

**MUESTREO**

**TÉCNICAS DE MUESTREO**

**ESTIMACIÓN PUNTUAL**

**ESTIMACIÓN POR INTERVALOS**

**INFERENCIA Y CONTRASTE DE HIPÓTESIS**

## **MUESTREO Y TÉCNICAS**

El proceso mediante el cual la información proveniente de datos de una muestra es utilizada para llegar a conclusiones sobre la población de la que proviene la muestra. Por tanto la inferencia consiste en obtener conclusiones poblacionales a partir de observaciones muestrales.

*Muestra* es una parte de la población, elegida convenientemente, para estudiar unos caracteres determinados de esta.

### *Tipos de muestreo*

Llamamos muestreo al procedimiento mediante el cual seleccionamos los individuos que han de formar la muestra. Todos los muestreos, han de ser aleatorios, es decir todos los individuos de la población tienen la misma probabilidad de ser elegidos para formar parte de la muestra.

Aunque existen muchas formas de realizar un muestreo aleatorio, nos centraremos en las tres que más se utilizan que son:

- *Muestreo aleatorio simple* se realizan numerando la población de tamaño  $N$ , y generando  $n$  números aleatorios, para formar parte de la muestra.
- *Muestreo sistemático* se obtiene el primer elemento al azar y los restantes de  $K$  en  $K$  si  $K = N/n$  con  $n$  tamaño de la muestra.
- *Muestreo estratificado* Se divide la población en estratos o grupos homogéneos, y de cada uno de ellos se elige una muestra aleatoria, simple o sistemática directamente proporcional a su tamaño.

### *Distribución muestral de medias*

Se demuestra que las muestras de tamaño  $n$ , procedentes de una población que sigue una distribución normal  $N(\mu, \sigma)$ , siguen a su vez una distribución normal de parámetros  $\mu_x = \mu$  y  $\sigma_x = \frac{\sigma}{\sqrt{n}}$ , por lo tanto la distribución muestral es  $N(\mu, \frac{\sigma}{\sqrt{n}})$

### **EJEMPLO:**

Las calificaciones de los alumnos de un instituto en las pruebas de selectividad, siguen una distribución normal  $N(5,3, 0,8)$ . Se pide:

- Probabilidad de que un estudiante elegido al azar tenga una nota superior a 5,7 puntos.
- Probabilidad de que en una muestra de tamaño 49, elegido un estudiante al azar la nota sea superior a 5,4 puntos.

### Distribución muestral de proporciones

Se demuestra que las muestras de tamaño  $n$  procedentes de una población de proporción poblacional  $p$ , siguen una distribución  $N(p, \sqrt{\frac{p(1-p)}{n}})$ , esta aproximación es mejor cuanto mayor sea  $n$  y cuanto más próximo esté  $p$  a  $0.5$ .

#### **EJEMPLO:**

El porcentaje de habitantes de Benavente que utiliza los servicios de Internet es del 30%. consideramos una muestra de tamaño 1000, queremos saber cuál es la probabilidad de que al menos el 27% tenga Internet.

### Distribución muestral de la diferencia de medias

Si disponemos de dos distribuciones  $X$ ,  $Y$  con sus respectivas medias y varianzas, que siguen ambas una distribución normal, y queremos comparar una con la otra, tomamos muestras que pueden ser de distinto tamaño en una y otra distribución, entonces la media de la variable aleatoria diferencia, es la media de una menos la de la otra es decir:

$$\mu_{x-y} = \mu_1 - \mu_2$$

La desviación típica viene dada por la expresión  $\sigma_{x-y} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

La diferencia de medias a medida que aumenta el tamaño de las muestras se aproxima a una distribución  $N(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$

#### **EJEMPLO:**

La altura media de los jugadores de baloncesto es 1.95 metros con una desviación típica de 0.2 metros, mientras que la de los jugadores de fútbol es 1.70 metros con una desviación típica de 0.25 metros. elegimos una muestra de 25 jugadores de baloncesto y 36 de fútbol ¿ cuál es la probabilidad de que la altura media de los jugadores de baloncesto supere a la de los futbolistas al menos en 30 cms?

ESTIMACIÓN PUNTUAL Y POR INTERVALOS DE CONFIANZA

Consideramos una población cuya distribución es conocida, pero desconocemos el valor de alguno de sus parámetros, (en este curso va a ser la media o la proporción poblacional). Podemos estimar el valor del parámetro desconocido a partir de una muestra representativa de la población, por medio de un proceso denominado estimación paramétrica.

