

Actividad Individual

FECHA DE ENTREGA : 4/3/2024

NOMBRE : STEVEN EGAS

ASIGNATURA : ESTADISTICA

NRC : 1270

TEMA:

Variable de Estudio

¿En qué porcentaje considera usted que en la Universidad de las Fuerzas Armadas - ESPE se usa ChatGPT? (valor decimal Ej: 67.5)



SANGOLQUI-ECUADOR 2024.



¿En qué porcentaje considera usted que en la Universidad de las Fuerzas Armadas - ESPE se usa ChatGPT? (valor decimal Ej: 67.5)

ANÁLISIS ESTADÍSTICOS REALIZADOS

1.1 ANOVA de un Factor

Planteamiento del Problema: Se desea determinar si existe una diferencia significativa entre los promedios de gasto en cuentas de pago de ChatGPT en los últimos seis meses entre las carreras de Software, Biotecnología y Mercadotecnia.

Datos y Resultados:

| Software | Biotecnología | Mercadotecnia |
|----------|---------------|---------------|
| 101 | 78 | 68 |
| 64 | 97 | 47 |
| 110 | 58 | 66 |
| 70 | 62 | 97 |



| n | Decisi | Valor P | Estadístico F | |
|------------|---|-------------|---------------|--|
| n | Se acepta H0; los promedios de dinero gastado s | | | |
| s . | igual | 0.510416949 | 0.72538206 | |

Conclusión: El análisis ANOVA permite determinar si las diferencias en los promedios de gasto son significativas o no.

1.2 ANOVA de Dos Factores

Planteamiento del Problema: Se investiga si la preferencia de los estudiantes por ciertas características de ChatGPT (facilidad de uso, rapidez, accesibilidad, calidad) influye en el gasto realizado en la plataforma.

Datos y Resultados:

| | | Respuesta | | |
|---------------|------------------|-----------|---------------|---------|
| | Facilidad de Uso | Rápida | Accesibilidad | Calidad |
| Software | 74 | 47 | 39 | 68 |
| Biotecnología | 43 | 47 | 46 | 57 |
| Mercadotecnia | 104 | 31 | 50 | 75 |

| Factor | Estadístico F | Valor P | Decisión |
|------------------|---------------|-------------|--------------|
| Facilidad de Uso | 0.85807392 | 0.406708215 | Se acepta H0 |
| Respuesta Rápida | 4.38861353 | 0.104251728 | Se acepta H0 |
| Accesibilidad | 4.093267449 | 0.113087035 | Se acepta H0 |
| Calidad | 1.934298593 | 0.236663189 | Se acepta H0 |

Conclusión: Se analiza si la preferencia en ciertas características afecta el gasto en ChatGPT de cada carrera.

1.3 Regresión Lineal Múltiple

Planteamiento del Problema: Se quiere determinar si el número de horas de estudio de los estudiantes afecta la cantidad de dinero gastado en cuentas premium de ChatGPT.

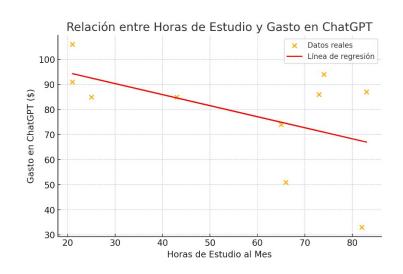


Datos y Resultados:

| Horas de Estudio | Gasto en ChatGPT | Gasto Predicho |
|------------------|------------------|----------------|
| 83 | 87 | \$67.00 |
| 25 | 85 | \$92.55 |
| 65 | 74 | \$74.93 |
| 43 | 85 | \$84.62 |
| 73 | 86 | \$71.40 |
| 74 | 94 | \$70.96 |
| 21 | 106 | \$94.31 |
| 21 | 91 | \$94.31 |
| 66 | 51 | \$74.49 |
| 82 | 33 | \$67.44 |

| Coeficiente | Coeficiente de | | | | |
|-------------|----------------|--------------|--------------|-------|-----------------|
| de | Determinación | | | Valor | |
| Correlación | (R^2) | Pendiente | Intersección | P | Decisión |
| | | | | | No hay relación |
| -0.51604722 | 0.266304733 | -0.440518869 | 103.5606934 | 0.127 | significativa |

Conclusión: Se evalúa si existe una relación lineal entre las horas de estudio y el gasto en ChatGPT.





