.

**22.-** Preguntadas 100 personas de cierta ciudad, elegidas al azar, si leen el periódico al menos una vez a la semana, sólo 40 han contestado que sí. Encuentra un intervalo de confianza, con nivel de confianza del 99%, para la proporción de personas de esa ciudad que leen el periódico al menos una vez a la semana.

**23.-** Con el fin de estimar la edad media de los habitantes de una gran ciudad, se tomó una muestra aleatoria de 300 habitantes que arrojó una edad media de 35 años y una desviación típica de 7 años.

1. Hallar el intervalo del 95% de confianza en el que se encontrará la edad media de la población.
2. ¿Qué nivel de confianza se debería usar para que el intervalo fuera 35 ± 0,44?

**24.-** La desviación típica del número de horas diarias que duermen los alumnos de cierta universidad es 3 horas. Se considera una muestra aleatoria de 40 estudiantes que revela una media de sueño de 7 horas. Hallar un intervalo de confianza de 95% para la media de horas de sueño de los estudiantes de esa universidad.

**25.-** Se desea hacer una estimación sobre la edad media de una determinada población. Calcula el tamaño de la muestra necesario para realizar dicha estimación con un error máximo de medio año a un nivel de confianza del 99,73%. Se conoce de estudios previos que la edad media de dicha población tiene una desviación típica de *σ* = 3.

**26.-** Deseamos conocer el número de personas mayores de edad que sería necesario incluir en una muestra nacional para estimar la clase de actividad en España con un error absoluto E = 0,04 y un nivel de confianza del 99,73%. Se dispone de un valor P = 0,45 del último censo.

**27.-** En cierto barrio se quiere hacer un estudio para conocer mejor el tipo de actividades de ocio que gustan más a sus habitantes. Para ello, van a ser encuestados 100 individuos elegidos al azar.

1. Explica qué procedimiento de selección sería más adecuado utilizar: muestreo con o sin reemplazamiento. ¿Por qué?
2. Como los gustos cambian con la edad y se sabe que en el barrio viven 2500 niños, 7000 adultos y 500 ancianos, posteriormente se decide elegir la muestra anterior utilizando muestreo estratificado.

b.1) Define los estratos.

b.2) Determina el tamaño muestral correspondiente a cada estrato.

**28.-** Se sabe que el cociente intelectual de los alumnos de una universidad se distribuye según la ley normal de media 100 y varianza 729.

1. Halla la probabilidad de que la muestra de 81 alumnos tenga un cociente intelectual medio inferior a 109.
2. Halla la probabilidad de que la muestra de 36 alumnos tenga un cociente intelectual medio superior a 109.

**29.-** Una muestra aleatoria de 100 alumnos que se presenta a las pruebas de selectividad revela que la media de edad es de 18,1 años. Halla un intervalo de confianza de 90% para la edad media de todos los estudiantes que se presentan a las pruebas, sabiendo que la desviación típica de la población es de 0,4.

**30.-** Se quiere conocer la permanencia media de pacientes en un hospital. Se tienen datos referidos a la estancia, expresada en días, de 800 pacientes, de donde se han sacado los resultados siguientes:

días, *s* = 9 días

Se pide obtener un intervalo de confianza del 95% para la estancia media.

**31.-** La duración de bombillas sigue una distribución normal de media desconocida y desviación típica de 50 horas. Para estimar la duración media, se experimenta con una muestra de tamaño *n*. Calcula el valor de *n* para que, con un nivel de confianza del 95%, se haya conseguido un error en la estimación inferior a 5 horas.

**32.-** La altura de los jóvenes andaluces se distribuye según una ley normal de media desconocida y varianza 25 cm2. Se ha seleccionado una muestra aleatoria y con una confianza del 95% se ha construido un intervalo para la media poblacional cuya amplitud es de 2,45 cm.

1. ¿Cuál ha sido el tamaño de la muestra seleccionada?
2. Determina el límite superior e inferior del intervalo de confianza si la muestra tomada dio una altura media de 170 cm.

**33.-** Se conoce que el número de días de permanencia de los enfermos en un hospital sigue una distribución normal de media 8,1 días y desviación típica 9 días. Se elige al azar una muestra de 100 enfermos:

1. Razona cuál es la distribución de la media muestral.
2. ¿Cuál es la probabilidad de que la media muestral está comprendida entre 8 y 10 días?

