UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

MAESTRÍA EN CIENCIAS DE LA INFORMACIÓN Y LAS TELECOMUNICACIONES

Herramientas matemáticas para el manejo de la información

Parcial #1

Salazar Ortiz, Jaiver
jesalazaro@correo.udistrital.edu.co
20221495012

FECHA: *16 de enero de 2024*
FECHA DE ENTREGA: *21 de mayo de 2022*

1. SOLUCIÓN

Dado que son 500 personas, 210 son hombres, 100 son mujeres, se puede encontrar el numero de niños:

$$N_n = 500 - 210 - 100 = 190 \tag{1.1}$$

Por tanto encontrando la probabilidad:

$$P(N_n) = \frac{190}{500} = 0{,}38 \tag{1.2}$$

La respuesta es B.

2. SOLUCIÓN

Se calcula la probabilidad de un trabajo mal:

$$\begin{aligned} Realizados\ por\ A = R_a = 40\,\% = \frac{2}{5} \\ Realizados\ por\ B = R_b = 60\,\% = \frac{3}{5} \end{aligned} \tag{2.1}$$

Probabilidad de un trabajo mal para cada opción:

$$\begin{aligned} P\left(\frac{T_m}{A}\right) = 0{,}01 \\ P\left(\frac{T_m}{B}\right) = 0{,}02 \end{aligned} \tag{2.2}$$

Por tanto para el trabajo mal T_m:

$$\begin{aligned} P(T_m) = P\left(\frac{T_m}{A}\right) \times R_a + P\left(\frac{T_m}{B}\right) \times R_b \\ P\left(\frac{A}{T_m}\right) = \frac{P(\frac{T_m}{A}) \times R_a}{P(T_m)} = \frac{0{,}01 \times \frac{2}{5}}{0{,}016} = 0{,}25 \end{aligned} \tag{2.3}$$

Se concluye que la respuesta es C

3. SOLUCIÓN

La variable nominal es X(colores preferidos) y su moda es rojo dado que es el valor con mayor frecuencia. Por tanto la respuesta es A

4. SOLUCIÓN

La variable ordinal es Y(estrato) y su moda es medio dado que es el valor con mayor frecuencia. Por tanto la respuesta es B

5. SOLUCIÓN

La variable discreta es Z(número de personas), para calcular el percentil 80 % se uso R mediante el código:

```
z  = c(3,3,4,4,5,5,4,3,2,4)
percentile80 = quantile(z, c(.80))
[1]4.2
```

Con lo cuál ese concluye que el percentil 80 % es 5 dado que los percentiles se aproximan al entero superior, por tanto la respuesta es D.

6. SOLUCIÓN

La variable continua es W(edad, vista como medida), para calcular el valor promedio y el error estándar del valor promedio se uso R junto con la biblioteca plotrix:

```
library(plotrix)
w  = c(18,20,21,18,20,20,23,19,20)
mean(w)
[1] 19.88889
std.error(w)
[1] 0.5121969
```

Por tanto se concluye que el resultado es 19.89 para el valor promedio y 0.5122 para el error estándar del valor promedio, en consecuencia la respuesta es C.

7. SOLUCIÓN

Haciendo uso de la respuesta del punto anterior y calculando la desviación estándar:

```
w  = c(18,20,21,18,20,20,23,19,20)
sd(w)
[1] 1.536591
```

Se puede concluir que la respuesta es E, dado que se tiene una desviación estándar de 1.536

8. SOLUCIÓN

Dado que se tiene una confianza del 99 % el valor del α es 0.01 por obvias razones, debido a que es necesario calcular el intervalo de confianza se va a hacer uso de R mediante el siguiente código:

```
library(distributions3)
n = length(z)
T_9 = StudentsT(df = 9)
mean(z) + quantile(T_9, 0.01 / 2) * sd(z) / sqrt(n)
[1] 2.725049
mean(z) - quantile(T_9,  0.01 / 2) * sd(z) / sqrt(n)
[1] 4.674951
```

Por comprobar también se puede realizar de la siguiente manera:

```
t.test(z, conf.level = 0.99)

        One Sample t-test

data:  z
t = 12.333, df = 9, p-value = 6.095e-07
alternative hypothesis: true mean is not equal to 0
99 percent confidence interval:
 2.725049 4.674951
sample estimates:
mean of x
      3.7
```

De lo anterior se tiene que los inviértalos de confidencia son 2.72 y 4.67 por tal razón la respuesta que más concuerda con los resultados es A.