1.4 Prueba del Signo

Planteamiento del Problema: Se analiza si una campaña publicitaria ha incrementado el uso de ChatGPT entre los estudiantes.

Datos y Resultados:

| Uso Antes | Uso Después | Diferencia | Signo |
|-----------|-------------|------------|-------|
| 31 | 16 | -15 | - |
| 35 | 23 | -12 | - |
| 52 | 56 | 4 | + |
| 62 | 72 | 10 | + |
| 33 | 41 | 8 | + |
| 42 | 46 | 4 | + |
| 94 | 81 | -13 | Ī |
| 81 | 62 | -19 | Ī |
| 53 | 47 | -6 | ļ |
| 42 | 56 | 14 | + |

| | | Signos | Signos |
|-----------------------------|------------|-----------|-----------|
| Decisión | Valor P | Negativos | Positivos |
| Se acepta H0; la campaña no | | | |
| aumenta el uso. | 0.62304688 | 5 | 5 |

Conclusión: Se determina si la publicidad influyó en el aumento del uso de ChatGPT.



Kruskal Wallis

Se realizó una encuesta en la Universidad de las Fuerzas Armadas "ESPE" acerca del uso de chat gpt entre estudiantes de Ing. en Software, Ing Biotecnologia , Mercadotecnia . En este ejercicio se quiere saber si la publicidad influye de manera diferente entre los estudiantes de las 3 carreras, para ello se toman muestras pequeñas independientes de cada carrera, tomando un nivel de significancia de 0,05 y la siguiente tabla.

| Ing. en Software | Ing.Biotecnologia | Lic. en Mercadotecnia |
|---------------------|-------------------|-----------------------|
| 78 | 85 | 30 |
| 72 | 6 | 73 |
| 20 | 88 | 45 |
| 78 | 58 | 33 |
| 89 | 68 | 55 |
| 80 | 6 | 7 |
| 6 | 80 | 87 |
| 70 | 75 | 40 |
| | 74 | 7 |
| | 6 | |

1 : Hipótesis

HO: No hay diferencias significativas entre los grupos (las distribuciones son iguales).

H1: Al menos un grupo tiene una distribución diferente.



2: reorganización, recopilación de datos y asignación de rango

| Ing. en Software | Rango |
|------------------|-------|
| 6 | 2,5 |
| 20 | 7 |
| 70 | 15 |
| 72 | 16 |
| 78 | 20,5 |
| 78 | 20,5 |
| 80 | 22,5 |
| 89 | 27 |
| | 131 |

| Ing.Biotecnologia | Rango |
|-------------------|-------|
| 6 | 2,5 |
| 6 | 2,5 |
| 6 | 2,5 |
| 58 | 13 |
| 68 | 14 |
| 74 | 18 |
| 75 | 19 |
| 80 | 22,5 |
| 85 | 24 |
| 88 | 26 |
| | 144 |
| | |

| Lic. en Mercadotecnia | Rango |
|-----------------------|-------|
| 7 | 5,5 |
| 7 | 5,5 |
| 30 | 8 |
| 33 | 9 |
| 40 | 10 |
| 45 | 11 |
| 55 | 12 |
| 73 | 17 |
| 87 | 25 |
| | 103 |

Paso 3: Calculo de Kruskal - Wallis utilizamos la siguiente formula

$$H = \frac{12}{n(n+1)} \left[\frac{(\sum r1)2}{n1} + \frac{(\sum r2)2}{n2} + \frac{(\sum r3)2}{n3} \right] - 3(n+1)$$

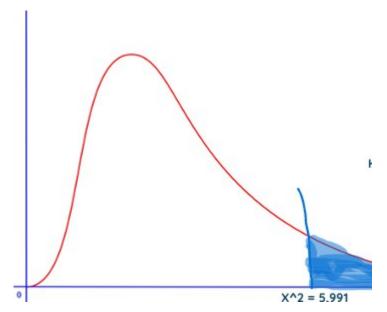
$$H = \frac{12}{27(27+1)} \left[\frac{(131)2}{8} + \frac{(144)2}{10} + \frac{(103)2}{9} \right] - 3(27+1)$$

$$H = 1.67$$

Paso 4: Valores Críticos. alpha = 0.05 En la tabla Ji cuadrada calculamos el grado de libertad calculando k-1 donde k es el numero de poblaciones entonces k-1=3-1=2