**34.-** Una variable aleatoria X tiene una distribución normal siendo su desviación típica igual a 3.

1. Si se consideran muestras de tamaño 16, ¿qué distribución sigue la variable aleatoria media muestral?
2. Si se desea que la media de la muestra no difiera en más de una unidad de la media de la población, con probabilidad de 0,99, ¿cuántos elementos, como mínimo, se deberían tomar en la muestra?

**35.-** Se ha tomado una muestra aleatoria de 100 individuos a los que se ha medido el nivel de glucosa en sangre, obteniéndose una media muestral de 110 mg/cc. Se sabe que la desviación típica de la población es de 20 mg/cc.

1. Obtén un intervalo de confianza al 90% para el nivel de glucosa en sangre de la población.
2. ¿Qué error máximo se comete con la estimación anterior?

**36.-** La media de edad de los alumnos que se presentan a las pruebas de acceso a la Universidad es de 18,1 años y la desviación típica 0,6 años.

1. De los alumnos anteriores se elige al azar una muestra de 100 individuos, ¿cuál es la probabilidad de que la media de la edad de la muestra esté comprendida entre 17,9 y 18,2 años?
2. ¿Qué tamaño debe tener una muestra de dicha población para que su media esté comprendida entre 17,9 y 18,3 años, con un nivel de confianza del 99,7%?

**37.-** Se sabe que el contenido de fructosa de cierto alimento sigue una distribución normal cuya varianza es conocida teniendo un valor de 0,25. Se desea estimar el valor de la media poblacional mediante el valor de la media de una muestra, admitiéndose un error máximo de 0,2, con una confianza del 95%. ¿Cuál debe ser el tamaño de la muestra?

**38.-** En un determinado barrio se seleccionó al azar una muestra de 100 personas cuya media de ingresos mensuales resultaba igual a 106.000 pesetas, con una desviación típica de 20.000 pesetas.

1. Si se toma un nivel de confianza del 95%, ¿cuál es el intervalo de confianza para la media de los ingresos mensuales de toda la población?
2. Si se toma un nivel de significación igual a 0,01, ¿cuál es el tamaño muestral necesario para estimar la media de ingresos mensuales con un error menor de 3.000 pesetas?

**39.-** Se desea estudiar el gasto semanal de fotocopias, en pesetas, de los estudiantes de bachillerato de Madrid. Para ello, se ha elegido una muestra aleatoria de 9 de estos estudiantes, resultando los valores siguientes para los gastos:

100 150 90 70 75 105 200 120 80

 Se supone que la variable aleatoria objeto de estudio sigue una distribución normal de media desconocida y desviación típica igual a 12.

 Determina un intervalo de confianza del 95% para la media del gasto semanal en fotocopias por estudiante.

**12.- SOLUCIONES.**

**1.-** 2000 impresoras: 300 matriciales; 1500 inyección de tinta; 200 láser

Suponiendo muestreo estratificado proporcional:



**2.-** 2100 vehículos: 600 monovolúmenes; 1200 furgonetas; 300 camiones



**3.-** 1–8–10–12

**a)** Muestras de tamaño 2:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1-1 | 1-8 | 1-10 | 1-12 | 8-1 | 8-8 | 8-10 | 8-12 | 10-1 | 10-8 | 10-10 | 10-12 | 12-1 | 12-8 | 12-10 | 12-12 |

**b)** Distribución ****:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mtra | 1-1 | 1-8 | 1-10 | 1-12 | 8-1 | 8-8 | 8-10 | 8-12 | 10-1 | 10-8 | 10-10 | 10-12 | 12-1 | 12-8 | 12-10 | 12-12 |
| Media | 1 | 4.5 | 5.5 | 6.5 | 4.5 | 8 | 9 | 10 | 5.5 | 9 | 10 | 11 | 6.5 | 10 | 11 | 12 |

**c) **

 ****

Se observa que: 

**d) **

 **=**

 **=**

**4.-** N(5, 2) *n* = 100

**a) ** sigue una distribución 



**b)** *n* = 1000 → **: **



**c)** A medida que aumenta el tamaño de la muestra, *n*, los datos se concentran más en torno a 

**5.-** p = 0.6

La distribución muestral de las proporciones sigue una normal =

= *N* (0.6; 0.098)

Nos piden 

**6.-** *p* = 0.6; *n* = 900; 1 – α = 0.95 → 

Intervalo de confianza para la proporción:

= = (0.568; 0.632)

**7.-** *N* = 20; *n* = 3 → Nº muestras: ; 1 – α = 0.95

**a)** NO: El número de intervalos que contienen la media poblacional es del 95%, es decir, el 5% no contienen dicha media.

