

Estadística (Química)
Práctica 5 - Estimación - Intervalos de Confianza

Comentario: En todos los ejercicios propuestos

- a) defina las variables aleatorias y los parámetros involucrados.
- b) de ser posible indique:
 - i. la distribución de las variables aleatorias
 - ii. el significado intuitivo de los parámetros.
- c) compare los resultados de hacer las cuentas a mano con las salidas obtenidas con el R, de manera de chequear las primeras y aprender a usar las segundas, en aquellos ejercicios en los que ambas cosas sean posibles.

1. Considere el ejercicio 4 de la Práctica de Estadística Descriptiva.

- (a) Usando las medias y desviaciones estándar calculadas en R, calcule un IC del 95% para la media de las mediciones hechas por cada uno de los grupos. ¿Los intervalos obtenidos tienen nivel de confianza exacto o aproximado? Para responder esta pregunta, recuerde lo realizado al analizar los datos en la práctica de descriptiva.
- (b) Observe que R también puede calcular automáticamente los ICs del inciso anterior. Considere la instrucción:

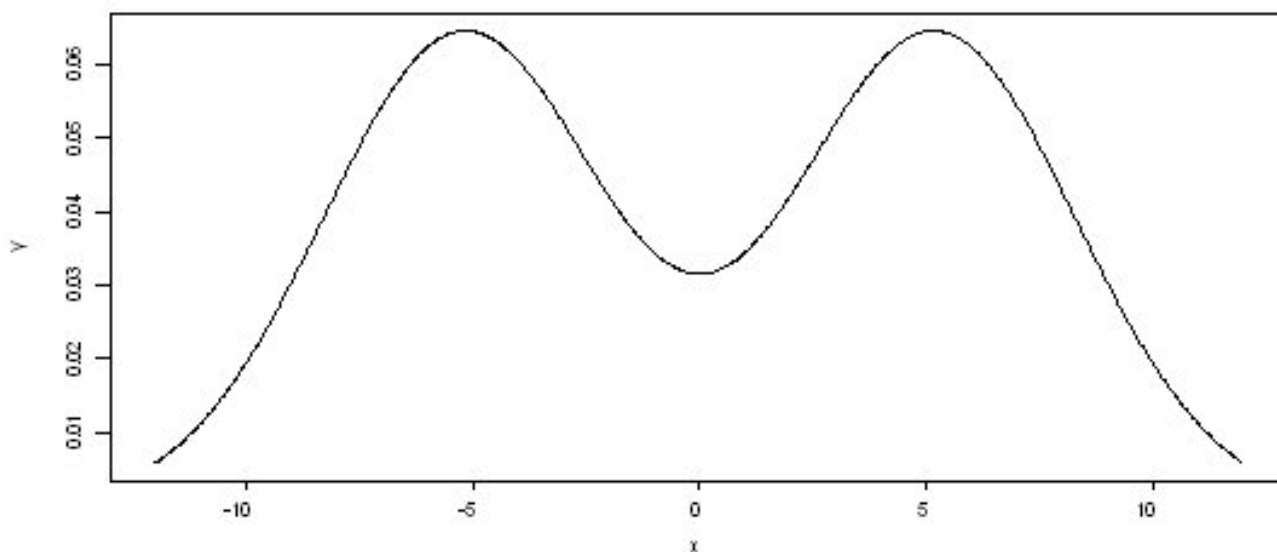
`t.test(GRUP01,alternative= "two.sided")`

- (c) ¿Es razonable pensar que ambos grupos de estudiantes están midiendo la concentración de ion nitrato sin sesgo? En caso contrario, ¿puede saberse si uno solo o ambos grupos hacen mediciones sesgadas?

2. Ciudad Gótica tiene numerosas casas alquiladas. Para realizar un estudio se eligen 400 al azar y se averigua el alquiler que pagaron sus ocupantes el mes anterior, resultando un promedio muestral $\bar{x} = \$184$ y un desvío estándar muestral $s = \$80$. Se dibuja un histograma con los alquileres registrados en la muestra y se ve que no sigue la curva normal.