Estudiamos dos tipos de estimación paramétrica:

**Estimación puntual** es un procedimiento mediante el que se obtiene un único valor del parámetro desconocido, en nuestro caso la media o la proporción.

**Estimación por intervalos de confianza** es un procedimiento mediante el que obtenemos un intervalo en el que se mueve el verdadero valor del parámetro desconocido, con un cierto nivel de confianza fijado de antemano.

**Para realizar una ESTIMACIÓN PUNTUAL**, tendremos que utilizar un estadístico que se denomina estimador puntual, en nuestro caso para la media y la proporción admitimos como cierto que la media y la proporción muestral son un estimador puntual de la media y la proporción poblacional.

Además tanto la media como la proporción muestral son estimadores que cumplen:

**Son Insesgados** es decir su valor se concentra alrededor del verdadero valor de la media y la proporción poblacional

**Son eficientes** es decir su desviación típica es pequeña, la eficiencia de un estimador es mayor cuanto mayor es el tamaño de la muestra.

Por lo tanto podemos afirmar:

La media muestral es un estimador insesgado de la media poblacional y más eficiente cuanto mayor es el tamaño de la muestra.

La proporción muestral es un estimador insesgado de la proporción poblacional y más eficiente cuanto mayor es el tamaño de la muestra.

**EJEMPLOS:**

1.- Se desea estudiar la altura de los pinos silvestres de la región de Sanabria, se sabe que la variable altura sigue una distribución normal, para determinar la media que es desconocida, tomamos una muestra aleatoria simple de 20 árboles, obteniendo los siguientes resultados:

3,10	4,32	2,86	3,72	3,50	4,31	5,41	4,04	4,09	3,75
4,40	4,06	3,55	4,40	3,84	3,14	4,91	3,93	3,39	3,51

Calculamos mediante estadística descriptiva la media y la varianza en este caso muestrales resultando ser:

$$\bar{X} = 3,9115 \quad \sigma^2 = 0,3638$$

Como la media muestral es un buen estimador puntual de la poblacional podemos concluir que la altura de los pinos de Sanabria sigue una distribución normal con una altura media de 3,9115 metros.

2.- Se desea estudiar la incidencia de la gripe A en los jóvenes en edad escolar en la comunidad de Castilla y León, para ello se toma una muestra aleatoria simple de 250 alumnos en edad escolar de la Comunidad de Castilla y León, de los cuales resultó que 18 estaban contagiados con la enfermedad, por lo tanto, la proporción muestral sería:

$$\bar{P} = \frac{18}{250} = 0,072$$

Como la proporción muestral es un buen estimador puntual de la proporción poblacional, podemos concluir que el 7,2% de los jóvenes en edad escolar de Castilla y León está contagiado con la gripe A.

Para realizar una **ESTIMACIÓN POR INTERVALOS DE CONFIANZA**, debemos de antemano fijar un nivel de confianza que denominamos  $1 - \alpha$ , donde  $\alpha$  es el nivel de riesgo o nivel de significación fijado de antemano, es decir la probabilidad de que el verdadero valor del parámetro desconocido no esté en el intervalo que construyamos. En este curso sólo trabajamos con poblaciones cuyas características siguen una distribución normal, como en la estimación puntual, tomamos una muestra aleatoria simple de la población y calculamos la media y la varianza muestrales, sabemos por el tema anterior que las medias muestrales siguen una distribución normal de media la de la población y desviación típica  $\frac{\sigma}{\sqrt{n}}$ .

Calculamos el denominado **valor crítico** para el nivel de confianza prefijado de antemano  $z_{\frac{\alpha}{2}}$ , tal que, si consideramos la curva de la distribución normal, deja a su

derecha un área igual a  $\frac{\alpha}{2}$

**Ejemplo:** Si  $1 - \alpha = 0,95$ , entonces  $\frac{\alpha}{2} = 0,025$ . Luego si  $z_{\frac{\alpha}{2}}$  deja a su derecha un área de 0,025 a su izquierda deja un área de  $1 - 0,025 = 0,975$ . buscamos en la tabla de la normal ese valor y corresponde a 1,96, por tanto el valor crítico par un nivel de confianza del 95% es  $z_{\frac{\alpha}{2}} = 1,96$ .

Sabemos que la distribución muestral de las medias se aproxima a una distribución normal  $N(\mu, \frac{\sigma}{\sqrt{n}})$  y por lo tanto al tipificar:

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Por lo tanto la probabilidad de que z esté en el área limitada por el 95% es:

$P(-z_{\frac{\alpha}{2}} \leq z \leq z_{\frac{\alpha}{2}}) = 1 - \alpha$ , sustituyendo z por el valor tipificado, se llega a la conclusión de que el **intervalo de confianza para un nivel del  $(1 - \alpha)\%$  es:**

$$\left( \bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

En el caso de que la desviación típica sea desconocida y  $n > 30$  podemos tomar como desviación típica la de la muestra.

