UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

Tercer examen parcial

Asignatura:

Estadística

P R E S E N T A

Celaya González David Alejandro	2
Garduño Pérez Ángel Isaac	15
Hernández García Aarón	19
Islas Pantoja Katia	22
Romero Vargas María Fernanda	40

Profesor

Barbosa Montes Héctor Ciro

Ciudad Universitaria, Cd. Mx., 2020

Índice

Objetivo: Aplicar los conocimientos adquiridos en la clase de estadística acerca de los temas de intervalos de confianza y pruebas de hipótesis mediante la resolución de seis problemas propuestos por los mismos alumnos, correspondientes a cada uno de los temas.

Intervalos de confianza

La estimación puntual es poco recomendable ya que no proporciona alguna información que indique la cercanía o lejanía del estadístico con respecto al parámetro a estimar, es por eso que una mejor opción para utilizar en la estadística es el uso de intervalos de confianza. Un intervalo de confianza se construye seleccionando primero un nivel o coeficiente de confianza que es representado por $1-\alpha$ y partiendo de in intervalo aleatorio con límite inferior L_1 y límite superior L_2 de manera que

$$P\left(L_1 \leq \theta \leq L_2\right) = 1-\alpha$$

En donde al menos uno de los límites es una variable aleatoria.

Cuando se desea construir un intervalo de confianza para la media aritmética en una población normal se tienen dos casos:

Caso 1

Intervalo de confianza para la media de una población normal cuando se conoce la varianza poblacional.

Utilizando la distribución normal estándar, se tiene

$$P\left(\acute{X}-z_{\frac{\alpha}{2}}\left(\frac{\sigma_x}{\sqrt{n}}\right)\leq\mu\leq\acute{X}+z_{\frac{\alpha}{2}}\left(\frac{\sigma_x}{\sqrt{n}}\right)\right)=1-\alpha$$

Caso 2

Intervalo de confianza para la media de una población normal cuando se desconoce la varianza poblacional.

Utilizando la distribución t de Student, se tiene

$$P\left(\acute{X}-t_{\frac{\alpha}{2},n-1}\left(\frac{S_{n-1}}{\sqrt{n}}\right)\leq\mu\leq\acute{X}+t_{\frac{\alpha}{2},n-1}\left(\frac{S_{n-1}}{\sqrt{n}}\right)\right)=1-\alpha$$

Cuando se desea construir un intervalo de confianza para comparar medias en dos poblaciones normales se tienen tres casos, los cuales comparan a las dos medias haciendo una diferencia entre ellas, para crear intervalos de confianza para diferencia de medias se tienen cuatro casos:

Caso 1

Intervalo de confianza para la diferencia de medias de dos poblaciones normales cuando se conocen las varianzas poblacionales.

$$\sigma_1^2,\sigma_2^2 n\leq 30\; o\; n>30$$

Utilizando la distribución normal estándar, se tiene que

$$P\left(\acute{X}_1-\acute{X}_2-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\leq\mu_1-\mu_2\leq\acute{X}_1-\acute{X}_2+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)=1-\alpha$$

Caso 2

Intervalo de confianza para la diferencia de dos poblaciones normales cuando se desconocen las varianzas poblacionales, pero se sabe que son iguales.

Utilizando la distribución t de Student, se tiene que

$P¿$

En donde $s_p=\sqrt{S_p^2}$

Y S_p^2 es el estimador combinado de la varianza y se calcula de la siguiente manera

$$S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

Caso 3

Intervalo de confianza para la diferencia de dos poblaciones normales cuando se desconocen las varianzas poblacionales, pero se sabe que son diferentes.

Utilizando la distribución t de Student, se tiene que

$$P\left(\acute{X}_1-\acute{X}_2-t_{\left(\frac{\alpha}{2},v\right)}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\leq\mu_1-\mu_2\leq\acute{X}_1-\acute{X}_2+t_{\left(\frac{\alpha}{2},v\right)}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right)=1-\alpha$$

En donde el cálculo de los grados de libertad se lleva a cabo de la siguiente manera

$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\left(\frac{S_1^2}{n_1}\right)^2 \left(\frac{1}{n_1 - 1}\right) + \left(\frac{S_2^2}{n_2}\right)^2 \left(\frac{1}{n_2 - 1}\right)}$$

Caso 4

Intervalo de confianza para la diferencia de medias de dos muestras pareadas. Las muestras pareadas son las que se toman al mismo sujeto.

$\mu_{antes} - \mu_{después}$

Cuando se desea construir un intervalo de confianza para la varianza, se hace uso de la distribución Ji cuadrada y el intervalo de confianza se construye de la siguiente manera

$$\frac{(n-1)\, s_{n-1}^2}{\chi^2_{\frac{\alpha}{2}(n-1)}} \leq \sigma^2 \leq \frac{(n-1)\, S_{n-1}^2}{\chi^2_{1-\frac{\alpha}{2},(n-1)}}$$

Y si se habla de la razón de dos varianzas, se utiliza el siguiente modelo

$$\frac{S_{n-11}^2}{S_{n-12}^2} \frac{1}{F_{\frac{\alpha}{2}(n_1-1, n_2-1)}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_{n-11}^2}{S_{n-12}^2} F_{\frac{\alpha}{2}(n_2-1, n_1-1)}$$

Para crear un intervalo de confianza para una proporción se hace uso de la distribución normal estándar y queda el intervalo de la siguiente forma

$$\hat{P} - z\frac{a}{2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \leq P \leq \hat{P} + z\frac{a}{2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

Mientras que para la diferencia de proporciones es el siguiente, utilizando también la distribución normal estándar

$$(\hat{P}_1 - \hat{P}_2) - z\frac{a}{2}\sqrt{\frac{\hat{P}, 11 - \hat{P}_2}{n_1}, + \frac{\hat{P}_2(1-\hat{P}_2)}{n_2}} \leq P_1 - P_2 \leq (\hat{P}_1 - \hat{P}_2) + z\frac{a}{2}\sqrt{\frac{\hat{P}1(1-\hat{P}1)}{n1}} + \sqrt{\frac{\widehat{P2}(1-\hat{P}2)}{n2}}$$

Ejercicio 1

Intervalo de confianza para la media con varianza desconocida, N=25 N.C.=95%

Una empresa manufacturera de ropa está realizando un estudio de mercado para determinar cuántas prendas se compran en México, se toma una muestra de 25 personas con un nivel de confianza de 95%, determine un intervalo de confianza para el público promedio si σ no se conoce

34	**25**	**25**	**20**	**20**
38	**15**	**30**	**25**	**25**
34	**24**	**23**	**20**	**20**
30	**20**	**20**	**25**	**30**
25	**15**	**30**	**25**	**25**

Usando

$$\acute{x}=\frac{1}{n}\sum_{i=1}^{25}x_i$$

$$\acute{x}_{MEX}=\frac{1}{25}(34+25+25+20+20+38+15+30+25+25+34+24+23+20+20+30+20+20+25+30+25+15+30+2$$

$$\acute{x}_{MEX}=24.92$$

Usando

$$S_{n-1}^2=\frac{1}{n-1}\sum_{i=1}^{25}\left(x_i-\acute{x}\right)^2$$

$$S_{n-1\,MEX}^2=\frac{1}{24}\left[(34-24.92)^2+(25-24.92)^2+(25-24.92)^2+(20-24.92)^2+(20-24.92)^2+(38-24.92)^2+(15-24.92)^2\right.$$

$S_{n-1\,MEX}^2=$¿32.74333

Usando

$$S_{n-1}=\sqrt{S_{n-1}^2}$$

$S_{n-1}=$¿5.72217;

Datos

n=25 $\acute{x}=24.92$; $S_{n-1}^{2}=¿32.74333$; $S_{n-1}=¿5.72217$; $1-\propto=0.95$; $\propto=0.05$; $\frac{\propto}{2}=0.025$

Usando

$$\acute{x}-t_{\frac{\propto}{2},(n-1)}\frac{S_{n-1}}{\sqrt{n}}\leq\mu\leq\acute{x}+t_{\frac{\propto}{2},(n-1)}\frac{S_{n-1}}{\sqrt{n}}$$

Valores de la probabilidad α área derecha de la distribución t-student (para los cuantiles se cambia el signo)

gl= n	0.0080	0.0085	0.0090	0.0095	0.010	0.015	0.020	0.025	0.030	0.035	0.040	0.045	0.050	0.055	0.060
1	39.780	37.439	35.359	33.496	31.821	21.205	15.894	12.706	10.579	9.058	7.916	7.026	6.314	5.730	5.242
2	7.810	7.572	7.353	7.151	6.965	5.643	4.849	4.303	3.896	3.578	3.320	3.104	2.920	2.760	2.620
3	4.930	4.821	4.721	4.628	4.541	3.896	3.482	3.182	2.951	2.763	2.605	2.471	2.353	2.249	2.156
4	4.010	3.937	3.870	3.806	3.747	3.298	2.999	2.776	2.601	2.456	2.333	2.226	2.132	2.048	1.971
5	3.573	3.516	3.462	3.412	3.365	3.003	2.757	2.571	2.422	2.297	2.191	2.098	2.015	1.941	1.873
6	3.320	3.272	3.226	3.183	3.143	2.829	2.612	2.447	2.313	2.201	2.104	2.019	1.943	1.874	1.812
7	3.157	3.113	3.073	3.034	2.998	2.715	2.517	2.365	2.241	2.136	2.046	1.966	1.895	1.830	1.770
8	3.043	3.003	2.965	2.930	2.896	2.634	2.449	2.306	2.189	2.090	2.004	1.928	1.860	1.797	1.740
9	2.958	2.921	2.886	2.853	2.821	2.574	2.398	2.262	2.150	2.055	1.973	1.899	1.833	1.773	1.718
10	2.894	2.859	2.825	2.794	2.764	2.527	2.359	2.228	2.120	2.028	1.948	1.877	1.812	1.754	1.700
11	2.843	2.809	2.777	2.747	2.718	2.491	2.328	2.201	2.096	2.007	1.928	1.859	1.796	1.738	1.686
12	2.801	2.769	2.738	2.709	2.681	2.461	2.303	2.179	2.076	1.989	1.912	1.844	1.782	1.726	1.674
13	2.767	2.736	2.706	2.677	2.650	2.436	2.282	2.160	2.060	1.974	1.899	1.832	1.771	1.715	1.664
14	2.739	2.708	2.678	2.651	2.624	2.415	2.264	2.145	2.046	1.962	1.887	1.821	1.761	1.706	1.656
15	2.714	2.684	2.655	2.628	2.602	2.397	2.249	2.131	2.034	1.951	1.878	1.812	1.753	1.699	1.649
16	2.693	2.663	2.635	2.609	2.583	2.382	2.235	2.120	2.024	1.942	1.869	1.805	1.746	1.692	1.642
17	2.675	2.645	2.618	2.592	2.567	2.368	2.224	2.110	2.015	1.934	1.862	1.798	1.740	1.686	1.637
18	2.658	2.630	2.603	2.577	2.552	2.356	2.214	2.101	2.007	1.926	1.855	1.792	1.734	1.681	1.632
19	2.644	2.616	2.589	2.564	2.539	2.346	2.205	2.093	2.000	1.920	1.850	1.786	1.729	1.677	1.628
20	2.631	2.603	2.577	2.552	2.528	2.336	2.197	2.086	1.994	1.914	1.844	1.782	1.725	1.672	1.624
21	2.620	2.592	2.566	2.541	2.518	2.328	2.189	2.080	1.988	1.909	1.840	1.777	1.721	1.669	1.621
22	2.610	2.582	2.556	2.532	2.508	2.320	2.183	2.074	1.983	1.905	1.835	1.773	1.717	1.665	1.618
23	2.600	2.573	2.547	2.523	2.500	2.313	2.177	2.069	1.978	1.900	1.832	1.770	1.714	1.662	1.615
24	2.592	2.565	2.539	2.515	2.492	2.307	2.172	2.064	1.974	1.896	1.828	1.767	1.711	1.660	1.612
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.970	1.893	1.825	1.764	1.708	1.657	1.610

Sustituyendo

$$24.92-2.064\frac{5.72217}{\sqrt{25}}\leq\mu\leq24.92+2.064\frac{5.72217}{\sqrt{25}}$$

$$22.55788\leq\mu\leq27.28211$$

El promedio de compra de prendas en México entre 22.55788 y 27.28211prendas con un nivel de confianza de 95% si la varianza es desconocida.

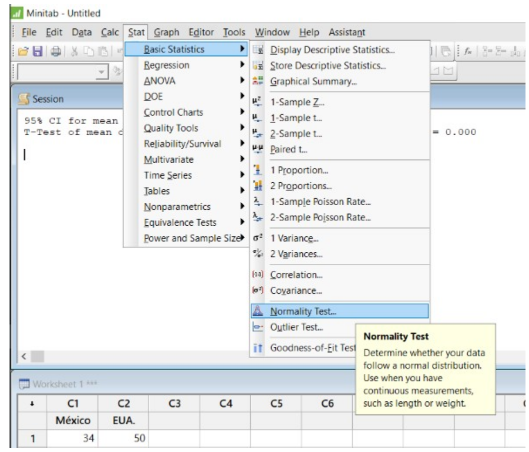
Solución por Minitab

Como primer paso necesitamos vaciar toda la información recopilada, asignando a esa columna un nombre para una mejor identificación.

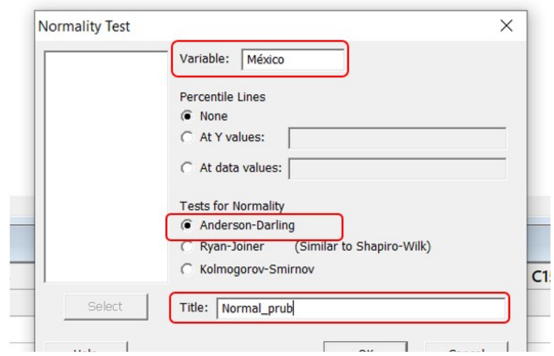
↓	C1
	México
1	34
2	25
3	25
4	20
5	20
6	38
7	15
8	30
9	25
10	25
11	34
12	24
13	23
14	20

15	20
16	30
17	20
18	20
19	25
20	30
21	25
22	15
23	30
24	25
25	25

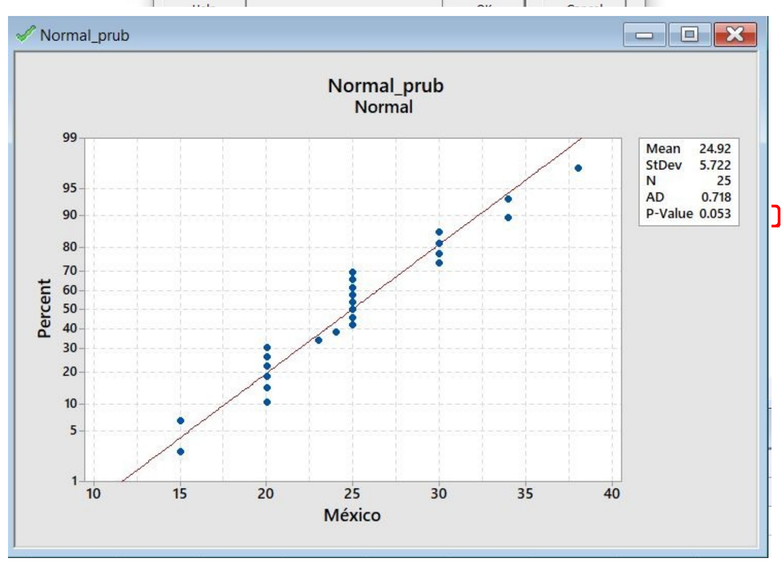
Es importante que verifiquemos que nuestros datos se comportan con una distribución normal. Entonces accediendo de la siguiente manera: Stat-> Basic Statistics-> Normaly Test.



En la ventana emergente, colocaremos la variable a evaluar, con el método 'Anderson_darling' y un título para su identificación y damos en la opción de 'Ok' para continuar.



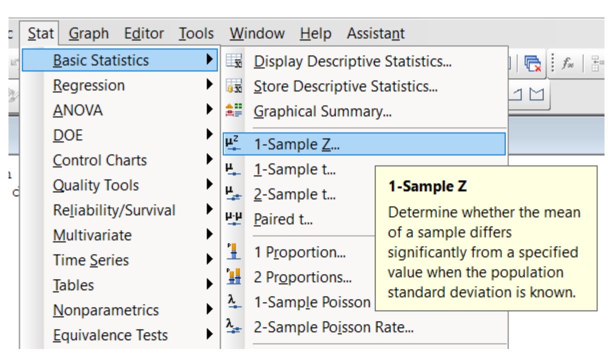
Nos fijaremos en el valor llamado P-Value, ya que este debe ser mayor a 0.05 para considerar una distribución normal. Nuestro caso es 0.053>0.05, por lo tanto, cumple.



Una vez asegurando una distribución normal de los datos, podemos continuar a realizar el intervalo de confianza.

Accederemos a la siguiente ruta:

Stat-> Basic Statistics->1-Sample-Z



Obtendremos los datos en la ventana de Session, en esta se nos mostrará la media y el intervalo de confianza, con un nivel de confianza del 95%.

En nuestro obtuvimos el intervalo de

$24.703 \leq \mu \leq 25.137$

One-Sample Z: México

```
The assumed standard deviation = 0.55377

Variable   N    Mean   StDev  SE Mean         95% CI
México    25  24.920   5.722    0.111  (24.703, 25.137)
```

Ejercicio 2

Intervalo de confianza para la diferencia de medias con varianza desconocida y diferentes tal que N_1=25 y N_2=25, nivel de confianza=96%

La misma empresa manufacturera de ropa vende tanto en México como en Estados Unidos, se desea comparar la diferencia de compras por año en cada país, para dicho estudio se obtuvieron las siguientes muestras:

México

34	**25**	**25**	**20**	**20**

38	**15**	**30**	**25**	**25**
34	**24**	**23**	**20**	**20**
30	**20**	**20**	**25**	**30**
25	**15**	**30**	**25**	**25**

Estados Unidos

50	**36**	**54**	**30**	**52**
30	**66**	**50**	**60**	**52**
8	**42**	**23**	**32**	**60**
37	**48**	**35**	**30**	**35**
56	**70**	**65**	**35**	**44**

Determine un intervalo de confianza para la diferencia promedio de compras de ropa por año de cada país si se sabe que sus varianzas son desconocidas y diferentes y con un nivel de confianza de 96%

Usando

$$\acute{x}=\frac{1}{n}\sum_{i=1}^{25} x_i$$

$$\acute{x}_{MEX}=\frac{1}{25}(34+25+25+20+20+38+15+30+25+25+34+24+23+20+20+30+20+20+25+30+25+15+30+2$$

$$\acute{x}_{MEX}=24.92$$

$$\acute{x}_{EUA}=\frac{1}{25}(50+36+54+30+52+30+66+50+60+52+8+42+23+32+60+37+48+35+30+35+56+70+65+35$$

$$\acute{x}_{EUA}=44$$

Usando

$$S_{n-1}^2=\frac{1}{n-1}\sum_{i=1}^{25}(x_i-\acute{x})^2$$

$$S_{n-1\,MEX}^2=\frac{1}{24}\left[(34-24.92)^2+(25-24.92)^2+(25-24.92)^2+(20-24.92)^2+(20-24.92)^2+(38-24.92)^2+(15-24.92)^2\right.$$

$$S_{n-1\,MEX}^2=¿32.74333$$

$$S_{n-1\,EUA}^{2}=\frac{1}{24}\left[(50-44)^{2}+(36-44)^{2}+(54-44)^{2}+(30-44)^{2}+(52-44)^{2}+(30-44)^{2}+(66-44)^{2}+(50-44)^{2}+(60-44)\right.$$

$S_{n-1\,EUA}^{2}=¿225.0833$

$n_1=25; n_2=25;\ 1-\propto=0.96;\ \propto=0.04;\ \frac{\propto}{2}=0.02$

Usando

$$\upsilon=\frac{\left(\frac{S_{n-1\,MEX}^{2}}{n_1}+\frac{S_{n-1\,EUA}^{2}}{n_2}\right)^{2}}{\left(\frac{S_{n-1\,MEX}^{2}}{n_1}\right)^{2}\frac{1}{n_1-1}+\left(\frac{S_{n-1\,EUA}^{2}}{n_2}\right)^{2}\frac{1}{n_2-1}}$$

$$\upsilon=\frac{\left(\frac{32.74333}{25}+\frac{225.0833}{25}\right)^{2}}{\left(\frac{32.7433}{25}\right)^{2}\frac{1}{24}+\left(\frac{225.0833}{25}\right)^{2}\frac{1}{24}}$$

$\upsilon=30.8379\approx30$

Usando

$t_{\left(\frac{\propto}{2},\upsilon\right)}$

gl= n	Valores de la probabilidad α área derecha de la distribución t-student (para los cuantiles se cambia el signo)														
	0.0080	0.0085	0.0090	0.0095	0.010	0.015	0.020	0.025	0.030	0.035	0.040	0.045	0.050	0.055	0.060
1	39.780	37.439	35.359	33.496	31.821	21.205	15.894	12.706	10.579	9.058	7.916	7.026	6.314	5.730	5.242
2	7.810	7.572	7.353	7.151	6.965	5.643	4.849	4.303	3.896	3.578	3.320	3.104	2.920	2.760	2.620
3	4.930	4.821	4.721	4.628	4.541	3.896	3.482	3.182	2.951	2.763	2.605	2.471	2.353	2.249	2.156
4	4.010	3.937	3.870	3.806	3.747	3.298	2.999	2.776	2.601	2.456	2.333	2.226	2.132	2.048	1.971
5	3.573	3.516	3.462	3.412	3.365	3.003	2.757	2.571	2.422	2.297	2.191	2.098	2.015	1.941	1.873
6	3.320	3.272	3.226	3.183	3.143	2.829	2.612	2.447	2.313	2.201	2.104	2.019	1.943	1.874	1.812
7	3.157	3.113	3.073	3.034	2.998	2.715	2.517	2.365	2.241	2.136	2.046	1.966	1.895	1.830	1.770
8	3.043	3.003	2.965	2.930	2.896	2.634	2.449	2.306	2.189	2.090	2.004	1.928	1.860	1.797	1.740
9	2.958	2.921	2.886	2.853	2.821	2.574	2.398	2.262	2.150	2.055	1.973	1.899	1.833	1.773	1.718
10	2.894	2.859	2.825	2.794	2.764	2.527	2.359	2.228	2.120	2.028	1.948	1.877	1.812	1.754	1.700
11	2.843	2.809	2.777	2.747	2.718	2.491	2.328	2.201	2.096	2.007	1.928	1.859	1.796	1.738	1.686
12	2.801	2.769	2.738	2.709	2.681	2.461	2.303	2.179	2.076	1.989	1.912	1.844	1.782	1.726	1.674
13	2.767	2.736	2.706	2.677	2.650	2.436	2.282	2.160	2.060	1.974	1.899	1.832	1.771	1.715	1.664
14	2.739	2.708	2.678	2.651	2.624	2.415	2.264	2.145	2.046	1.962	1.887	1.821	1.761	1.706	1.656
15	2.714	2.684	2.655	2.628	2.602	2.397	2.249	2.131	2.034	1.951	1.878	1.812	1.753	1.699	1.649
16	2.693	2.663	2.635	2.609	2.583	2.382	2.235	2.120	2.024	1.942	1.869	1.805	1.746	1.692	1.642
17	2.675	2.645	2.618	2.592	2.567	2.368	2.224	2.110	2.015	1.934	1.862	1.798	1.740	1.686	1.637
18	2.658	2.630	2.603	2.577	2.552	2.356	2.214	2.101	2.007	1.926	1.855	1.792	1.734	1.681	1.632
19	2.644	2.616	2.589	2.564	2.539	2.346	2.205	2.093	2.000	1.920	1.850	1.786	1.729	1.677	1.628
20	2.631	2.603	2.577	2.552	2.528	2.336	2.197	2.086	1.994	1.914	1.844	1.782	1.725	1.672	1.624
21	2.620	2.592	2.566	2.541	2.518	2.328	2.189	2.080	1.988	1.909	1.840	1.777	1.721	1.669	1.621
22	2.610	2.582	2.556	2.532	2.508	2.320	2.183	2.074	1.983	1.905	1.835	1.773	1.717	1.665	1.618
23	2.600	2.573	2.547	2.523	2.500	2.313	2.177	2.069	1.978	1.900	1.832	1.770	1.714	1.662	1.615
24	2.592	2.565	2.539	2.515	2.492	2.307	2.172	2.064	1.974	1.896	1.828	1.767	1.711	1.660	1.612
25	2.584	2.557	2.532	2.508	2.485	2.301	2.167	2.060	1.970	1.893	1.825	1.764	1.708	1.657	1.610
26	2.577	2.550	2.525	2.501	2.479	2.296	2.162	2.056	1.967	1.890	1.822	1.761	1.706	1.655	1.608
27	2.570	2.544	2.519	2.495	2.473	2.291	2.158	2.052	1.963	1.887	1.819	1.758	1.703	1.653	1.606
28	2.564	2.538	2.513	2.490	2.467	2.286	2.154	2.048	1.960	1.884	1.817	1.756	1.701	1.651	1.604
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.045	1.957	1.881	1.814	1.754	1.699	1.649	1.602
30	2.553	2.527	2.503	2.479	2.457	2.278	2.147	2.042	1.955	1.879	1.812	1.752	1.697	1.647	1.600

$t_{(0.02,30)}$=2.147

Usando

$$\left(\acute{x}_{MEX}-\acute{x}_{EUA}\right)-t_{\left(\frac{\alpha}{2},\upsilon\right)}\sqrt{\frac{S^2_{n\text{-}1MEX}}{n_1}+\frac{S^2_{n\text{-}1EUA}}{n_2}}\le\mu_1-\mu_2\le\left(\acute{x}_{MEX}-\acute{x}_{EUA}\right)+t_{\left(\frac{\alpha}{2},\upsilon\right)}\sqrt{\frac{S^2_{n\text{-}1MEX}}{n_1}+\frac{S^2_{n\text{-}1EUA}}{n_2}}$$

$$\left(24.92\text{-}44\right)\text{-}2.147\sqrt{\frac{32.74333}{25}+\frac{225.0833}{25}}\le\mu_1-\mu_2\le\left(24.92\text{-}44\right)+2.147\sqrt{\frac{32.74333}{25}+\frac{225.0833}{25}}$$

$$-25.9748\le\mu_1-\mu_2\le-12.1851$$

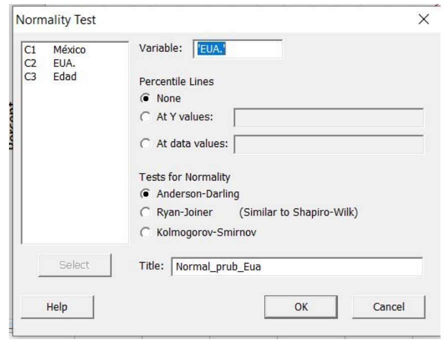
Ambos límites son negativos por lo cual con un nivel de confianza de 98% podemos estimar que $\mu_1<\mu_2$.