- (a) Si es posible, halle un intervalo de confianza del 68% aproximadamente para el alquiler medio de las casas de alquiler ocupadas en Ciudad Gótica. Si no es posible, justifique.
- (b) Indique si es verdadero o falso: Para alrededor del 68% del total de casas de alquiler ocupadas en la ciudad, el alquiler fue de entre \$180 y \$188.

3. Una balanza comete errores aleatorios que siguen la densidad que se muestra en el gráfico:



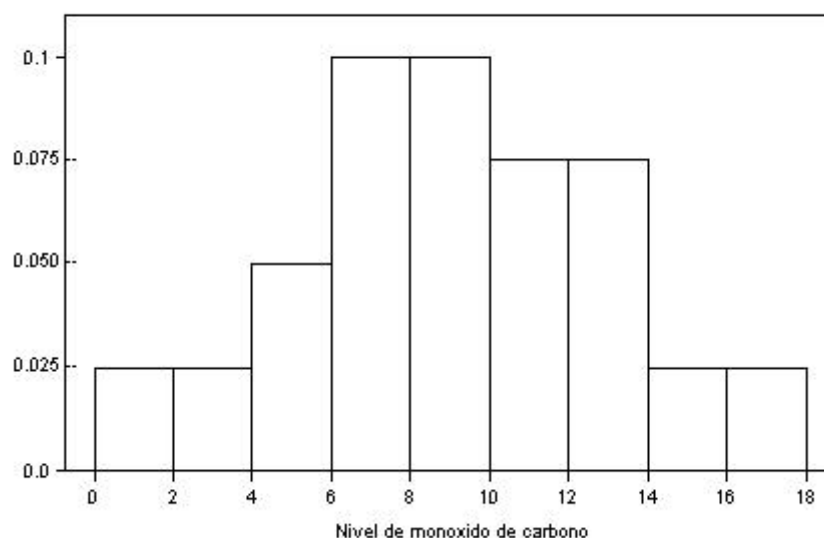
La esperanza de esta densidad es 0 microgramos y su desviación estándar es 6 microgramos. Se hacen 4 mediciones de un peso y se requiere un intervalo de confianza (exacto o aproximado) para el peso verdadero.

- (a) Para hallar el intervalo de confianza, ¿puede usarse la curva normal? ¿Y la t de Student? Justifique.
- (b) Ídem con 100 mediciones. Dé el intervalo correspondiente para un nivel del 95% e indique si es exacto o aproximado.
- (c) Se hacen 100 mediciones, en 57 de las cuales se comete un error positivo. Hallar un intervalo de confianza de nivel 95% aproximadamente para la proporción de mediciones en las que se cometen errores positivos. Investigar las siguientes instrucciones de R, y compararlas con la resolución a mano del ejercicio.

```
prop.test(x = 57,n = 100,correct=TRUE)
prop.test(x = 57,n = 100,correct=FALSE)
binom.test(x = 57,n = 100)
```

4. Para un estudio de mercado 277 personas degustan un nuevo licor y 69 de ellas desaprueban el nuevo sabor. Construya un intervalo de aproximadamente 95% de confianza para la verdadera proporción poblacional p de personas que aprueban el nuevo licor.
5. Un fabricante asegura a una compañía (que le compra un producto en forma regular) que el porcentaje de productos defectuosos no es mayor del 5%. La compañía decide comprobar la afirmación del comerciante seleccionando al azar de su inventario 200 unidades de este producto y probándolas. En la muestra encuentran 19 unidades defectuosas.
 - (a) Construya un intervalo del 95% aproximadamente de confianza para la verdadera proporción de unidades defectuosas.
 - (b) ¿Tiene razones la compañía para sospechar de la afirmación del fabricante? Justifique.
6. Se desea construir un intervalo de confianza para la media μ del puntaje de cierta prueba de lectura diseñada para los alumnos de tercer grado de la ciudad de Buenos Aires. Supongamos que, de estudios anteriores, se puede inferir que el desvío estándar poblacional es $\sigma = 12$.