**Error de muestreo** es  $E = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

**Intervalo de confianza para la proporción**

$$\left( \bar{p} - z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot q}{n}}, \bar{p} + z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot q}{n}} \right)$$

**Error de muestreo** es  $E = z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot q}{n}}$

**EJEMPLOS:**

1.- El pulso cardiaco de las personas se sabe que sigue una distribución Normal con una desviación típica conocida de 12,5. Para estimar el pulso medio seleccionamos una muestra aleatoria simple de 40 personas, obteniendo un pulso promedio de 76,3 latidos por minuto. Obtener un intervalo de confianza al 95% para el verdadero valor del pulso medio. ¿Qué error cometemos en la estimación?

Hay que construir el intervalo de confianza para el número medio de latidos con los datos que tenemos que son :

$$n = 40 \quad 1 - \alpha = 0,95 \quad \frac{\alpha}{2} = 0,025 \quad \bar{X} = 76,3$$

y por tanto, el intervalo de confianza será aplicando la fórmula expresa en teoría:

$$IC = \left[ 76,3 - 1,96 \cdot \frac{12,5}{\sqrt{40}}, 76,3 + 1,96 \cdot \frac{12,5}{\sqrt{40}} \right] = [76,3 - 3,8738, 76,3 + 3,8738] =$$

$$[72,4262, 80,1738]$$

Como conclusión se ha obtenido una estimación por intervalos del pulso medio de las personas que con una confianza del 95% y con un error máximo de 3,87 latidos por minuto, nos permite afirmar que el verdadero pulso medio de las personas se encuentra en el intervalo [72,43 , 80,17]

2.- Se sabe que 18 de cada 250 escolares de una muestra de escolares de Castilla y León padecen la gripe, determinar con una confianza del 99% el intervalo en el que se encuentra la verdadera proporción de todos los escolares de Castilla y León que padecen la enfermedad. ¿Cuál debería ser el tamaño de la muestra necesario para obtener una estimación de la proporción por intervalos con una confianza del 99% y un error máximo inferior al 2%?

Hay que construir el intervalo de confianza para la proporción, los datos que tenemos son:

$$n = 250 \quad 1 - \alpha = 0,99 \quad \frac{\alpha}{2} = 0,005 \quad \bar{p} = \frac{18}{250} = 0,072$$

Por tanto el intervalo de confianza para la proporción poblacional será:

$$IC = \left[ 0,072 - 2,57 \cdot \sqrt{\frac{0,072 \cdot (1 - 0,072)}{250}}, 0,072 + 2,57 \cdot \sqrt{\frac{0,072 \cdot (1 - 0,072)}{250}} \right] =$$

$$|0,072 - 0,042, 0,072 + 0,042| = |0,030, 0,114|$$

Así pues se ha estimado que la proporción de individuos contagiados, con la gripe se encuentra entre el 3% y el 11,4%, y el error máximo que presenta esta estimación es del 4,2% con una confianza del 99%

Si queremos que el error máximo sea inferior al 2%=0,02, tendremos que aumentar el tamaño de la muestra, entonces:

$$E = 2,57 \cdot \sqrt{\frac{0,072 \cdot (1 - 0,072)}{n}} \leq 0,02 \Rightarrow n \geq \left(\frac{2,57}{0,02}\right)^2 \cdot 0,072 \cdot (1 - 0,072) = 1103,28$$

Por lo tanto el tamaño mínimo de la muestra para que el error cometido no supere el 2% será de 1104 individuos.

## INFERENCIA ESTADÍSTICA

En el capítulo anterior se estudió cómo estimar los parámetros de una población a través de los resultados obtenidos a partir de una muestra aleatoria (solamente se hizo para el caso de la media con varianza conocida y de la proporción). Pero con frecuencia sucede que sólo estamos interesados en comprobar una afirmación que se hace acerca del valor de un parámetro y no nos interesa nada más que poder aceptar o rechazar la hipótesis de la afirmación sobre el valor del parámetro, por ejemplo, una persona puede afirmar que en su empresa el contenido de las latas de sardinas por término medio es de 8 piezas, pero eso puede ser cierto o no, afirmarlo no le costó nada, pero hay que realizar un procedimiento que llamamos *Contraste de Hipótesis o Test de Hipótesis* que nos permita afirmar o rechazar la afirmación del empresario con un cierto nivel de error o significación prefijado de antemano.