5% · 1140 = 57 intervalos no contienen la media.

**b)** NO: Contienen la media 95% · 1140 = 1083.

**8.-** σ = 12; *n* = 120; 

**a)** 90%; 1 – α = 0.9; 

 95%; 1 – α = 0.95; 

Intervalo de confianza: 

90%: 

95%: 

Si aumentamos el nivel de confianza aumenta el tamaño del intervalo.

**b) **

90%: **** individuos. No tendríamos que aumentar el tamaño de la muestra.

95%: **** individuos. Tendríamos que aumentar el tamaño de la muestra en 19 individuos (139 – 120 = 19).

**9.-** *n* = 400; *p* = 0.48, referido a los niños.



Nos piden: 

**10.-** *p* = 0.7 ; *n* = 400 → 

**a)** 

**b)** En este caso 



**11.-** 1 – α = 0.9545 → α = 0.0455 → 



Intervalo de confianza: =

= 

**12.-** *n* = 100; μ = 1; 

Es un problema de contraste de hipótesis. La zona de aceptación es:

=

= (0.997; 1.003).

Como  podemos afirmar que el peso medio no es de 1 kg. con un nivel de confianza del 99%.

***De otra forma:***

Intervalo de confianza:

= (0.960; 0.966).

Como , el peso medio no es de 1 kg.

1

0.966

0.960

**13.-** *p* = 0.65

**a)**  E = 0.05; α = 0.05 → 

 estudiantes.

**b)** *p* = 0.5

 estudiantes.

**14.-** *n* = 225; 

**a)** 1 – α = 0.99 → 

Intervalo de confianza:



= (174.97; 177.03)

**b)** 1 – α = 0.95 → ; E = 1

**** individuos.

**15.-** σ = 2

**a)** *n* = 64

(7.26; 8.14) =  



 



**b)** 1 – α = 0.98 → α = 0.02 → ; E = 0.75

**** estudiantes.

**16.-** *n* = 100; *p* = 0.62

**a)** 1 – α = 0.95 → 

=

**b)**  alumnos.

**17.-** *n* = 300; (0.22; 0.28); 1 – α = 0.99 → 

**a)** Intervalo de confianza: = (0.22; 0.28)



La estimación puntual será *p* = 0.25.

**b)**  estudiantes.

**18.- **; *n* = 400; 1 – α = 0.95 → 

Intervalo de confianza: =

=

**19.-** σ = 8; *n* = 100; 

**a)** 1 – α = 0.99 → 



**=** (29.939; 34.061)

**b)** α = 0.05 → ; E = 3

**** enfermos.

**20.-** *n* = 600; 

**a)** 1 – α = 0.95 → 

Intervalo de confianza: =

=

**b)** α = 0.01 → 1 – α = 0.99 → ; E = 0.02

 personas.

**21.-** σ = 10; *n* = 50; ; 1 – α = 0.99 → 

 

= (31.357; 38.643)

**22.-** *n* = 100; ; 1 – α = 0.99 → 

= =

**23.-** σ = 7; *n* = 300; 

**a)** 1 – α = 0.95 → 

 

= (34.21; 35.79)

**b)** ¿1 – α?



0.86

Nivel de confianza: 1 – α = 0.72 = 72%



0.72

**24.-** σ = 3; *n* = 40; ; 1 – α = 0.95 → 

Intervalo de confianza:  (6.07; 7.93).

**25.-** 1 – α = 99.73% → ; σ = 3; E = 0.5

****

La muestra debe estar compuesta de, al menos, 324 personas.

**26.-** 1 – α = 99.73% → ; E = 0.04; *p* = 0.45



La muestra estará compuesta de, al menos, 1393 personas.

**27.-**

**a)** Es preferible un muestreo sin reemplazamiento; de esta forma se evitaría que una misma persona sea elegida más de una vez.

