SITUACIÓN 1. Para contrastar la hipótesis de si la edad media de inicio en el consumo de alcohol de los jóvenes de una determinada comunidad es más tardía que la media de la población general establecida en 13 años, un investigador utiliza una muestra de 25 jóvenes encontrando que la edad media en su comunidad es de 14 años con una desviación típica insesgada de 2,8. Asumimos que la variable edad de inicio en el consumo de alcohol se distribuye normalmente en la población.

1. A partir de la información de la muestra utilizada el intervalo de confianza más aproximado para la varianza poblacional con un nivel de confianza del 95% es: A) 5,70; 14,41; **B) 4,77; 15,17**; C) 5,23; 15,13.

   El intervalo de confianza para la varianza poblacional viene dado por la siguiente expresión:

$$\frac{(n-1)S_{n-1}^2}{\chi^2_{1-\alpha/2}} < \sigma^2 < \frac{(n-1)S_{n-1}^2}{\chi^2_{\alpha/2}}$$

   Para un nivel de confianza del 95%

$$\frac{(n-1)S_{n-1}^2}{\chi^2_{0,975}} < \sigma^2 < \frac{(n-1)S_{n-1}^2}{\chi^2_{0,025}}$$

   Buscando en la tabla y sustituyendo:

$$\frac{24 \cdot 2{,}8^2}{39{,}3641} < \sigma^2 < \frac{24 \cdot 2{,}8^2}{12{,}4012} \rightarrow (4{,}77 < \sigma^2 < 15{,}17)$$

2. La hipótesis nula para contrastar la hipótesis sobre la edad media en el consumo de alcohol es: **A) $\boldsymbol{H_0: \mu \leq 13}$**; B) $H_0: \mu = 13$; C) $H_0: \mu \geq 13$

3. Con los datos de la situación 1, el estadístico de contraste para poner a prueba la media es: A) 2,102; **B) 1,78**; C) 1,59.

$$T = \frac{\bar{Y} - \mu_0}{S_{n-1}/\sqrt{n}} = \frac{14-13}{2{,}8/\sqrt{25}} = 1{,}786$$

4. Conforme a los objetivos del investigador y un "alfa"= 0'05, la máxima diferencia que puede producirse por simple azar, si H_0 es cierta, entre la media observada en la muestra y la media planteada en la hipótesis nula, expresado en unidades de desviación típica es: A) 2,064; **B) 1,711**; C) 2,79.

   Nos piden el valor crítico. El contraste de hipótesis es unilateral derecho, luego tenemos que buscar en la tabla "t" de Student para 24 grados de libertad $(n-1)$, la puntuación que supera una proporción igual a 0'95. La respuesta correcta es "B".

5. La interpretación de los resultados obtenidos es que la edad media de inicio en el consumo de alcohol de los jóvenes de la comunidad: **A) es más tardía que la media general con un nivel de confianza del 95%**; B) es más tardía que la media general con un nivel de confianza del 99%; C) no es más tardía que la media general con un nivel de confianza del 95%.

   El estadístico de contraste es significativo a un nivel de confianza del 95% (1'786>1'711), pero no es significativo al nivel de confianza del 99% (1'786<2'492).

SITUACIÓN 2: La empresa SND's de sondeos electorales ha pronosticado que el nivel de apoyo que recibirá el partido X en las próximas elecciones será del 40%. Desde el propio partido X se promueve un nuevo sondeo con el fin de contrastar la veracidad de esta afirmación. Se elige al azar una muestra aleatoria de 400 personas, con derecho a voto, de los cuales 128 manifiestan su intención de votar al partido X.

1- Con un nivel de confianza del 95%, el Intervalo de confianza aproximado de la proporción de personas que votarán al partido X, es: **A) 0,2743; 0,3657**; B) 0,352; 0,424; C) 0,318; 0,322.

La proporción en la muestra es: $p = \frac{128}{400} = 0{,}32$
Al nivel de confianza del 95%, las puntuaciones típicas son: $\pm 1{,}96$

$$p \pm Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

$$0{,}32 \pm 1{,}96\sqrt{\frac{0{,}32(1-0{,}32)}{400}} \rightarrow (0{,}2743; 0{,}3657)$$

2- La hipótesis alternativa es; A) $H_1: \pi < 0{,}40$; **B)** $\boldsymbol{H_1: \pi \neq 0{,}40}$; C) $H_1: \pi > 0{,}40$.

El contraste es bilateral porque el enunciado no especifica si el partido X espera unos resultados inferiores o superiores a los obtenidos por la empresas SND's. Simplemente, el partido X pretende contrastar la veracidad de los datos pronosticados por SND's.

3- El estadístico de contraste es: A) -3,43; **B) -3,266**; C) -6,67.

$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0{,}32 - 0{,}40}{\sqrt{\frac{0{,}40 \cdot 0{,}60}{400}}} = -3{,}266$$

4- Con un nivel de confianza del 99%, si los dirigentes del partido X considerasen que la proporción de apoyos no alcanza el valor pronosticado por la empresa SND's, el valor crítico para rechazar la hipótesis nula, es: **A) -2,58**; B) 1,64; **C) -2,33**.