Solución por Minitab

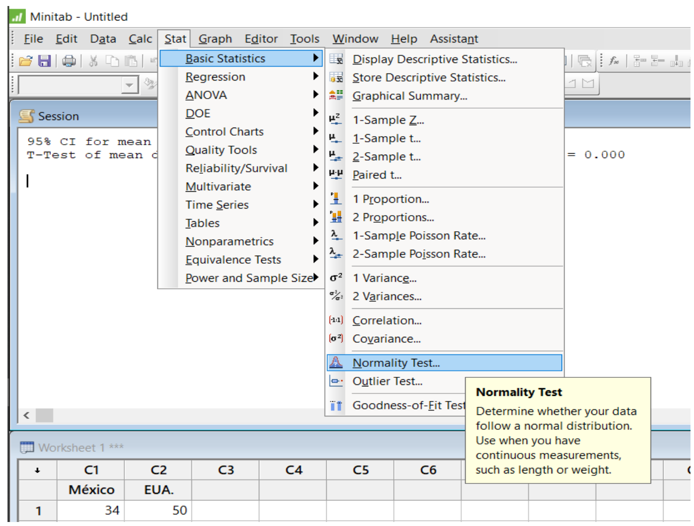
Como primer paso necesitamos vaciar toda la información recopilada, asignando a cada columna un nombre para una mejor identificación.

↓	C1	C2	C3
	México	EUA.	Edad
1	34	50	20
2	25	36	21
3	25	54	20
4	20	30	21
5	20	52	20
6	38	30	20
7	15	66	19
8	30	50	20
9	25	60	21
10	25	52	20

15	20	60	21
16	30	37	20
17	20	48	20
18	20	35	20
19	25	30	20
20	30	35	19
21	25	56	20
22	15	70	20
23	30	65	20
24	25	35	20
25	25	44	21

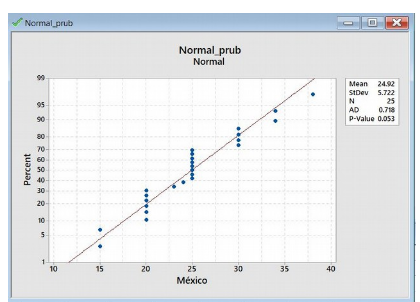


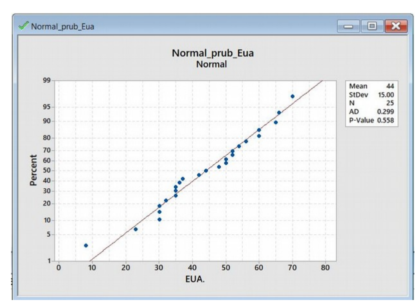
Es importante que verifiquemos que nuestros datos se comportan con una distribución normal. Entonces accediendo de la siguiente manera: Stat-> Basic Statistics-> Normaly Test.

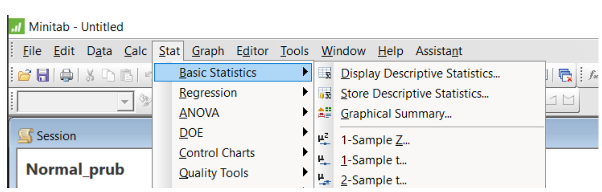


En la ventana emergente, colocaremos la variable a evaluar, con el método 'Anderson_darling' y un título para su identificación y damos en la opción de 'Ok' para continuar.

Nos fijaremos en el valor llamado P-Value, ya que este debe ser mayor a 0.05 para considerar una distribución normal. En el caso de los datos de México obtenemos 0.053>0.05 y en el caso de E.U.A tenemos 0.558>0.05, por lo tanto, ambos casos cumplen.





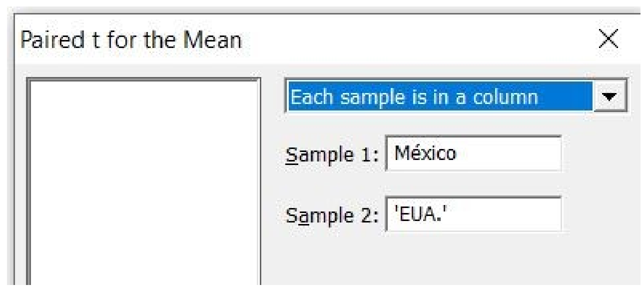


Una vez asegurando una distribución normal de los datos, podemos continuar a realizar el intervalo de confianza.

Accederemos a la siguiente ruta:

Stat-> Basic Statistics->Paired t…

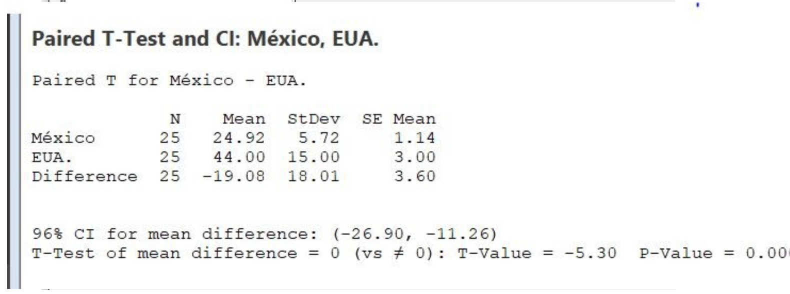
De nuevo se desplegará una ventana emergente donde seleccionaremos las poblaciones de estudio, en este caso ‘México’ y ‘EUA’.



Obtendremos los datos en la ventana de Session, en esta se mostrará la diferencia de media, con un nivel de confianza del 96%.

En nuestro obtuvimos el intervalo de

$$-26.90 \leq \mu_1 - \mu_2 \leq -11.26$$



Paired T-Test and CI: México, EUA.

```
Paired T for México - EUA.

             N    Mean  StDev  SE Mean
México      25   24.92   5.72     1.14
EUA.        25   44.00  15.00     3.00
Difference  25  -19.08  18.01     3.60


96% CI for mean difference: (-26.90, -11.26)
T-Test of mean difference = 0 (vs ≠ 0): T-Value = -5.30  P-Value = 0.000
```

Ejercicio 3

Intervalo de confianza para la varianza tal que N_1=25 y un nivel de confianza=98%

Una empresa manufacturera de ropa está realizando un estudio de mercado para determinar cuántas prendas se compran en México, se toma una muestra de 25 personas con un nivel de confianza de 98%, determine un intervalo de confianza de dos lados para la varianza, considere los datos siguientes:

34	**25**	**25**	**20**	**20**
38	**15**	**30**	**25**	**25**

34	**24**	**23**	**20**	**20**
30	**20**	**20**	**25**	**30**
25	**15**	**30**	**25**	**25**

Usando

$$\acute{x}=\frac{1}{n}\sum_{i=1}^{25}x_i$$

$$\acute{x}_{MEX}=\frac{1}{25}(34+25+25+20+20+38+15+30+25+25+34+24+23+20+20+30+20+20+25+30+25+15+30+2$$

$$\acute{x}_{MEX}=24.92$$

Usando

$$S_{n-1}^2=\frac{1}{n-1}\sum_{i=1}^{25}(x_i-\acute{x})^2$$

$$S_{n-1MEX}^2=\frac{1}{24}[(34-24.92)^2+(25-24.92)^2+(25-24.92)^2+(20-24.92)^2+(20-24.92)^2+(38-24.92)^2+(15-24.92)^2$$

$S_{n-1MEX}^2=¿$32.74333

$$1-\propto=0.98;\ \propto=0.02;\ \frac{\propto}{2}=0.01$$

$$x_{(\frac{\propto}{2},n-1)}^2\ ;\ x_{(1-\frac{\propto}{2},n-1)}^2$$

Gl=ν	Valores de la distribución Ji-Cuadrada para una probabilidad α de área derecha (intercambiados son los cuantiles)											
	0.005	0.995	0.010	0.990	0.015	0.985	0.020	0.980	0.025	0.975	0.030	0.970
1	7.8794	0.0000	6.6349	0.0002	5.9165	0.0004	5.4119	0.0006	5.0239	0.0010	4.7093	0.0014
2	10.5965	0.0100	9.2104	0.0201	8.3994	0.0302	7.8241	0.0404	7.3778	0.0506	7.0131	0.0609
3	12.8381	0.0717	11.3449	0.1148	10.4651	0.1516	9.8374	0.1848	9.3484	0.2158	8.9473	0.2451
4	14.8602	0.2070	13.2767	0.2971	12.3391	0.3682	11.6678	0.4294	11.1433	0.4844	10.7119	0.5351
5	16.7496	0.4118	15.0863	0.5543	14.0978	0.6618	13.3882	0.7519	12.8325	0.8312	12.3746	0.9031
6	18.5475	0.6757	16.8119	0.8721	15.7774	1.0160	15.0332	1.1344	14.4494	1.2373	13.9676	1.3296
7	20.2777	0.9893	18.4753	1.2390	17.3984	1.4184	16.6224	1.5643	16.0128	1.6899	15.5091	1.8016
8	21.9549	1.3444	20.0902	1.6465	18.9738	1.8603	18.1682	2.0325	17.5345	2.1797	17.0105	2.3101
9	23.5893	1.7349	21.6660	2.0879	20.5125	2.3348	19.6790	2.5324	19.0228	2.7004	18.4796	2.8485
10	25.1881	2.1558	23.2093	2.5582	22.0206	2.8372	21.1608	3.0591	20.4832	3.2470	19.9219	3.4121
11	26.7569	2.6032	24.7250	3.0535	23.5028	3.3634	22.6179	3.6087	21.9200	3.8157	21.3416	3.9972
12	28.2997	3.0738	26.2170	3.5706	24.9628	3.9103	24.0539	4.1783	23.3367	4.4038	22.7418	4.6009
13	29.8193	3.5650	27.6882	4.1069	26.4034	4.4757	25.4715	4.7654	24.7356	5.0087	24.1249	5.2210
14	31.3194	4.0747	29.1412	4.6604	27.8268	5.0573	26.8727	5.3682	26.1189	5.6287	25.4931	5.8556
15	32.8015	4.6009	30.5780	5.2294	29.2349	5.6534	28.2595	5.9849	27.4884	6.2621	26.8480	6.5032
16	34.2671	5.1422	31.9999	5.8122	30.6292	6.2628	29.6332	6.6142	28.8453	6.9077	28.1908	7.1625
17	35.7184	5.6973	33.4087	6.4077	32.0111	6.8842	30.9950	7.2550	30.1910	7.5642	29.5227	7.8324
18	37.1564	6.2648	34.8052	7.0149	33.3817	7.5165	32.3462	7.9062	31.5264	8.2307	30.8447	8.5120
19	38.5821	6.8439	36.1908	7.6327	34.7419	8.1589	33.6874	8.5670	32.8523	8.9065	32.1577	9.2004
20	39.9969	7.4338	37.5663	8.2604	36.0926	8.8105	35.0196	9.2367	34.1696	9.5908	33.4623	9.8971
21	41.4009	8.0336	38.9322	8.8972	37.4345	9.4708	36.3434	9.9145	35.4789	10.2829	34.7593	10.6013
22	42.7957	8.6427	40.2894	9.5425	38.7681	10.1390	37.6595	10.6000	36.7807	10.9823	36.0491	11.3125
23	44.1814	9.2604	41.6383	10.1957	40.0941	10.8147	38.9683	11.2926	38.0756	11.6885	37.3323	12.0303
24	45.5584	9.8862	42.9798	10.8563	41.4129	11.4974	40.2703	11.9918	39.3641	12.4011	38.6093	12.7543
25	46.9280	10.5196	44.3140	11.5240	42.7252	12.1867	41.5660	12.6973	40.6465	13.1197	39.8804	13.4839
26	48.2898	11.1602	45.6416	12.1982	44.0312	12.8821	42.8558	13.4086	41.9231	13.8439	41.1461	14.2190
27	49.6450	11.8077	46.9628	12.8785	45.3311	13.5833	44.1399	14.1254	43.1945	14.5734	42.4066	14.9592
28	50.9936	12.4613	48.2782	13.5647	46.6255	14.2900	45.4188	14.8475	44.4608	15.3079	43.6622	15.7042
29	52.3355	13.1211	49.5878	14.2564	47.9147	15.0019	46.6926	15.5745	45.7223	16.0471	44.9132	16.4538
30	53.6719	13.7867	50.8922	14.9535	49.1988	15.7188	47.9618	16.3062	46.9792	16.7908	46.1600	17.2076

$x^2_{(0.01,24)}=42.9798$

$x^2_{(0.99,24)}=10.8563$

Usando

$$\frac{(n-1)\,S^2_{n-1}}{x^2_{\left(\frac{\alpha}{2},n-1\right)}}\leq\sigma^2_x\leq\frac{(n-1)\,S^2_{n-1}}{x^2_{\left(1-\frac{\alpha}{2},n-1\right)}}$$

Sustituyendo

$$\frac{(24)\,32.74333}{42.9798}\leq\sigma^2_x\leq\frac{(24)\,32.74333}{10.8563}$$

$$18.28393\leq\sigma^2_x\leq72.38554$$

Con un nivel de confianza del 98% la varianza de los datos se encuentra entre

18.28393 y 72.38554

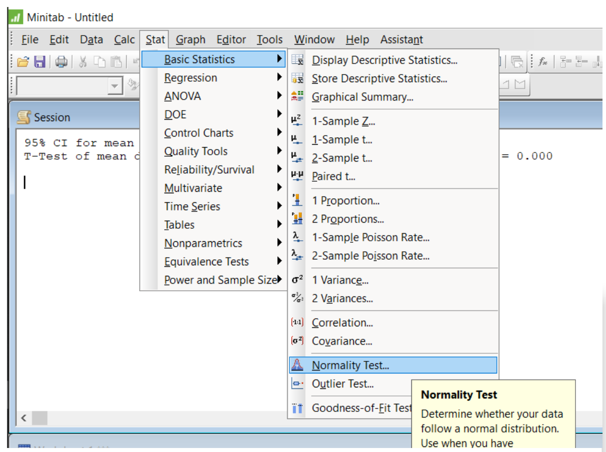
Solución por Minitab:

Como primer paso necesitamos vaciar toda la información recopilada, asignando a esa columna un nombre para una mejor identificación

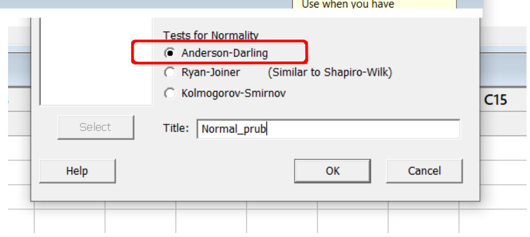
↓	C1
	México
1	34
2	25
3	25
4	20
5	20

15	20
16	30
17	20
18	20
19	25
20	30

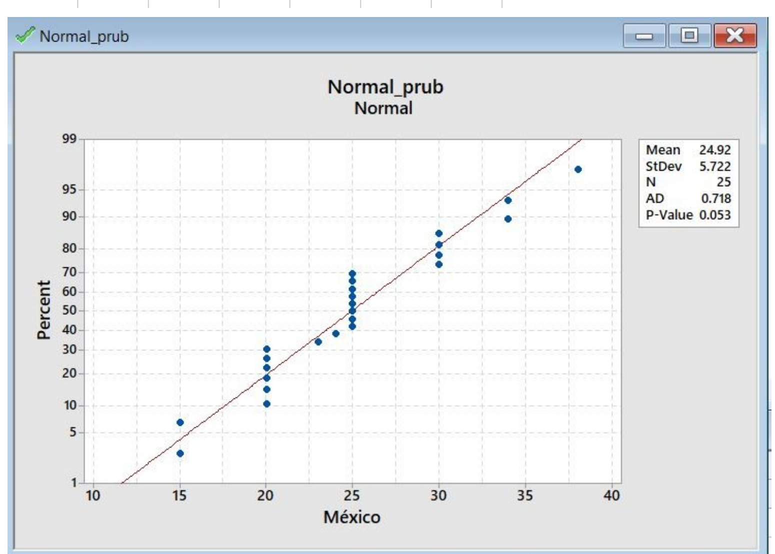
Es importante que verifiquemos que nuestros datos se comportan con una distribución normal. Entonces accediendo de la siguiente manera: Stat-> Basic Statistics-> Normaly Test.

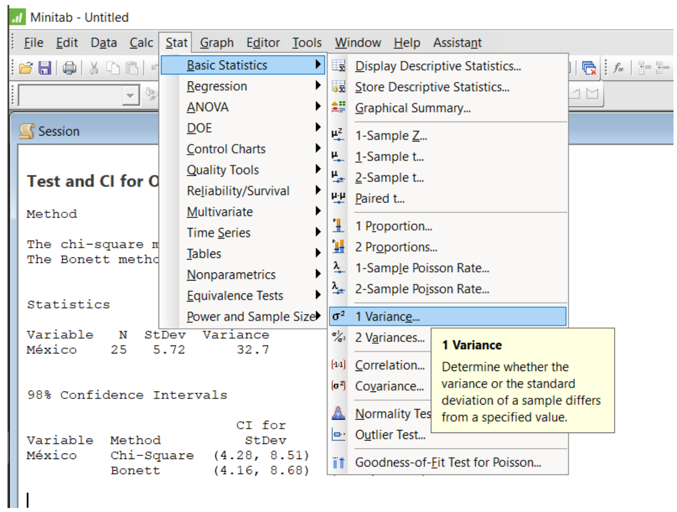


En la ventana emergente, colocaremos la variable a evaluar, con el método ‘Anderson_darling’ y un título para su identificación y damos en la opción de ‘Ok’ para continuar.



Nos fijaremos en el valor llamado P-Value, ya que este debe ser mayor a 0.05 para considerar una distribución normal. Nuestro caso es 0.053>0.05, por lo tanto, cumple.





De nuevo se desplegará una ventana emergente donde seleccionaremos la población de estudio, en este caso ‘México’ y un intervalo de confianza que del 98% que se encontrar en el botón ‘Options’.

Obtendremos los datos en la ventana de Session, en esta se nos mostrará el intervalo de confianza, con un nivel de confianza del 98%.

En este caso Minitab nos proporciona la solución por dos métodos distintos, en este caso Chi-square es el que se acerca más a nuestros resultados por el primer método.

Test and CI for One Variance: México

```
Method

The chi-square method is only for the normal distribution.
The Bonett method is for any continuous distribution.


Statistics

Variable   N  StDev  Variance
México    25   5.72      32.7


98% Confidence Intervals

                           CI for          CI for
Variable  Method           StDev          Variance
México    Chi-Square   (4.28, 8.51)   (18.3, 72.4)
          Bonett       (4.16, 8.68)   (17.3, 75.3)
```

Ejercicio 4

Intervalo de confianza para la razón de varianzas tal que N_1=25 y N_2=25, nivel de confianza=98%

La misma empresa manufacturera de ropa vende tanto en México como en Estados Unidos, se desea comparar la diferencia en promedio de compras por año en cada país, para dicho estudio se obtuvieron las siguientes muestras:

México

34	**25**	**25**	**20**	**20**
38	**15**	**30**	**25**	**25**
34	**24**	**23**	**20**	**20**
30	**20**	**20**	**25**	**30**
25	**15**	**30**	**25**	**25**

Estados Unidos

50	**36**	**54**	**30**	**52**
30	**66**	**50**	**60**	**52**
8	**42**	**23**	**32**	**60**
37	**48**	**35**	**30**	**35**
56	**70**	**65**	**35**	**44**

Determine un intervalo de confianza para la razón entre varianzas para la diferencia promedio de compras de ropa por año de cada país con un nivel de confianza de 98%

Usando

$$\acute{x}=\frac{1}{n}\sum_{i=1}^{25} x_i$$

$$\acute{x}_{MEX}=\frac{1}{25}(34+25+25+20+20+38+15+30+25+25+34+24+23+20+20+30+20+20+25+30+25+15+30+2$$

$$\acute{x}_{MEX}=24.92$$

$$\acute{x}_{EUA}=\frac{1}{25}(50+36+54+30+52+30+66+50+60+52+8+42+23+32+60+37+48+35+30+35+56+70+65+35$$

$$\acute{x}_{EUA}=44$$

Usando

$$S_{n-1}^2=\frac{1}{n-1}\sum_{i=1}^{25}(x_i-\acute{x})^2$$

$$S_{n-1\,MEX}^2=\frac{1}{24}[(34-24.92)^2+(25-24.92)^2+(25-24.92)^2+(20-24.92)^2+(20-24.92)^2+(38-24.92)^2+(15-24.92)^2$$

$$S_{n-1\,MEX}^2=¿32.74333$$

$$S_{n-1\,EUA}^2=\frac{1}{24}[(50-44)^2+(36-44)^2+(54-44)^2+(30-44)^2+(52-44)^2+(30-44)^2+(66-44)^2+(50-44)^2+(60-44)$$

$$S_{n-1\,EUA}^2=¿225.0833$$

$$n_1=25;\ n_2=25;\ 1-\propto=0.98;\ \propto=0.02;\ \frac{\propto}{2}=0.01$$

En este caso

$$F_{\frac{\propto}{2}(n_1-1,n_2-1)}=F_{\frac{\propto}{2}(n_2-1,n_1-1)}$$

$$F_{0.01(24,24)}$$

ν_2	ν_1 grados de libertad del numerador y ν_2 grados de libertad del denominador; ν_1 y $\alpha = 0.01$																		
	20	21	22	23	24	25	26	27	28	29	30	35	40	45	50	55	60	70	80
1	6209	6216	6223	6229	6235	6240	6245	6249	6253	6257	6261	6276	6287	6296	6303	6308	6313	6317	6321
2	99.4	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5	99.5
3	26.69	26.66	26.64	26.62	26.60	26.58	26.56	26.55	26.53	26.52	26.50	26.45	26.41	26.38	26.35	26.33	26.32	26.30	26.29
4	14.02	13.99	13.97	13.95	13.93	13.91	13.89	13.88	13.86	13.85	13.84	13.79	13.75	13.71	13.69	13.67	13.65	13.64	13.63
5	9.55	9.53	9.51	9.49	9.47	9.45	9.43	9.42	9.40	9.39	9.38	9.33	9.29	9.26	9.24	9.22	9.20	9.19	9.18
6	7.40	7.37	7.35	7.33	7.31	7.30	7.28	7.27	7.25	7.24	7.23	7.18	7.14	7.11	7.09	7.07	7.06	7.04	7.03
7	6.16	6.13	6.11	6.09	6.07	6.06	6.04	6.03	6.02	6.00	5.99	5.94	5.91	5.88	5.86	5.84	5.82	5.81	5.80
8	5.36	5.34	5.32	5.30	5.28	5.26	5.25	5.23	5.22	5.21	5.20	5.15	5.12	5.09	5.07	5.05	5.03	5.02	5.01
9	4.81	4.79	4.77	4.75	4.73	4.71	4.70	4.68	4.67	4.66	4.65	4.60	4.57	4.54	4.52	4.50	4.48	4.47	4.46
10	4.41	4.38	4.36	4.34	4.33	4.31	4.30	4.28	4.27	4.26	4.25	4.20	4.17	4.14	4.12	4.10	4.08	4.07	4.06
11	4.099	4.077	4.057	4.038	4.021	.005	3.990	3.977	3.964	3.952	3.941	3.895	3.860	3.832	3.810	3.791	3.776	3.763	3.752
12	3.858	3.836	3.816	3.798	3.780	.765	3.750	3.736	3.724	3.712	3.701	3.654	3.619	3.592	3.569	3.551	3.535	3.522	3.511
13	3.665	3.643	3.622	3.604	3.587	.571	3.556	3.543	3.530	3.518	3.507	3.461	3.425	3.398	3.375	3.357	3.341	3.328	3.317
14	3.505	3.483	3.463	3.444	3.427	.412	3.397	3.383	3.371	3.359	3.348	3.301	3.266	3.238	3.215	3.197	3.181	3.168	3.157
15	3.372	3.350	3.330	3.311	3.294	.278	3.264	3.250	3.237	3.225	3.214	3.167	3.132	3.104	3.081	3.063	3.047	3.034	3.022
16	3.259	3.237	3.216	3.198	3.181	.165	3.150	3.137	3.124	3.112	3.101	3.054	3.018	2.990	2.967	2.949	2.933	2.920	2.908
17	3.162	3.139	3.119	3.101	3.084	.068	3.053	3.039	3.026	3.014	3.003	2.956	2.920	2.892	2.869	2.851	2.835	2.821	2.810
18	3.077	3.055	3.035	3.016	2.999	.983	2.968	2.955	2.942	2.930	2.919	2.871	2.835	2.807	2.784	2.765	2.749	2.736	2.724
19	3.003	2.981	2.961	2.942	2.925	.909	2.894	2.880	2.868	2.855	2.844	2.797	2.761	2.732	2.709	2.690	2.674	2.661	2.649
20	2.938	2.916	2.895	2.877	2.859	.843	2.829	2.815	2.802	2.790	2.778	2.731	2.695	2.666	2.643	2.624	2.608	2.594	2.582
21	2.880	2.857	2.837	2.818	2.801	.785	2.770	2.756	2.743	2.731	2.720	2.672	2.636	2.607	2.584	2.565	2.548	2.535	2.523
22	2.827	2.805	2.785	2.766	2.749	.733	2.718	2.704	2.691	2.679	2.667	2.620	2.583	2.554	2.531	2.511	2.495	2.481	2.469
23	2.781	2.758	2.738	2.719	2.702	.686	2.671	2.657	2.644	2.632	2.620	2.572	2.535	2.506	2.483	2.463	2.447	2.433	2.421
24	2.738	2.716	2.695	2.676	2.659	.643	2.628	2.614	2.601	2.589	2.577	2.529	2.492	2.463	2.440	2.420	2.403	2.389	2.377
25	2.699	2.677	2.657	2.638	2.620	2.604	2.589	2.575	2.562	2.550	2.538	2.490	2.453	2.424	2.400	2.380	2.364	2.350	2.337

$$F_{0.01(24,24)} = 2.659$$

Usando

$$\frac{S^2_{n-1MEX}}{S^2_{n-1EUA}} \frac{1}{F_{\frac{\alpha}{2}(n_1-1,n_2-1)}} \leq \frac{\sigma^2_{MEX}}{\sigma^2_{EUA}} \leq \frac{S^2_{n-1MEX}}{S^2_{n-1EUA}} F_{\frac{\alpha}{2}(n_2-1,n_1-1)}$$