- (a) Se extrae una muestra al azar de 100 niños de tercer grado de la ciudad de Buenos Aires, a los que se les toma la prueba. Calcule la longitud del intervalo de confianza de nivel 95% basado en dicha muestra. El intervalo de confianza, ¿es exacto o aproximado? ¿Que hipótesis deben satisfacer las variables involucradas?
- (b) Ídem (a) pero si el presupuesto permitiese sólo evaluar a 10 niños.
- (c) Halle el menor valor de n para el cual se satisface el requerimiento de los investigadores de obtener un intervalo del 95% de confianza con una longitud de a lo sumo 2. Especifique los supuestos utilizados.
7. En un debate sobre contaminación ambiental, se está evaluando la concentración de monóxido de carbono en las esquinas de la ciudad en los momentos de mayor tráfico (los viernes a la tarde). Interesa estimar p , la proporción de esquinas que presentan valores de monóxido inferiores a 12 ppm (partes por millón). Para ello, se recolecta una muestra de 120 valores correspondientes al nivel de monóxido de carbono en esquinas de la ciudad aleatoriamente elegidas un viernes a la tarde. Los datos obtenidos se representan en el siguiente histograma de densidad.



- (a) En base a ellos, la estimación de p resulta ser
- (b) Para completar el análisis de los datos del ítem anterior se decide calcular un intervalo de confianza de nivel aproximado 0.95 para p . Calcúlelo, y brinde los extremos numéricos de dicho intervalo: (.....).
8. Considere el ejercicio 6 de la Práctica de Estadística Descriptiva. Trabajaremos con las nubes Tratadas.
- (a) ¿Puede calcular un intervalo de confianza de nivel 0.95 (exacto o aproximado) para la cantidad esperada de agua caída de una nube tratada con un bombardeo de átomos? Para responder este ítem, recuerde lo realizado al analizar los datos en la práctica de descriptiva.
- (b) ¿Puede calcular un intervalo de confianza de nivel 0.95 para el valor esperado del logaritmo de la cantidad de agua caída de una nube tratada con un bombardeo de átomos? Resuelva usando R. A partir del intervalo calculado, ¿puede ahora responder positivamente al ítem (a)?
9. En un estudio se le pregunta a un grupo de personas si está de acuerdo con la frase *existe un único verdadero amor en la vida*. Se entrevista a 2625 individuos; 372 de los 1213 hombres están de acuerdo

con la frase mientras que 363 de las 1412 mujeres concuerdan con el enunciado. Determine si estos datos muestran diferencia entre la opinión, a favor de la afirmación, de hombres y mujeres.

10. Un estudio examina los efectos del chocolate a partir de los vasos sanguíneos de personas sanas. En el estudio aleatorizado, 31 personas recibieron 46 gramos de chocolate negro (que es naturalmente rico en flavonoides) todos los días durante dos semanas (grupo tratados). Por otra parte, 30 personas (grupo control) recibieron un placebo que consiste en chocolate negro con bajo contenido en flavonoides. A los participantes se les midió su salud vascular (por medio de dilatación mediada por flujo) antes y después de las dos semanas de estudio. Para el grupo que recibió el buen chocolate negro, el promedio de las diferencias (después-antes) de la dilatación medida fue de 1,3, con una desviación estándar de 2,32. El grupo control tuvo un cambio promedio de 0,96 (después-antes) con una desviación estándar de 1,58.
- (a) Identifique cuáles son las variables involucradas en el problema y el parámetro de interés.
 - (b) Calcule un intervalo de confianza asintótico estimado del 95% para el parámetro de interés. Interprete.
 - (c) ¿Es posible que no haya “diferencia alguna” entre los dos tipos de chocolate? Justifique su respuesta usando el intervalo de confianza encontrado en el ítem anterior.