Marcamos como correctas las opciones A y C, porque esta pregunta puede plantear problemas de interpretación.

Si nos ceñimos al enunciado planteado para esta situación, el contraste es bilateral. Es decir, ***inicialmente*** *el partido X no sabe si el nivel de apoyo que recibirá será superior o inferior al pronosticado por la empresa SND's, pero "después" de realizar el estudio, observa que la proporción es inferior a 0,40. En este caso el valor crítico es -2,58.*

Por otro lado, podríamos interpretar que "antes" de realizar el estudio, el partido X considera que la proporción de apoyos será inferior a la proporción pronosticada por SND's, por lo que el contraste sería unilateral y por lo tanto la respuesta correcta sería C.

5- Para un contraste bilateral con un nivel de significación de 0,01, la conclusión es: **A) Rechazar la hipótesis nula con un nivel de significación del 0,01**; B) No hay evidencias para rechazar la hipótesis nula con un nivel de significación 0,05; C) No hay evidencias para rechazar la hipótesis nula con un nivel de significación 0,01.

Rechazamos la hipótesis nula al nivel de confianza del 99% porque el estadístico de contraste es menor que el valor crítico $(-3{,}266 < -2{,}58)$*, luego la opción A es correcta, siendo las opciones B y C claramente incorrectas.*

SITUACIÓN 3: "La formación alcanzada por la población adulta española ha mejorado, de forma continua, en los últimos 10 años. Desde 1998 el porcentaje de españoles de 25 a 64 años que poseen estudios superiores a los obligatorios ha pasado de 33% al 51% en 2008. En la misma proporción ha disminuido, por tanto, el porcentaje de españoles que sólo poseen estudios obligatorios, que ha pasado del 67% en 1998 al 49% en 2008" (Panorama de la Educación. Informe OCDE 2010). Suponga que usted quiere estudiar si estos datos son los que realmente existen en su Comunidad, para lo que utiliza una muestra aleatoria de 900 adultos con edades comprendidas entre 25 y 64 años, encontrando que 378 de ellos tienen solo los estudios obligatorios.

Los datos son:

	Estudios superiores	*Estudios obligatorios*	*Muestra:*
1998	33%	67%	$\left.\begin{matrix} n = 900 \\ 378\ con\ estudios\ obligatorios \end{matrix}\right\} \rightarrow p = \frac{378}{900} = 0{,}42$
2008	51%	49%	

1- Si decide trabajar con un nivel de confianza del 95% y se establece el error máximo de estimación en el 3%, ¿cuál debe ser el tamaño de la muestra?: **A) 1040**; B) 544; C) 789.

$$n = p(1-p)\frac{Z^2_{1-\alpha/2}}{E^2} = 0{,}42 \cdot 0{,}58 \cdot \frac{1{,}96^2}{0{,}03^2} = 1039{,}79 \approx 1040$$

2- Trabajando con un nivel de confianza del 95%, el intervalo de confianza de la proporción de adultos de su Comunidad que solo tienen estudios obligatorios, se encuentra entre: A) 0,378 y 0,462; **B) 0,388 y 0,452**; C) 0,376 y 0,461.

$$p \pm Z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}} \rightarrow 0{,}42 \pm 1{,}96\sqrt{\frac{0{,}42 \cdot 0{,}58}{900}} \rightarrow \begin{cases} 0{,}388 \\ 0{,}452 \end{cases}$$

3- Si el investigador quiere confirmar que la proporción de adultos de su Comunidad que solo tienen los estudios obligatorios es significativamente inferior que el indicador de España en 2008, la hipótesis alternativa es: **A)** $\boldsymbol{H_1: \pi < 0{,}49}$; B) $H_1: \pi > 0{,}49$; C) $H_1: \pi \leq 0{,}42$.

La respuesta correcta es "A", puesto que las hipótesis que debemos plantear son:

$$H_0: \pi \geq 0{,}49$$
$$H_1: \pi < 0{,}49$$

4- Si el investigador cree que la proporción de adultos que solo tienen los estudios obligatorios es inferior que el indicador de España en 2008, el estadístico de contraste y el nivel crítico p para analizar si la diferencia es estadísticamente significativa es: A) 3,34 (p=0,0008); **B) -4,2 (p<0,0002)**; C) -2,33 (p=0,0198);

$$Z = \frac{0,42 - 0,49}{\sqrt{\frac{0,49 \cdot 0,51}{900}}} = -4,20 \xrightarrow{Tabla\ curva\ normal} p < 0,0002$$

5- Respecto a la hipótesis del investigador analizada en las preguntas anteriores, el investigador concluye que la proporción de adultos que solo tienen estudios obligatorios en su Comunidad, es: **A) inferior a la del resto de España con un nivel de significación de 0,01**; B) la misma que la del resto de España con un nivel de significación de 0,05; C) superior a la del resto de España con un nivel de significación de 0,01.