Sustituyendo

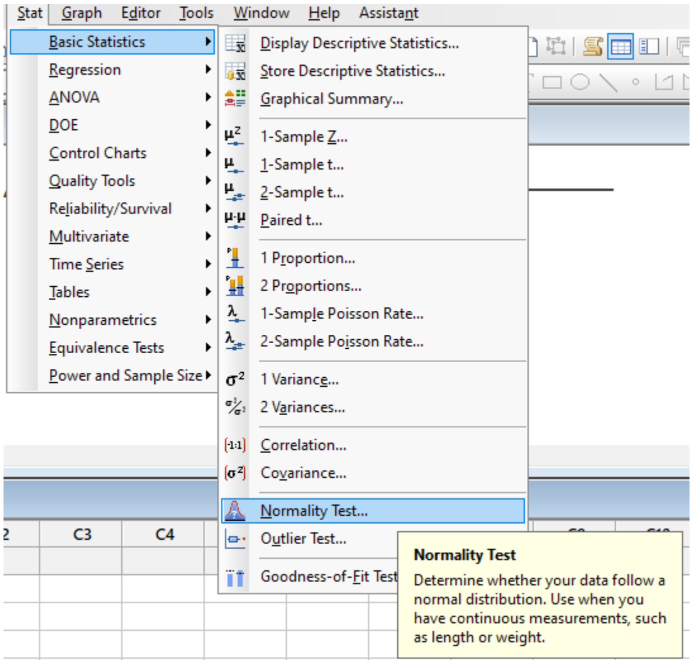
$$\frac{32.74333}{225.0833} \frac{1}{2.659} \leq \frac{\sigma^2_{MEX}}{\sigma^2_{EUA}} \leq \frac{32.74333}{225.0833} 2.659$$

$$0.054709 \leq \frac{\sigma^2_{MEX}}{\sigma^2_{EUA}} \leq 0.38681$$

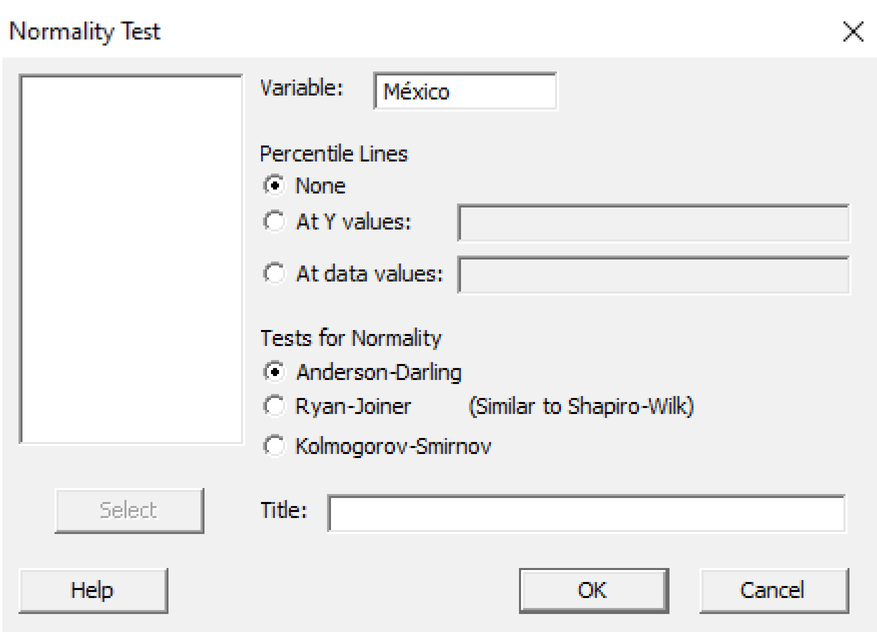
Con un nivel de confianza de 98% la razón de varianzas se encuentra entre 0.054709 y 0.38681 y el intervalo al no contener a 1 se concluye que las varianzas no son iguales

Solución con Minitab

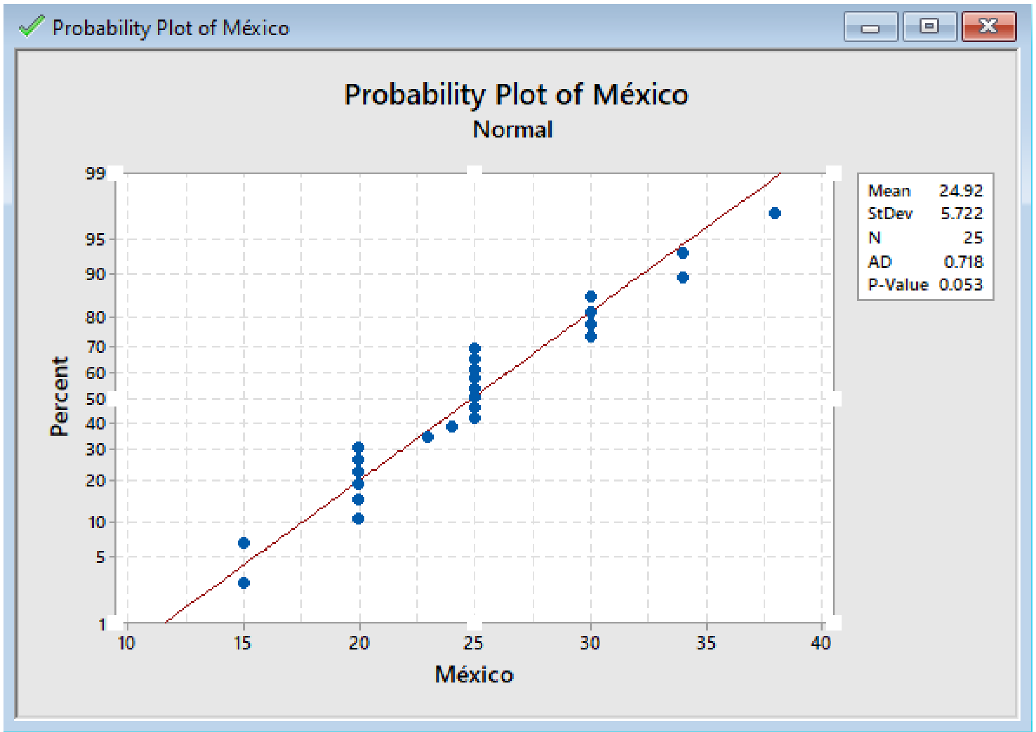
Primero verificamos que los datos tengan una distribución normal, esto lo hacemos entrando al menú Stat → Basic statistics → Normality test

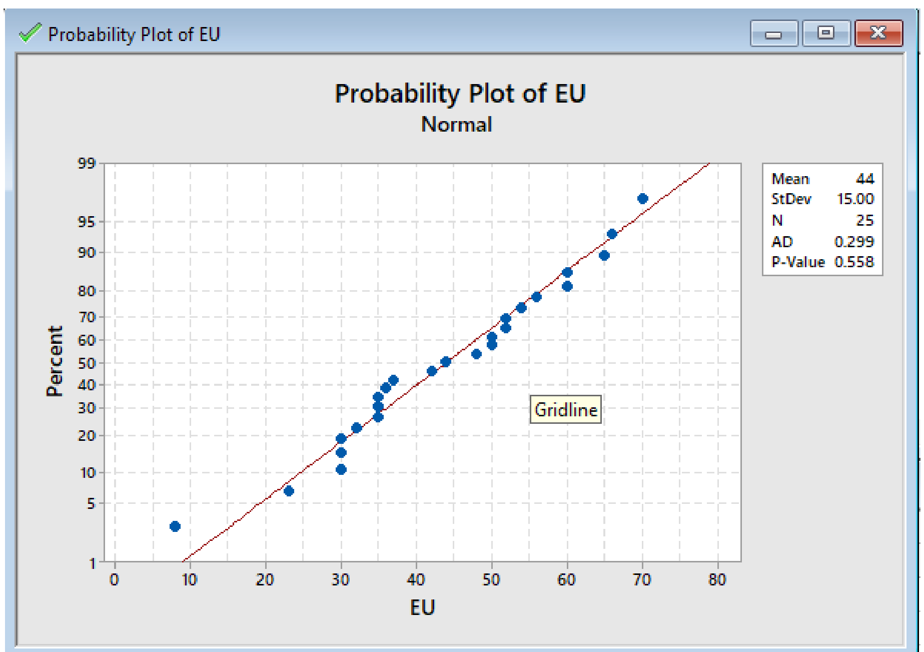


Una vez seleccionada la prueba de normalidad seleccionamos la variable a analizar, elegimos la prueba Anderson-Darling y presionamos “Ok”

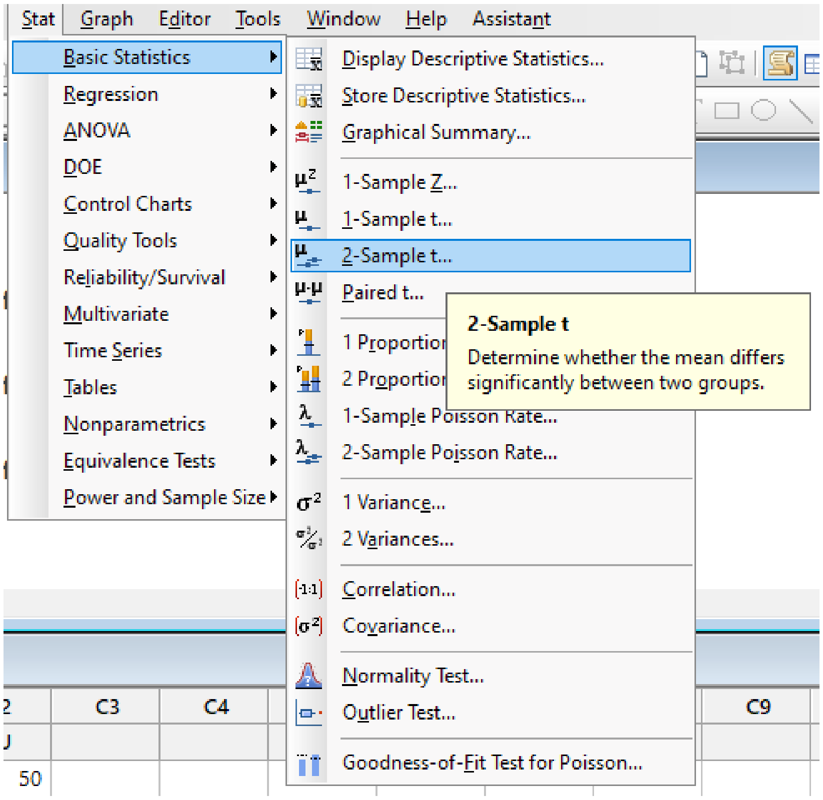


En la siguiente ventana el valor que nos interesa es el de P, para que nuestra muestra tenga una distribución normal P tendrá que ser mayor que α que en nuestro caso vale 0.04 por tener un nivel de confianza del 96% en nuestro caso, ambas distribuciones salieron normales

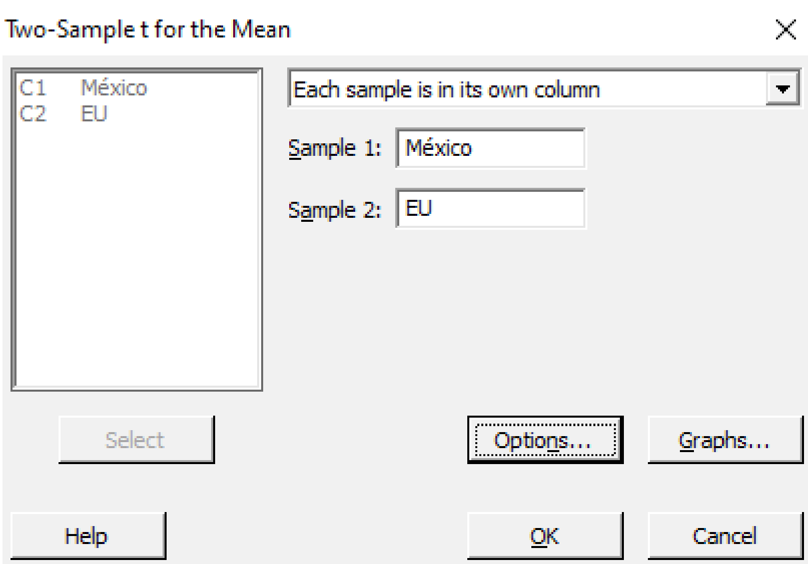




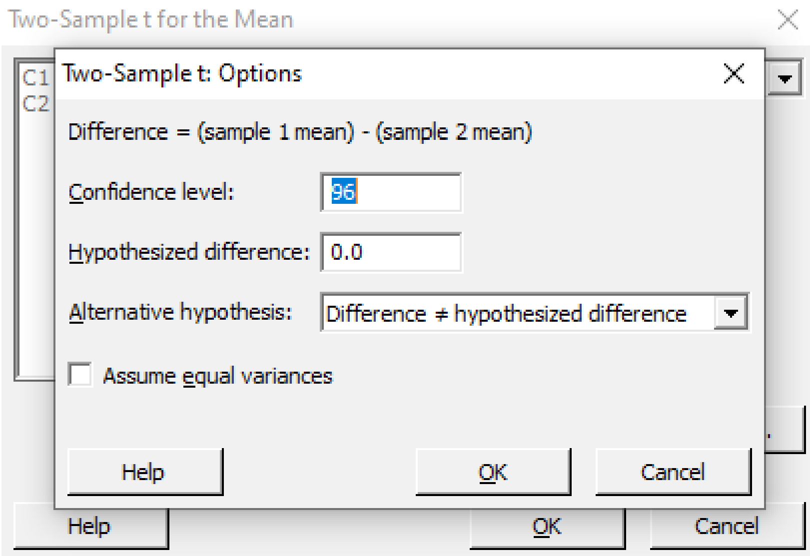
Ahora debemos entrar al menú Stat → Basic statistics → 2-sample t… esto debido a que desconocemos las varianzas y queremos encontrar un intervalo de confianza para una diferencia de medias.



Posteriormente en el cuadro mostrado elegiremos “each sample is its own column”, luego, elegiremos cada una de las muestras



Tras presionar “Options” se abrirá una nueva ventana, en ella, habrá que escribir el nivel de confianza de 96, posteriormente, presionamos “Ok” en ambas ventanas



En la siguiente ventana, aparecen ambas medias y desviaciones estándar y 3 renglones abajo aparece nuestro intervalo de confianza con un nivel de confianza del 96% que en este caso es (-25.97,-12.19)

```
         N   Mean  StDev  SE Mean
México  25  24.92   5.72      1.1
EU      25   44.0   15.0      3.0


Difference = µ (México) - µ (EU)
Estimate for difference:  -19.08
96% CI for difference:  (-25.97, -12.19)
T-Test of difference = 0 (vs ≠): T-Value = -5.94  P-Value = 0.000  DF = 30
```

Dado que ambos límites son negativos, podemos estimar con un 96% de confianza que $\mu_1<\mu_2$

Ejercicio 5

Intervalo de confianza para una proporción tal que N_1=25 y un nivel de confianza=99%

En una encuesta al azar de 25 alumnos que se encuentran estudiando en área 1 de fisicomatemáticas y de las ingenierías, 10 de ellos se encuentran estudiando en una FES en la UNAM. Construir un intervalo de confianza del 99% para la proporción de los alumnos que se encuentran estudiando en una FES.

PASO 1. Se usa la fórmula para intervalos de confianza para una proporción.

FORMULA:

$$\hat{P}-z\frac{a}{2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\leq P\leq\hat{P}+z\frac{a}{2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

PASO 2. Se hace la recopilación de datos que nos da el problema y se busca el valor de Z en las tablas de distribución normal.

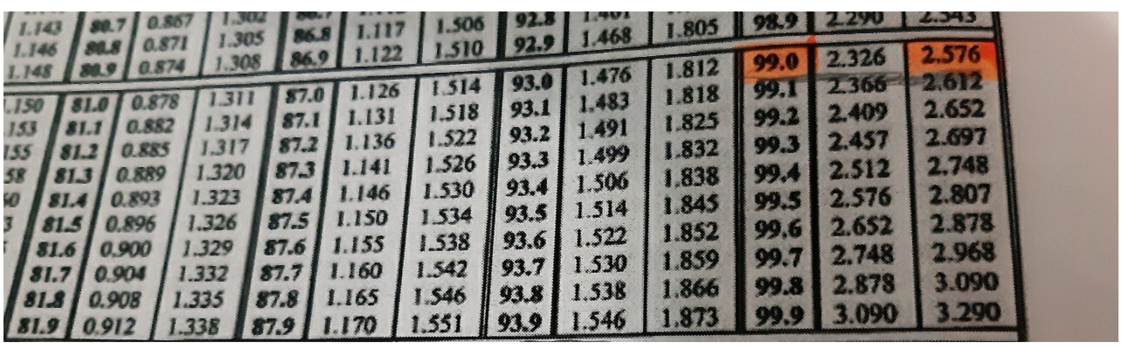
Datos:

N.C.= 99%

1-a= 0.99

a= 0.01 …. a/2 = 0.005

$z\frac{a}{2}$ = Z 0.005 = 2.576 (Buscado en tablas de distribución Normal, p.8)



1.143	80.7	0.867	1.302	86.7	[illegible]							
1.146	80.8	0.871	1.305	86.8	1.117	1.506	92.8	1.461				
1.148	80.9	0.874	1.308	86.9	1.122	1.510	92.9	1.468	1.805	98.9	2.290	2.543
.150	81.0	0.878	1.311	87.0	1.126	1.514	93.0	1.476	1.812	99.0	2.326	2.576
153	81.1	0.882	1.314	87.1	1.131	1.518	93.1	1.483	1.818	99.1	2.366	2.612
155	81.2	0.885	1.317	87.2	1.136	1.522	93.2	1.491	1.825	99.2	2.409	2.652
58	81.3	0.889	1.320	87.3	1.141	1.526	93.3	1.499	1.832	99.3	2.457	2.697
60	81.4	0.893	1.323	87.4	1.146	1.530	93.4	1.506	1.838	99.4	2.512	2.748
3	81.5	0.896	1.326	87.5	1.150	1.534	93.5	1.514	1.845	99.5	2.576	2.807
	81.6	0.900	1.329	87.6	1.155	1.538	93.6	1.522	1.852	99.6	2.652	2.878
	81.7	0.904	1.332	87.7	1.160	1.542	93.7	1.530	1.859	99.7	2.748	2.968
	81.8	0.908	1.335	87.8	1.165	1.546	93.8	1.538	1.866	99.8	2.878	3.090
	81.9	0.912	1.338	87.9	1.170	1.551	93.9	1.546	1.873	99.9	3.090	3.290

PASO 3. Se sustituye los datos en la formula.

$$\frac{10}{25}-2.576\sqrt{\frac{\frac{10}{25}\left(1-\frac{10}{25}\right)}{25}}\leq p\leq\frac{10}{25}+2.576\sqrt{\frac{\frac{10}{25}\left(1-\frac{10}{25}\right)}{25}}$$

Por lo tanto = 0.1678 ≤ p ≤ 0.6703

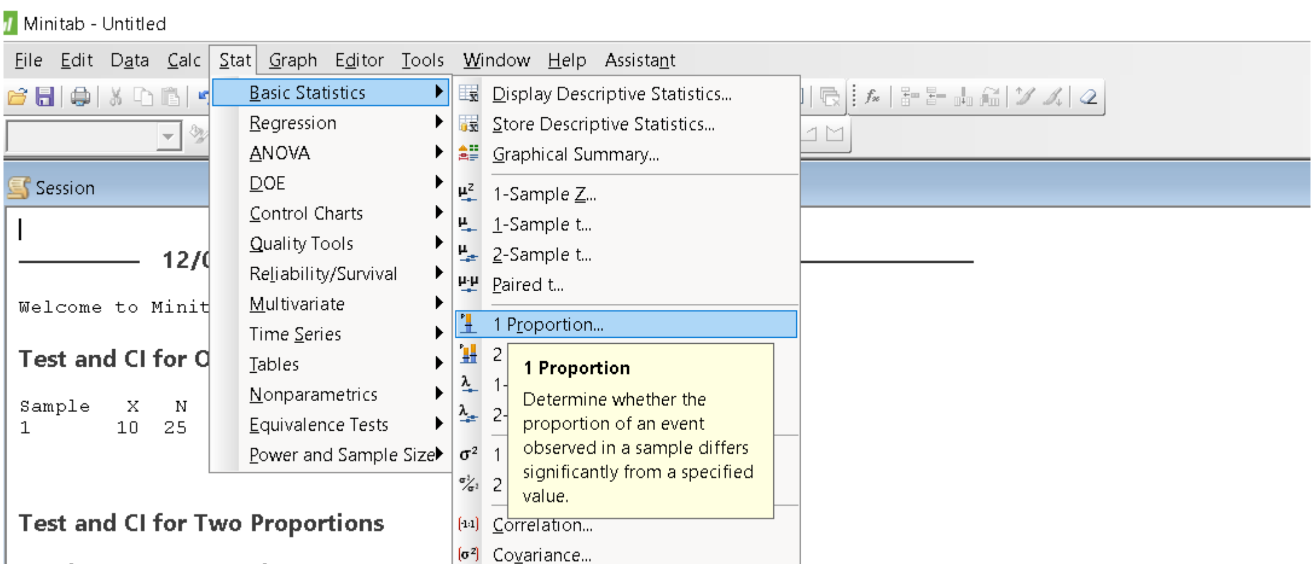
= 16.78 % ≤ p ≤ 67.03 %

Método en minita

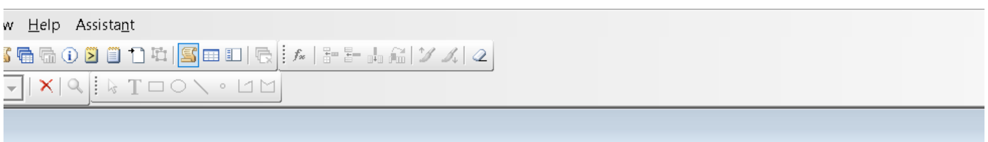
Ahora se resolverá el problema en MINITAB, para obtener el mismo intervalo de confianza con un nivel de confianza de 0.99.

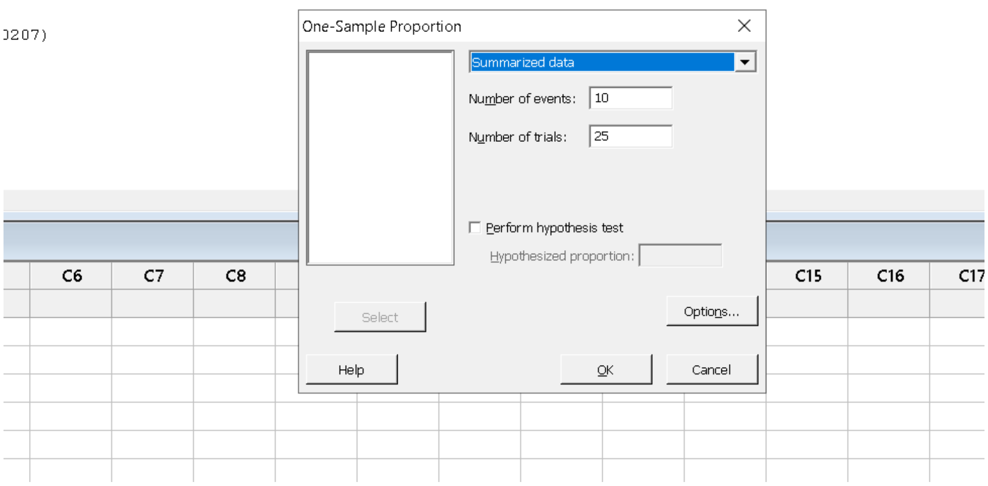
Paso 1: No es necesario anotar los datos.

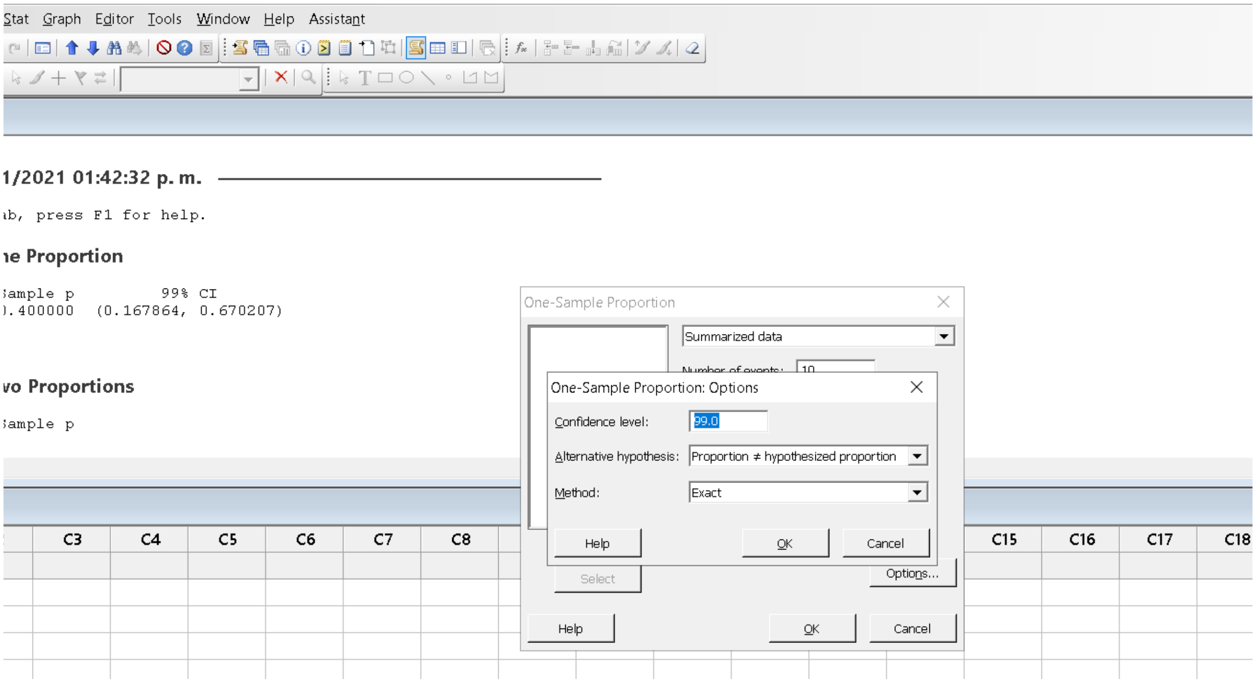
Selecciono la parte de estadística básica y escojo la pestaña de ´´1 proporción´´.