En estadística una **Hipótesis** es una afirmación que se hace acerca de una propiedad de la población (en nuestro curso sólo sobre la media o la proporción poblacional).

Un **Contraste de Hipótesis** es un procedimiento para comprobar la certeza de la hipótesis.

En cualquier problema de contraste de hipótesis estadístico hay dos hipótesis contradictorias bajo consideración. El objetivo es determinar, en base a la información muestral, cuál de las dos hipótesis es la correcta.

En un juicio hay siempre dos hipótesis, el acusado es inocente o es culpable, pero siempre se parte de que la hipótesis de inocencia es cierta, mientras no se demuestre con evidencias de lo contrario, entonces el jurado rechaza la inocencia a favor de la alternativa de que el acusado es culpable.

La afirmación que se hace sobre el valor del parámetro inicialmente se denomina **Hipótesis Nula y se denota por  $H_0$** . La otra hipótesis que contradice a la hipótesis nula se denomina **Hipótesis Alternativa y se denota por  $H_1$  o también  $H_a$** . En un contraste de hipótesis la hipótesis nula se considera cierta y sólo se rechazará a favor de la hipótesis alternativa si existe una fuerte evidencia proporcionada por los datos muestrales.

La alternativa a la hipótesis nula  $H_0 : \mu = \theta$  será una de las siguientes afirmaciones:

- 1.-  $H_1 : \mu \neq \theta$  lo que nos da lugar a un contraste bilateral o test significativo.
- 2.-  $H_1 : \mu < \theta$  lo que nos da lugar a un contraste unilateral.
- 3.-  $H_1 : \mu > \theta$  lo que nos da lugar a un contraste unilateral.

Una vez que tenemos establecidas tanto la hipótesis nula como la alternativa se necesita un **Estadístico de Contraste**, es decir un estadístico que a partir de su valor, obtenido sobre la muestra, nos indique si debemos, o no, rechazar la hipótesis nula.

Se denomina **Región Crítica** al conjunto de valores del estadístico de contraste para los que la hipótesis nula es rechazada.

**Resumiendo**, un problema de contraste de hipótesis se compone de dos hipótesis, nula y alternativa, un estadístico de contraste y una región crítica. Si el valor del estadístico en los datos muestrales, se encuentran la región crítica se rechaza la hipótesis nula, si, por el contrario, el valor del estadístico de contraste no se encuentra en la región crítica, entonces no se rechaza la hipótesis nula.

Cuando realizamos un contraste de hipótesis se pueden cometer dos tipos de error. Es posible que la hipótesis nula sea rechazada siendo cierta, y es posible que sea aceptada cuando es falsa, entonces los tipos de error son:

**Error de tipo I** consiste en rechazar la hipótesis nula cuando es verdadera

**Error de tipo II** consiste en aceptar la hipótesis nula cuando es falsa.

Un error de tipo I (en un juicio condenar al inocente) es por lo general más serio que un error de tipo II (no condenar al culpable). Entonces tendemos a fijar el valor de nivel de rechazo del contraste  $\alpha$  y encontrar la región crítica que tenga ese valor como probabilidad de error de tipo I.

El valor  $\alpha$  se conoce con el nombre de nivel de significación del contraste y es la probabilidad de cometer error de tipo I. Los niveles tradicionales de significación son: 0,10 , 0,05 y 0,01, aunque el nivel en cualquier ejercicio puede ser otro.

**Determinación del estadístico de contraste:**

Si la hipótesis nula es sobre la media poblacional :

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Si la hipótesis nula es sobre la proporción poblacional:

$$Z = \frac{\bar{p} - p}{\sqrt{\frac{p \cdot (1 - p)}{n}}}$$

Si la hipótesis nula es sobre la diferencia de medias:

$$Z = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Como vimos en la anterior unidad, todos estos estadísticos se aproximan a una distribución normal  $N(0,1)$ , entonces determinamos la región de aceptación o rechazo dependiendo de lo que nos indique la hipótesis alternativa es decir :

**Si el contraste es bilateral** la región de aceptación es  $(-z_{\frac{\alpha}{2}}, z_{\frac{\alpha}{2}})$ , por lo tanto si  $Z$  está en el intervalo aceptamos la hipótesis nula y si no está rechazamos, con una probabilidad de error de tipo I del  $\alpha\%$  .

**Si el contraste es unilateral de hipótesis alternativa  $H_1 : \mu < \theta$** , la región de aceptación es  $(-z_{\alpha}, \infty)$  por lo tanto si  $Z$  está en el intervalo aceptamos la hipótesis nula y si no está la rechazamos.