SITUACIÓN 4. "En Educación Primaria, la media de alumnos por clase en los centros públicos de España (19,7) es más baja que en la OCDE (21,6) y que en la Unión Europea (20,3). En los centros privados ocurre lo contrario, pues la media en España es de 24,4 frente a 20,8 de media de la OCDE y 19,1 de la UE". (Panorama de la Educación. Informe OCDE 2010). Suponga que usted quiere estudiar si estos datos difieren significativamente de los que existen en su Comunidad. Asumiendo que la variable número de alumnos por clase se distribuye normalmente, selecciona una muestra aleatoria de 100 aulas de educación primaria en centros públicos, encontrando que la media de alumnos por clase es de 20,9 con una desviación típica poblacional de 5,8.

Los datos del enunciado son:

	Media de alumnos por clase		
	España	*OCDE*	*UE*
Públicos	*19,7*	*21,6*	*20,3*
Privados	*24,4*	*20,8*	*19,1*

Datos de la muestra: $n = 100$, $\bar{Y} = 20,9$, $\sigma = 5,8$

1- Trabajando con un nivel de confianza del 95%, el intervalo de confianza para la media de alumnos por clase en los centros públicos de su Comunidad es: A) 19,98; 21,8; B) 19,4; 22,39; **C) 19,76; 22,04**.

$$20,9 \pm 1,96 \frac{5,8}{\sqrt{100}} \rightarrow (19,76; 22,04)$$

2- El procedimiento que aplicaría para contrastar si la media de los datos de su Comunidad es significativamente menor que de la media de la OCDE es: A) la prueba t con n-1 grados de libertad; **B) la prueba z**; C) la prueba de chi-cuadrado con *n* grados de libertad.

La distribución en la población es normal, el nivel de medida de la variable dependiente es de razón y conocemos la varianza de la población, luego la respuesta correcta es B.

3- Si se decide trabajar con un nivel de confianza del 95% y se establece un error máximo de estimación en 1 punto, ¿Cuál debe ser el tamaño de la muestra? **A) 129**; B) 12923; C) 378.

$$n = \sigma^2 \frac{\left(Z_{1-\alpha/2}\right)^2}{E^2} = 5,8^2 \frac{1,96^2}{1^2} = 129,23 \approx 129$$

4- El estadístico de contraste para analizar si la media de alumnos en los centros públicos de su Comunidad es significativamente mayor que la media de España, toma el valor: A) 1,056; B) 1,261**; C) 2,069**.

$$Z = \frac{\bar{Y} - \mu_0}{\sigma/\sqrt{n}} = \frac{20{,}9 - 19{,}7}{5{,}8/10} = 2{,}069 \approx 2{,}07$$

5- Si sospecha que la media de los centros públicos de su Comunidad es significativamente mayor que la media de España, el nivel crítico p es: A) p=0,1446; B) p< 0,05; **C) p=0,0192**.

Acudiendo a la tabla de curva normal, deducimos que la probabilidad de obtener valores iguales o más extremos que el estadístico de contraste (2,07) es 0,0192, luego la respuesta correcta es C.

6- Teniendo en cuenta el estadístico de contraste obtenido ¿cuál de las siguientes sentencias es verdadera?: A) no hay evidencia suficiente para rechazar que la media de alumnos por clase en los colegios públicos de la Comunidad es igual a la media de España; **B) La media de alumnos por clase en los colegios públicos de la Comunidad es mayor que la media de España (NC= 95%)**; C) rechazamos la hipótesis nula con un nivel de confianza del 99%.

Los resultados obtenidos son significativos a un nivel de confianza del 95% (2,069 > 1,64), pero no a un nivel de confianza del 99% (2,069 < 2,33) luego la respuesta correcta es B.

7- Trabajando con un nivel de confianza del 95%, la interpretación del intervalo de confianza para la media de alumnos por clase en los centros públicos de su Comunidad, le permite afirmar que la media de su comunidad es**: A) significativamente distinta que la de España pero igual que la de la UE**; B) significativamente distinta a la de la OCDE pero igual que la de España; C) significativamente distinta que la de la UE pero igual que la de la OCDE.

El intervalo de confianza $(19{,}76; 22{,}04)$ *incluye a la media de la OCDE (21,6) y la media de la UE (20,3), pero no incluye la media de España (19,7). Luego la respuesta correcta es A.*

SITUACIÓN 5. Según el último estudio del Observatorio Español sobre Drogas (2009), realizado en estudiantes de Secundaria de 14 a 18 años, el 5,1% ha consumido cocaína alguna vez en la vida y el 2,7% éxtasis. Además, el inicio en el consumo de cocaína y éxtasis tiene lugar cada vez a edades más tempranas. Así, mientras que en el año 2004 la edad media de inicio para la cocaína era de 15,9 años en los hombres y 15,7 en las mujeres, en el año 2008 disminuyó a 15,3 años y 15,2, respectivamente (ENCUESTA ESTATAL SOBRE USO DE DROGAS EN ENSEÑANZAS SECUNDARIAS, 2009). Imagine que los datos disponibles por un equipo de atención primaria que cubre a un determinado sector de su municipio indican que en la muestra de 37 jóvenes varones y 41 mujeres atendidos el pasado año, el 8% habían consumido cocaína, al menos un vez, siendo la edad media de inicio en el consumo de cocaína de 15,4 años en los hombres y de 15,2 años en las mujeres con una desviación típica de 1,3 para los hombres y 1,1 en las mujeres.