Paso 2. Ya que fue seleccionada la pestaña, se pone en la parte de ´´numero de eventos´´ lo que nosotros consideramos como error, y en la casilla de abajo, se pone el tamaño de la muestra, que en este caso es 25.

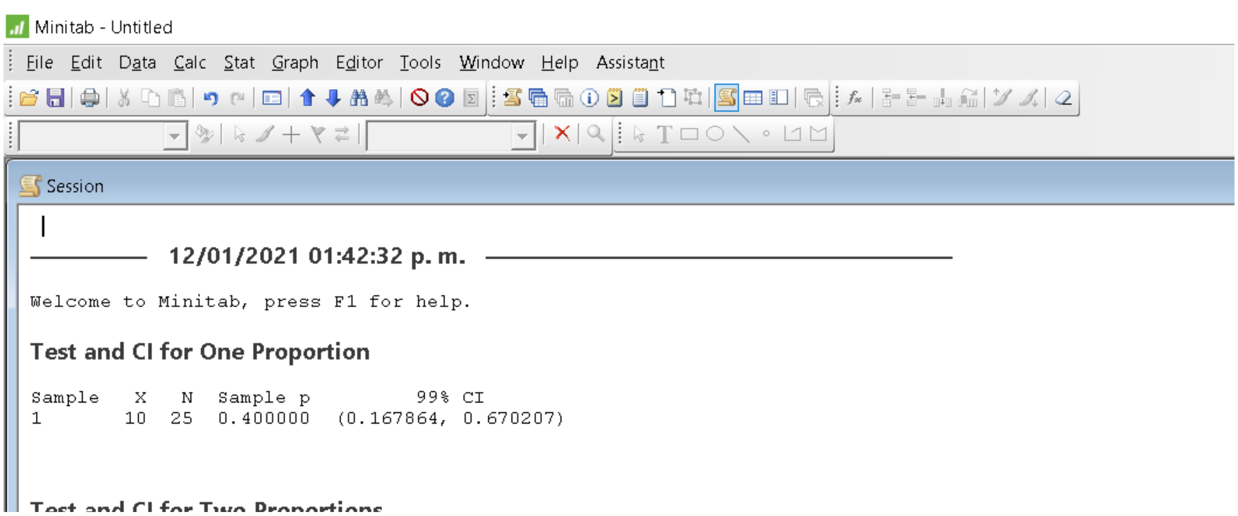






Paso 3. Finalmente, el software nos arroja que: 

En un tamaño de muestra de 25 (N=25), encontré 10 eventos (X=10), en seguida me dan el valor de la proporción que es de 0.40000 o del 40% y el intervalo de confianza que va del 16.7864 % al 67.0307%con un nivel de confianza del 99%.



Ejercicio 6

Intervalo de confianza para una diferencia de proporciones con N_1=25, N_2=25 y un nivel de confianza=94%

Dos grupos de 25 alumnos de área 1 y área 2 fueron parte de una encuesta en donde el objetivo era recabar datos para saber que tanto interés tienen estas dos áreas de la UNAM en Ciencias sociales e historia. En el grupo de alumnos de área 1, 15 personas contestaron que su interés era moderado, mientras que en el grupo de alumnos de área 2, solo 20 personas contestaron que su interés era moderado. Obtener un I.C. para la diferencia de proporciones, N1=25 Y N2=25, con in nivel de confianza del 94%.

Ciencias sociales e historia (interés)	Área 1
Moderado	Nulo
Moderado	Nulo
Moderado	Nulo
Moderado	Nulo
Moderado	Nulo
Moderado	Nulo
Moderado	Nulo
Moderado	Nulo
Moderado	Nulo
Moderado	Nulo
Moderado	
Moderado	
Moderado	
Moderado	
Moderado	
Moderado	
Área 2	Ciencias sociales e historia(interés)
Nulo	Moderado
Nulo	Moderado

Nulo	Moderado
Nulo	Moderado
Nulo	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado
	Moderado

PASO 1: Se hace uso de la fórmula de intervalos de confianza para una diferencia de proporciones.

FORMULA:

$$\left(\hat{P}_1-\hat{P}_2\right)-z\frac{a}{2}\sqrt{\frac{\hat{P},11-\hat{P}_2}{n_1}},+\frac{\hat{P}_2\left(1-\hat{P}_2\right)}{n_2}\leq P_1-P_2\leq\left(\hat{P}_1-\hat{P}_2\right)+z\frac{a}{2}\sqrt{\frac{\hat{P}1\left(1-\hat{P}1\right)}{n1}}+\sqrt{\frac{\widehat{P2}\left(1-\hat{P}2\right)}{n2}}$$

PASO 2: Se hace una recopilación de los datos que nos proporciona el enunciado y se va sustituyendo para que se puedan introducir los datos en la formula con mayor facilidad.

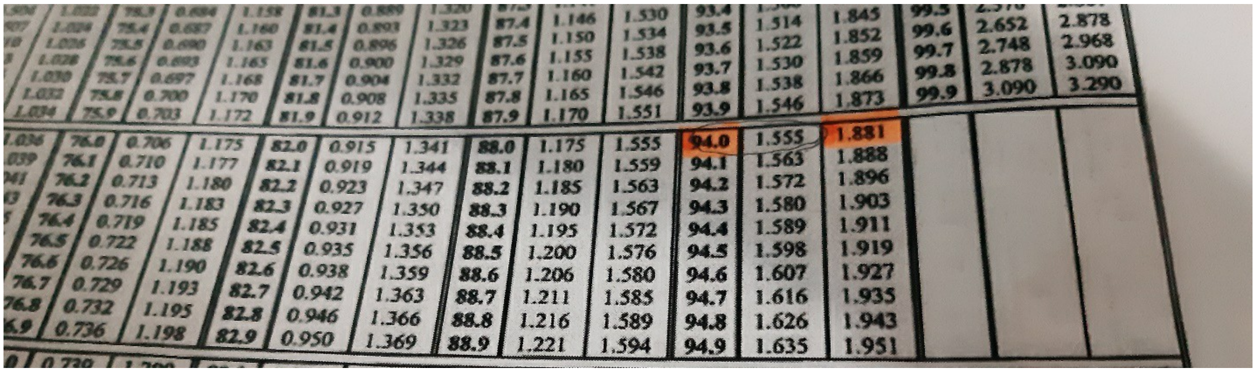
$\hat{P}_1 = \frac{15}{25}$ $\hat{P}_2 = \frac{Y}{n2} = \frac{20}{25}$

Datos:

N.C. = 94%

1-a= 0.94

PASO 3: Se busca en las tablas de distribución normal estándar la Z de acuerdo con nuestro nivel de confianza, que en este caso es de un 94%.



75.4	0.687	1.160	81.4	0.893	1.323	87.4	1.146	1.530	93.4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
75.5	0.690	1.163	81.5	0.896	1.326	87.5	1.150	1.534	93.5	1.514	1.845	99.5	[illegible]	[illegible]
75.6	0.693	1.165	81.6	0.900	1.329	87.6	1.155	1.538	93.6	1.522	1.852	99.6	2.652	2.878
75.7	0.697	1.168	81.7	0.904	1.332	87.7	1.160	1.542	93.7	1.530	1.859	99.7	2.748	2.968
75.8	0.700	1.170	81.8	0.908	1.335	87.8	1.165	1.546	93.8	1.538	1.866	99.8	2.878	3.090
75.9	0.703	1.172	81.9	0.912	1.338	87.9	1.170	1.551	93.9	1.546	1.873	99.9	3.090	3.290
76.0	0.706	1.175	82.0	0.915	1.341	88.0	1.175	1.555	94.0	1.555	1.881			
76.1	0.710	1.177	82.1	0.919	1.344	88.1	1.180	1.559	94.1	1.563	1.888			
76.2	0.713	1.180	82.2	0.923	1.347	88.2	1.185	1.563	94.2	1.572	1.896			
76.3	0.716	1.183	82.3	0.927	1.350	88.3	1.190	1.567	94.3	1.580	1.903			
76.4	0.719	1.185	82.4	0.931	1.353	88.4	1.195	1.572	94.4	1.589	1.911			
76.5	0.722	1.188	82.5	0.935	1.356	88.5	1.200	1.576	94.5	1.598	1.919			
76.6	0.726	1.190	82.6	0.938	1.359	88.6	1.206	1.580	94.6	1.607	1.927			
76.7	0.729	1.193	82.7	0.942	1.363	88.7	1.211	1.585	94.7	1.616	1.935			
76.8	0.732	1.195	82.8	0.946	1.366	88.8	1.216	1.589	94.8	1.626	1.943			
76.9	0.736	1.198	82.9	0.950	1.369	88.9	1.221	1.594	94.9	1.635	1.951			

a= 0.06 …. $\frac{a}{2}$ = 0.03 por lo tanto… $z\frac{a}{2}$ = Z 0.03= 1.555 (Buscado en tablas de distribución normal, p.8)

PASO 4: Se sustituye los datos en la formula.

Sustituyendo en la fórmula:

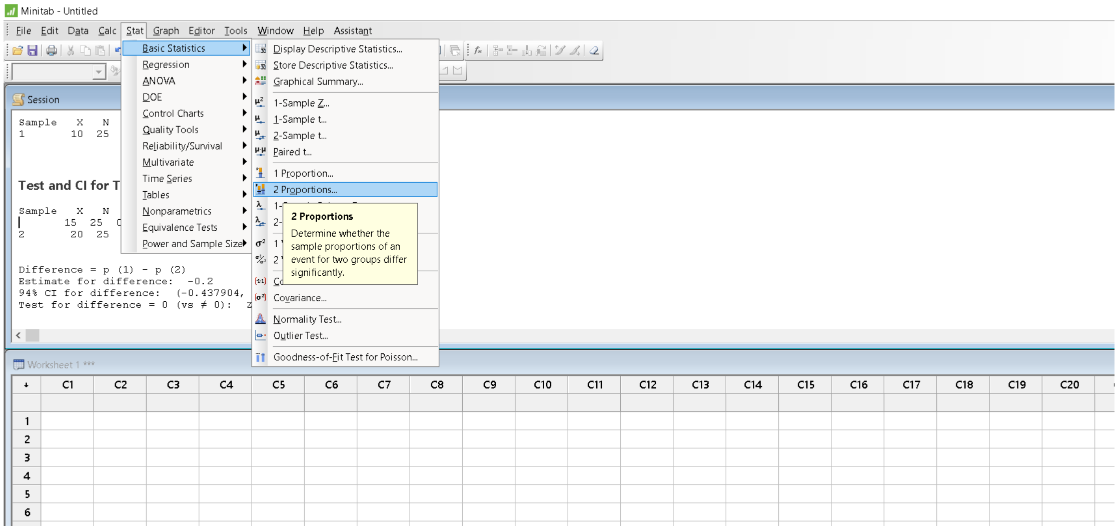
$$\left(\frac{15}{25}-\frac{20}{25}\right)-1.881\sqrt{\frac{\frac{15}{25}\left(1-\frac{15}{25}\right)}{25}+\frac{\frac{20}{25}\left(1-\frac{20}{25}\right)}{25}}\le P_1-P_2\le\left(\frac{15}{25}-\frac{20}{25}\right)+1.881\sqrt{\frac{\frac{15}{25}\left(1-\frac{15}{25}\right)}{25}+\frac{\frac{20}{25}\left(1-\frac{20}{25}\right)}{25}}$$

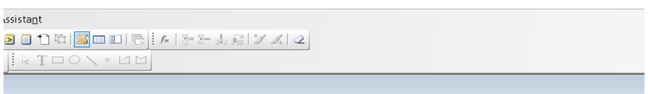
Da como resultado: -0.4379 $\le P_1-P_2 \le$ 0.0379

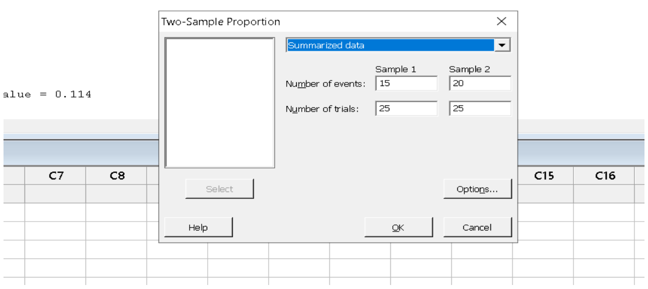
Comprobación con Minitab

Paso 1. Para esta comprobación lo primero que se hizo fue:

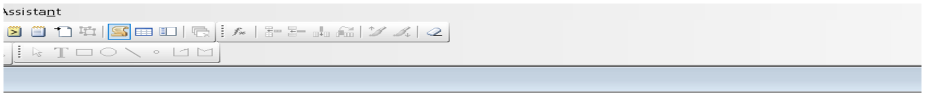
- Dirigirnos a la parte de estadística básica
- De ahí se van a desplegar las pestañas y nos vamos a dirigir a la parte de ´´2 proporciones´´
- Se da clic a esa pestaña y se va a abrir un cuadro en donde va a aparecer el número de eventos y el tamaño de cada muestra.
- En la parte de ´´ Numero de eventos ´´ se va a poner el error tipo 1 (x1) y en el error tipo2 (x2).
- En las casillas de abajo se va a colocar el tamaño de la muestra 1 (n1)) que en este caso es 25 y en la otra casilla, la muestra 2 (n2) que sería igual 25.

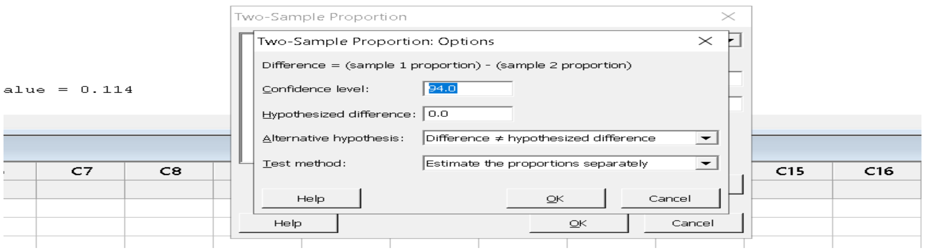






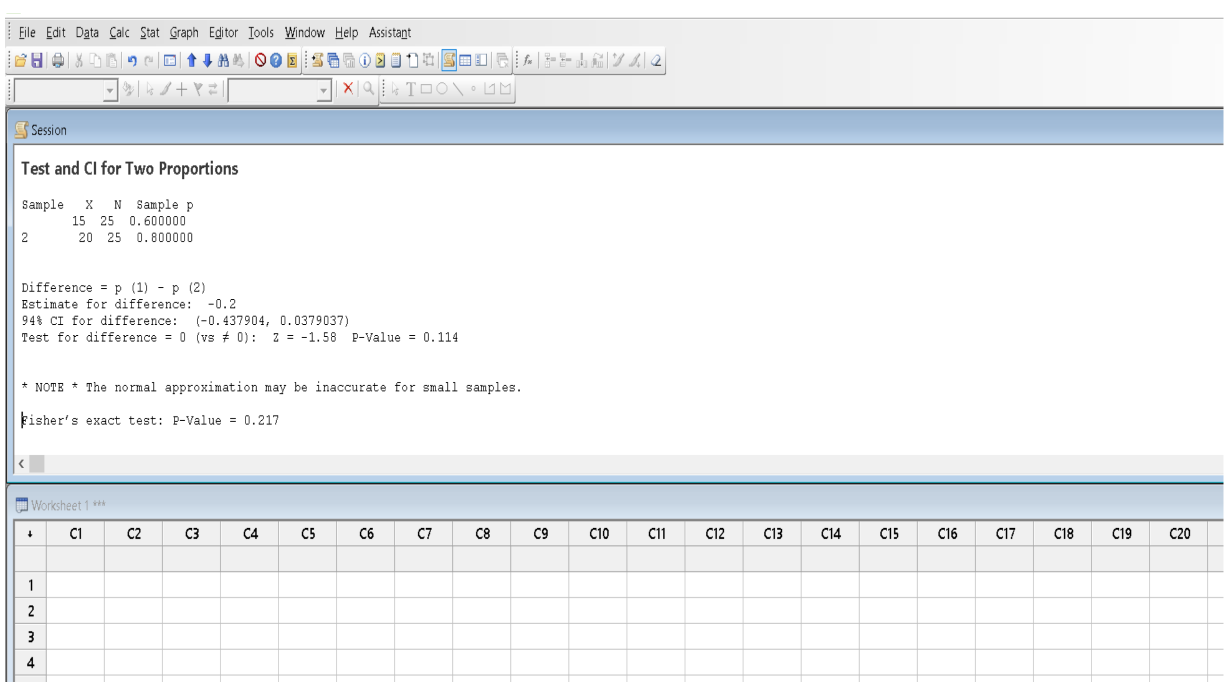
Paso 2. En la parte de **opciones** se cambia el nivel de confianza en este caso al 94 % y se da **ok**.





Paso 3. Después de meter los datos, el software nos da un resumen, en donde en donde nos da la proporción de cada muestra.

También nos da el resultado de la resta entre las dos proporciones y abajo nos da el nivel de confianza que en este caso es del 94% para una diferencia que va de -0.437904 al 0.0379079; y al final nos hace una pequeña recomendación.



Pruebas de hipótesis

Crear una hipótesis es una manera de predecir el comportamiento de una variable aleatoria. Las pruebas de hipótesis son parte de la inferencia estadística y consisten en el planteamiento de una hipótesis nula y una hipótesis alterna

$H_0 : Hipótesis\, nula\, H_1 : Hipótesis\, alterna$

seguido de un desarrollo diferente para cada parámetro a estimar de manera que un resultado cuantificable acepte o rechace la hipótesis nula y rechace o acepte la alterna, respectivamente

Ejercicio 1

Prueba de hipótesis unilateral para la media con varianza desconocida,

$N=25, N.S.=3\%[\mu_0 \neq \acute{X}]$

Los ejecutivos de una marca de ropa suponen que sus clientes en México, en promedio, compran más prendas de ropa al año que sus clientes en Colombia debido a que sus ganancias en este país son mayores. Tras realizar una encuesta al azar a 25 de sus clientes sobre cuantas prendas compran de su marca al año obtienen los siguientes datos.

34
38
34
30
25
25
15
24
20
15
25
30
23
20
30
20
25
20
25
25
20
25
20
30

Siendo el promedio de ventas de la marca en Colombia de 27 prendas al año ¿hay evidencia para afirmar que los clientes en México compran más? Utilizar un nivel de significancia del 3%

DATOS	$n=25$	$\mu=27$	$\alpha=0.03$

Calculando la media muestral:

$$\acute{x}=\frac{1}{n}\sum_{i=1}^{25} x_i=\frac{1}{25}(34+38+34+30+25+25+15+24+20+15+25+30+23+20+30+20+25+20+25+25+20+25+$$

$$\bar{x}=24.92$$

Calculando la desviación estándar muestral:

$$S_{n-1}^2=\frac{1}{24}((34-29.5)^2+(38-29.5)^2+(34-29.5)^2+(30-29.5)^2+(25-29.5)^2+(25-29.5)^2+(15-29.5)^2+(24-29.5$$

$$S_{n-1}=\sqrt{S_{n-1}^2}=\sqrt{32.743333}=5.7222$$

$$S_{n-1}=5.7222$$

Solución

Ya que los ejecutivos solo quieren saber si la gente en México consume más su marca que en Colombia la hipótesis estadística es la siguiente:

$H_o: \mu \leq 26$

$H_1: \mu > 26$

Con el estadístico de prueba:

$$T=\frac{\acute{x}-\mu}{\frac{S}{\sqrt{n}}}$$

Y con valor de significancia:

$$\alpha=0.03$$

Dado que la prueba es unilateral, solo debemos buscar el valor $t_{0.03,24}$ en las tablas de distribución T de student

$$t_{0.03,24}=1.974$$

Entonces el valor del estadístico de prueba T es:

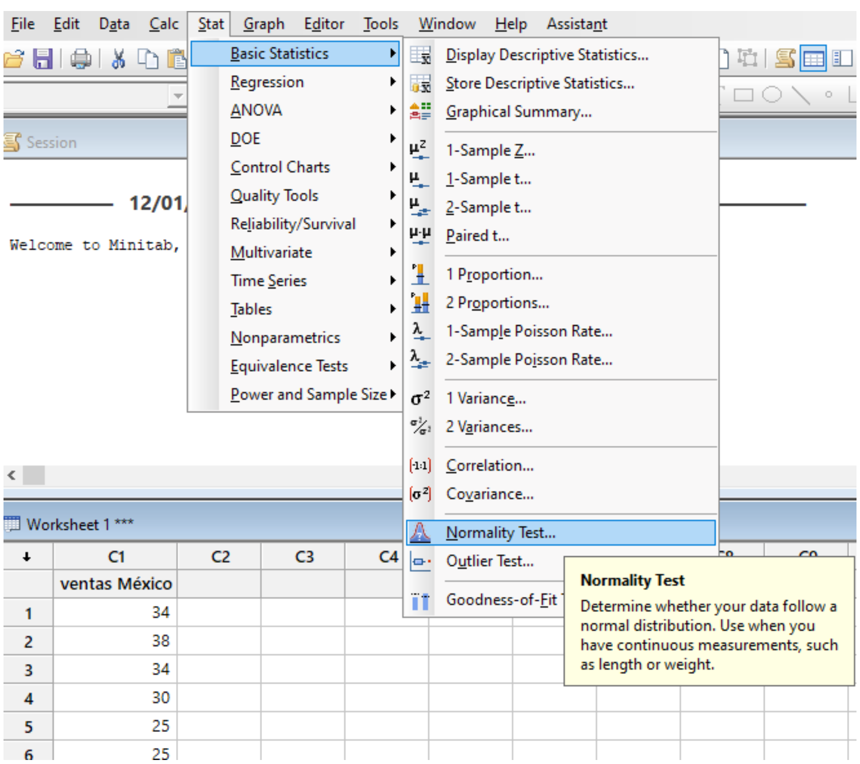
$$T=\frac{\acute{x}-\mu}{\frac{S}{\sqrt{n}}}=\frac{24.92-26}{\frac{5.7222}{\sqrt{25}}}=-0.0377$$

1.974>-0.037

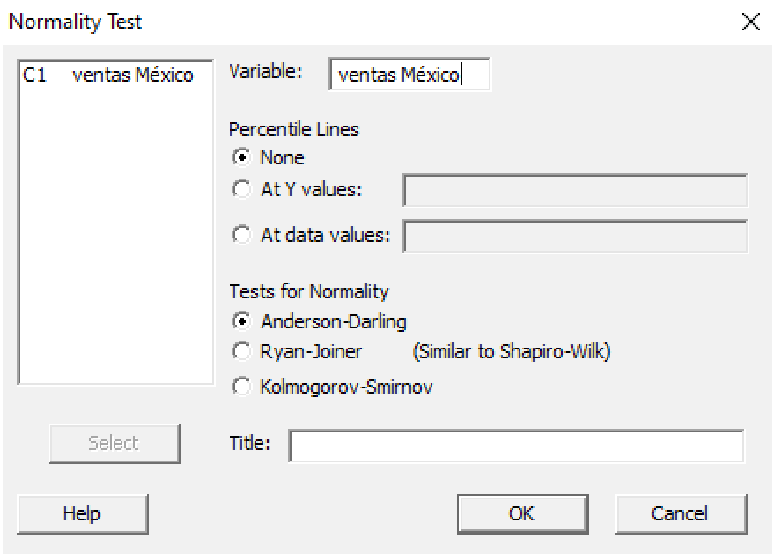
El valor del estadístico se encuentra en la zona de aceptación lo que significa que H_0 NO se rechaza y la afirmación de los ejecutivos es incorrecta.