9. SOLUCIÓN

Planteando la diferencia de valores absolutos en R:

```
w = c(18,20,21,18,20,20,23,19,20)
e = c(18,20,20,21,22,20)
abs(mean(w)-mean(e))
[1] 0.2777778
```

Por tanto la respuesta es 0.277 - A.

10. SOLUCIÓN

Debido a que se tienen una confianza del 95 % el valor de α es 0.05, por tanto se puede usar la función defecto t.test de R

```
w <- c(18,20,21,18,20,20,23,19,20)
e <- c(18,20,20,21,22,20)
t.test(w, e, alternative = "two.sided", var.equal=FALSE)

        Welch Two Sample t-test

data:  w and e
t = -0.37226, df = 11.95, p-value = 0.7162
alternative hypothesis: true difference in means
is not equal to 0
95 percent confidence interval:
 -1.904327  1.348771
sample estimates:
mean of x mean of y
 19.88889  20.16667
```

Como se puede observar $p - valor > \alpha$ por tanto se acepta la hipótesis nula, es decir se puede afirmar que las proporciones son prácticamente iguales, y afirmar que la respuesta correcta es A.

11. SOLUCIÓN

Dado que es una distribución binomial es conveniente usar la función pbinom de R:

```
hereford = 3
extraen = 3
#probabilidad de obtener un hereford
probabilidad = hereford/12
#la probabilidd de obtener 1 perro en los 3 intentos:
pbinom(1, 3 , prob=probabilidad)
[1] 0.84375
```

Como se puede observar la respuesta es 0.84, debido a que existen dos opciones con el resultado se opta por la expresión que más sentido tiene, es decir la B.

12. SOLUCIÓN

Cuando se realizan pruebas de hipótesis y se asume un nivel de significación, dado nivel nos indica que existe una probabilidad de que podamos equivocarnos, en este caso como lo menciona la pregunta el nivel de significación es de 0.05, por tanto pasándolo al convertirlo a porcentaje se tiene un 5 % de probabilidad de equivocarse si se concluye que las poblaciones son iguales, por tal razón la respuesta es A.

13. SOLUCIÓN

Los parámetros son medidas descriptivas de toda una población. Los parámetros son constantes fijas, es decir, no varían como las variables. Sin embargo, sus valores generalmente se desconocen porque no es factible medir una población completa, si se conocen los parámetros para una distribución dada es posible conocer completamente la distribución, Por dichas razones la respuesta es F.

14. SOLUCIÓN

Debido a que la distribución es simétrica, la media, moda y mediana toman el mismo valor, por tal razón la respuesta es la B

15. SOLUCIÓN

Dado que se tiene la función de densidad es necesario integrar para encontrar la distribución de probabilidad acumulada:

$$\begin{aligned} f(x) &= 4x^3 \quad [0, b] \\ \int_0^b 4x^3 dx &= \left[4\frac{x^4}{4}\right] = \left[x^4\right]_0^b \end{aligned} \tag{15.1}$$

Por tanto la respuesta es A.

16. SOLUCIÓN

Primero:

$$\begin{aligned} \mu_{\bar{x}} &= \mu = 381 \\ \sigma_{\bar{x}} &= \frac{\sigma}{\sqrt{250}} = 5{,}3758 \end{aligned} \tag{16.1}$$

Debido a teorema central del calculo se puede realizar:

$$\begin{aligned} P(375 \leq \bar{X} \leq 385) = P\left(\frac{375-\mu_{\bar{x}}}{\sigma_{\bar{x}}} \leq Z \leq \frac{385-\mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) \\ = P\left(\frac{375-381}{5,3758} \leq Z \leq \frac{385-381}{5,3758}\right) \\ = P(-1{,}116 \leq Z \leq 0{,}744) \\ = P(Z \leq 0{,}74) - P(Z \leq -1{,}1) \\ = 0{,}7704 - 0{,}1562 \\ = 0{,}6142 \end{aligned} \tag{16.2}$$

Por tanto la respuesta es la opción E.