Valores Críticos. Por la tabla entendemos que el valor critico teniendo en cuenta una significancia de 0,05 es de 5,99





Paso 5: Regla de decisión Ho se acepta si H =< 5,991

Paso 6: Toma de decisión HO se acepta; No hay diferencias significativas entre los grupos

Prueba de Spearman

Se realizó una encuesta en la Universidad de las Fuerzas Armadas "ESPE" para conocer la relación entre las recomendaciones del uso de relojes de mano por parte de estudiantes de dos carreras: Ing. Software y Ing. Biotecnología Se seleccionó una muestra de 11 estudiantes de cada carrera, y se les pidió que calificaran, en una escala del 1 al 100, qué tan recomendable consideran el uso de chat gpt.

| Estudiante | Nivel de recomendación Software | Nivel de recomendación Biotecnologia |
|------------|---------------------------------|--------------------------------------|
| 1 | 65 | 75 |
| 2 | 83 | 80 |
| 3 | 95 | 55 |
| 4 | 55 | 66 |
| 5 | 90 | 88 |
| 6 | 70 | 45 |
| 7 | 65 | 60 |
| 8 | 64 | 90 |
| 9 | 50 | 62 |
| 10 | 80 | 78 |
| 11 | 45 | 50 |

Paso 1: Asignar rangos a cada conjunto de dato.



| Nivel de recomendación Software | Rangos |
|---------------------------------|--------|
| 45 | 1 |
| 50 | 2 |
| 55 | 3 |
| 64 | 4 |
| 65 | 5 |
| 65 | 6 |
| 70 | 7 |
| 80 | 8 |
| 83 | 9 |
| 90 | 10 |
| 95 | 11 |
| | |

| livel de recomendación Biotecnologia | Danges |
|---------------------------------------|--------|
| ilver de recomendación biolectiológia | Rangos |
| 45 | 1 |
| 50 | 2 |
| 55 | 3 |
| 60 | 4 |
| 62 | 5 |
| 66 | 6 |
| 75 | 7 |
| 78 | 8 |
| 80 | 9 |
| 88 | 10 |
| 90 | 11 |
| | |

Paso 2: Calcular diferencias de rangos y su cuadrado

| Nivel de | | Nivel de | | | |
|---------------|----------|---------------|---------------|----|-----|
| recomendación | | recomendación | | | |
| Software | Software | Biotecnologia | Biotecnologia | d | d2 |
| 65 | 5 | 75 | 7 | -2 | 4 |
| 83 | 9 | 80 | 9 | 0 | 0 |
| 95 | 11 | 55 | 3 | 8 | 64 |
| 55 | 3 | 66 | 6 | -3 | 9 |
| 90 | 10 | 88 | 10 | 0 | 0 |
| 70 | 7 | 45 | 1 | 6 | 36 |
| 65 | 6 | 60 | 4 | 2 | 4 |
| 64 | 4 | 90 | 11 | -7 | 49 |
| 50 | 2 | 62 | 5 | -3 | 9 |
| 80 | 8 | 78 | 8 | 0 | 0 |
| 45 | 1 | 50 | 2 | -1 | 1 |
| | | | | | 176 |

Paso 3: Aplicar la fórmula de Spearman



$$r_{\rm s}=1-\frac{6\Sigma d^2}{n(n^2-1)}$$

$$r_s = 1 \cdot \frac{6(176)}{11(11^2 \cdot 1)}$$

$$r_s = 0.20$$

Interpretación

El coeficiente de calificación de Spearman obtenido es 0.20, lo que indica una calificación positiva pero débil entre la recomendación del uso de chat gpt por parte de los estudiantes de Software y Biotecnología. Esto sugiere que, aunque existe una relación entre ambas carreras, no es lo suficientemente fuerte como para afirmar que las recomendaciones están altamente alineadas.

Paso 1: Calcular el coeficiente de correlación de Spearman.

$$r_s = 1 - \frac{6(176)}{11(11^2 - 1)}$$

$$r_s = 0.20$$

Paso 2: Planteamiento de hipótesis

$$H_0: r_s \leq 0$$

$$H_1: r_s > 0$$

Distribución de prueba: α=0.05



t de Student con n-2 grados de libertad (df=11-2=9).