Solución por minitab

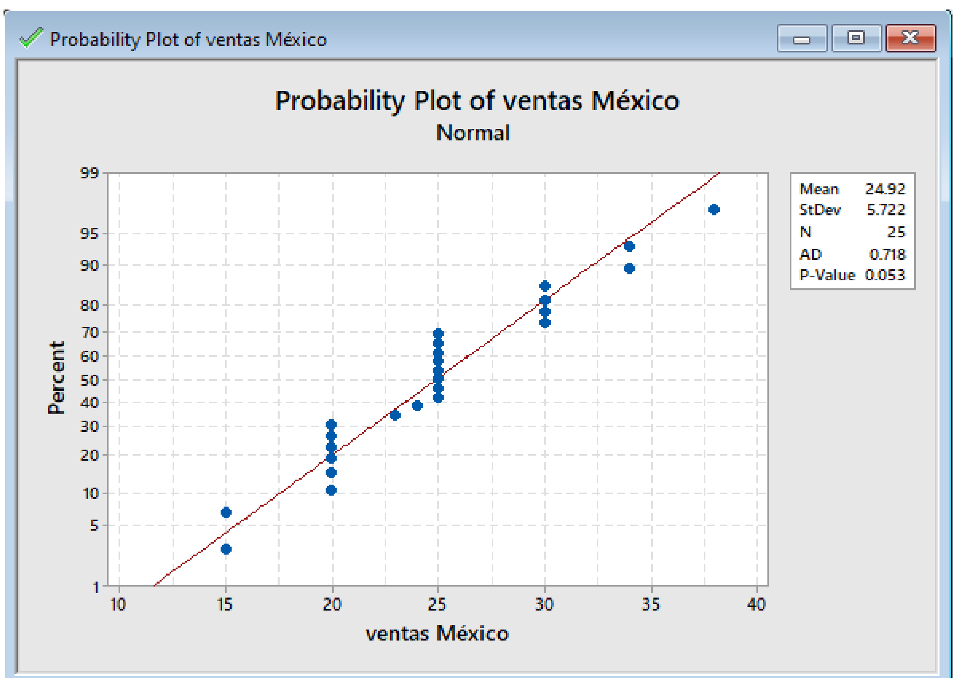
Una vez copiados los datos en Minitab comprobamos que estos tengan una distribución normal esto se hace entrando en: Stat → Basic statistics → Normality test



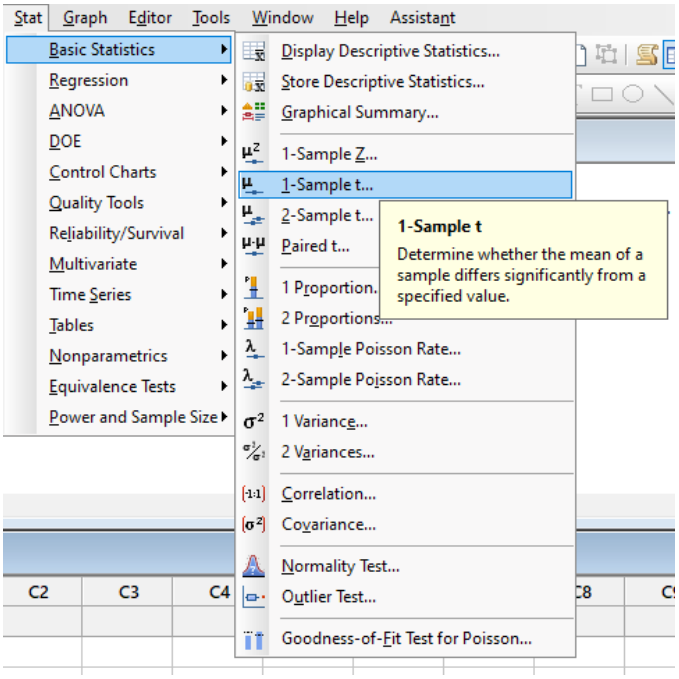
Una vez seleccionado Normality test, escribimos el nombre de nuestra variable y seleccionamos la prueba Anderson-Darling



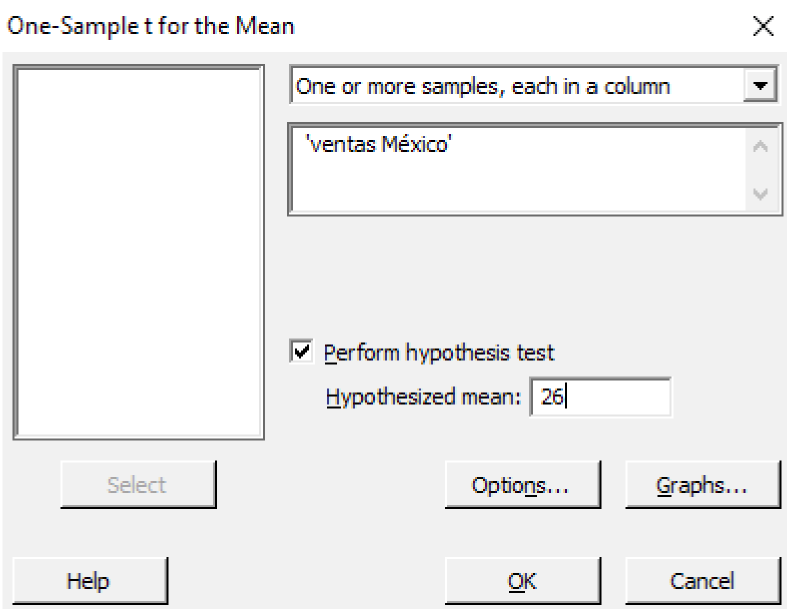
Al presionar ok, mostrará la siguiente pantalla, a nosotros nos interesa que "P-value" sea mayor a 0.05 para confirmar que la distribución de los datos es normal. En nuestro caso, observamos que P es 0.053, por lo que nuestros datos son apenas normales.



Una vez confirmado que nuestros datos son normales, entraremos a Stat → Basic Statistics → 1-Sample t…



Nos aparecerá la siguiente ventana, en ella, seleccionaremos “one or more samples each column”, introduciremos el nombre de la variable y haremos click en la casilla “perform hypothesis test”, tras hacer esto escribiremos el valor de la media poblacional a utilizar, finalmente damos click a “options”



En la siguiente ventana escribimos el nivel de confianza (100-nivel de significancia), seleccionamos la hipótesis alterna y hacemos click en “Ok” en ambas ventanas

Probability Plot of ventas México

One-Sample T: ventas México

```
Test of μ = 26 vs > 26


Variable        N   Mean  StDev  SE Mean  97% Lower Bound      T      P
ventas México  25  24.92   5.72     1.14            22.66  -0.94  0.823
```

|

En la ventana aparecerán el nombre de la variable, tamaño de la muestra, valor de la media y valor de la desviación estándar. En el extremo derecho están el valor del estadístico de prueba y la probabilidad(P)

El valor que nos interesa es P, pues si $P<\alpha$ la hipótesis nula se rechaza

Como P es 0.823 y α es 0.03 entonces podemos concluir que H_0 NO se rechaza

Ejercicio 2

Prueba de hipótesis bilateral para la igualdad de medias con varianzas desconocidas y diferentes, $N_1=25, N_2=25, N.S.=2\%, \left[\mu_1 \neq \acute{X}_1, \mu_2 \neq \acute{X}_2\right]$

Los ejecutivos de misma marca del ejercicio anterior, al ver que sus ganancias en México y Estados Unidos son prácticamente iguales, decidieron hacer también la misma encuesta al azar a 25 clientes en EU, obteniendo los siguientes resultados

50
30
8
37
56
36
66
42
48
70
54
50
23
35
65
30

60
32
30
35
52
52
60
35
44

Ellos suponen que, al ser las ventas muy similares, la media de ambos países debe ser igual. Utilizando los datos del ejercicio anterior y los proporcionados en esta tabla ¿Es posible afirmar que los ejecutivos están en lo correcto? Utilizar un nivel de significancia del 2%

Calculando la media muestral de ambas muestras:

$$\acute{x}_1=\frac{1}{n}\sum_{i=1}^{25} x_i=\frac{1}{25}(34+38+34+30+25+25+15+24+20+15+25+30+23+20+30+20+25+20+25+25+20+25$$

$$X_1^{-}=24.92$$

$$\acute{x}_2=\frac{1}{n}\sum_{i=1}^{25} x_i=\frac{1}{25}(50+30+8+37+56+36+66+42+48+70+54+50+23+35+65+30+60+32+30+35+52+52+$$

$$X_2^{-}=44$$

Calculando la varianza muestral de ambas muestras:

$$S_{n-1}^2=\frac{1}{24}\left((34-29.5)^2+(38-29.5)^2+(34-29.5)^2+(30-29.5)^2+(25-29.5)^2+(25-29.5)^2+(15-29.5)^2+(24-29.5\right.$$

$$S_1^2=32.7433$$

$$S_{n-1}^2=\frac{1}{24}\left[(50-44)^2+(36-44)^2+(54-44)^2+(30-44)^2+(52-44)^2+(30-44)^2+(66-44)^2+(50-44)^2+(60-44)^2+(\right.$$

$S_{n-1}^2=¿$225.0833

Solución

Los ejecutivos desean determinar la equivalencia entre las medias de ambos países con las varianzas poblacionales desconocidas pero consideradas diferentes, por lo tanto, las hipótesis resultan en:

$H_0 : \mu_1 - \mu_2 = 0$

$H_1 : \mu_1 - \mu_2 \neq 0$

Con el estadístico de prueba:

$$T_0^{¿}=\frac{\acute{x}_1-\acute{x}_2}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$$

Y con grados de libertad dados por:

$$v=\frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1+1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2+1}}$$

Utilizando los datos de las tablas obtenemos el valor de $T_0^{¿}$ t_v y de v

$$T_0^{¿}=\frac{24.92-44}{\sqrt{\frac{32.7433}{25}+\frac{225.0833}{25}}}$$

$$T_0^{¿}=-5.9413$$

$$v=\frac{\left(\frac{32.7433}{25}+\frac{225.0833}{25}\right)^2}{\frac{\left(\frac{32.7433}{25}\right)^2}{25+1}+\frac{\left(\frac{225.0833}{25}\right)^2}{25+1}}$$

$$v=33.40777$$

Redondeando al entero más cercano

$$v\approx 33$$

Con $\alpha=0.02$ y los grados de libertad buscamos el valor de $t_{\left(\frac{\alpha}{2},v-2\right)}$ en las tablas.

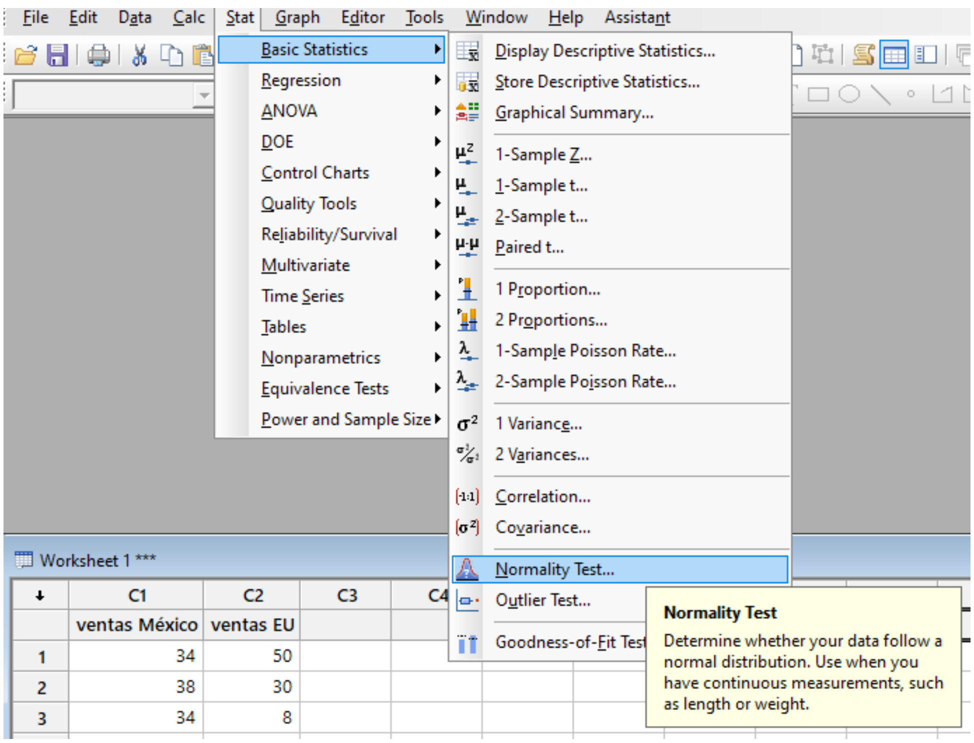
$$t_{(0.01,30)}=2.457$$

2.457>-5.9413

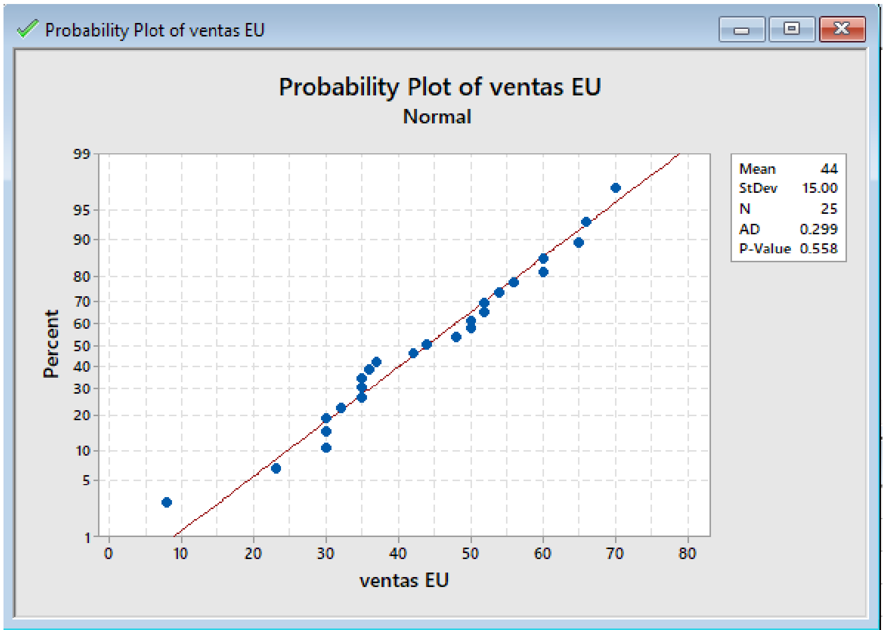
El valor del estadístico se encuentra en la zona de rechazo lo que significa que H_0 se rechaza y la afirmación de los ejecutivos es falsa.

Solución por minitab

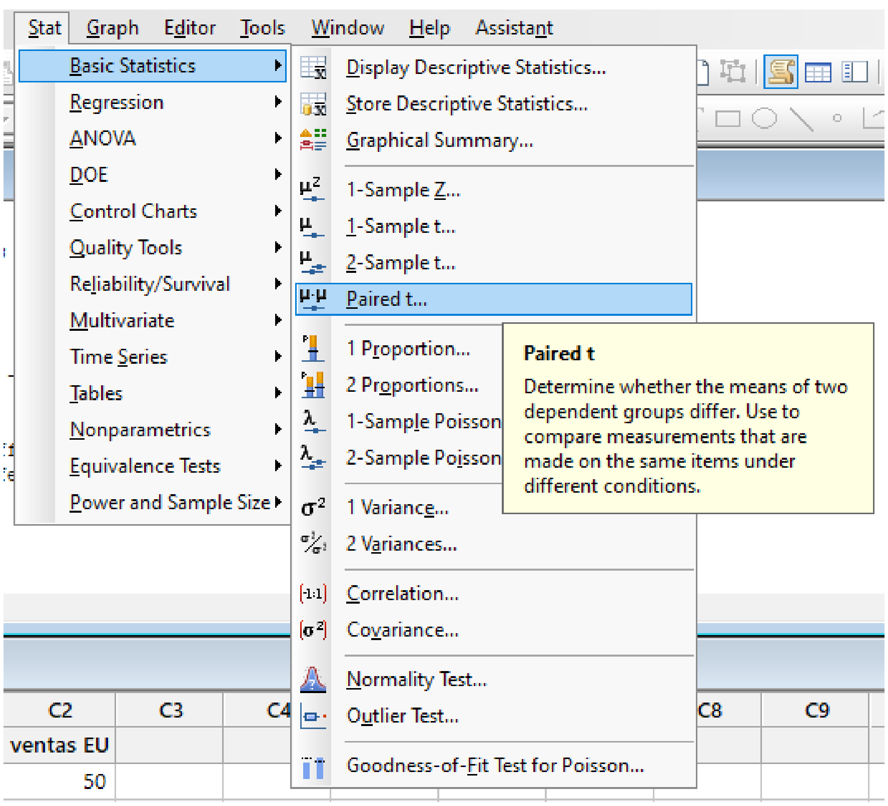
Primero hay que copiar los 2 grupos de datos. Al igual que en el ejercicio anterior, comprobaremos que la distribución de las ventas en EU es normal, esto lo haremos entrando al menú Stat → Basic statistics → Normality test



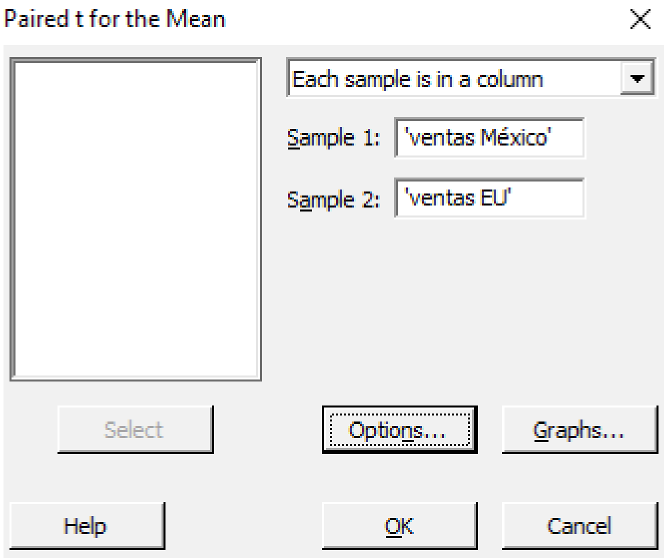
Nos aparecerá la siguiente ventana en la que nos fijaremos que el valor de P sea mayor a 0.05, en nuestro caso es 0.558 por lo que nuestros datos son normales



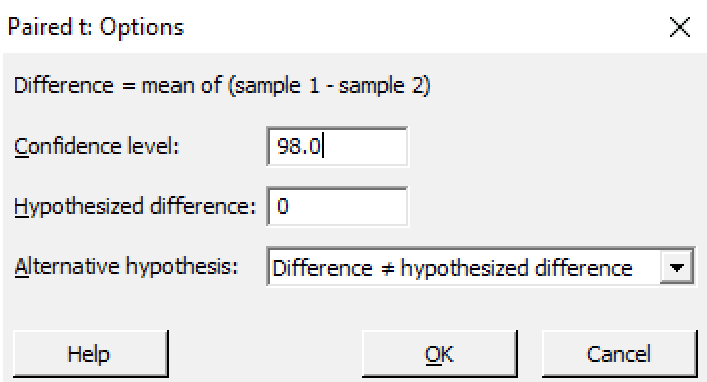
Posteriormente, para comparar ambas medias, iremos al menú Stat ⟶ Basic statistics ⟶ Paired t…



Al abrirse la ventana seleccionamos “each sample is in a columna”, procedemos a elegir las muestras 1 y 2 y presionamos “Options”



En la nueva ventana escribimos el valor del nivel de confianza (100-nivel de significancia), en el segundo recuadro escribimos la diferencia hipotética, en este caso 0 porque suponemos que las medias son iguales, finalmente seleccionamos la hipótesis alternativa y presionamos “Ok” en ambas ventanas



En el recuadro siguiente aparecerán el valor de ambas medias y ambas desviaciones estándar, así como el intervalo de confianza correspondiente para la diferencia de medias. En el último renglón aparecen el valor del estadístico y la probabilidad(P), dado que el valor de P es 0 concluimos que no es posible que las medias sean iguales y por lo tanto H_0 SE RECHAZA

```
Paired T for ventas México - ventas EU

                 N     Mean  StDev  SE Mean
ventas México   25    24.92   5.72     1.14
ventas EU       25    44.00  15.00     3.00
Difference      25   -19.08  18.01     3.60


98% CI for mean difference: (-28.06, -10.10)
T-Test of mean difference = 0 (vs ≠ 0): T-Value = -5.30  P-Value = 0.000
```

Ejercicio 3

Prueba de hipótesis unilateral para la varianza $N=25, N.S.=1\%[\sigma_0^2 \neq S_{n-1}^2]$

Un ejecutivo de una empresa manufacturera de ropa que realizo un estudio de mercado sobre las compras en México de ropa y afirma que la varianza es menor o igual a 20, si el nivel de significancia es de 1%, el tamaño de la muestra es 25 ¿Es posible contradecir al ejecutivo en su afirmación?

México

34	**25**	**25**	**20**	**20**
38	**15**	**30**	**25**	**25**
34	**24**	**23**	**20**	**20**
30	**20**	**20**	**25**	**30**
25	**15**	**30**	**25**	**25**

Usando

$$\acute{x}=\frac{1}{n}\sum_{i=1}^{25} x_i$$

$$\acute{x}_{MEX}=\frac{1}{25}(34+25+25+20+20+38+15+30+25+25+34+24+23+20+20+30+20+20+25+30+25+15+30+2$$

$$\acute{x}_{MEX}=24.92$$

Usando

$$S_{n-1}^2=\frac{1}{n-1}\sum_{i=1}^{25}\left(x_i-\acute{x}\right)^2$$

$$S_{n-1MEX}^2=\frac{1}{24}\left[\left(34-24.92\right)^2+\left(25-24.92\right)^2+\left(25-24.92\right)^2+\left(20-24.92\right)^2+\left(20-24.92\right)^2+\left(38-24.92\right)^2+\left(15-24.92\right)^2\right.$$

$S_{n-1MEX}^2=$¿32.74333

Establecemos nuestra prueba de hipótesis:

$$H_0:\sigma^2\leq 20$$

$$H_1:\sigma^2>20$$

Nuestro estadístico de prueba para la varianza es:

$$x_0^2=\frac{\left(n-1\right)S^2}{\sigma^2}$$

Sustituyendo

$$x_0^2=\frac{\left(25-1\right)32.74333}{20}$$

$$x_0^2=39.29199$$

Como $H_1:\sigma^2>20$ la zona de rechazo se encuentra a la derecha por lo tanto nuestra región de aceptación será si: $x_0^2\leq x_{(\propto,n-1)}^2$

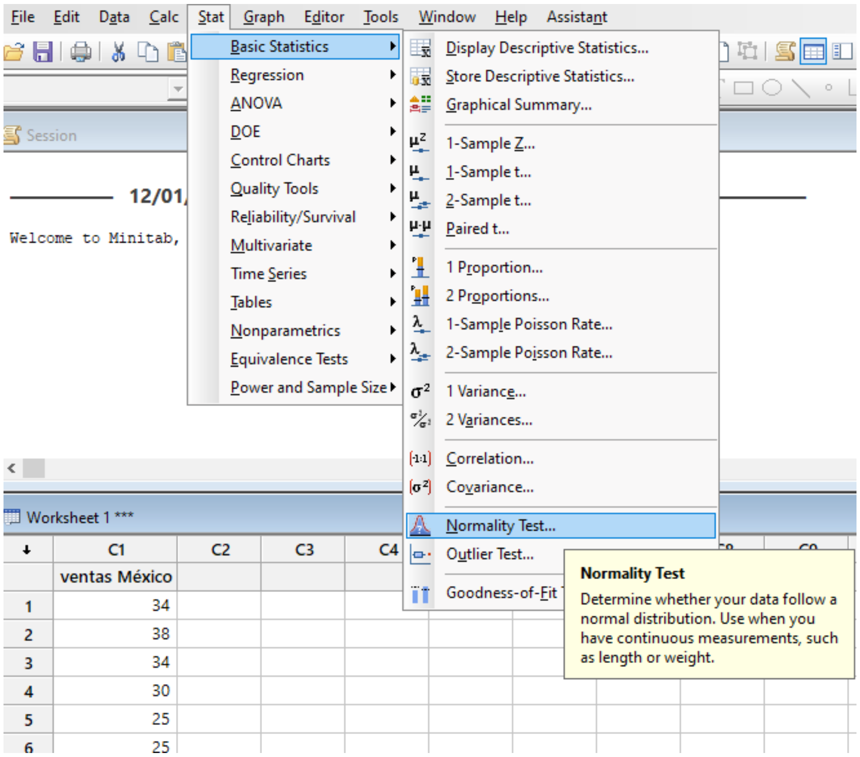
Gl=v	Valores de la distribución Ji-Cuadrada para una probabilidad α de área derecha (intercambiados son los cuantiles)											
	0.005	**0.995**	**0.010**	**0.990**	**0.015**	**0.985**	**0.020**	**0.980**	**0.025**	**0.975**	**0.030**	**0.970**
1	7.8794	0.0000	6.6349	0.0002	5.9165	0.0004	5.4119	0.0006	5.0239	0.0010	4.7093	0.0014
2	10.5965	0.0100	9.2104	0.0201	8.3994	0.0302	7.8241	0.0404	7.3778	0.0506	7.0131	0.0609
3	12.8381	0.0717	11.3449	0.1148	10.4651	0.1516	9.8374	0.1848	9.3484	0.2158	8.9473	0.2451
4	14.8602	0.2070	13.2767	0.2971	12.3391	0.3682	11.6678	0.4294	11.1433	0.4844	10.7119	0.5351
5	16.7496	0.4118	15.0863	0.5543	14.0978	0.6618	13.3882	0.7519	12.8325	0.8312	12.3746	0.9031
6	18.5475	0.6757	16.8119	0.8721	15.7774	1.0160	15.0332	1.1344	14.4494	1.2373	13.9676	1.3296
7	20.2777	0.9893	18.4753	1.2390	17.3984	1.4184	16.6224	1.5643	16.0128	1.6899	15.5091	1.8016
8	21.9549	1.3444	20.0902	1.6465	18.9738	1.8603	18.1682	2.0325	17.5345	2.1797	17.0105	2.3101
9	23.5893	1.7349	21.6660	2.0879	20.5125	2.3348	19.6790	2.5324	19.0228	2.7004	18.4796	2.8485
10	25.1881	2.1558	23.2093	2.5582	22.0206	2.8372	21.1608	3.0591	20.4832	3.2470	19.9219	3.4121
11	26.7569	2.6032	24.7250	3.0535	23.5028	3.3634	22.6179	3.6087	21.9200	3.8157	21.3416	3.9972
12	28.2997	3.0738	26.2170	3.5706	24.9628	3.9103	24.0539	4.1783	23.3367	4.4038	22.7418	4.6009
13	29.8193	3.5650	27.6882	4.1069	26.4034	4.4757	25.4715	4.7654	24.7356	5.0087	24.1249	5.2210
14	31.3194	4.0747	29.1412	4.6604	27.8268	5.0573	26.8727	5.3682	26.1189	5.6287	25.4931	5.8556
15	32.8015	4.6009	30.5780	5.2294	29.2349	5.6534	28.2595	5.9849	27.4884	6.2621	26.8480	6.5032
16	34.2671	5.1422	31.9999	5.8122	30.6292	6.2628	29.6332	6.6142	28.8453	6.9077	28.1908	7.1625
17	35.7184	5.6973	33.4087	6.4077	32.0111	6.8842	30.9950	7.2550	30.1910	7.5642	29.5227	7.8324
18	37.1564	6.2648	34.8052	7.0149	33.3817	7.5165	32.3462	7.9062	31.5264	8.2307	30.8447	8.5120
19	38.5821	6.8439	36.1908	7.6327	34.7419	8.1589	33.6874	8.5670	32.8523	8.9065	32.1577	9.2004
20	39.9969	7.4338	37.5663	8.2604	36.0926	8.8105	35.0196	9.2367	34.1696	9.5908	33.4623	9.8971
21	41.4009	8.0336	38.9322	8.8972	37.4345	9.4708	36.3434	9.9145	35.4789	10.2829	34.7593	10.6013
22	42.7957	8.6427	40.2894	9.5425	38.7681	10.1390	37.6595	10.6000	36.7807	10.9823	36.0491	11.3125
23	44.1814	9.2604	41.6383	10.1957	40.0941	10.8147	38.9683	11.2926	38.0756	11.6885	37.3323	12.0303
24	45.5584	9.8862	42.9798	10.8563	41.4129	11.4974	40.2703	11.9918	39.3641	12.4011	38.6093	12.7543
25	46.9280	10.5196	44.3140	11.5240	42.7252	12.1867	41.5660	12.6973	40.6465	13.1197	39.8804	13.4839
26	48.2898	11.1602	45.6416	12.1982	44.0312	12.8821	42.8558	13.4086	41.9231	13.8439	41.1461	14.2190
27	49.6450	11.8077	46.9628	12.8785	45.3311	13.5833	44.1399	14.1254	43.1945	14.5734	42.4066	14.9592
28	50.9936	12.4613	48.2782	13.5647	46.6255	14.2900	45.4188	14.8475	44.4608	15.3079	43.6622	15.7042
29	52.3355	13.1211	49.5878	14.2564	47.9147	15.0019	46.6926	15.5745	45.7223	16.0471	44.9132	16.4538
30	53.6719	13.7867	50.8922	14.9535	49.1988	15.7188	47.9618	16.3062	46.9792	16.7908	46.1600	17.2076