**b)**

 **b.1)** Debemos considerar los estratos formados por niños, adultos y ancianos.

 **b.2)**



**28.-** μ = 100; σ2 = 729; σ = 27; *n =* 81

**a)** Las medias de las muestras de tamaño *n* se distribuyen según la normal 



**b)** Ahora *n* = 36 → 



**29.-** *n* = 100; ; 1 – α = 0.90 → ; σ = 0.4



= (18.03; 18.17)

**30.-** *n* = 800; ; 1 – α = 0.95 → 

*s* = 9, desviación típica de la muestra. Como *n* = 800 > 30, hacemos σ = *s* = 9



= (7.476; 8.724)

**31.-** σ = 50; 1 – α = 0.95 → ; E = 5

**** bombillas**.**

**32.-** σ2 = 25; σ = 5; 1 – α = 0.95 → 

**a)**

2.45

Amplitud del intervalo:

*límite superior – límite inferior* = = 

 jóvenes.

**b)** *límite superior*: 

*límite inferior*: 

**33.-**

**a)** La distribución de la media muestral sigue una ley: 

**b)** 

**34.-** σ = 3

**a)** *n* = 16

La media muestral sigue una distribución 

**b)** 1 – α = 0.99 → ; E = 1

**** elementos.

**35.-** *n* = 100; ; σ = 20

**a)** 1 – α = 0.90 → 



= (106.71; 113.29)

**b) **

**36.-** μ =18.1; σ = 0.6; *n* = 100

Población:  → Medias: 

**a) **

**b)** 1 – α = 0.997 → 

E

18.1

18.3

17.9

**37.-** σ2 = 0.25; σ = 0.5; E = 0.2; 1 – α = 0.95 → 

****

Tamaño de la muestra *n* = 25.

**38.-** *n =* 100; ; σ = 20000

**a)** 1 – α = 0.95 → 

(102080; 109920)

**b)** α = 0.01 → 1–α = 0.99 → ; E = 3000

**** personas**.**

**39.-** *n* = 9; σ = 12; 1 – α = 0.95 → 





= (102.16; 117.84)

**13.-P(x) DE LA DISTRIBUCIÓN NORMAL.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **α** | **0'00**  | **0'01** | **0'02** | **0'03** | **0'04** | **0'05** | **0'06** | **0'07** | **0'08** | **0'09** |
| **0.0** | ∞ | 2.576 | 2.326 | 2.170 | 2.054 | 1.960 | 1.881 | 1.812 | 1.751 | 1.695 |
| **0.1** | 1.645 | 1.598 | 1.555 | 1.514 | 1.476 | 1.440 | 1.405 | 1.372 | 1.341 | 1.311 |
| **0.2** | 1.282 | 1.254 | 1.227 | 1.200 | 1.175 | 1.150 | 1.126 | 1.103 | 1.080 | 1.058 |
| **0.3** | 1.036 | 1.015 | 0.994 | 0.974 | 0.954 | 0.935 | 0.915 | 0.896 | 0.878 | 0.860 |
| **0.4** | 0.842 | 0.824 | 0.806 | 0.789 | 0.772 | 0.755 | 0.739 | 0.722 | 0.706 | 0.690 |
| **0.5** | 0.674 | 0.659 | 0.643 | 0.628 | 0.613 | 0.598 | 0.583 | 0.568 | 0.553 | 0.539 |
| **0.6** | 0.524 | 0.510 | 0.496 | 0.482 | 0.468 | 0.454 | 0.440 | 0.426 | 0.412 | 0.399 |
| **0.7** | 0.385 | 0.372 | 0.358 | 0.345 | 0.332 | 0.319 | 0.305 | 0.292 | 0.279 | 0.266 |
| **0.8** | 0.253 | 0.240 | 0.228 | 0.215 | 0.202 | 0.189 | 0.176 | 0.164 | 0.151 | 0.138 |
| **0.9** | 0.126 | 0.113 | 0.100 | 0.088 | 0.075 | 0.063 | 0.050 | 0.038 | 0.025 | 0.013 |

 |
| Tabla para los pequeños valores de alfa |
|

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **α** | 0'002 | 0'001 | 0'000 1 | 0'000 01 | 0'0000 001 | 0'000 000 1 |
|  | 3.090 | 3.291 | 3.891 | 4.417 | 4.892 | 5.327 |

 |

Para calcular el valor de  sólo tienes que buscar el valor de α en la tabla, la parte entera y la primera cifra decimal en la columna de la izquierda y la segunda cifra decimal en la fila horizontal, donde se cortan es el valor de .  |
|  |