(**Apéndice F**) t= 1.833

Paso 3: Cálculo del estadístico de prueba

$$t = r_s \sqrt{\frac{n-2}{1-r_s^2}}$$
 $t = 0.20 \sqrt{\frac{11-2}{1-0.20^2}}$ $t = 0.56$

Paso 4: Comparación con el valor crítico

Se acepta H_1 si t > 1,833. En este caso, t = 1.833 (según la tabla), lo que indica que estamos en el límite de significancia estadística.

Paso 5

Por lo tanto, hay indicios de una activación positiva entre los niveles de recomendación del uso de ChatGPT en los estudiantes de Software y Biotecnología en la Universidad de las Fuerzas Armadas "ESPE", aunque la evidencia no es contundente.

Chi Cuadrado

Frecuencias esperadas iguales

En la Universidad de las Fuerzas Armadas ESPE, se tiene un favoritismo por el uso de chat gpt, se recogió una muestra de 120 estudiantes para demostrar que las 4 versiones de chat gpt obtienen una alta calificación a la hora de recomendarlas en una escala del 1 al 100. Se espera que las 4 versiones tengan una alta calificación. Pero al revisar las calificaciones los encuestadores se dieron cuenta que no todos las versiones tenían una alta recomendación entonces ¿La diferencia de calificación de recomendación se debe al azar o existe algún favoritismo a algúna versión en particular? Trabaje con un nivel de significancia del 5%.

| Versión de ChatGPT | Personas que prefieren esta versión | Personas esperadas que prefieren esta versión. |
|-----------------------|-------------------------------------|--|
| ChatGPT-3 | 45 | 30 |
| ChatGPT- 3.5 | 33 | 30 |
| ChatGPT- | 28 | 30 |
| ChatGPT-4 Turbo | 14 | 30 |
| Total | 120 | 120 |



1. Plantear las hipótesis:

HO = NO existen diferencia entre el numero de frecuencias observadas y esperadas

Hl = Existe diferencia entre el numero de frecuencias observadas y esperadas.

2. Puntos críticos:

 α = 0.05 grados de libertad: k - 1 = 4- 1 = 3

3. Hallar el estadístico de prueba Ji- Cuadrado:

$$X^2 = \sum \left[\frac{(f_0 - f_\theta)^2}{f_\theta} \right] con \text{ k-1 } grados \text{ } de \text{ } libertad$$

| Versión de ChatGPT | Personas que prefieren esta versión | Personas esperadas que prefieren esta versión. | $f_0 - f_\theta$ | $(f_0 - f_\theta)^2$ | $\frac{(f_0 - f_\theta)^2}{f_\theta}$ |
|-----------------------|---|--|------------------|----------------------|---------------------------------------|
| ChatGPT-3 | 45 | 30 | 15 | 225 | 7,5 |
| ChatGPT-3.5 | 33 | 30 | 3 | 9 | 0,3 |
| Charlar | 28 | 30 | -2 | 4 | 0,13 |
| ChatGPT- | 14 | 30 | -16 | 256 | 8.53 |
| Total | 120 | 120 | | | 16.46 |

4. Realizar el gráfico de Ji-Cuadrado, colocar el punto crítico y el valor de Ji-Cuadrado (0.55)



5. Regla de decisión:

Ho se acepta si; $X^2 \le 7.815$



6. Decisión: Ho se rechaza; Las calificaciones otorgadas por los estudiantes de la Universidad de las Fuerzas Armadas "ESPE" a las versiones de chat gpt no tienden a tener una probabilidad igual de alta recomendacion en cada uno de sus versiones.