$x^2_{(0.01,24)}$=42.9798

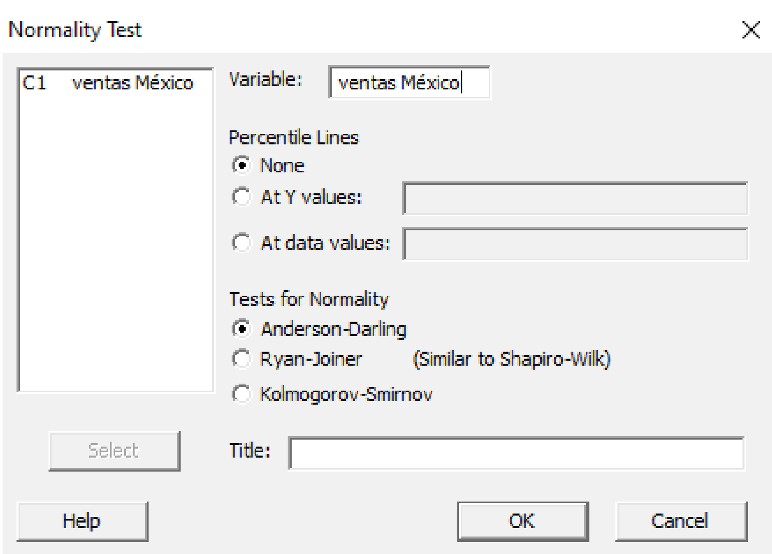
Se cumple que $x_0^2 \leq x^2_{(\propto,n-1)}$ por lo tanto está dentro de la región de aceptación, se puede decir que con un nivel de significancia del 1% la varianza poblacional es menor que 20 esto indica que se rechaza la hipótesis alterna H_1acepta la hipótesis nula H_0

Solución por minitab

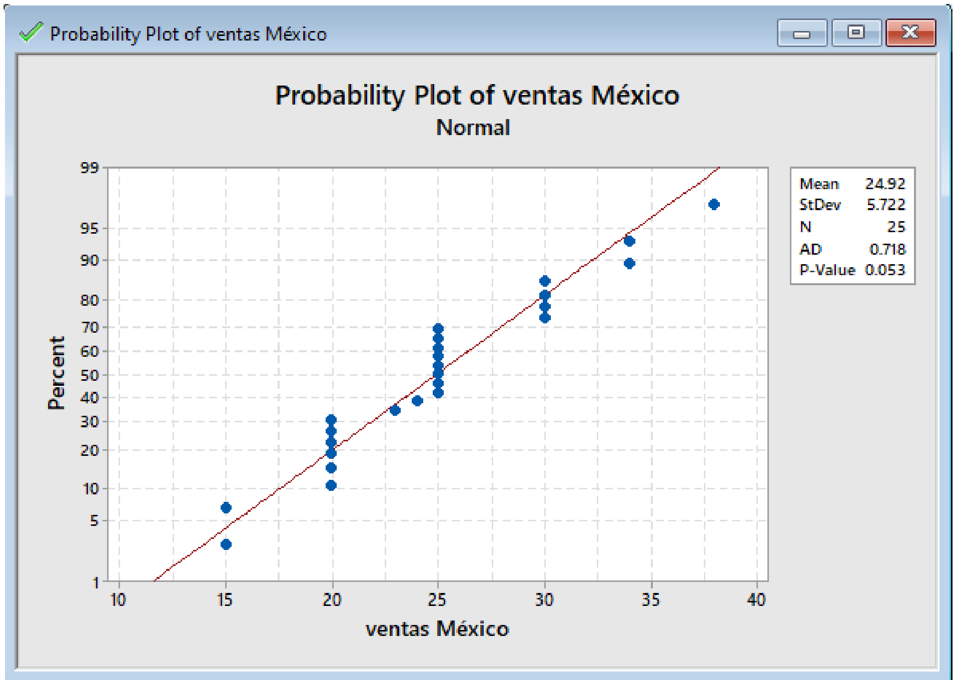
Una vez copiados los datos en Minitab comprobamos que estos tengan una distribución normal esto se hace entrando en: Stat $\longrightarrow$ Basic statistics $\longrightarrow$ Normality test



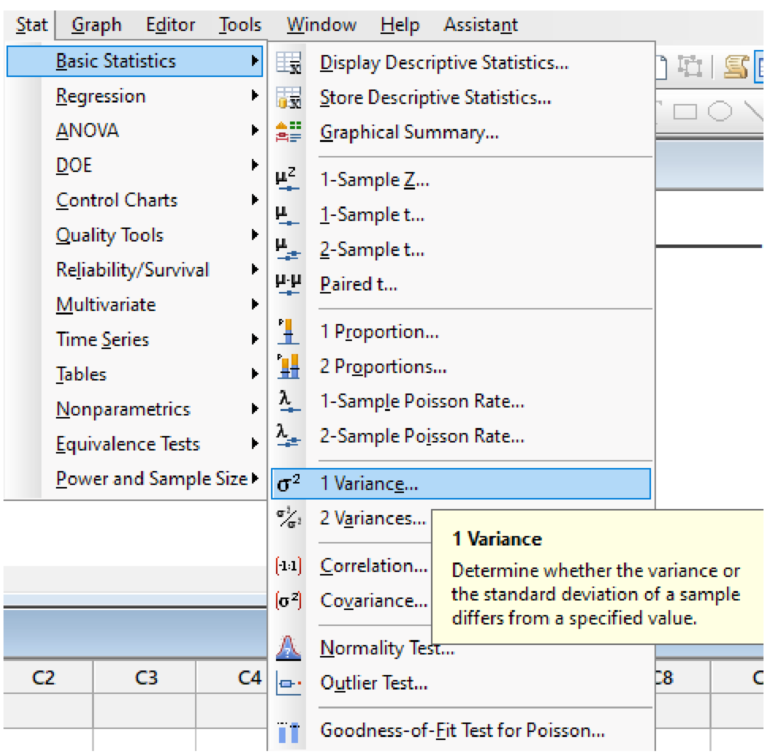
Una vez seleccionado Normality test, escribimos el nombre de nuestra variable y seleccionamos la prueba Anderson-Darling



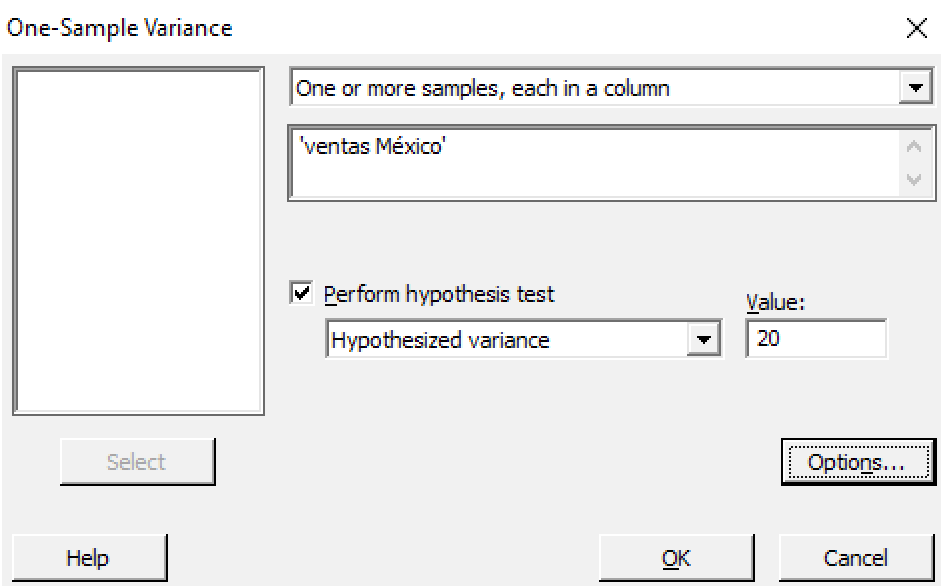
Al presionar ok, mostrará la siguiente pantalla, a nosotros nos interesa que “P-value” sea mayor a 0.05 para confirmar que la distribución de los datos es normal. En nuestro caso, observamos que P es 0.053, por lo que nuestros datos son apenas normales.



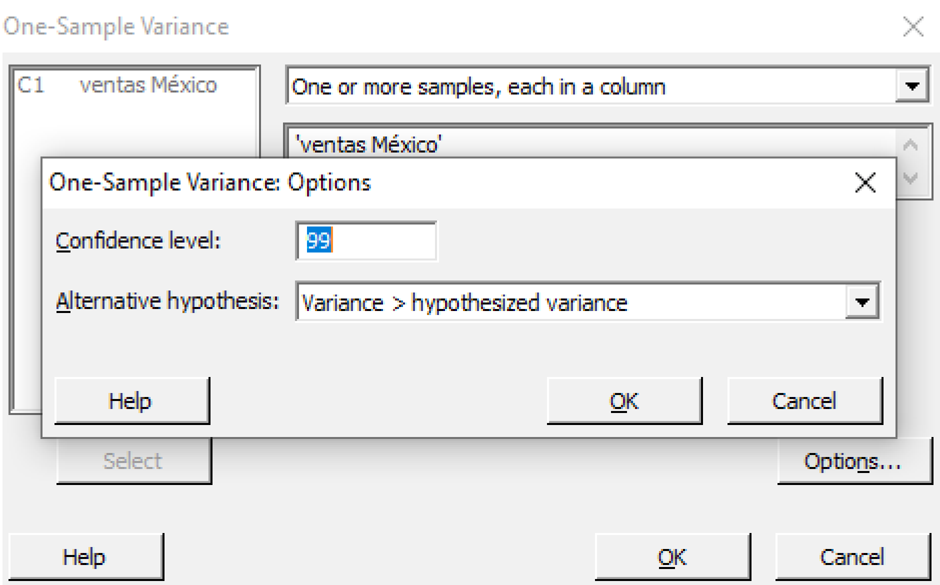
Una vez confirmado que nuestros datos son normales, entraremos a Stat → Basic Statistics → 1 variance



Nos aparecerá la siguiente ventana, en ella, seleccionaremos "one or more samples each column", introduciremos el nombre de la variable y haremos click en la casilla “perform hypothesis test”, tras hacer esto seleccionamos “hypothesized variance” y escribiremos el valor de la varianza hipotética, finalmente damos click a “options”



En la siguiente ventana escribimos el nivel de confianza (100-nivel de significancia), seleccionamos la hipótesis alterna y hacemos click en “Ok” en ambas ventanas



En la siguiente ventana tendremos que fijarnos en el valor de la probabilidad (P value), pues si $P > \alpha$ la hipótesis nula no se rechaza

Como P es 0.025 y αes 0.03 entonces podemos concluir que H_0 NO se rechaza

Session

```
Variable       Method      StDev  for Variance
ventas México  Chi-Square   4.28          18.3
               Bonett       4.16          17.3


Tests

                                Test
Variable       Method      Statistic  DF  P-Value
ventas México  Chi-Square      39.29  24    0.025
               Bonett              —   —    0.033
```

Ejercicio 4

Prueba de hipótesis bilateral para la igualdad de varianzas

$N_1=25, N_2=25, N.S.=5\%[\sigma_1^2 \neq S_{n_1-1}^2, \sigma_2^2 \neq S_{n_2-1}^2]$

Un ejecutivo de una empresa manufacturera de ropa afirma que las varianzas de la ropa vendida en México y en Estados Unidos son iguales, con un nivel de significancia de 0.05 y un tamaño de muestra de 25 para ambas, ¿Es posible contradecir al ejecutivo en su afirmación?

Usando

$$\acute{x}=\frac{1}{n}\sum_{i=1}^{25} x_i$$

$$\acute{x}_{MEX}=\frac{1}{25}(34+25+25+20+20+38+15+30+25+25+34+24+23+20+20+30+20+20+25+30+25+15+30+2$$

$$\acute{x}_{MEX}=24.92$$

$$\acute{x}_{EUA}=\frac{1}{25}(50+36+54+30+52+30+66+50+60+52+8+42+23+32+60+37+48+35+30+35+56+70+65+35$$

$\acute{x}_{EUA} = 44$

Usando

$$S_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^{25} (x_i - \acute{x})^2$$

$$S_{n-1MEX}^2 = \frac{1}{24} [(34-24.92)^2 + (25-24.92)^2 + (25-24.92)^2 + (20-24.92)^2 + (20-24.92)^2 + (38-24.92)^2 + (15-24.92)^2$$

$S_{n-1MEX}^2 = ¿32.74333$

$$S_{n-1EUA}^2 = \frac{1}{24} [(50-44)^2 + (36-44)^2 + (54-44)^2 + (30-44)^2 + (52-44)^2 + (30-44)^2 + (66-44)^2 + (50-44)^2 + (60-44)$$

$S_{n-1EUA}^2 = ¿225.0833$

$n_1 = 25; n_2 = 25; \propto = 0.05; \frac{\propto}{2} = 0.025; S_1^2 = 32.74333; S_2^2 = ¿225.0833$

Establecemos nuestra prueba de hipótesis:

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

Nuestro estadístico de prueba es:

$$F_0 = \frac{S_1^2}{S_2^2}$$

Sustituyendo

$$F_0 = \frac{32.74333}{225.0833}$$

$F_0 = 0.14547$

Para rechazar H_0debemos definir nuestra región de aceptación y si H_0 no se encuentra en la región de aceptación entonces H_0se rechaza.

$f_0 > f_{\frac{\propto}{2}(n_1-1, n_2-1)}$

$f_0 < f_{1-\frac{\alpha}{2}(n_1-1, n_2-1)}$

para definir la región de aceptación usamos:

$$f_{\frac{\alpha}{2}(n_1-1, n_2-1)} = f_{0.025(24,24)}$$

$$f_{1-\frac{\alpha}{2}(n_1-1, n_2-1)} = f_{0.975(24,24)} = \frac{1}{f_{0.025(24,24)}}$$

	ν_1 grados de libertad del numerador y ν_2 grados de libertad del denominador; ν_1 y $\alpha = 0.025$																		
ν_2	20	21	22	23	24	25	26	27	28	29	30	35	40	45	50	55	60	70	80
1	993	994	995	996	997	998	999	1000	1000	1001	1001	1004	1006	1007	1008	1009	1010	1010	1011
2	39.4	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5	39.5
3	14.17	14.16	14.14	14.13	14.12	14.12	14.11	14.10	14.09	14.09	14.08	14.06	14.04	14.02	14.01	14.00	13.99	13.99	13.98
4	8.56	8.55	8.53	8.52	8.51	8.50	8.49	8.48	8.48	8.47	8.46	8.43	8.41	8.39	8.38	8.37	8.36	8.35	8.35
5	6.33	6.31	6.30	6.29	6.28	6.27	6.26	6.25	6.24	6.23	6.23	6.20	6.18	6.16	6.14	6.13	6.12	6.11	6.11
6	5.17	5.15	5.14	5.13	5.12	5.11	5.10	5.09	5.08	5.07	5.07	5.04	5.01	4.99	4.98	4.97	4.96	4.95	4.94
7	4.47	4.45	4.44	4.43	4.41	4.40	4.39	4.39	4.38	4.37	4.36	4.33	4.31	4.29	4.28	4.26	4.25	4.25	4.24
8	4.00	3.98	3.97	3.96	3.95	3.94	3.93	3.92	3.91	3.90	3.89	3.86	3.84	3.82	3.81	3.79	3.78	3.78	3.77
9	3.67	3.65	3.64	3.63	3.61	3.60	3.59	3.58	3.58	3.57	3.56	3.53	3.51	3.49	3.47	3.46	3.45	3.44	3.43
10	3.42	3.40	3.39	3.38	3.37	3.35	3.34	3.34	3.33	3.32	3.31	3.28	3.26	3.24	3.22	3.21	3.20	3.19	3.18
11	3.226	3.211	3.197	3.184	3.173	3.162	3.152	3.142	3.133	3.125	3.118	3.086	3.061	3.042	3.027	3.014	3.004	2.994	2.987
12	3.073	3.057	3.043	3.031	3.019	3.008	2.998	2.988	2.979	2.971	2.963	2.931	2.906	2.887	2.871	2.859	2.848	2.839	2.831
13	2.948	2.932	2.918	2.905	2.893	2.882	2.872	2.862	2.853	2.845	2.837	2.805	2.780	2.760	2.744	2.731	2.720	2.711	2.703
14	2.844	2.828	2.814	2.801	2.789	2.778	2.767	2.758	2.749	2.740	2.732	2.699	2.674	2.654	2.638	2.625	2.614	2.605	2.597
15	2.756	2.740	2.726	2.713	2.701	2.689	2.679	2.669	2.660	2.652	2.644	2.610	2.585	2.565	2.549	2.535	2.524	2.515	2.506
16	2.681	2.665	2.651	2.637	2.625	2.614	2.603	2.594	2.584	2.576	2.568	2.534	2.509	2.488	2.472	2.458	2.447	2.437	2.429
17	2.616	2.600	2.585	2.572	2.560	2.548	2.538	2.528	2.519	2.510	2.502	2.468	2.442	2.422	2.405	2.392	2.380	2.370	2.362
18	2.559	2.543	2.529	2.515	2.503	2.491	2.481	2.471	2.461	2.453	2.445	2.410	2.384	2.364	2.347	2.333	2.321	2.312	2.303
19	2.509	2.493	2.478	2.465	2.452	2.441	2.430	2.420	2.411	2.402	2.394	2.359	2.333	2.312	2.295	2.281	2.270	2.260	2.251
20	2.464	2.448	2.434	2.420	2.408	2.396	2.385	2.375	2.366	2.357	2.349	2.314	2.287	2.266	2.249	2.235	2.223	2.213	2.205
21	2.425	2.409	2.394	2.380	2.368	2.356	2.345	2.335	2.325	2.317	2.308	2.273	2.246	2.225	2.208	2.194	2.182	2.172	2.163
22	2.389	2.373	2.358	2.344	2.331	2.320	2.309	2.299	2.289	2.280	2.272	2.237	2.210	2.188	2.171	2.157	2.145	2.134	2.125
23	2.357	2.340	2.325	2.312	2.299	2.287	2.276	2.266	2.256	2.247	2.239	2.204	2.176	2.155	2.137	2.123	2.111	2.100	2.091
24	2.327	2.311	2.296	2.282	2.269	2.257	2.246	2.236	2.226	2.217	2.209	2.173	2.146	2.124	2.107	2.092	2.080	2.069	2.060
25	2.300	2.284	2.269	2.255	2.242	2.230	2.219	2.209	2.199	2.190	2.182	2.146	2.118	2.096	2.079	2.064	2.052	2.041	2.032
26	2.276	2.259	2.244	2.230	2.217	2.205	2.194	2.184	2.174	2.165	2.157	2.120	2.093	2.071	2.053	2.038	2.026	2.015	2.006
27	2.253	2.237	2.222	2.208	2.195	2.183	2.171	2.161	2.151	2.142	2.133	2.097	2.069	2.047	2.029	2.014	2.002	1.991	1.982
28	2.232	2.216	2.201	2.187	2.174	2.161	2.150	2.140	2.130	2.121	2.112	2.076	2.048	2.025	2.007	1.992	1.980	1.969	1.959
29	2.213	2.196	2.181	2.167	2.154	2.142	2.131	2.120	2.110	2.101	2.092	2.056	2.028	2.005	1.987	1.972	1.959	1.948	1.939
30	2.195	2.178	2.163	2.149	2.136	2.124	2.112	2.102	2.092	2.083	2.074	2.037	2.009	1.986	1.968	1.953	1.940	1.929	1.920

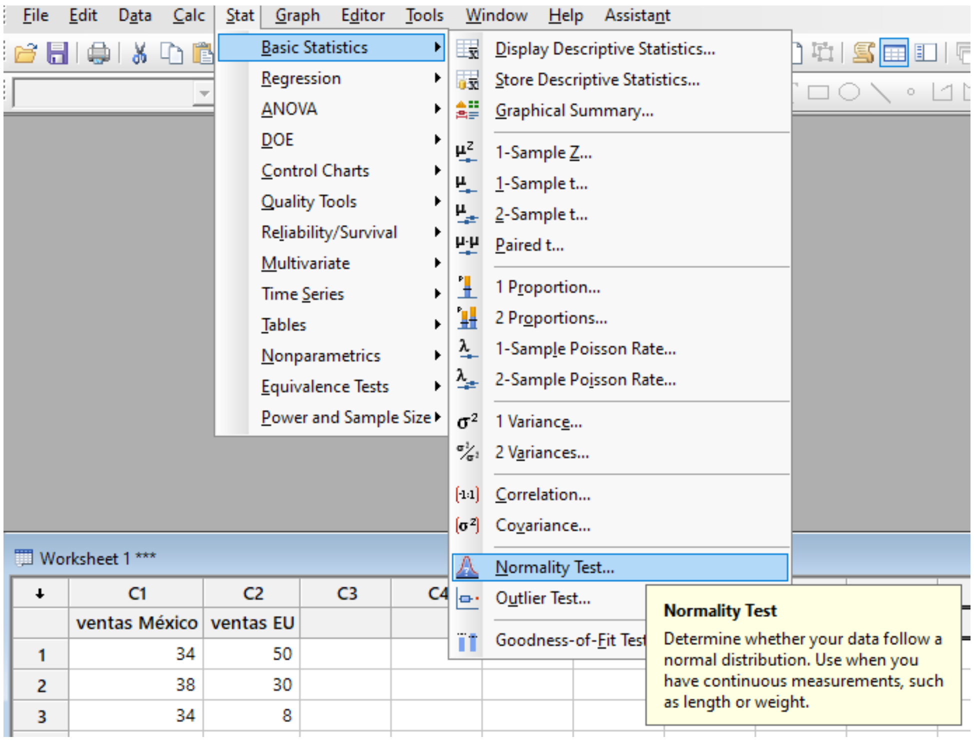
$$f_{0.025(24,24)} = 2.269$$

$$f_{0.975(24,24)} = \frac{1}{f_{0.025(24,24)}} = \frac{1}{2.269} = 0.44072$$

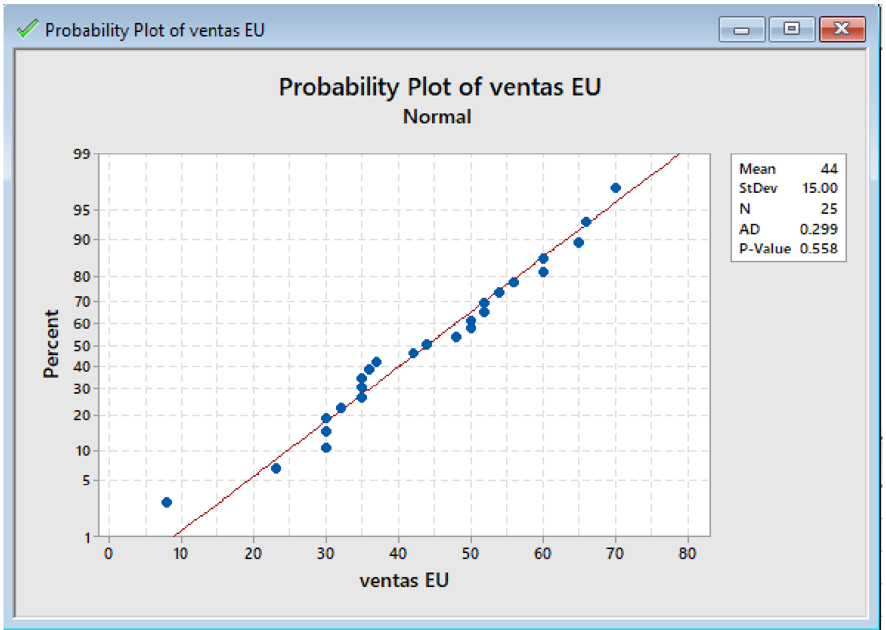
La región de aceptación se encuentra entre 0.44072 y 2.269 y $F_0 = 0.14547$ es decir no se encuentra en la región de aceptación, por lo tanto H_0 se rechaza y H_1 se acepta para un nivel de significancia de 5%, las varianzas de las compras de prendas de ropa en México y Estados Unidos no son iguales.

