Universidad del Valle de Guatemala
Departamento de Matemática
Licenciatura en Matemática Aplicada

Estudiante: Rudik Roberto Rompich
E-mail: rom19857@uvg.edu.gt
Carné: 19857

MM2040 - Estadística 2 - Catedrático: Eugenio Aristondo
24 de mayo de 2021

HT 5

1. Problema 1

1. M&M/MARS, fabricante de los chocolates M&M, realizó un sondeo nacional en el que más de 10 millones de personas dieron su preferencia para un nuevo color. El resultado de este sondeo fue el remplazo del color café claro por uno azul. En el folleto "Colors", distribuido por el área de Asuntos del Consumidor de M&M/Mars, la distribución de los colores de las lunetas (chocolates en forma de gragea) es la siguiente.

Café	Amarillo	Rojo	Naranja	Verde	Azul
30%	20%	20%	10%	10%	10%

En un estudio posterior se emplearon como muestras bolsas de 1 libra para determinar si los porcentajes reportados eran válidos. En una muestra de 506 lunetas se obtuvieron los siguientes resultados.

Café	Amarillo	Rojo	Naranja	Verde	Azul
177	135	79	41	36	38

Use α=0.05 para determinar si estos datos coinciden con los porcentajes reportados por la empresa.

Solución. Comenzamos planteando las hipótesis:

$$
\begin{aligned}
H_0 &: \text{la población tiene una distribución multinomial con la} \\
&\quad\ \text{probabilidad específica de cada una de las } k \text{ categorías.} \\
H_a &: \text{la población no tiene una distribución multinomial con la} \\
&\quad\ \text{probabilidad específica de cada una de las } k \text{ categorías.}
\end{aligned}
$$

Hacemos el análisis en Geogebra:

1. EL valor-p es 0.

2. Por la prueba del valor-p, 0<0.05. Por lo tanto la H_0 se rechaza, es decir que la población no tiene una distribución multinomial con la probabilidad específica de cada uno de los colores respectivos.

□

2. Problema 2

2. La National Sleep Foundation utilizó una encuesta para determinar si las horas de sueño por noche son independientes de la edad (Newsweek, 19 de enero de 2004). Las siguientes son las horas de sueño entre semana en una muestra de personas de 49 años de edad o menos, y en otra muestra de personas de 50 años de edad o más.

	HORAS DE SUEÑO				
EDAD	menos de 6	6 a 6.9	7 a 7.9	8 o mas	Total
49 o menos	38	60	77	65	240
50 o mas	36	57	75	92	260

Solución. Se identifica como un problema de independencia. Se consideran las hipótesis:

H_0 : la variable de las columnas es independiente de la variable de las filas.

H_a : la variable de las columnas no es independiente de la variable de las filas.

Haciendo el análisis en Geogebra:

Distribution Statistics

ChiSquared Test

Rows 2

Columns 4

Row % Column % Expected Count X² Contribution

	menos de 6	6 a 6.9	7 a 7.9	8 o más
49 o menos	38 35.52 0.1732 15.8333% 51.3514%	60 56.16 0.2626 25% 51.2821%	77 72.96 0.2237 32.0833% 50.6579%	65 75.36 1.4242 27.0833% 41.4013%
50 o mas	36 38.48 0.1598 13.8462% 48.6486%	57 60.84 0.2424 21.9231% 48.7179%	75 79.04 0.2065 28.8462% 49.3421%	92 81.64 1.3147 35.3846% 58.5987%
	74 14.8%	117 23.4%	152 30.4%	157 31.4%

Result

ChiSquared Test

df	3
X²	4.007
P	0.2607

1. El valor-p es 0.2607.
2. Considerando la prueba del valor-p, 0.2607>0.05, por lo tanto, H_0 no se puede rechazar; las horas del sueño son independientes a la edad.