Datos de encuestas estatales:

2009	*5,1% Cocaína*	*2,7% Éxtasis*
	Edad media consumo de cocaína	
	Hombres	*Mujeres*
2004	*15,9*	*15,7*
2008	*15,3*	*15,2*

Datos de la muestra:

	n	$\bar{Y}$	S_n	
Hombres	*37*	*15,4*	*1,3*	*8% Cocaína*
Mujeres	*41*	*15,2*	*1,1*	

1- Bajo el supuesto de que la edad media de los hombres que han consumido cocaína al menos una vez, es de 15,9 años, el error típico de la distribución muestral de la media es: A) 0,453; **B) 0,2166**; C) 0,334.

En el caso de los hombres, el error típico de la distribución cuando no conocemos la varianza poblacional viene dado por:

$$\frac{S_n}{\sqrt{n-1}} = \frac{1{,}3}{\sqrt{37-1}} = 0{,}2167$$

2- Si la hipótesis del investigador es que la edad media de inicio en el consumo de cocaína en los hombres de su municipio ha disminuido significativamente respecto a los datos del 2004, el estadístico de contraste es: A) -2,43; B) -2,7; **C) -2,3.**

Contraste unilateral de medias para una muestra (varones) varianza poblacional desconocida:

$$T = \frac{\bar{Y} - \mu_0}{\frac{S_n}{\sqrt{n-1}}} = \frac{15{,}4 - 15{,}9}{\frac{1{,}3}{\sqrt{37-1}}} = -2{,}31$$

3- A partir de los datos obtenidos en la muestra y fijando un nivel de confianza del 95% ¿cuál es el tamaño de la muestra necesario para estimar la proporción poblacional de jóvenes de su municipio que han consumido cocaína al menos una vez con un error máximo de estimación del 2%?: **A) 707;** B) 1536; C) 978.

$$n = p(1-p)\frac{z^2_{1-\alpha/2}}{E^2_{max}} = 0{,}08 \cdot (1-0{,}08)\frac{1{,}96^2}{0{,}02^2} = 706{,}85 \approx 707$$

4- A partir de los datos obtenidos en la muestra y fijando un nivel de confianza del 95%, ¿entre qué valores se encontrará la proporción poblacional de jóvenes de su municipio que han consumido cocaína al menos una vez?: A) 0,71; 0,88; **B) 0,02; 0,14**; C) 0,04; 0,12.

$$E_{max} = z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}} = 1{,}96\sqrt{\frac{0{,}08(1-0{,}08)}{37+41}} = 0{,}06$$

$$p \pm E_{max} \rightarrow 0{,}08 \pm 0{,}06 \begin{cases} 0{,}14 \\ \\ 0{,}02 \end{cases}$$

5- Con un nivel de confianza del 95%, el intervalo de confianza de la varianza poblacional de la edad de las mujeres de su municipio que han consumido al menos una vez cocaína es: **A) 0,83; 2,03**; B) 0,89; 1,87; C) 0,95; 1,23.

$$l_i = \frac{nS_n^2}{\chi^2_{n-1;1-\alpha/2}} = \frac{41 \cdot 1{,}1^2}{59{,}3417} = 0{,}836 \qquad l_s = \frac{nS_n^2}{\chi^2_{n-1;\alpha/2}} = \frac{41 \cdot 1{,}1^2}{24{,}4330} = 2{,}03$$

6- Si desea contrastar la hipótesis de que la varianza poblacional de la edad de las mujeres de su municipio es mayor que 1, el nivel crítico p que obtiene para tomar una decisión respecto a la H_0, es: **A) p > 0,10**; B) p < 0,05; C) p > 0,01.

$$\chi^2 = \frac{nS_n^2}{\sigma_0^2} = \frac{41 \cdot 1{,}1^2}{1} = 49{,}61$$

Buscando en la Tabla de Chi-cuadrado observamos que para 40 grados de libertad, el valor 49,61 se encuentra entre las puntuaciones:

29,05 que deja por encima de sí el 0,90 de la distribución

51,80 que deja por encima de sí el 0,10 de la distribución.

Luego sabemos que 49,61 deja por encima una proporción superior a 0,10.

SITUACIÓN 6. Según el último estudio del Observatorio Español sobre Drogas (2009), realizado en estudiantes de Secundaria de 14 a 18 años, el 5,1% ha consumido cocaína alguna vez en la vida y el 2,7% éxtasis. Además, el inicio en el consumo de cocaína y éxtasis tiene lugar cada vez a edades más tempranas. Así, mientras que en el año 2004 la edad media de inicio para la cocaína era de 15,9 años en los hombres y 15,7 en las mujeres, en el año 2008 disminuyó a 15,3 años y 15,2, *respectivamente* (ENCUESTA ESTATAL SOBRE EL USO DE DROGAS EN ENSEÑANZAS SECUNDARIAS, 2009). Imagine que los datos disponibles por un equipo de atención primaria que cubre un determinado sector de su municipio indican que con 41 jóvenes varones y 51 mujeres atendidos el pasado año, el 8% habían consumido cocaína, al menos un vez, siendo la edad media de inicio en el consumo de cocaína de 15,4 años en los hombres y de 15,2 años en las mujeres con una desviación típica de 1,3 años y 1,1 años, respectivamente.