Frecuencias esperadas Diferentes

En la Universidad de las Fuerzas Armadas ESPE, se consumen papas sin marca algunas veces a la semana pero varian de carrera en carrera. Se analizo el porcentaje de chicas/os que estudian mecanica cuantas veces a la semana consumen papas sin marca

| Veces a la | | |
|------------|---|-------------|
| semana | | % del total |
| | 1 | 35 |
| | 2 | 20 |
| | 3 | 15 |
| | 4 | 13 |
| | 5 | 10 |
| | 6 | 7 |
| Total | | 100 |

¿Cómo puedo comprar valores porcentuales con valores numericos concetros?

| Veces a la semana | Número observado de uso (f _o) | Número esperado de uso (f ₆) |
|-------------------|---|--|
| 1 | 42 | 47 |
| 2 | 30 | 27 |
| 3 | 23 | 20 |
| 4 | 19 | 18 |
| 5 | 12 | 14 |
| 6 | 9 | 9 |

135

1) Hipótesis

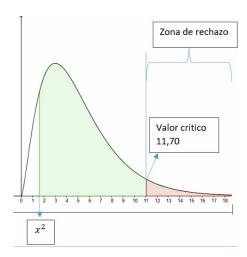
Ho= No existe diferencia entre el numero de compra a la semana de chat gpt entre los chicos de mecanica y de administración de empresas

H1= Si existe diferencia entre el numero de compra a la semana de chat gpt entre los chicos de mecanica y de administración de empresas



2) Para determinar la regla de decision se utiliza el apendice I. Existen 6 categorias de admision, por lo que los grados de libertad son gl=k-1= 6-1 =5 entonces el valor critico es: 11.070

| Veces a la semana | Número observado de uso (f _o) | Número esperado de uso (f _e) | f0-fef0 -fe | (f0-fe)2(f0 -fe)2 | (f0-fe)2fefe (f0-fe)2 |
|----------------------|---|--|----------------|----------------------|--------------------------|
| 1 | 42 | 47 | -5 | 25 | 0,532 |
| 2 | 30 | 27 | 3 | 9 | 0,333 |
| 3 | 23 | 20 | 3 | 9 | 0,45 |
| 4 | 19 | 18 | 1 | 1 | 0,056 |
| 5 | 12 | 14 | -2 | 4 | 0,286 |
| 6 | 9 | 9 | 0 | 0 | 0 |
| Σ | | 135 | | | 1,657 |



3) El valor de ji cuadrada queda a la izquierda del valor critico por lo cual se acepta la hipotesis nula etonces: Ho se acepta; No existe diferencia entre el numero de compra a la semana de papitas entre los chicos de mecanica y de administracion de empresas

Ajuste para probar Normalidad

En una encuesta a estudiantes de la Universidad de las Fuerzas Armadas ESPE se realizo la siguiente pregunta: "En los últimos seis meses, ¿cuántas veces aproximadamente utilizaron chat gpt ?" Se encuestó a 150 estudiantes de Ingeniería de Software, Ingeniería Biotecnologia , Mercadotecnia . Los datos se agruparon en intervalos.



| Veces que se usa ChatGPT | Frecuencia observada (fo) |
|-----------------------------|---------------------------|
| 0 a 10 | 12 |
| 10 a 20 | 28 |
| 20 a 30 | 45 |
| 30 a 40 | 38 |
| 40 a 50 | 20 |
| 50 a 60 | 7 |
| Total | 150 |

Paso 1: Calcular frecuencias esperadas bajo normalidad Supongamos que la media muestral (μ) es 28.5 y la desviación estándar (σ) es 12.3. Calcular valores z para los límites de clase:

$$\mathbf{z} = \frac{\mathbf{X} - \mathbf{\mu}}{\mathbf{\sigma}}$$
 $z = \frac{10 - 28.5}{12.3} \approx -1.50.$

Determinar áreas bajo la curva normal:

- 1. Usando tablas z, se calcula el área entre los límites de cada intervalo.
- 2. Calcular frecuencias esperadas (fe)

| Intervalo de uso de | | | |
|---------------------|-------------------------|--------|---------------------|
| ChatGPT | zz (inf sup.) | Área | fe≈fe≈ |
| | | | 0.0668 * 150 = |
| 0 a 10 | -∞ a-1.50-∞a-1.50 | 0.0668 | 10.02 |
| | | | 0.1783 * 150 = |
| 10 a 20 | -1.50 a-0.69-1.50a-0.69 | 0.1783 | 26.75 |
| | | | 0.3027 * 150 = |
| 20 a 30 | -0.69 a0.12-0.69a0.12 | 0.3027 | 45.41 |
| | | | 0.2760 * 150 = |
| 30 a 40 | 0.12 a 0.93 | 0.2760 | 41.40 |
| | | | 0.1361 * 150 = |
| 40 a 50 | 0.93 a 1.75 | 0.1361 | 20.42 |
| 50 a 60 | 1.75 a +∞+∞ | 0.0401 | 0.0401 * 150 = 6.02 |
| Total | | 10.000 | 150 |

