Parcial 3 - Probabilidad y Estadística Fundamental

Wullfredo Javier Barco Godoy
Ángela Zoraya Cortés Yepes
Juan Manuel Garcia Mejia
Andrés Felipe Patiño Nivia
Jafeth Paz Cortés

Universidad Nacional de Colombia, Bogotá, Colombia

26 de enero de 2022

---

1. Saturación de gas residual

... Suponiendo normalidad de la población de donde se extrae la muestra, encuentre un intervalo de confianza de 98 % para la cantidad promedio verdadera de saturación de gas residual.

Contando con sólo 18 (n) datos, y asumiendo normalidad (que se puede comprobar con una gráfica cuantil-cuantil), se puede utilizar una distribución t. Calculamos primero la desviación estándar y el promedio:

$$\bar{x} = 38.661$$

$$\delta = 8.473$$

Con lo que se puede calcular (o consultar) el t crítico, sabiendo que tenemos 17 grados de libertad y buscamos un intervalo de confianza del 98 %:

$$t_c = 2.11$$

Así pues, calculamos nuestro intervalo:

$$\left(\bar{x} - t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}}\right) \tag{1}$$

$$(34.447, 42.875)$$

Para calcular el intervalo, en realidad, se utilizó una sola línea en R. Sin embargo, valía explicar el procedimiento.

2. Medios inalámbricos

... Calcule e interprete un intervalo de confianza del 95 % para la proporción de todos los adultos estadounidenses que en el momento de la encuesta habían usado medios inalámbricos para el acceso en línea.

Sabiendo que la muestra completa es de 2253 (n) adultos, tenemos:

$$p = \frac{1262}{2253} = 0.56$$

$$q = 1 - p = 0.44$$

De tal modo que podemos consultar en una tabla de distribución normal:

$$z = 1.96$$

Y así, el intervalo de confianza estará dado por

$$p \pm z \cdot \sqrt{\frac{p \cdot q}{n}}$$

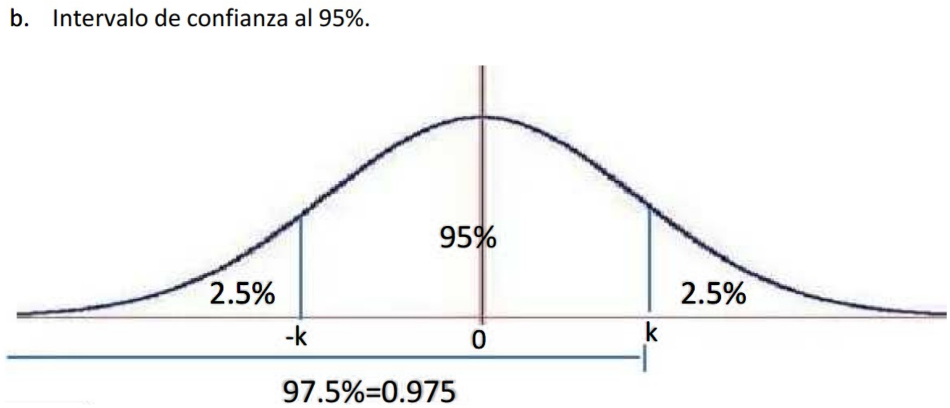
Resultando en:

$$(0.5395, 0.5804)$$

3. Pasajeros de aerolínea

... Encuentre un intervalo de confianza del 95 % para el verdadero promedio de viajeros entre Amsterdam y Viena.

$$\bar{x} = 136.22$$



$$P(Z \leq k) = 0.975$$
$$P(Z \leq k) = P(Z \leq 1.96) = 0.9750$$
$$k = \frac{X - \overline{x}}{s} = 1.96$$

RTA Intervalo de confianza al 95 % es : (-1.96; 1.96)

4. Ruptura de circuitos eléctricamente sobrecargados

En total tenemos 17 datos (n), además del promedio y la desviación estandar:

$$\bar{x} = 2126.47$$

$$\sigma = 370.575$$

Como vamos estamos buscando el 95 % de confianza recurrimos a la tabla de valores, donde vemos que para 95 % el valor más cercano es 1.96

$$z = 1.96$$

Luego usaremos la ecuación de para hallar el error estándar:

$$errorestandar = \frac{desviacion}{\sqrt{n}}$$

$$errorestandar = \frac{1.045}{\sqrt{17}}$$
$$errorestandar = 0.2534$$

Y luego con esta hallaremos los límites del intervalo:

$$Liminf = media - (z * errorestandar)$$

$$Liminf = 2126.47 - (1.96 * 89.87) = 1950.31$$
$$Limsup = 2126.47 + (1.96 * 89.87) = 2302.62$$

Concluimos que los valores del límite inferior y superior son respectivamente 1950.31 y 2302.62