Datos de encuestas estatales:

2009	*5,1% Cocaína*	*2,7% Éxtasis*
	Edad media consumo de cocaína	
	Hombres	*Mujeres*
2004	*15,9*	*15,7*
2008	*15,3*	*15,2*

Datos de la muestra:

	n	$\bar{Y}$	S_n	
Hombres	*41*	*15,4*	*1,3*	*8% Cocaína*
Mujeres	*51*	*15,2*	*1,1*	

1. Bajo el supuesto de que el 8% de los jóvenes ha consumido cocaína al menos una vez, el error típico de la distribución muestral de la proporción es: A) 0,901; **B) 0,0282**; C) 0,0417.

$$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0{,}08(1-0{,}08)}{41+51}} = 0{,}02828$$

2. Para comprobar si la edad media del inicio en el consumo de cocaína ha disminuido respecto a los datos del 2004, el estadístico de contraste que aplicaría, es: A) Z porque utilizamos una muestra grande proveniente de una variable que se distribuye normalmente en la población; **B) T porque desconocemos la varianza y la forma de la distribución poblacional;** C) El intervalo de confianza porque proporciona más información sobre la seguridad al tomar una decisión.

3. Si fijamos un nivel de confianza del 95%, ¿entre qué valores se encontrará la edad media del consumo de cocaína en las mujeres de su municipio?: **A) 14,887; 15,512**; B) 14,941; 15,962; C) 14,682; 15,621.

   Nos están pidiendo el intervalo de confianza (con $\alpha = 0{,}05$*) para la edad media del consumo de cocaína en las mujeres. Este viene dada por las fórmulas:*

$$E_{max} = t_{n-1;1-\alpha/2} \frac{S_{n-1}}{\sqrt{n}} = 2{,}009 \frac{1{,}1}{\sqrt{51-1}} = 0{,}3125$$

$$\bar{Y} \pm E_{max} \rightarrow 15{,}2 \pm 0{,}3125 \rightarrow \begin{cases} 15{,}5125 \\ \\ 14{,}8875 \end{cases}$$

4. Si su hipótesis es que la edad media del inicio en el consumo de cocaína en los hombres de su municipio ha disminuido respecto a los datos de 2004, el valor del estadístico de contraste que obtiene es: A) 2,34; B) -3,11; **C) -2,43.**

   Para el grupo de hombres el estadístico de contraste es:

$$T = \frac{\bar{Y} - \mu_0}{\frac{S_n}{\sqrt{n-1}}} = \frac{15{,}4 - 15{,}9}{\frac{1{,}3}{\sqrt{41-1}}} = -2{,}4325$$

5. Si desea contrastar la hipótesis de que la varianza poblacional en la edad de inicio del consumo de cocaína en los hombres de su municipio es significativamente mayor que 1,1 el nivel crítico p que obtiene para tomar una decisión respecto a la H_0, es: A) p> 0,10; **B) p< 0,025**; C) p< 0,01.

$$\chi^2 = \frac{nS_n^2}{\sigma_0^2} = \frac{4{,}1 \cdot 1{,}3^2}{1{,}1} = 62{,}99$$

   En las tablas $(g.l. = n - 1 = 40)$ *no podemos encontrar el valor exacto, pero observamos que el estadístico de contraste (62,99) se encuentra entre las puntuaciones 59,3417 y 63,6907 que dejan por encima de sí, respectivamente unas probabilidades iguales a 0,025 y 0,01, luego la opción correcta es B.*

SITUACIÓN 7. El Centro de Investigaciones Sociológicas (CIS) realiza constantes estudios sobre la ideología e intención de voto de los españoles. Uno de los *ítems* que incluye en sus cuestionarios es el de la "*identificación ideológica*", para lo que se le pide al entrevistado que se ubique ideológicamente en una escala de 1 a 10 (donde 1 significa extrema izquierda y 10 extrema derecha). En su última encuesta a la población española la media aritmética obtenida para la variable identificación ideológica fue de un 4,86. Un psicólogo social interesado en estudiar esta temática en su localidad selecciona una muestra aleatoria formada por 31 sujetos a los que aplica el ítem de identificación ideológica, obteniendo una media de 5,40 y una cuasi-desviación típica de 1,2. Con esta información, responda a las siguientes preguntas:

1- Con un nivel de confianza del 95%, el tamaño de la muestra necesario para que el error de estimación de la media en la escala de Identificación ideológica no supere 0,2 puntos, es**: A) 150;** B) 102; C) 138.

$$n = S_{n-1}^2 \frac{t_{n-1;1-\alpha/2}^2}{E_{max}^2} = 1{,}2^2 \frac{2{,}042^2}{0{,}2^2} = 150{,}11 \approx 150$$

2- Con un nivel de confianza del 95%, el intervalo de confianza aproximado para la media en la escala de identificación ideológica de los residentes en su localidad es: A) 4,98; 5,82; B) 5,03; 5,76; **C) 4,952; 5,847**.

$$E_{max} = t_{n-1;1-\alpha/2} \frac{S_{n-1}}{\sqrt{n}} = 2{,}042 \frac{1{,}2}{\sqrt{31}} = 0{,}44$$

$$\bar{Y} \pm E_{max} \rightarrow 5{,}40 \pm 0{,}44 \rightarrow \begin{cases} 5{,}84 \\ \\ 4{,}96 \end{cases}$$

3- De acuerdo a los objetivos del psicólogo, la hipótesis nula es: **A)** $\boldsymbol{H_0: \mu = 4{,}86}$; B) $H_0: \mu \geq 4{,}86$; C) $H_0: \mu \leq 4{,}86$

El contraste es bilateral, porque el psicólogo no parte de una hipótesis previa sobre los resultados de la muestra.