**Si el contraste es unilateral de hipótesis alternativa  $H_1 : \mu > \theta$** , la región de aceptación es  $(-\infty, z_{\alpha})$  por lo tanto si  $Z$  está en el intervalo aceptamos la hipótesis nula y si no está rechazamos.

**EJEMPLO:**

Una compañía de televisión por cable, hace una promoción para un nuevo canal de pago, dando acceso libre a los posibles usuarios durante dos meses. Una vez finalizado el plazo dado, se eligió una muestra aleatoria de 400 espectadores suscritos a la compañía de televisión por cable y se les preguntó si pagarían 5 € extras al mes para continuar viendo el nuevo canal. Solamente respondieron que sí 25. Con un nivel de significación del 1% ¿ puede la compañía concluir que más del 5% de sus abonados pagarán el nuevo canal?

Se desea contrastar una hipótesis sobre la proporción de abonados que pagarían el nuevo canal .

Establecemos entonces la hipótesis nula y la alternativa.

$$\begin{cases} H_0 : p = 0,05 \\ H_1 : p > 0,05 \end{cases}$$

Con este planteamiento la compañía de televisión podrá asegurar que más del 5% de sus abonados pagarán el nuevo canal si se rechaza la hipótesis nula, con una probabilidad de error de tipo I del 1%.

Los datos de que disponemos son:

$$n = 400 \quad \bar{p} = \frac{25}{400} = 0,0625 \quad \alpha = 0,01 \quad z_{\alpha} = 2,32$$

Tengamos en cuenta que el contraste es unilateral de hipótesis alternativa mayor.

Estadístico de Contraste:

$$Z = \frac{0,0625 - 0,05}{\sqrt{\frac{0,05 \cdot (1 - 0,05)}{400}}} = 1,1471$$

Como la Región Crítica por ser la alternativa mayor es  $(2,32, \infty)$  y el estadístico de contraste  $1,1471 \notin (2,32, \infty)$  entonces aceptamos la hipótesis nula y rechazamos la alternativa, por tanto, los datos no ofrecen evidencia significativa a favor de la hipótesis de la compañía y no pueden asegurar que más del 5% de sus abonados pagarán el nuevo canal.

**EJERCICIOS MUESTREO****ESTIMACIÓN E INFERENCIA****EJERCICIOS MUESTREO**

18.- En una gran ciudad española , la altura media de sus habitantes tiene una desviación típica de 8 cms. Se pide:

- a) Si la altura media de dichos habitantes fuera de 175 cms. ¿cuál sería la probabilidad de que la altura media de una muestra de 100 individuos tomada al azar fuera superior a 176 cms?
- b) Si la muestra tomada anteriormente tiene una altura media de 178cms. Determina un intervalo de confianza del 98% para la altura media de los habitantes de esta ciudad

**EJERCICIOS ESTIMACIÓN POR INTERVALOS**

1 - Se realizó una encuesta a 350 familias de Benavente preguntando si tenían ordenador en casa, encontrándose que 75 de ellas lo poseían. Estima el intervalo en que se moverá la proporción de familias que disponen de ordenador con nivel de confianza del 95%.

2.- Se sabe que 25 de cada 100 objetos elaborados por una empresa son defectuosos ¿ De qué tamaño conviene tomar una muestra para que la proporción estimada de defectuosos no difiera de la verdadera en más de un 5% con un nivel de confianza de un:

1. 95%
2. 99%.

3.- En una ciudad se toma una muestra al azar de 256 personas. De esta muestra, el 20% lleva gafas graduadas y el resto no.

Calcula el intervalo de confianza aproximado para la proporción poblacional de las personas que llevan gafas graduadas para un nivel de confianza del 95%.

4.- Se sabe que en un estudio sobre 40 aviones comerciales, la antigüedad media de éstos es de 13´41 años con una desviación típica muestral de 8´28 años. Se pide:

1. Entre qué valores, con una confianza del 90% , se encuentra la antigüedad media de la flota comercial.
2. Si se quisiera obtener un nivel de confianza del 95% cometiendo el mismo error de estimación que en el apartado anterior y suponiendo la misma desviación típica,¿cuántos aviones deberán componer la muestra?

5.- Se desea hacer un estudio de mercado para conocer el precio medio de los libros de matemáticas. Para ello, se elige una muestra aleatoria formada por 34 libros y se determina que la media muestral es de 30 €con una desviación típica de 5 €.

Halla el intervalo de confianza para el precio medio de los libros de matemáticas al nivel del 99%.