Solución por minitab

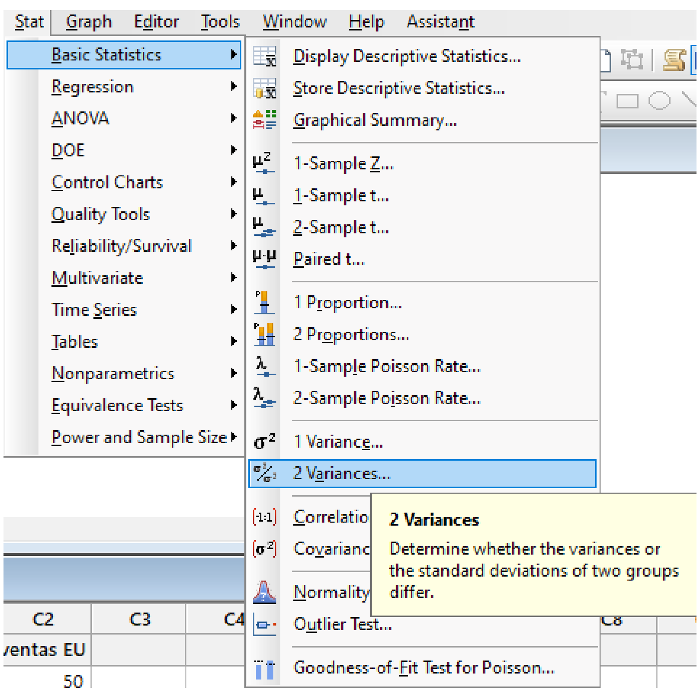
Al igual que en el ejercicio anterior, comprobaremos que la distribución de los datos es normal, sabemos que los datos de ventas en México lo son así que solo comprobaremos los datos de EU. Esto lo haremos entrando al menú Stat → Basic statistics → Normality test



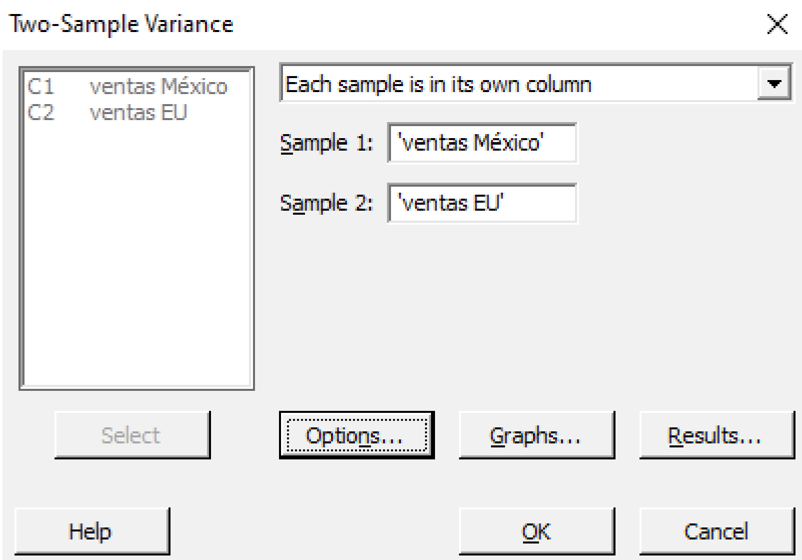
Nos aparecerá la siguiente ventana en la que nos fijaremos que el valor de P sea mayor a 0.05, en nuestro caso es 0.558 por lo que nuestros datos son normales



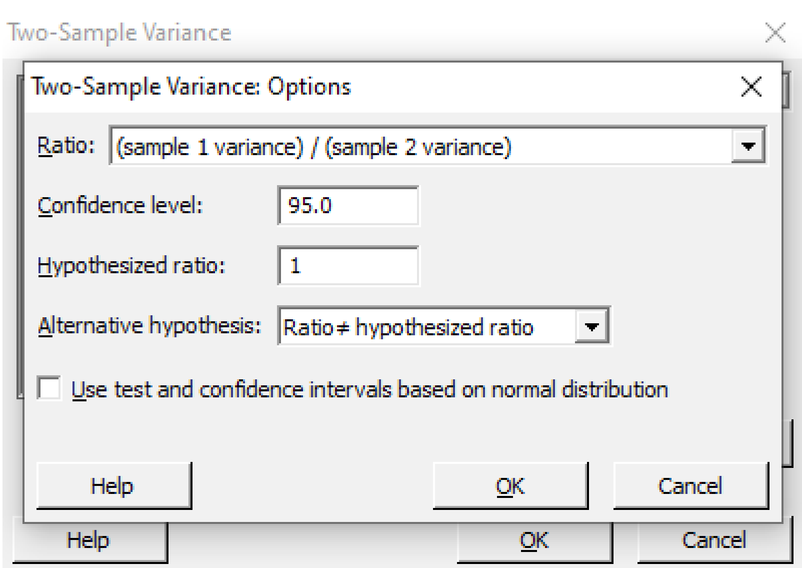
Posteriormente, para comparar ambas medias, iremos al menú Stat ⟶ Basic statistics ⟶ 2 variances



Al abrirse la ventana seleccionamos “each sample is in its own column”, procedemos a elegir las muestras 1 y 2 y presionamos “Options”



En la nueva ventana seleccionaremos en la pestaña ratio “sample 1 variance/sample 2 variance” escribimos el valor del nivel de confianza (100-nivel de significancia), en el segundo recuadro escribimos el radio hipotético, en este caso 1 porque suponemos que las medias son iguales, finalmente seleccionamos la hipótesis alternativa y presionamos “Ok” en ambas ventanas



En el recuadro siguiente buscaremos el valor la probabilidad(P), pues si P>α entonces H_0 no se rechaza

```
Tests

                              Test
Method  DF1  DF2  Statistic  P-Value
Bonett    1    —      13.62    0.000
Levene    1   48      21.31    0.000
```

Test and CI for Two Variances: ventas México, ventas EU

|

Dado que 0<0.05(P<α) entonces H_0 SE RECHAZA

Ejercicio 5

Prueba de hipótesis unilateral para una proporción $N=25, N.S.=4\%, [p_0 \neq \hat{p}]$

El Instituto Nacional de Estadística y Geografía, por sus siglas INEGI, asegura que, dentro de la población mexicana, de un grupo de 9,652,596 personas con edades entre los 18 y los 29 años, más del 40% tienen un interés grande o muy grande por nuevos inventos, descubrimientos científicos y desarrollo tecnológico.

Se realizó una encuesta a 25 personas con edades dentro del mismo rango su interés por este tema con opciones de respuesta: Nulo, moderado, grande o muy grande que son las mismas opciones de respuesta utilizadas por el INEGI y se obtuvieron los siguientes resultados

Muy grande	**Grande**	**Muy grande**	**Muy grande**	**Grande**
Grande	**Grande**	**Grande**	**Moderado**	**Muy grande**
Grande	**Grande**	**Moderado**	**Moderado**	**Grande**
Moderado	**Muy grande**	**Grande**	**Grande**	**Grande**
Grande	**Grande**	**Muy grande**	**Muy grande**	**Grande**

Partiendo de los resultados de la encuesta, y con un nivel de significancia del 4%, se quiere saber si la afirmación puede afirmarse o rechazarse.

Solución

Primero se debe calcular la proporción de la muestra tomada, en la siguiente tabla se resaltan las respuestas de Grande y Muy grande

Muy grande	**Grande**	**Muy grande**	**Muy grande**	**Grande**
Grande	**Grande**	**Grande**	**Moderado**	**Muy grande**
Grande	**Grande**	**Moderado**	**Moderado**	**Grande**
Moderado	**Muy grande**	**Grande**	**Grande**	**Grande**
Grande	**Grande**	**Muy grande**	**Muy grande**	**Grande**

Se tiene entonces que, de 25 personas, 21 tienen un interés grande o muy grande por nuevos inventos, descubrimientos científicos y desarrollo tecnológico de manera que la proporción queda

$$\hat{p}=\frac{21}{25}=0.84$$

Se plantea la hipótesis de la siguiente manera

$$H_0: p \leq 0.4 \quad H_1: p > 0.4$$

Tomando en cuenta que el rechazo de la hipótesis nula significaría que la afirmación inicial es correcta.

Se utiliza la distribución normal estándar, es decir, el parámetro

$$Z=\frac{\hat{p}-p_0}{\sqrt{\frac{p_o(1-p_o)}{n}}}$$

La región de rechazo es en la que el estadístico obtenido sea mayor al valor de la distribución normal estándar para el nivel de significancia, es decir, cuando

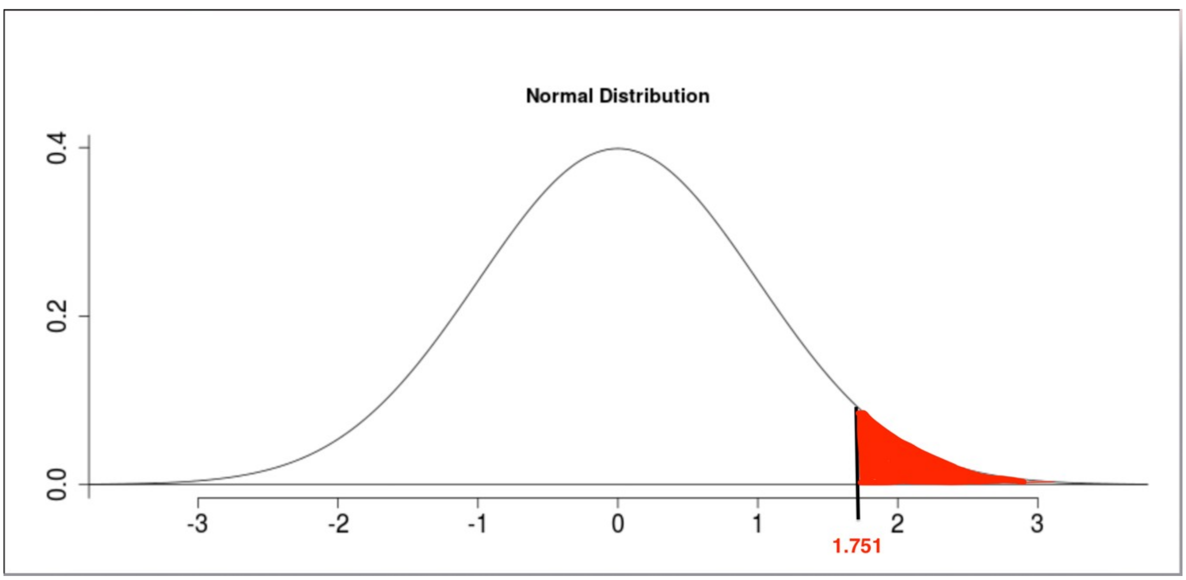
$$z_0 > z_\alpha$$

Sabiendo que el nivel de significancia es del 4%, se busca el valor de $z_{0.04}$ en las tablas y se encuentra que

$$z_{0.04}=1.751$$

%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$
0.0	$-\infty$	0	**5.0**	-1.645	0.063	**10.0**	-1.282	0.126	**15.0**	-1.036	0.189	**20.0**	-0.842	0.253	**25.0**	-0.674	0.319
0.1	-3.090	0.001	**5.1**	-1.635	0.064	**10.1**	-1.276	0.127	**15.1**	-1.032	0.190	**20.1**	-0.838	0.255	**25.1**	-0.671	0.320
0.2	-2.878	0.003	**5.2**	-1.626	0.065	**10.2**	-1.270	0.128	**15.2**	-1.028	0.192	**20.2**	-0.834	0.256	**25.2**	-0.668	0.321
0.3	-2.748	0.004	**5.3**	-1.616	0.066	**10.3**	-1.265	0.129	**15.3**	-1.024	0.193	**20.3**	-0.831	0.257	**25.3**	-0.665	0.323
0.4	-2.652	0.005	**5.4**	-1.607	0.068	**10.4**	-1.259	0.131	**15.4**	-1.019	0.194	**20.4**	-0.827	0.259	**25.4**	-0.662	0.324
0.5	-2.576	0.006	**5.5**	-1.598	0.069	**10.5**	-1.254	0.132	**15.5**	-1.015	0.196	**20.5**	-0.824	0.260	**25.5**	-0.659	0.325
0.6	-2.512	0.008	**5.6**	-1.589	0.070	**10.6**	-1.248	0.133	**15.6**	-1.011	0.197	**20.6**	-0.820	0.261	**25.6**	-0.656	0.327
0.7	-2.457	0.009	**5.7**	-1.580	0.071	**10.7**	-1.243	0.135	**15.7**	-1.007	0.198	**20.7**	-0.817	0.262	**25.7**	-0.653	0.328
0.8	-2.409	0.010	**5.8**	-1.572	0.073	**10.8**	-1.237	0.136	**15.8**	-1.003	0.199	**20.8**	-0.813	0.264	**25.8**	-0.650	0.329
0.9	-2.366	0.011	**5.9**	-1.563	0.074	**10.9**	-1.232	0.137	**15.9**	-0.999	0.201	**20.9**	-0.810	0.265	**25.9**	-0.646	0.331
1.0	-2.326	0.013	**6.0**	-1.555	0.075	**11.0**	-1.227	0.138	**16.0**	-0.994	0.202	**21.0**	-0.806	0.266	**26.0**	-0.643	0.332
1.1	-2.290	0.014	**6.1**	-1.546	0.077	**11.1**	-1.221	0.140	**16.1**	-0.990	0.203	**21.1**	-0.803	0.268	**26.1**	-0.640	0.333
1.2	-2.257	0.015	**6.2**	-1.538	0.078	**11.2**	-1.216	0.141	**16.2**	-0.986	0.204	**21.2**	-0.800	0.269	**26.2**	-0.637	0.335
1.3	-2.226	0.016	**6.3**	-1.530	0.079	**11.3**	-1.211	0.142	**16.3**	-0.982	0.206	**21.3**	-0.796	0.270	**26.3**	-0.634	0.336
1.4	-2.197	0.018	**6.4**	-1.522	0.080	**11.4**	-1.206	0.143	**16.4**	-0.978	0.207	**21.4**	-0.793	0.272	**26.4**	-0.631	0.337
1.5	-2.170	0.019	**6.5**	-1.514	0.082	**11.5**	-1.200	0.145	**16.5**	-0.974	0.208	**21.5**	-0.789	0.273	**26.5**	-0.628	0.338
1.6	-2.144	0.020	**6.6**	-1.506	0.083	**11.6**	-1.195	0.146	**16.6**	-0.970	0.210	**21.6**	-0.786	0.274	**26.6**	-0.625	0.340
1.7	-2.120	0.021	**6.7**	-1.499	0.084	**11.7**	-1.190	0.147	**16.7**	-0.966	0.211	**21.7**	-0.782	0.275	**26.7**	-0.622	0.341
1.8	-2.097	0.023	**6.8**	-1.491	0.085	**11.8**	-1.185	0.148	**16.8**	-0.962	0.212	**21.8**	-0.779	0.277	**26.8**	-0.619	0.342
1.9	-2.075	0.024	**6.9**	-1.483	0.087	**11.9**	-1.180	0.150	**16.9**	-0.958	0.213	**21.9**	-0.776	0.278	**26.9**	-0.616	0.344
2.0	-2.054	0.025	**7.0**	-1.476	0.088	**12.0**	-1.175	0.151	**17.0**	-0.954	0.215	**22.0**	-0.772	0.279	**27.0**	-0.613	0.345
2.1	-2.034	0.026	**7.1**	-1.468	0.089	**12.1**	-1.170	0.152	**17.1**	-0.950	0.216	**22.1**	-0.769	0.281	**27.1**	-0.610	0.346
2.2	-2.014	0.028	**7.2**	-1.461	0.090	**12.2**	-1.165	0.154	**17.2**	-0.946	0.217	**22.2**	-0.765	0.282	**27.2**	-0.607	0.348
2.3	-1.995	0.029	**7.3**	-1.454	0.092	**12.3**	-1.160	0.155	**17.3**	-0.942	0.219	**22.3**	-0.762	0.283	**27.3**	-0.604	0.349
2.4	-1.977	0.030	**7.4**	-1.447	0.093	**12.4**	-1.155	0.156	**17.4**	-0.938	0.220	**22.4**	-0.759	0.285	**27.4**	-0.601	0.350
2.5	-1.960	0.031	**7.5**	-1.440	0.094	**12.5**	-1.150	0.157	**17.5**	-0.935	0.221	**22.5**	-0.755	0.286	**27.5**	-0.598	0.352
2.6	-1.943	0.033	**7.6**	-1.433	0.095	**12.6**	-1.146	0.159	**17.6**	-0.931	0.222	**22.6**	-0.752	0.287	**27.6**	-0.595	0.353
2.7	-1.927	0.034	**7.7**	-1.426	0.097	**12.7**	-1.141	0.160	**17.7**	-0.927	0.224	**22.7**	-0.749	0.288	**27.7**	-0.592	0.354
2.8	-1.911	0.035	**7.8**	-1.419	0.098	**12.8**	-1.136	0.161	**17.8**	-0.923	0.225	**22.8**	-0.745	0.290	**27.8**	-0.589	0.356
2.9	-1.896	0.036	**7.9**	-1.412	0.099	**12.9**	-1.131	0.162	**17.9**	-0.919	0.226	**22.9**	-0.742	0.291	**27.9**	-0.586	0.357
3.0	-1.881	0.038	**8.0**	-1.405	0.100	**13.0**	-1.126	0.164	**18.0**	-0.915	0.228	**23.0**	-0.739	0.292	**28.0**	-0.583	0.358
3.1	-1.866	0.039	**8.1**	-1.398	0.102	**13.1**	-1.122	0.165	**18.1**	-0.912	0.229	**23.1**	-0.736	0.294	**28.1**	-0.580	0.360
3.2	-1.852	0.040	**8.2**	-1.392	0.103	**13.2**	-1.117	0.166	**18.2**	-0.908	0.230	**23.2**	-0.732	0.295	**28.2**	-0.577	0.361
3.3	-1.838	0.041	**8.3**	-1.385	0.104	**13.3**	-1.112	0.167	**18.3**	-0.904	0.231	**23.3**	-0.729	0.296	**28.3**	-0.574	0.362
3.4	-1.825	0.043	**8.4**	-1.379	0.105	**13.4**	-1.108	0.169	**18.4**	-0.900	0.233	**23.4**	-0.726	0.298	**28.4**	-0.571	0.364
3.5	-1.812	0.044	**8.5**	-1.372	0.107	**13.5**	-1.103	0.170	**18.5**	-0.896	0.234	**23.5**	-0.722	0.299	**28.5**	-0.568	0.365
3.6	-1.799	0.045	**8.6**	-1.366	0.108	**13.6**	-1.098	0.171	**18.6**	-0.893	0.235	**23.6**	-0.719	0.300	**28.6**	-0.565	0.366
[illegible]	[illegible]	0.046	**8.7**	-1.359	0.109	**13.7**	-1.094	0.173	**18.7**	-0.889	0.237	**23.7**	-0.716	0.302	**28.7**	-0.562	0.368
[illegible]	[illegible]	[illegible]	**8.8**	-1.353	0.111	**13.8**	-1.089	0.174	**18.8**	-0.885	0.238	**23.8**	-0.713	0.303	**28.8**	-0.559	0.369
3.9	-1.762	[illegible]	**8.9**	-1.347	0.112	**13.9**	-1.085	0.175	**18.9**	-0.882	0.239	**23.9**	-0.710	0.304	**28.9**	-0.556	0.371
4.0	-1.751	[illegible]	**9.0**	-1.341	0.113	**14.0**	-1.080	0.176	**19.0**	-0.878	0.240	**24.0**	-0.706	0.305	**29.0**	-0.553	0.372
4.1	-1.739	[illegible]	**9.1**	-1.335	0.114	**14.1**	-1.076	0.178	**19.1**	-0.874	0.242	**24.1**	-0.703	0.307	**29.1**	-0.550	0.373
[illegible]	[illegible]	[illegible]	**9.2**	-1.329	0.116	**14.2**	-1.071	0.179	**19.2**	-0.871	0.243	**24.2**	-0.700	0.308	**29.2**	-0.548	0.375
[illegible]	[illegible]	[illegible]	**9.3**	-1.323	0.117	**14.3**	-1.067	0.180	**19.3**	-0.867	0.244	**24.3**	-0.697	0.309	**29.3**	-0.545	0.376

Se grafica entonces la región de rechazo



Sustituyendo los valores con los obtenidos de la muestra

$$z_0 = \frac{0.84 - 0.4}{\sqrt{\frac{(0.4)(0.6)}{25}}} = 4.4907$$

Ya que

$$4.4907 > 1.751 z_0 > z_\alpha$$

Se dice que el estadístico encontrado está en el área de rechazo, por lo que se rechaza la hipótesis nula H_0

$$H_0: p \leq 0.4$$
$$H_1: p > 0.4$$

Se concluye que con base en los datos de la muestra y un nivel de significancia del 4%, existe evidencia suficiente para afirmar que más del 40% de la población mexicana con edades entre 18 y 29 años está interesada en grande o muy grande medida por nuevos inventos, descubrimientos científicos y desarrollo tecnológico.

Solución con minitab

El primer paso para la solución en minitab es registrar los datos que se obtuvieron en la muestra en una sola columna como se muestra en la siguiente imagen, en donde el número 1 representa que el encuestado contestó que su interés es grande o muy grande y el número 0 representa que no fue así.

↓	C1
	Interes
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	0
10	1
11	1
12	1
13	0
14	0
15	1
16	0
17	1
18	1
19	1
20	1
21	1

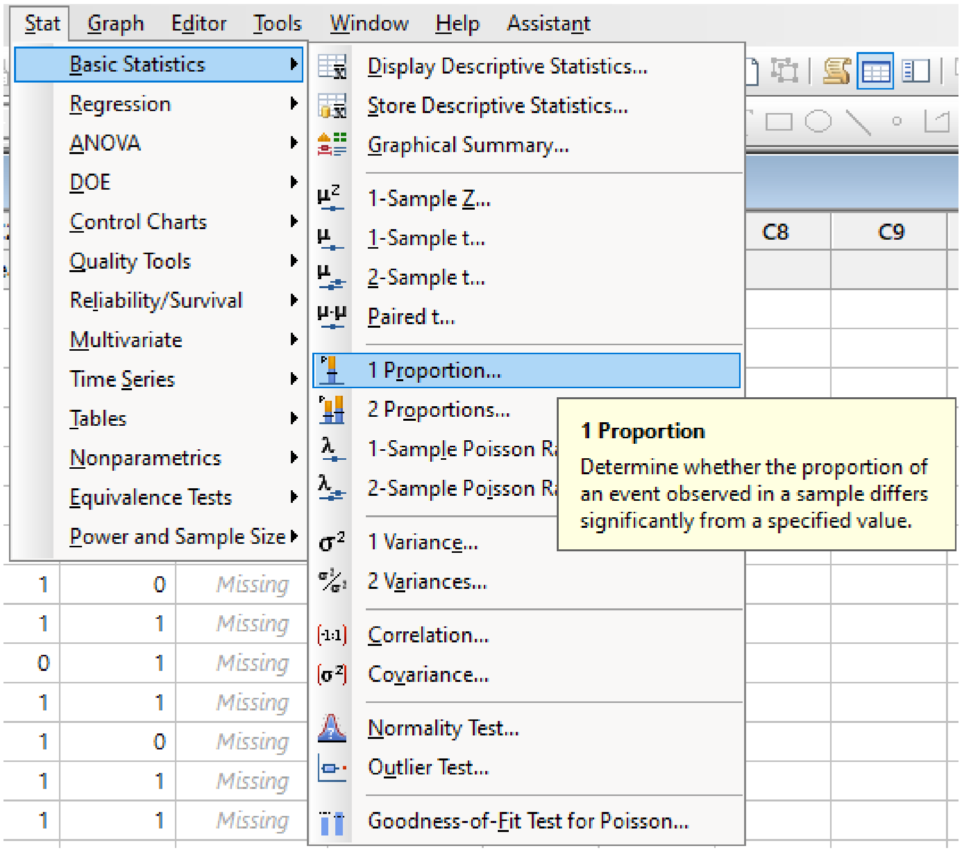
El segundo es establecer la hipótesis nula y la alterna, es recomendable que se escriban en el mismo documento para tener fácil acceso a ellas.

1 proporcion
H0: P<=0.4
H1: P>0.4

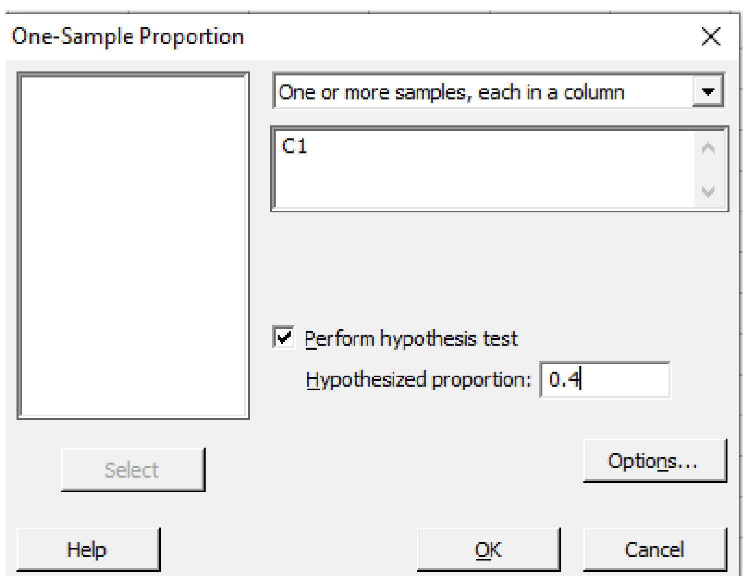
Después en el menú superior se seleccionan las siguientes

$Stat \rightarrow Basic\ Statistics \rightarrow 1\ proportion$

Como se observa en la siguiente imagen



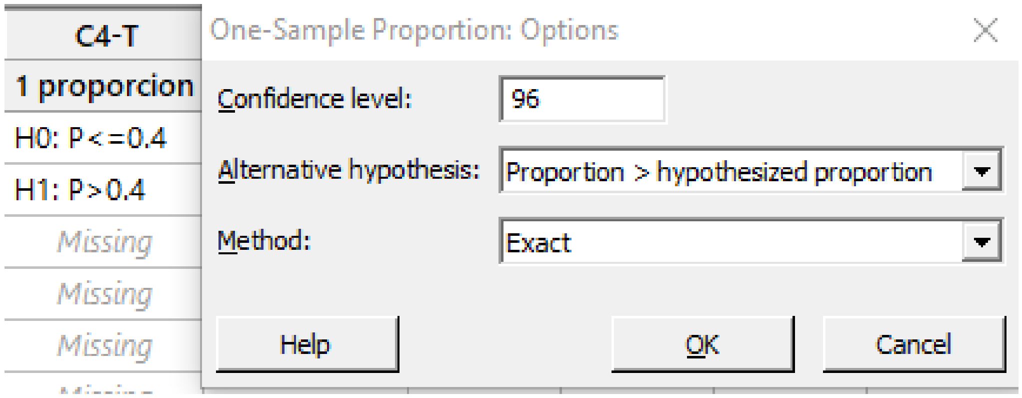
Lo que hará que aparezca la ventana que se muestra a continuación



En el primer campo se debe seleccionar la opción de *One or more* samples, *each in a column*. Ya que los datos se encuentran en una columna.
En el segundo campo se selecciona la columna en la que se anotaron los datos de la muestra, en este caso la columna es C1.