4- El procedimiento que aplicaría para probar la hipótesis del psicólogo, es: A) la Prueba Z; **B) la prueba t con** $\boldsymbol{n-1}$ **grados de libertad**; C) la prueba Chi cuadrado con $n-1$ grados de libertad.

Tenemos un contraste sobre una media con una muestra de 31 sujetos y de la población no conocemos la varianza.

5- El estadístico de contraste para la hipótesis del psicólogo es: A) 2,352; B) 1,697; **C) 2,5**.

$$T = \frac{\bar{Y} - \mu_0}{\frac{S_{n-1}}{\sqrt{n}}} = \frac{5{,}40 - 4{,}86}{\frac{1{,}2}{\sqrt{31}}} = 2{,}51$$

6- Si el psicólogo fija un nivel de significación de 0,05 el valor crítico para tomar una decisión respecto a la hipótesis nula es: A) 1, 697; B) 1,96; **C) 2,042**.

Buscando en las tablas para un contraste bilateral: $g.l. = n - 1 = 30$ *y* $\alpha/2 = 0{,}025$

7- Bajo el supuesto de que la H_0 es verdadera, la probabilidad de obtener una muestra con una media tan alejada o más de 5,39 es: A) p > 0,10; B) p < 0,001; **C) p < 0,02**.

El valor en las tablas más próximo al estadístico de contraste ($T = 2{,}51$) *corresponde a la puntuación* $2{,}457$*, que deja por debajo de sí una proporción igual a* $0{,}990$ *(*$1 - 0{,}990 = 0{,}01$ *por encima). Como el contraste es bilateral, se multiplica por dos la probabilidad obtenida en la tabla.*

8- Con los datos originales obtenidos en nuestra muestra, la decisión que toma respecto a la hipótesis nula es: **A) Rechazarla porque, bajo el supuesto de que fuera cierta, la probabilidad de obtener una muestra con una media de 5,40 o superior es muy pequeña**; B) Mantenerla porque el valor absoluto del estadístico de contraste es menor que el valor absoluto del valor crítico; C) No se puede rechazar la hipótesis nula.

SITUACIÓN 8: El *"Síndrome Jubilación"* es la situación que experimentan ciertas personas ante esta nueva etapa vital con manifestaciones somáticas, psíquicas y sociales negativas que afectan la calidad de vida del jubilado. Un estudio publicado por el GIE (Grupo de Investigación del Envejecimiento) mediante una encuesta realizada en el 2006 utilizando una muestra de jubilados con edad media de 77,6 años y desviación típica de 8,79 años encuentra que las manifestaciones psíquicas más frecuentes eran la ansiedad (82%), el pesimismo (13,3%) y la depresión (4,7%) y que para el 32% de los expertos consultados el apoyo psicológico constituye la estrategia de intervención más adecuada para superar estos estados. Suponga que usted quiere estudiar la situación de los jubilados de su localidad respecto a este "síndrome", para lo que utiliza una muestra aleatoria de 362 jubilados, con una edad media de 71,2 años y una desviación típica de 12,5 y de los cuales, el 59,8% presenta signos de ansiedad, el 35% pesimismo y el 5,2% depresión.

Datos del enunciado:

		Proporción de:			
	Edad	Ansiedad	Pesimismo	Depresión	Apoyo psicológico como Mejor estrategia
GIE	$\bar{Y} = 77,6 \ S_n = 8,79$	0,82	0,133	0,047	0,32
Muestra	$\bar{Y} = 71,2 \ S_n = 12,5$ $n = 362$	0,598	0,35	0,052	-

1- Utilizando los datos del GIE y con un nivel de confianza del 95%: si le dicen que el intervalo de confianza de la proporción de jubilados con manifestaciones de depresión, es un valor comprendido entre 0,0387 (3,87%) y 0,0553 (5,53%), ¿Cuál es el tamaño más aproximado de la muestra que se ha utilizado?: A) 1348; **B) 2500**; C) 3481.