5. Visita de animales domésticos al veterinario en un año

Recopilando los datos que nos da el enunciado tenemos:

$$\mu = 3.59$$

$$\sigma = 1.045$$

$$x_1 = 3.49$$

$$x_2 = 3.69$$

Por lo que podemos encontrar los valores z que les corresponde a cada uno de los límites del intervalo de confianza facilmente usando la ecuación:

$$z = \frac{x - \mu}{\sigma}$$

Usaremos el valor "3.69"pues sabemos que si el valor de es menor que el de μ el resultado será negativo, valor que no encontraremos en la tabla de frecuencias. La ecuación entonces será:

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{3.69 - 3.59}{1.045}$$

$$z = 0.095$$

El porcentaje al que le corresponde este valor en la tabla de frecuencias es 0.5359, como se muestra a continuación:

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	[illegible]
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319

Ahora, sabiendo que nos referimos al límite superior la suma del primer alpha, mas el intervalo de confianza menos 100 % es igual a b(donde a y b son las dos regiones $\alpha/2$):

$$a + IC = 53.59\,\%$$

$$b = 46.41\,\%$$

a y b son iguales, por lo que podemos concluir que:
% Intervalo de confianza (3.49,3.69)= 7.18 %
Estos datos fueron corroborados en Rstudio.

6. Asbesto y elasticidad pulmonar

... Suponiendo normalidad en la población, construya un intervalo de confianza de 99 % para la elasticidad pulmonar promedio verdadera después de la exposición.

Siendo similar al primer problema, con 16 datos, se siguió el mismo procedimiento. Sin embargo, se mencionarán todas las cantidades relevantes.

$$\bar{x} = 209.75$$

$$\delta = 24.156$$

$$t_c = 2.131$$

Usando la ecuación 1, tenemos:

$$(191.955, 227.545)$$

7. Secado de pintura

... Suponga que las mediciones representan una muestra aleatoria de una población normal y con base en esto calcule el intervalo de predicción del 95 % para el tiempo de secado de la siguiente prueba de pintura.

Tenemos, entonces, una muestra de 15 (n):

$$\bar{x} = 3.78$$

$$\delta = 0.97$$

Y, consultando una tabla de distribución normal:

$$z = 1.96$$

Así pues, tenemos que el intervalo estará dado por:

$$\bar{x} \pm z \cdot \sqrt{\frac{\delta}{\sqrt{n}}}$$

Resultando en:

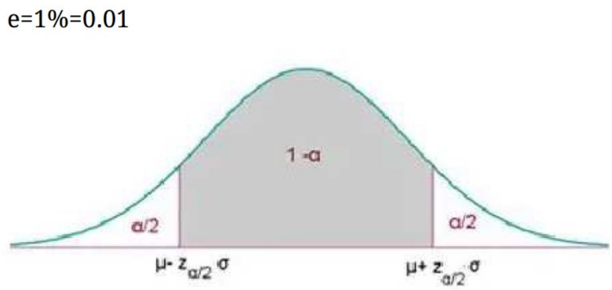
$$(4.270, 3.289)$$

8. Conjetura

... ¿Qué tan grande debería ser la muestra para estimar el porcentaje de habitantes de cierta ciudad que están a favor de tener agua fluorada, si se desea tener al menos un 99 % de confianza en que el estimado esta dentro del 1 % del porcentaje verdadero?

si utilizamos p como un estimado de p, podemos tener al menos un $100(1 - alpha)\,\%$ de confianza en que el errores no excediera una cantidad especifica e cuando el tamaño de la muestra sea

$$n = \frac{Z^2\alpha/2}{4e^2}$$



$$Z = \alpha/2 = Z0.005 = 2.575$$

$$n = \frac{2.575^2\alpha/2}{4 * 0.01^2} = 16576.7$$

la muestra debe ser aproximadamente de 16577

9. Dureza de Rockwell en cabeza de alfileres

Suponiendo una distribución normal, y con una muestra (n) de 12, el procedimiento es esencialmente igual al del primer problema, donde ya se explicó el método utilizado. Como tal, se calcularon el promedio, la desviación y el valor t-crítico. Así, utilizando la fórmula 1, tenemos el resultado:

$$(47.69, 48.89)$$

Vale la pena reiterar que este resultado se obtuvo utilizando la función *t.test* de Rstudio.

10. Rockwell, parte 2

Ahora necesitamos averiguar el intervalo de confianza del 99 % de σ (desviación estándar). Para ello, basta con sacar el intervalo para la varianza (σ^2) y sacarle raíz cuadrada a cada punto extremo.
Para ello, primero utilizamos la función *qchisq* que indica el valor crítico de chi cuadrada. Esta necesita:

- n=Número de datos
- df=Grados de libertad, que son n-1
- lower.tail=Determinar si es cola izquierda o derecha

Ya con esto, se crea una función para calcular tanto el límite inferior como el superior, con los valores que obtuvimos de *qhisq* y la multiplicación entre los grados de libertad y la varianza de la muestra (σ^2)
Sacando raíz, tenemos que el intervalo es:

$$(0.8106, 2.0102)$$