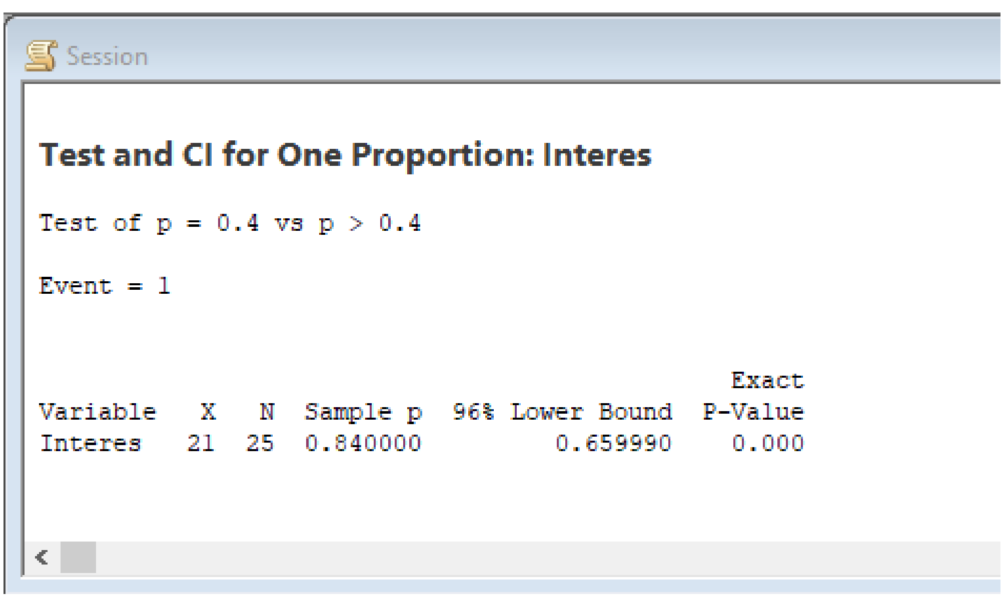
En el tercer campo se escribe el nivel de significancia del ejercicio como decimal que en este caso es 0.4.
Se debe seleccionar la casilla de *Perform hypothesis test*

Luego, en esa misma ventana se hace clic en el botón de *Options* lo que abrirá una ventana como la siguiente



En esta ventana, el primer campo se refiere al nivel de confianza, que para nuestro caso es 96, resultado del 100% menos el 4% del nivel de significancia.
En el segundo campo se selecciona la opción que corresponde a la hipótesis alterna, en este caso, es que la proporción sea mayor a la proporción de la hipótesis.
El último campo simplemente se deja como *Exact*

Por último, se da clic en
$OK \rightarrow OK$
Se obtiene entonces el siguiente resultado



Lo que es de nuestro interés en este ejercicio es lo que el software es el valor denominado *Exact P-Value* y ya que este valor es menor al nivel de significancia, se rechaza la hipótesis nula.

Se concluye que con base en los datos de la muestra y un nivel de significancia del 4%, existe evidencia suficiente para afirmar que más del 40% de la población mexicana con edades entre 18 y 29 años está interesada en grande o muy grande medida por nuevos inventos, descubrimientos científicos y desarrollo tecnológico.

Ejercicio 6

Prueba de hipótesis bilateral para la igualdad de proporciones

$N_1=25, N_2=25, N.S.=3\%, [p_1 \neq \widehat{p}_1, p_2 \neq \widehat{p}_2]$

En la Universidad Nacional Autónoma de México, las carreras universitarias se dividen en cuatro áreas que son las siguientes:

1. Ciencias físico-matemáticas y de las ingenierías
2. Ciencias biológicas y de la salud
3. Ciencias sociales
4. Humanidades y de las artes

Y en cada una de ellas se encuentran alumnos con intereses afines al área en la que se encuentran, sin embargo, todos los alumnos tienen intereses que no se relacionan demasiado con la carrera que estudian. En la siguiente encuesta se preguntó a 25 alumnos de las carreras del área 1 y a 25 alumnos de carreras del área 2 cuál es su interés en política un tema que no se relaciona con ninguna de las dos áreas con opciones de respuesta: Nulo, Moderado, Grande y Muy grande de donde se obtuvieron los siguientes datos

De área 1

Nulo	Moderado	Grande	Nulo	Moderado
Grande	Moderado	Moderado	Moderado	Grande
Moderado	Moderado	Moderado	Moderado	Grande
Moderado	Moderado	Moderado	Moderado	Moderado
Grande	Nulo	Moderado	Grande	Moderado

De área 2

Nulo	**Grande**	**Muy grande**	**Nulo**	**Moderado**
Moderado	**Nulo**	**Grande**	**Moderado**	**Moderado**
Moderado	**Grande**	**Moderado**	**Moderado**	**Nulo**
Nulo	**Moderado**	**Moderado**	**Grande**	**Grande**
Moderado	**Moderado**	**Moderado**	**Grande**	**Moderado**

Y se desea saber si las dos áreas tienen el mismo porcentaje de alumnos con un interés moderado en la política.

Solución

Primero se debe calcular la proporción de los estudiantes cuya respuesta fue "Moderado" de ambas áreas

Para el área 1

Nulo	**Moderado**	**Grande**	**Nulo**	**Moderado**
Grande	**Moderado**	**Moderado**	**Moderado**	**Grande**
Moderado	**Moderado**	**Moderado**	**Moderado**	**Grande**
Moderado	**Moderado**	**Moderado**	**Moderado**	**Moderado**
Grande	**Nulo**	**Moderado**	**Grande**	**Moderado**

Se obtiene que, de 25 alumnos encuestados, 16 de ellos tienen un interés moderado por la política.

$$\hat{p}_1 = \frac{16}{25} = 0.64$$

Para el área 2

Nulo	**Grande**	**Muy grande**	**Nulo**	**Moderado**
Moderado	**Nulo**	**Grande**	**Moderado**	**Moderado**
Moderado	**Grande**	**Moderado**	**Moderado**	**Nulo**
Nulo	**Moderado**	**Moderado**	**Grande**	**Grande**
Moderado	**Moderado**	**Moderado**	**Grande**	**Moderado**

Se obtiene que, de 25 alumnos encuestados, 13 tienen un interés moderado por la política.

$$\hat{p}_2 = \frac{13}{25} = 0.52$$

También se debe obtener la proporción de alumnos que están moderadamente interesados en la política en ambas áreas de manera conjunta calculada de la siguiente manera

$$\hat{p} = \frac{16+13}{50} = \frac{29}{50} = 0.58$$

Si lo que se desea saber es si las proporciones son las mismas, es decir

$$p_1 = p_2$$

Se plantea la hipótesis

$$H_0 : p_1 - p_2 = 0 H_I : p_1 - p_2 \neq 0$$

Tomando en cuenta que, si la hipótesis nula H_0 es aceptada, las proporciones de los alumnos moderadamente interesados en la política en ambas áreas son iguales.

Se utiliza la distribución normal estándar, es decir, el parámetro

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

La región de rechazo es en la que el estadístico obtenido sea mayor al valor de la distribución normal estándar para el nivel de significancia dividido entre dos o cuando es menor que el valor de la distribución normal estándar para el negativo de este, es decir, cuando

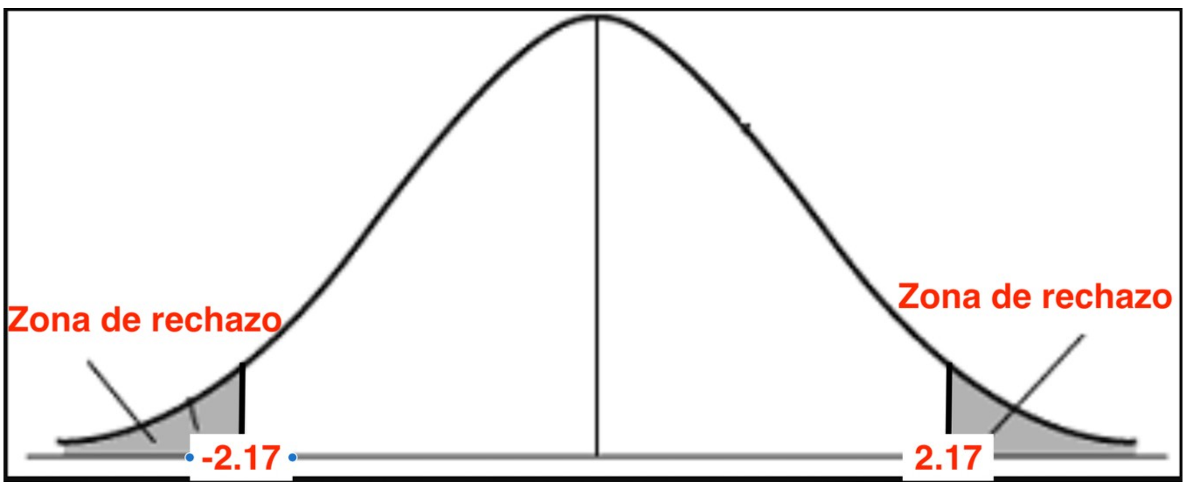
$$z_0 > z_{\frac{\alpha}{2}} o\, z_o \leftarrow z_{\frac{\alpha}{2}}$$

Sabiendo que el nivel de significancia es del 3%, se busca el valor de $z_{0.015}$ en las tablas y se encuentra que

$$z_{0.015} = \pm 2.17$$

%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$
0.0	$-\infty$	0	**5.0**	-1.645	0.063	**10.0**	-1.282	0.126	**15.0**	-1.036	0.189	**20.0**	-0.842	0.253	**25.0**	-0.674	0.319
0.1	-3.090	0.001	**5.1**	-1.635	0.064	**10.1**	-1.276	0.127	**15.1**	-1.032	0.190	**20.1**	-0.838	0.255	**25.1**	-0.671	0.320
0.2	-2.878	0.003	**5.2**	-1.626	0.065	**10.2**	-1.270	0.128	**15.2**	-1.028	0.192	**20.2**	-0.834	0.256	**25.2**	-0.668	0.321
0.3	-2.748	0.004	**5.3**	-1.616	0.066	**10.3**	-1.265	0.129	**15.3**	-1.024	0.193	**20.3**	-0.831	0.257	**25.3**	-0.665	0.323
0.4	-2.652	0.005	**5.4**	-1.607	0.068	**10.4**	-1.259	0.131	**15.4**	-1.019	0.194	**20.4**	-0.827	0.259	**25.4**	-0.662	0.324
0.5	-2.576	0.006	**5.5**	-1.598	0.069	**10.5**	-1.254	0.132	**15.5**	-1.015	0.196	**20.5**	-0.824	0.260	**25.5**	-0.659	0.325
0.6	-2.512	0.008	**5.6**	-1.589	0.070	**10.6**	-1.248	0.133	**15.6**	-1.011	0.197	**20.6**	-0.820	0.261	**25.6**	-0.656	0.327
0.7	-2.457	0.009	**5.7**	-1.580	0.071	**10.7**	-1.243	0.135	**15.7**	-1.007	0.198	**20.7**	-0.817	0.262	**25.7**	-0.653	0.328
0.8	-2.409	0.010	**5.8**	-1.572	0.073	**10.8**	-1.237	0.136	**15.8**	-1.003	0.199	**20.8**	-0.813	0.264	**25.8**	-0.650	0.329
0.9	-2.366	0.011	**5.9**	-1.563	0.074	**10.9**	-1.232	0.137	**15.9**	-0.999	0.201	**20.9**	-0.810	0.265	**25.9**	-0.646	0.331
1.0	-2.326	0.013	**6.0**	-1.555	0.075	**11.0**	-1.227	0.138	**16.0**	-0.994	0.202	**21.0**	-0.806	0.266	**26.0**	-0.643	0.332
1.1	-2.290	0.014	**6.1**	-1.546	0.077	**11.1**	-1.221	0.140	**16.1**	-0.990	0.203	**21.1**	-0.803	0.268	**26.1**	-0.640	0.333
1.2	-2.257	0.015	**6.2**	-1.538	0.078	**11.2**	-1.216	0.141	**16.2**	-0.986	0.204	**21.2**	-0.800	0.269	**26.2**	-0.637	0.335
[illegible]	[illegible]	0.016	**6.3**	-1.530	0.079	**11.3**	-1.211	0.142	**16.3**	-0.982	0.206	**21.3**	-0.796	0.270	**26.3**	-0.634	0.336
1.4	-2.197	0.018	**6.4**	-1.522	0.080	**11.4**	-1.206	0.143	**16.4**	-0.978	0.207	**21.4**	-0.793	0.272	**26.4**	-0.631	0.337
1.5	-2.170	0.019	**6.5**	-1.514	0.082	**11.5**	-1.200	0.145	**16.5**	-0.974	0.208	**21.5**	-0.789	0.273	**26.5**	-0.628	0.338
1.6	-2.144	0.020	**6.6**	-1.506	0.083	**11.6**	-1.195	0.146	**16.6**	-0.970	0.210	**21.6**	-0.786	0.274	**26.6**	-0.625	0.340
[illegible]	[illegible]	0.021	**6.7**	-1.499	0.084	**11.7**	-1.190	0.147	**16.7**	-0.966	0.211	**21.7**	-0.782	0.275	**26.7**	-0.622	0.341
[illegible]	[illegible]	0.023	**6.8**	-1.491	0.085	**11.8**	-1.185	0.148	**16.8**	-0.962	0.212	**21.8**	-0.779	0.277	**26.8**	-0.619	0.342
1.9	-2.075	0.024	**6.9**	-1.483	0.087	**11.9**	-1.180	0.150	**16.9**	-0.958	0.213	**21.9**	-0.776	0.278	**26.9**	-0.616	0.344

Se grafica entonces la región de rechazo



Sustituyendo los valores con los obtenidos de la muestra

$$z_0 = \frac{0.64 - 0.52}{\sqrt{0.58(0.42)\left(\frac{1}{25} + \frac{1}{25}\right)}} = 0.8596$$

Ya que

$$-2.17 < 0.8596 < 2.17$$

Se dice que el estadístico encontrado no está en el área de rechazo, por lo que se acepta la hipótesis nula H_0

$$H_0: p_1 - p_2 = 0$$
$$H_I: p_1 - p_2 \neq 0$$

Se concluye que con base en los datos de la muestra y un nivel de significancia del 3%, existe evidencia suficiente para afirmar que la proporción de los alumnos moderadamente interesados es igual en área 1 y área 2.

Solución con minitab

El primer paso para la solución en minitab es registrar los datos que se obtuvieron en cada muestra en una sola columna como se muestra en la siguiente imagen, en donde el número 1 representa que el encuestado contestó que su interés es moderado y el número 0 representa que no fue así.

C2	C3
Area 1	Area 2
0	0
0	0
1	1
0	1
1	0
1	0
1	1
0	1
1	1
1	0
1	1
1	1
0	0
1	0
1	1
1	1
1	0
1	0
0	1

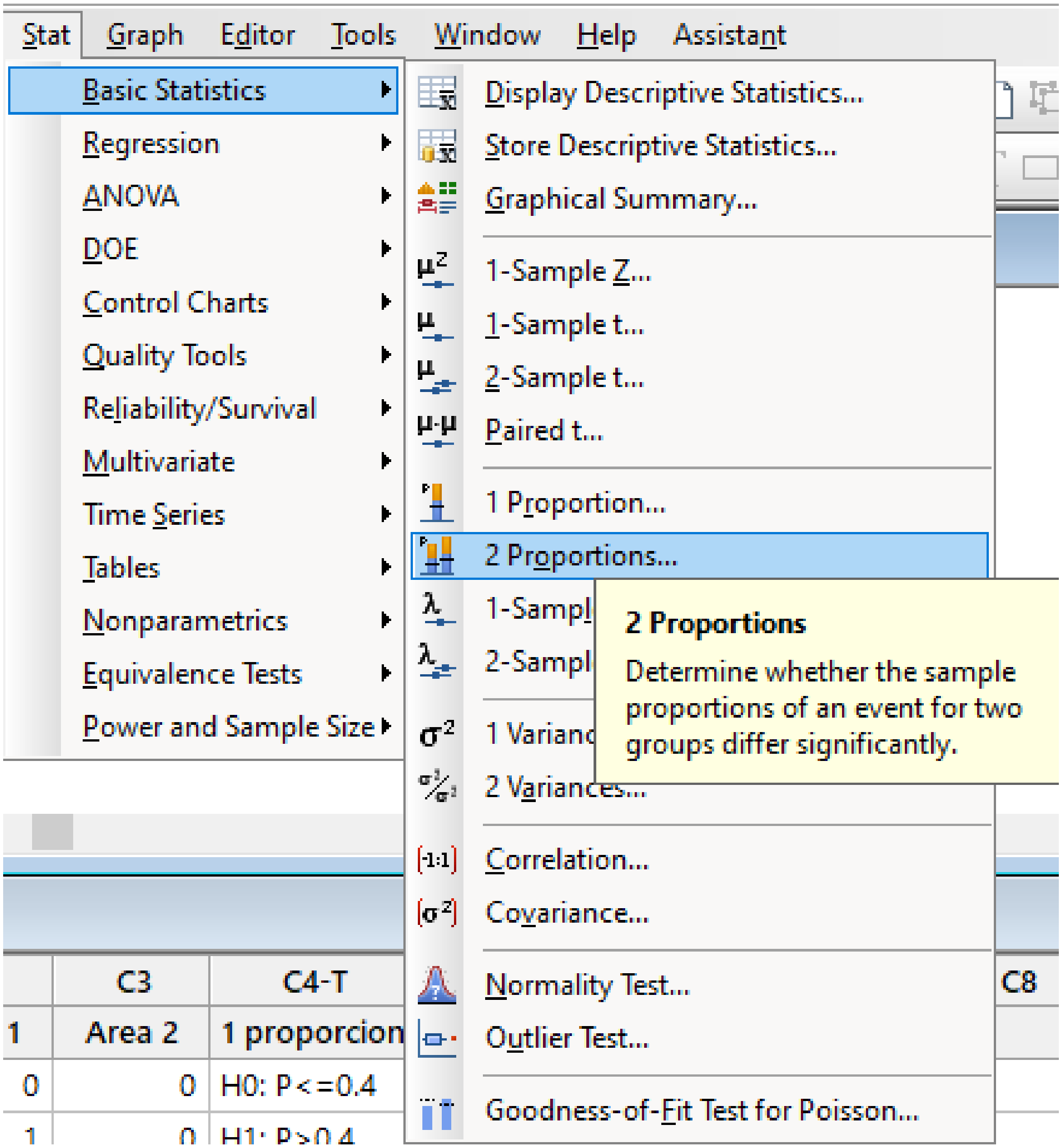
El segundo es establecer la hipótesis nula y la alterna, es recomendable que se escriban en el mismo documento para tener fácil acceso a ellas.

C5-T
2 proporcion
H0: P1-P2=0
H1: P1-P2=/0

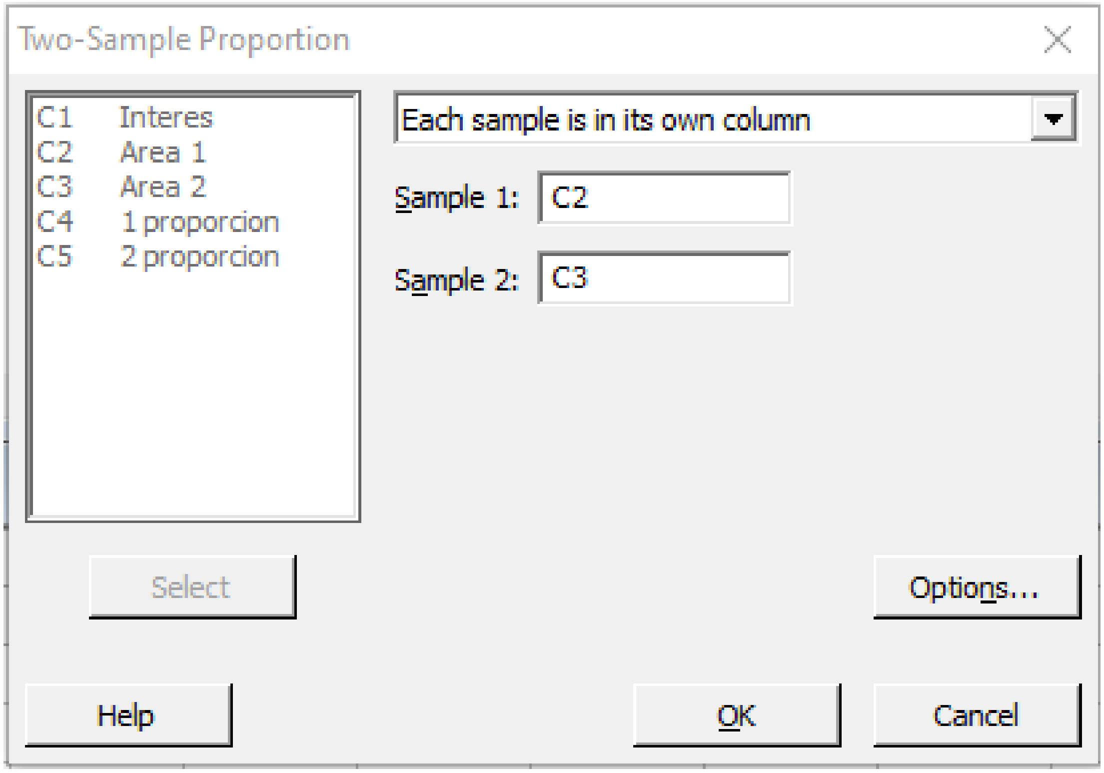
Después en el menú superior se seleccionan las siguientes

$Stat \rightarrow Basic\ Statistics \rightarrow 2\ proportions$

Como se observa en la siguiente imagen

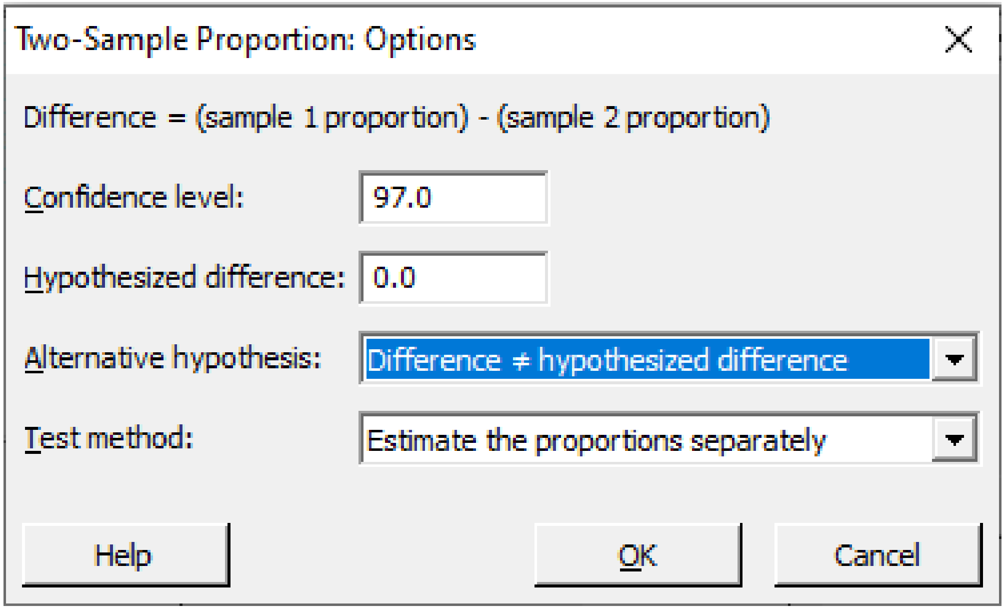


Lo que hará que aparezca la ventana que se muestra a continuación



En el primer campo se debe seleccionar la opción de *Each simple is in its own column*. Ya que los datos se encuentran en una columna por muestra.
En el segundo campo se selecciona la columna en la que se anotaron los datos de la muestra número 1, en este caso la columna es C2.
En el tercer campo se selecciona la columna en la que se anotaron los datos de la muestra número 2, en este caso la columna es C3.

Luego, en esa misma ventana se hace clic en el botón de *Options* lo que abrirá una ventana como la siguiente



En esta ventana, el primer campo se refiere al nivel de confianza, que para nuestro caso es 97, resultado del 100% menos el 3% del nivel de significancia.

El segundo campo sirve para escribir el resultado de la diferencia de proporciones en nuestra hipótesis que para este caso es igual a 0.
En el tercer campo se selecciona la opción que corresponde a la hipótesis alterna, en este caso, es que la diferencia de proporciones sea igual a 0, que es la diferencia de la hipótesis.
El último campo simplemente se deja como *Estimate the proportions separately*

Por último, se da clic en
OK → *OK*
Se obtiene entonces el siguiente resultado

```
Session

Area 1    16   25   0.640000
Area 2    13   25   0.520000


Difference = p (Area 1) - p (Area 2)
Estimate for difference:  0.12
97% CI for difference:  (-0.180697, 0.420697)
Test for difference = 0 (vs ≠ 0):  Z = 0.87  P-Value = 0.386

Fisher's exact test: P-Value = 0.567
```

Lo que es de nuestro interés en este ejercicio es lo que el software es el valor denominado *P-Value* y ya que este valor es mayor al nivel de significancia, se acepta la hipótesis nula.

Se concluye que con base en los datos de la muestra y un nivel de significancia del 3%, existe evidencia suficiente para afirmar que la proporción de los alumnos moderadamente interesados es igual en área 1 y área 2.

Conclusiones

Debido al estudio práctico de los temas vistos en clase nos fue posible comprender, analizar, justificar e interpretar el comportamiento de diferentes poblaciones, esto con el propósito de mejorar algunos procesos, y así poder optimizarlos por otro lado logramos comprender la estadística como una herramienta en la ingeniería y cómo podemos utilizarla para llegar a predecir comportamientos de una o más poblaciones a través de pequeñas muestras.

Finalmente es importante mencionar que el manejo y compresión de los datos es importante, ya que muchas herramientas como el Deep Learning o Machine Learning tienen como base la estadística para poder entender el comportamiento de diferentes poblaciones como animales, humanos, virus, procesos, y con ello buscar estrategias para su mejora, la estadística es una herramienta que nos ayudará a tomar decisiones bien fundamentadas gracias a la interpretación de los datos.

Bibliografía

- Devore, Jay L. *Probabilidad y estadística para ingeniería y ciencias.* Cengage Learning, México, 2011.
- Trula, E. (5 de marzo de 2019)*150.000 millones de prendas de ropa al año (y otras cifras en las que las tiendas no quieren que pienses).* Magnet. URL: https://magnet.xataka.com/preguntas-no-tan-frecuentes/150-000-millones-prendas-ropa-al-ano-otras-cifras-que-tiendas-no-quieren-que-pienses
- INEGI. (2018). *Encuesta sobre la Percepción Pública de la Ciencia y la Tecnología en México*.URL: https://www.inegi.org.mx/contenido/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825006793.pdfç