Conociendo los límites superior e inferior del intervalo de confianza, deducimos fácilmente el error máximo:

$$\left.\begin{array}{l} l_i = p - E \\ l_s = p + E \end{array}\right\} \rightarrow E = \frac{l_s - l_i}{2} = \frac{0,0553 - 0,0387}{2} = 0,0083$$

El tamaño de la muestra será igual a:

$$n = p(1-p)\frac{z_{\alpha/2}^2}{E^2} = 0,047(1-0,047)\frac{1,96^2}{0,0083^2} = 2497,73 \approx 2500$$

2- Utilizando los datos de su estudio y con un nivel de confianza del 95%, la edad media de los jubilados de su localidad es un valor comprendido, aproximadamente, entre: **A) 69,9 y 72,5 años**; B) 69,5 y 72,9 años; C) 70,2 y 73,1 años.

$$71,2 \pm 1,96\frac{12,5}{\sqrt{362-1}} \rightarrow 71,2 \pm 1,29 \rightarrow (69,91; 72,49)$$

3- Con los datos de su estudio y un nivel de confianza del 99%, el intervalo de confianza para la proporción de jubilados con manifestaciones psíquicas de ansiedad, es: A) 0,557 y 0,639; B) 0,547 y 0,648; **C) 0,531 y 0,664**.

$$0{,}598 \pm 2{,}58\sqrt{\frac{0{,}598(1-0{,}598)}{362}} \rightarrow (0{,}5315; 0{,}6645)$$

4- Si desea comprobar que la proporción de jubilados de su localidad con manifestaciones psíquicas de pesimismo es significativamente mayor que el valor 0,133 facilitado por el GIE en el 2006, ¿cuál es, aproximadamente, el valor del estadístico de contraste que obtendría?: A) 8,77; B) 10,3; **C) 12,18**.

$$H_0: \pi \leq 0{,}133$$
$$H_1: \pi > 0{,}133$$

$$z = \frac{p-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0{,}35-0{,}133}{\sqrt{\frac{0{,}133(1-0{,}133)}{362}}} = 12{,}158$$

5- Utilizando la información de su trabajo, si desea comprobar que la proporción de jubilados de su localidad con manifestaciones psíquicas de pesimismo es significativamente mayor que el valor 0,133 facilitado por el GIE en el 2006, el nivel crítico que se obtiene: **A) es menor que 0,0002**; B) es mayor que 0,001; C) depende del nivel de confianza.

El estadístico de contraste es un valor muy extremo (z = 12,158). El valor máximo que podemos consultar en las tablas de curva normal es z = 3,59, que deja por encima de sí una proporción: p = 0,0002

SITUACIÓN 9: El barómetro del CIS de marzo de 2012 realizado en 240 municipios de 48 provincias señalaba que el 23,4% estaba en situación de paro y de éstos, a la pregunta *¿Y cree Ud. que es muy probable, bastante, poco o nada probable que durante los próximos doce meses encuentre Ud. trabajo?*, el 22,6% manifestaba que "bastante probable", frente al 43,1% que creía que "poco probable" y el 19,2% que "nada probable" y el resto "NS/NC". Suponga que quiere estudiar si estos resultados se reproducen actualmente en su localidad, para lo que realiza una encuesta sobre una muestra de 100 personas en situación de paro con una edad media de 39 años y desviación típica de 8,6 años de los cuales 25 le responden que "bastante probable", 35 responden que "poco probable" y 20 que "nada probable" mientras que el resto "no saben o no contestan".

Datos de la Situación 1		Pregunta: ¿probabilidad de encontrar trabajo en los próximos doce meses?			
		Bastante	Poco	Nada	NS/NC
CIS	0,234 en paro	0,226	0,431	0,192	0,151
Encuesta investigador	$\bar{Y} = 39;\ S_n = 8{,}6; n = 100$	0,25	0,35	0,20	0,20

1- Trabajando con un nivel de confianza del 95% y apoyándose en la distribución chi-cuadrado, el intervalo de confianza de la varianza de la edad de las personas en situación de paro de su localidad es un valor comprendido entre: **A) 57,08 y 99,65**; B) 6,64 y 11,59; C) 59,48 y 100,01.

$$\frac{nS_n^2}{\chi^2_{1-\alpha/2,n-1}} < \sigma^2 < \frac{nS_n^2}{\chi^2_{\alpha/2,n-1}} \rightarrow \frac{100 \cdot 8{,}6^2}{129{,}5612} < \sigma^2 < \frac{100 \cdot 8{,}6^2}{74{,}2219} \rightarrow 57{,}085 < \sigma^2 < 99{,}65$$

Con los datos de su encuesta desea comprobar que la proporción de personas en paro que consideran "poco probable" encontrar trabajo en los próximos doce meses en su localidad es significativamente menor que el valor 43,1% proporcionados en el estudio del CIS:

2- ¿Cuál sería la hipótesis nula?: A) $H_o\!: \pi \leq 0{,}431$; **B)** $\boldsymbol{H_o\!: \pi \geq 0,431}$; C) $H_o\!: \pi = 0{,}431$.

3- ¿Cuál sería el estadístico de contraste que utilizaría?: A) Z para dos muestras independientes con varianzas poblacionales conocidas; **B) Z como aproximación de la distribución binomial**; C) t por tratarse de una muestra grande con varianza poblacional desconocida.

4- ¿Cuál es el valor aproximado del estadístico de contraste de su hipótesis?: A) -1,70; B) -1,873; **C) -1,63**

$$z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0{,}35 - 0{,}431}{\sqrt{\frac{0{,}431(1-0{,}431)}{100}}} = -1{,}636$$

5- ¿Cuál es el valor aproximado del nivel crítico?: **A) 0,0516**; B) 0,0734; C) 0,0307.