17. SOLUCIÓN

La desviación estándar es :

$$SE = \sqrt{\frac{p(1-p)}{n}} \tag{17.1}$$

Por tanto por analogía la respuesta debería ser:

$$std(\pi) = \sqrt{\frac{\pi_0(1-\pi_0}{n}} \tag{17.2}$$

por tanto la respuesta es C.

18. SOLUCIÓN

Dado que se pueden encontrar los grados de libertad basados en el tamaño de la muestra y que se conoce el valor de α, es conveniente usar la función qt de R:

```
qt(p=0.05, df=5, lower.tail=FALSE)
[1]2.015048
```

Por tanto la respuesta es A.

19. SOLUCIÓN

Para un nivel de confianza del 95 % se sabe que $\alpha = 0{,}05$, por tanto:

$$\begin{aligned} \frac{\alpha}{2} &= 0{,}025 \\ Z_{\frac{\alpha}{2}} &= 1{,}96 \\ P(Z > Z_{\frac{\alpha}{2}} &= \frac{\alpha}{2} \\ P(Z > 1{,}96) &= 0{,}025 \end{aligned} \tag{19.1}$$

Por tanto dado que la solución no esta se concluye que la respuesta es E.

20. SOLUCIÓN

Como se puede observar en el grafico las plantas con altura de inserción entre 70 y 80 cm son menos del 15 % del total, esto se puede intuir por el alto de la barra dado que o llega a la mitad de 10 % y 20 %, por tanto la respuesta es C.

21. SOLUCIÓN

Es necesario entender los valores:

$$\begin{aligned} 60\,\%\ &\textit{del total es de la clase A} \\ 40\,\%\ &\textit{del total es de la clase B} \\ 20\,\%\ &\textit{de A no reacciona con nitratos} \\ 40\,\%\ &\textit{de B no reacciona con nitratos} \end{aligned} \tag{21.1}$$

Por tanto encontrando los valores de probabilidad de no reacción para cada clase de bacteria y sumando:

$$\begin{aligned} P_a &= 0{,}6 \times 0{,}2 = 0{,}12 \\ P_b &= 0{,}4 \times 0{,}4 = 0{,}16 \\ P_a + P_b &= 0{,}12 + 0{,}16 = 0{,}28 \end{aligned} \tag{21.2}$$

Por tanto la respuesta es B

22. SOLUCIÓN

Simplemente sera la suma de la probabilidad de que sea B junto con la probabilidad total de que reaccione con nitratos :

$$0{,}28 + 0{,}4 = 0{,}68 \tag{22.1}$$

Por consiguiente la respuesta es A

23. SOLUCIÓN

Dado que se quiere observar diferencias entre las tasaciones es conveniente realizar un t.test two.sided en R con el 99 % de confianza:

```
tasadora = c(87, 107, 65, 79, 110, 128, 59, 150, 110, 75)
tasadorb = c(93, 101, 80, 70, 114, 132, 55, 160, 112, 71)
sdesviation = 7.550
t.test(tasadora, tasadorb, alternative = "two.sided",
var.equal = FALSE, conf.level = 0.99, sd=sdesviation)

        Welch Two Sample t-test

data:  tasadora and tasadorb
t = -0.13194, df = 17.841, p-value = 0.8965
alternative hypothesis: true difference in means
is not equal to 0
99 percent confidence interval:
 -41.10866  37.50866
sample estimates:
mean of x mean of y
     97.0      98.8
```

como se puede observar los valores de intervalos de confidencia son -41.10 y 37.50, por tal razón la respuesta es D.

24. SOLUCIÓN

La variable predictora es el consumo de sal debido a que se puede establecer y modificar, además de ello debería alterar la variable respuesta en este caso la tensión arterial dado que es la que se quiere medir. Por dicha razón la respuesta es A.