Paso 2: Combinar categorías con fe

| Intervalo de uso de ChatGPT | f0f0 | fcfc | (f0-fc)2fcfc(f0-fc)2 |
|-----------------------------------|------|-------|--|
| 0 a 10 | | 10.02 | (12-10.02)210.02≈0.3910.02(12-10.02)2 ≈0.39 |
| 10 a 20 | 28 | 26.75 | (28-26.75)226.75≈0.0626.75(28-26.75)2 ≈0.06 |



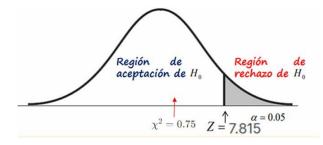
| 20 a 30 | 45 | 45.41 | (45-45.41)245.41≈0.0045.41(45-45.41)2 ≈0.00 |
|---------|----|--------|--|
| 30 a 40 | 38 | 41.40 | (38-41.40)241.40≈0.2841.40(38-41.40)2 ≈0.28 |
| 30 a 40 | 30 | 41.40 | 3.23 |
| 40 a 50 | 20 | 20.42 | (20-20.42)220.42≈0.0120.42(20-20.42)2 ≈0.01 |
| 50 a 60 | 7 | 6.02 | (7-6.02)26.02≈0.166.02(7-6.02)2≈0.16 |
| Total | | 150.00 | χ2=0.90χ2=0.90 |

Paso 4: Determinar grados de libertad y valor crítico Grados de libertad: k-p-1=6-2-1=3 (donde k=6 categorías y p=2 parámetros estimados: μ y σ). Valor crítico ($\alpha=0.05$): 7.815 (tabla ji-cuadrada)

H0: El consumo sigue una distribución normal

H1: El consumo no sigue una distribución normal

Conclusión El valor calculado ($\chi 2$ =0.90) es menor que el valor crítico (7.815). No se rechaza H0 . El consumo de papas sin marca sigue una distribución normal.



Tablas de Contingencia

En la Universidad de las Fuerzas Armadas ESPE se realizo una encuensta acerca del consumo de papas sin marca. se desea investigar si existe una relacion entre la edad de los encuestados y el porque prefieren consumir papas sin marca. para lo cual tenemos estos datos:

| n | edad | Porque prefiere utilizar |
|---|------|--------------------------|
| 1 | 20 | Calidad |
| 2 | 24 | Eficiente |
| 3 | 21 | Facilidad |
| 4 | 25 | Calidad |
| 5 | 23 | Calidad |
| 6 | 22 | Eficiente |
| 7 | 21 | Facilidad |
| 8 | 19 | Calidad |
| 9 | 20 | Facilidad |

Ho: No existe relacion entre la edad y la preferencia al momento de utilizar chat gpt.

H1: Existe relacion entre la edad y la peferencia al momento de utilizar chat gpt

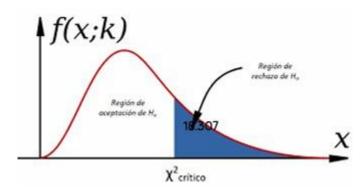


Segundo paso: identificamos filas y columnas en este caso 6 filas y 3 columnas. Ahora podemos determinar el valor crítico y regla de decisión. para una prueba de significancia de ji cuadrada donde se clasifican 2 características en una tabla de este tipo. se determinan grados de libertad por medio de:

$$gl = (número de rengiones - 1)(número de columnas - 1) = (r - 1)(c - 1)$$

En este problema seria: gl=(6-1)(3-1)=10 para el valor critico como el ejercicio no nos dice un valor de significancia tomamos 0.05

Planteamiento de la regla de desicion: Se acepta Ho si el valos calculado de ji cuadrado es menor que 18.307



$$F.esperada = \frac{(2)(3)}{9} = \frac{2}{3} = 0.66$$

$$\chi^2 = 10.000$$

Por último, como 10< 18.307 Ho se acepta; No hay relación significativa entre las variables.