Buscando en las tablas de curva normal la probabilidad de encontrar valores inferiores a: z = -1,63, comprobamos que la respuesta correcta es A.

6- Su decisión respecto a la hipótesis formulada, es: A) rechazar la H_0 con un nivel de confianza del 99%; B) rechazar la H_0 con un nivel de confianza del 95%; **C) no se puede rechazar la $\boldsymbol{H_0}$ nula con un nivel de confianza del 95%**.

Los valores críticos para los niveles de confianza del 95 y 99% son, respectivamente: z = -1,64 y z = -2,33, por lo que no podemos rechazar la hipótesis nula.

7- Con los datos de su encuesta y con un nivel de confianza del 95%, el intervalo de confianza para la proporción de personas en situación de paro que considera "poco o nada probable" encontrar trabajo en los próximos doce meses, es un valor comprendido entre: A) 0,256 y 0,443; B) 0,122 y 0,278; **C) 0,452 y 0,647**.

La proporción de personas que consideran "poco o nada probable" encontrar trabajo es: p = 0,35 + 0,20 = 0,55

$$0{,}55 \pm 1{,}96\sqrt{\frac{0{,}55(1-0{,}55)}{100}} \rightarrow (0{,}4525; 0{,}6475)$$

8- Si desea contrastar la hipótesis nula de que la edad media de los parados de su localidad es menor de 40 años, el estadístico de contraste que utilizaría es: A) Z por tratarse de una variable cuantitativa con distribución normal; B) t por tratarse de una variable normalmente distribuida; **C) t por tratarse de una variable cuantitativa con varianza desconocida en la población**.

SITUACIÓN 10. Un psicólogo que investiga la percepción del tiempo, considera que dicha habilidad se encuentra deteriorada en los fumadores durante la retirada de nicotina, para lo que toma una muestra aleatoria de 41 fumadores a los que somete a una abstinencia de tabaco con una duración de 24 horas, pidiéndoles que estimen el tiempo había transcurrido (en segundos) en un periodo que, objetivamente, fue de 45 segundos. La media aritmética de la muestra fue igual a 51 segundos con una cuasivarianza igual a 164. El psicólogo desea determinar si la abstinencia tiene un impacto negativo sobre la percepción temporal, provocando que el tiempo se sobre-estime. Nivel de confianza 95%.

1- La hipótesis nula que ha de plantear el psicólogo es: **A) $\boldsymbol{H_0: \mu \leq 45}$**; B) $H_0: \mu = 45$; C) $H_0: \mu \geq 45$.

Dado que el psicólogo desea comprobar si el tiempo se sobre-estima, hemos de plantear un contraste unilateral derecho, cuyas hipótesis son:

$$H_0: \mu \leq 45$$
$$H_1: \mu > 45$$

2- El estadístico de contraste es igual a: A) 1; B) 2; **C) 3**.

$$T = \frac{\bar{Y} - \mu_0}{\frac{S_{n-1}}{\sqrt{n}}} = \frac{51 - 45}{\sqrt{\frac{164}{41}}} = 3$$

3- El valor crítico es igual a: A) 1,303; **B) 1,684**; C) 2,2021.

4- Con los datos recogidos, se rechaza la hipótesis nula: **A) Sí, porque el estadístico de contraste es superior al valor crítico**; B) Sí, porque el nivel de significación es menor que el nivel crítico; C) No, porque el estadístico de contraste es menor que el valor crítico.

5- ¿Es significativo el estadístico de contraste?: A) No; B) Si para un nivel de confianza del 95%, pero no a un nivel de confianza del 99%; **C) Es significativo a un nivel de confianza del 99**%.

El estadístico de contraste ($T = 3$) supera un nivel de confianza del 95% ($3 > 1{,}684$), pero también supera un nivel de confianza del 99% ($3 > 2{,}423$), luego la respuesta correcta es C.

6- ¿Cuál es el valor del nivel crítico p?: A) 0,005; **B) menor que 0,005**; C) 0,005 < p < 0,01.

El valor más extremo que podemos consultar en las tablas t de Student, para 40 grados de libertad, es 2,704, que deja por encima de sí una probabilidad igual a: 0,005, luego el estadístico de contraste ($T = 3$) que es todavía más extremo, dejará por encima de sí una probabilidad menor que 0,005

7- Si hemos realizado un estudio donde el nivel crítico ha sido superior a α. **A) Mantenemos la hipótesis nula**; B) Rechazamos la hipótesis nula; C) No tenemos información para mantener o rechazar la hipótesis nula.

8- La probabilidad de mantener H_0 falsa se denomina: A) nivel de significación; **B) error de tipo II**; C) potencia del contraste.

La probabilidad pedida es el error de tipo II, que también podemos definir como la probabilidad de rechazar H_1 verdadera.

9- Un estadístico cuyo valor medio es igual al valor de esta característica en la población, se dice que es; A) eficiente; B) suficiente; **C) insesgado**.

10- Dados varios estadísticos insesgados para estimar la misma característica poblacional, la mejor elección es utilizar el estadístico: A) con la desviación típica mas grande; **B) con la desviación típica más pequeña**; C) con la media mas grande.

La respuesta correcta es B, puesto que el estimador con la desviación típica más pequeña es el más eficiente.