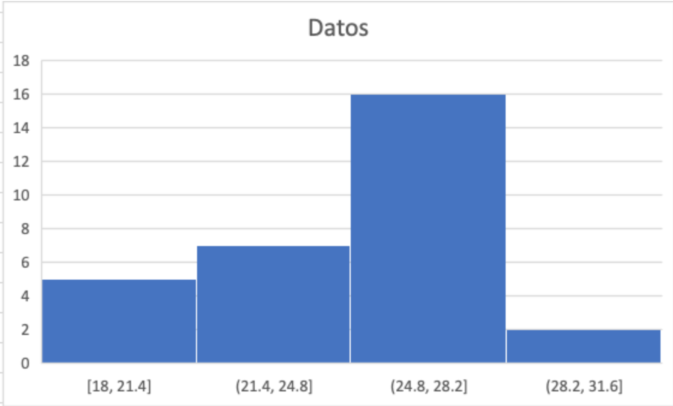
□

3. Problema 3

Se tiene la percepción de que la demanda semanal de un producto tiene una distribución normal. Aplique una prueba de bondad de ajuste y los datos siguientes para probar este supuesto. Use α=0.1.

18	20	22	27	22
25	22	27	25	24
26	23	20	24	26
27	25	19	21	25
26	25	31	29	25
25	28	26	28	24

Solución. En un vistazo inicial, la distribución se ve así:



Las hipótesis se formulan como:

$$H_0 : \text{la población tiene una distribución normal.}$$
$$H_a : \text{la población no tiene una distribución normal.}$$

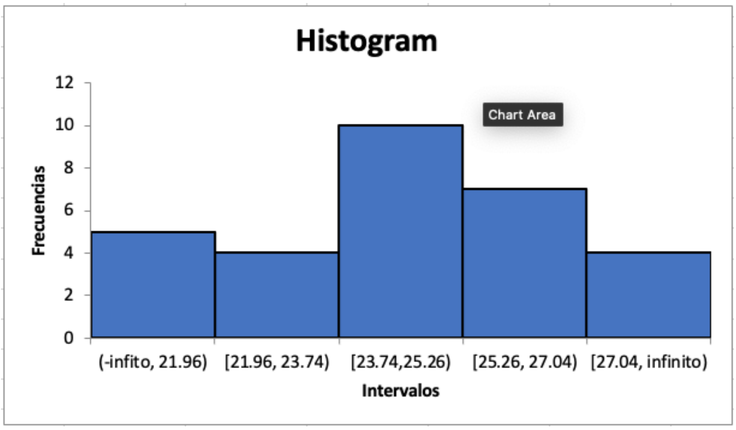
Se propone dividir los datos en 5 secciones de 20 % cada una. Es decir, que ahora tenemos:

Porcentaje	Separadores
0.2	21.96
0.4	23.74
0.6	25.26
0.8	27.04
Media	24.5
DEV	3.014

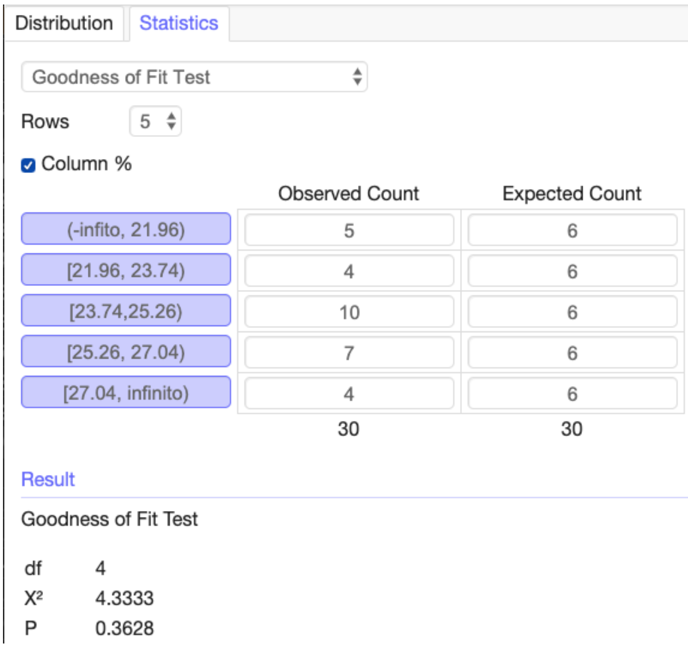
Es decir, que se tiene:

Intervalos	*Frecuencia observada*	*Frecuencia esperada*
(-infito, 21.96)	5	6
[21.96, 23.74)	4	6
[23.74,25.26)	10	6
[25.26, 27.04)	7	6
[27.04, infinito)	4	6

Gráficamente:



Haciendo el análisis en Geogebra:



Distribution Statistics

Goodness of Fit Test

Rows 5

Column %

	Observed Count	Expected Count
(-infito, 21.96)	5	6
[21.96, 23.74)	4	6
[23.74,25.26)	10	6
[25.26, 27.04)	7	6
[27.04, infinito)	4	6
	30	30

Result

Goodness of Fit Test

df	4
X²	4.3333
P	0.3628

1. El valor-p es 0.3628.
2. Por lo tanto, considerando la prueba del valor-p: 0.3628>0.1. Es decir que H_0 se acepta, entonces los datos tienen una distribución normal con una significancia de 0.1.

□

4. Problema 4

Se cree que el número de llamadas telefónicas que llegan por minuto al conmutador de una empresa tiene una distribución de Poisson. Use α=0.1 y los datos de la página siguiente para probar este supuesto.

Llamadas por minuto	Frecuencia observada
0	15
1	31
2	20
3	15
4	13
5	4
6	2

Solución. Primero, se hacen los cálculos pertinentes con Excel, usando la distribución de Poisson definida como:

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

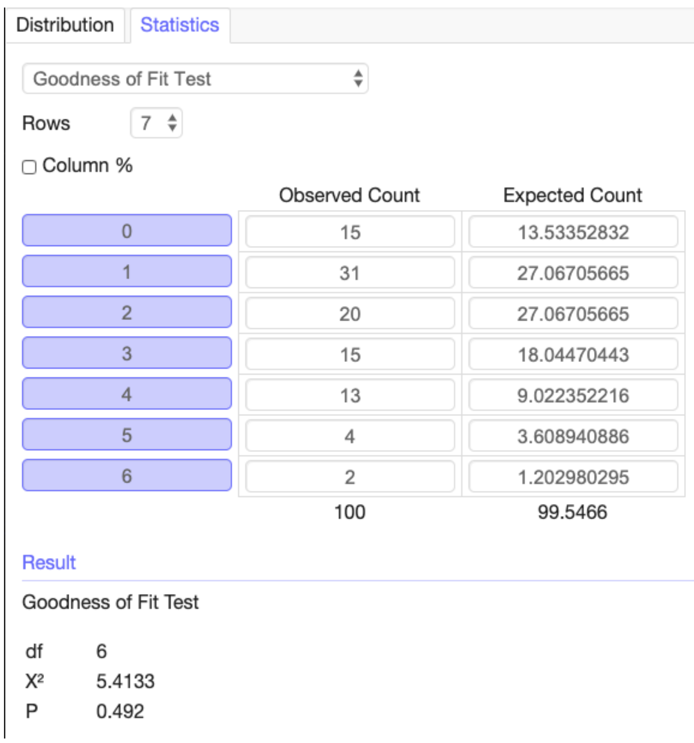
Por lo cual, se tiene:

Llamadas por minuto	Frecuencia observada	Llamadas	Probabilidad	Frecuencia esperada
0	15	0	0.135335283	13.53352832
1	31	31	0.270670566	27.06705665
2	20	40	0.270670566	27.06705665
3	15	45	0.180447044	18.04470443
4	13	52	0.090223522	9.022352216
5	4	20	0.036089409	3.608940886
6	2	12	0.012029803	1.202980295
Datos	100			
Media		2		

Definimos las hipótesis:

$$H_0 : \text{la población tiene una distribución de Poisson.}$$
$$H_a : \text{la población no tiene una distribución de Poisson.}$$

Usando Geogebra:



1. El valor-p es 0.492.
2. Por el método del valor-p: $0{,}492 > 0{,}1$; por lo que H_0 no se rechaza. Por lo tanto, podemos concluir que los datos sí se ajustan a una distribución de Poisson.

□