6.- Un fabricante de electrodomésticos sabe que la vida media de éstos sigue una ley normal de media 100 meses y desviación típica 12 meses.

Hallar el mínimo tamaño de la muestra que garantiza, con una probabilidad de 0'98, que la vida media de los electrodomésticos en dicha muestra e encuentra entre 90 y 110 meses.

7- La desviación típica de la altura de los habitantes de un país es de 10 cm. Calcular el tamaño mínimo de una muestra para que el error cometido al estimar la altura media sea inferior a 1 cm con un nivel de confianza del 99%.

8.- Una muestra aleatoria de 100 alumnos que se presentan a las PAU, revela que la media de edad es 18'1 años. Hallar el intervalo de confianza al 90% para la edad media de todos los estudiantes que se presenten, sabiendo que la desviación típica de la población es 0'4.

### EJERCICIOS CONTRASTE DE HIPÓTESIS

1.- Un dentista afirma que el 40% de los niños de 10 años presentan indicios de caries dental. Tomada una muestra de 100 niños, se observó que 30 presentaban indicios de caries. Comprueba si el resultado proporciona evidencia al nivel de significación del 5% que permita rechazar la afirmación del dentista.

2 .- El peso de las bolsas de azúcar fabricadas en la Azucarera de Benavente sigue una distribución normal de media 800 gramos y desviación típica 120 gramos, un consumidor habitual sospecha que el peso medio es menor, para tomar una decisión compra un lote de 50 bolsas y comprueba que el peso medio es de 750 gramos. Plantea el contraste para tomar una decisión.

1. Se rechazaría lo que dice el fabricante a un nivel de significación del 1%.
2. Qué tamaño habría de tener la muestra para que se acepte lo dicho por el fabricante con una confianza del 95%.

3.- Las tensiones de ruptura de los cables fabricados por una empresa , tiene media de 1800 Nw y una desviación típica de 100 Nw.

Se desea comprobar si un nuevo proceso modifica dicha tensión de ruptura. Para ello se toman una muestra de 50 cables y se encuentra que su tensión media de ruptura es de 1850 Nw.

¿ Se puede afirmar que el nuevo proceso ha modificado la tensión media de ruptura con un nivel de significación del 5% ?.

4.- Un experto, basado en anteriores comicios, sostiene que en las próximas elecciones generales, tan sólo acudirán a votar el 48% de la población. No obstante, en un sondeo electoral realizado recientemente entre 1500 personas, 800 tienen intención de votar. ¿Supone esto con un nivel de confianza del 99%, que el experto se equivoca y la intención de votar es mayor?

5.- Se ha llevado a cabo un estudio en diferentes países de la Unión Europea del porcentaje de la población que accede a estudios superiores.

En los países escogidos, se han obtenido los siguientes valores en tanto por ciento: 23'5, 35'0, 29'5, 31'0, 23'0, 27'0, 28'0, 30'5, 23'0.

Se supone que estos porcentajes siguen una distribución normal con desviación típica el 5%.

Se desea contrastar con un nivel de significación del 5% si los datos anteriores son compatibles con un valor medio del porcentaje de la población que cursa estudios superiores igual al 28%.

- a) Plantéense el contraste diciendo hipótesis nula y alternativa.
- b) Determínese la región crítica del contraste.
- c) Es posible aceptar ¿ la hipótesis con el nivel de significación indicado?.

6.- En los últimos tiempos, las ventas medias de un comercio rondaban los 1200 € diarios. Sin embargo, hace unos meses se abrió una superficie comercial cerca del comercio. El establecimiento defiende que las ventas medias se mantienen o, incluso, han aumentado, pero que no han disminuido.

Para contrastar este hecho, el hijo del dueño que estudiaba economía, selecciona una muestra de las ventas diarias realizadas después de la apertura de la superficie comercial.

- a) Establecer la hipótesis nula y la alternativa.
- b) ¿ Qué nombre recibe la probabilidad de que el establecimiento, concluya erróneamente que las ventas medias han disminuido?. Explica cómo se denomina y en qué consiste el otro error posible.

7.- El salario medio correspondiente a una muestra de 1600 personas de Benavente es de 930 €. Se sabe que la desviación típica de los salarios de la población es de 155 €. ¿ Se puede afirmar con un nivel de significación del 1% que el salario medio en Benavente es de 950 €?.

8.- El tiempo necesario para armar una pieza es una variable aleatoria normal de media desconocida y desviación típica 0'6 minutos. Se toma una muestra de 20 piezas, dado el tiempo medio para armar cada una de ellas de 10'2 minutos. El fabricante asegura que el tiempo medio de armarlas es 10 minutos. Se sospecha que el tiempo medio es mayor. Se pide:

- a.- Plantear el contraste
- b.- Tomar una decisión para nivel de significación 0'05.

9.- En los folletos de propaganda de una empresa ubicada en el centro de transportes de Benavente, se asegura que los carteles publicitarios que coloca tienen una duración media de 1600 días. Un cliente sospecha que es menor y toma una muestra aleatoria de 49 carteles, preguntando cuando fueron puestos, obteniendo una duración media de 1570 días, con una desviación típica de 120 días. ¿Puede aceptarse la información de los folletos con un nivel de confianza del 96%?

IES LEÓN FELIPE  
 Dpto Matemáticas  
 Examen Estadística Matemática

1.- Sean A y B dos sucesos aleatorios, de los que se sabe que  $p(A) = 0.4$  y  $p(A \cup B) = 0.7$ .

- a.- Calcular la probabilidad de B para que A y B sean incompatibles. **(0.5 puntos)**
- b.- Calcular la probabilidad de B para que A y B sean independientes. **(0.5 puntos)**

2.- En una empresa muy grande los ascensos son muy solicitados. La probabilidad de ascender por oposición es 0.2, la de ascender por méritos es 0.8, y la de ascender por enchufe es 1. Si los presentados para un posible ascenso el 70% lo hace por oposición, el 25% por méritos y el resto son enchufados, se pregunta:

¿ Si una persona ascendió por qué método es más probable que lo consiguiera?. **(2 puntos)**

3.- Una variable aleatoria tiene por función de densidad  $f(x) = \begin{cases} \frac{c \cdot x^2}{36} & \text{si } 0 \leq x \leq 6 \\ 0 & \text{en otro caso} \end{cases}$

- a.- Hallar c para que sea una función de densidad. **(0.5 puntos)**
- b.- Hallar la función de distribución de esta variable. **(0.5 puntos)**
- c.- Calcular  $p(0 \leq x \leq 1)$ ,  $p(x > 3)$ . **(0.5 puntos)**
- d.- Calcular la esperanza y la mediana. **(0.5 puntos)**

4.- El Ayuntamiento de una ciudad muy noble trata de investigar el grado de aceptación que tiene la instalación de un parking. Para ello toma una muestra de 100 individuos y obtiene que hay 76 en contra.

- a.- Calcular el error de muestreo y el error máximo admisible con una confianza del 96%. **(1 punto)**
- b.- Dar una estimación por intervalo con nivel de confianza 0.99, e interpretarlo. **(1 punto)**
- c.- Encontrar el tamaño de la muestra necesario para que con un nivel 0.99 el error sea menor que 0.02. **(0.5 puntos)**

5.- Los resultados de una prueba de una población estudiantil determinan una variable aleatoria que se supone distribuida normalmente con desviación típica 5. Un sociólogo afirma que la puntuación media de esta población es 79, mientras que un compañero cree que es inferior.

- a.- Plantear el contraste, si tomada muestra de tamaño 40 la media es 78. **(1 punto)**
- b.- Definir el error de tipo I en esta situación. **(0.5 puntos)**
- c.- Tomar una decisión para nivel de significación 0.1. **(1 punto)**

IES LEÓN FELIPE  
Dpto Matemáticas  
Examen Estadística Matemática

1.- Dados los sucesos A y B de un mismo espacio muestral, se sabe que:  
 $P(A) = 0,4$ ;  $P(A \cup B) = 0,8$  y  $P(\overline{A} \cup \overline{B}) = 0,7$

- Comprobar si los sucesos A y B son independientes
- Calcular la probabilidad de que sólo se verifique uno de los dos sucesos

2.- Un dentista afirma que el 40% de los niños presentan indicios de caries dental. Tomada una muestra de 100 niños se observó que 30 presentaban indicios de caries. Plantea razonadamente un contraste para esta situación. Encuentra el p valor. Toma una decisión con nivel de significación del 8%. Define e interpreta en este caso el error de tipo I.

3.- Se sabe que una moneda está trucada de modo que la probabilidad de sacar cara es  $\frac{3}{5}$ . Se lanza la moneda 8 veces. Hallar:

- Probabilidad de sacar 5 caras
- Probabilidad de sacar al menos una cruz.

4.- El 60% de los habitantes jubilados de la muy noble villa de Benavente están satisfechos con los jardines de la Mota, y el 70% de estos habitantes pasean diariamente por ellos. De los no satisfechos sólo el 20% pasean diariamente.

- ¿Qué tanto por ciento pasean diariamente por la Mota?
- ¿Qué porcentaje de los que pasearán el 31 de Marzo, están satisfechos con los jardines de la Mota?

5.- Sea una variable aleatoria continua de función de densidad:

$$f(x) = \begin{cases} mx & \text{si } 0 \leq x \leq 2 \\ 0 & \text{en otros casos} \end{cases}$$

- Calcular:
- el valor de m
  - función de distribución
  - $P(1 < x < 1,5)$



IES LEÓN FELIPE  
Dpto Matemáticas  
Examen Estadística Matemática

1.- Sean A y B dos sucesos aleatorios. Supóngase que  $p(A) = 0'4$  y  $p(A \cup B) = 0'7$ . Sea  $p(B) = p$ .

¿Para qué valor de p son A y B sucesos incompatibles?. ¿Para qué valor de p son A y B sucesos independientes?

2.- En nuestro instituto se imparten sólo dos idiomas inglés y francés. El **90 por ciento** de los alumnos estudian inglés y el resto francés. El **30 por ciento** de los alumnos que estudian inglés pertenecen a la asociación de alumnos, y de los que estudian francés son socios el **40 por ciento**. Se elige un alumno al azar.

A) Hallar la probabilidad de que pertenezca a la asociación de alumnos.

B) Si pertenece a la asociación de alumnos, cuál es la probabilidad de que estudie francés.

3.- Con motivo del año cultural de Benavente y comarca, se han programado conciertos en el Teatro Reina Sofía el **80 por ciento** de los sábados. Del total de conciertos, un **25 por ciento** son de piano, un **15 por ciento** de orquesta y un **60 por ciento** de rock. Un joven benaventano que estudia en Salamanca, que desconoce el programa de conciertos, acude a Benavente un sábado cualquiera.

¿Cuál es la probabilidad de que pueda asistir a un concierto de rock?

4.- De una baraja española de **40 cartas** se extraen simultáneamente **3**. Hallar la probabilidad de que sean:

A) Las tres oros

B) Al menos una oros

C) Primero un oro y luego dos copas

5.- Una encuesta revela que el **30 por ciento** de los alumnos de 2º de Bachillerato C.C.S.S. son partidarios de las Pruebas de Acceso a la Universidad y el resto están en contra. En nuestra clase se eligen dos al azar. Calcula la probabilidad de que:

A) Los dos sean partidarios de dichas pruebas

B) Uno esté a favor y el otro en contra

IES LEÓN FELIPE  
Dpto Matemáticas  
Examen Estadística Matemática

1., Dado un experimento aleatorio y el espacio muestral asociado a él, se conoce que  $p(A) = 0,3$   $p(B) = 0,4$  y  $p(A \cup B) = 0,5$ . Se pide calcular:

- Probabilidad de que sólo ocurra A
- Probabilidad de que no ocurra ninguno de los dos
- Probabilidad de que si no ha ocurrido A ocurra B.
- Son incompatibles A y B, son independientes?

2.- Se sabe que el 5% de los hombres son daltónicos y el 25 de cada 10.000 mujeres lo son, se elige al azar una persona de una población que tiene el doble de mujeres que de hombres y resultó ser daltónica ¿Cuál es la probabilidad de que la persona elegida sea mujer?.

3.- La longitud de la ballena azul se distribuye según una ley normal con desviación típica 7,5 m. En un estudio estadístico realizado a 25 ejemplares se ha obtenido un intervalo de confianza (21,06; 26,94) para la longitud media .

- Calcula la longitud media de los 25 ejemplares de la muestra
- Calcula el nivel de confianza con el que se ha construido dicho intervalo.

4.- El 25% de las viviendas de una determinada región están conectadas a Internet . Se eligen 80 viviendas de esa región, se pide:

- Probabilidad de que al menos 20 viviendas estén conectadas a Internet
- Número esperado de viviendas no conectadas a Internet
- Probabilidad de que el número de viviendas conectadas a Internet esté entre 10 y 30.

5.- Un fabricante de electrodomésticos sabe que la vida media en meses de éstos sigue una ley  $N(100, 12)$ . Determinése el tamaño de la muestra que garantiza, con una probabilidad de 0,98, que la vida media de los electrodomésticos está entre 90 y 110 meses.

6.- La publicidad de una determinada marca de productos lácteos dice que su duración es como máximo 15 días después de la fecha de fabricación. Se toma una muestra de 64 envases y se observa que la duración media ha sido 16 días con una desviación típica de 2 días .

- Se puede decir que la publicidad es correcta con un nivel de significación del 5%?
- Se llegaría a la misma conclusión si la desviación típica fuera 4 días y el nivel de confianza el 99%?