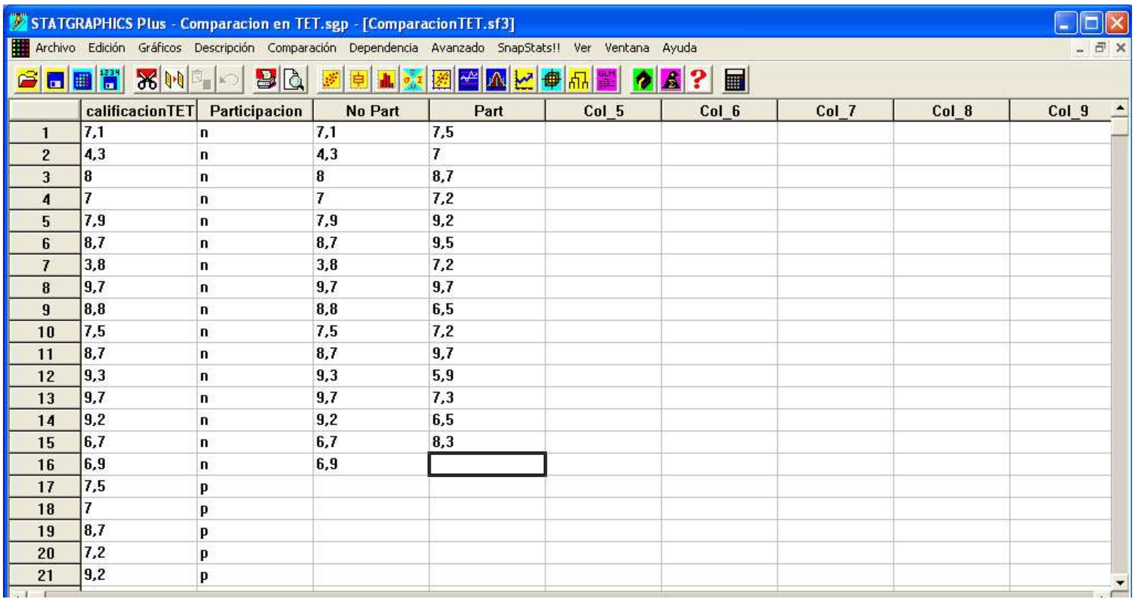
Estudio estadístico de la asignatura

TÉCNICAS ESTADÍSTICA PARA TURISMO

La asignatura se estructura en dos grupos, en los que el alumnado matriculado se distribuye de una forma aleatoria. Dentro de esta asignatura, se selecciona el Grupo 2 para la aplicación del contenido del Proyecto. Este Grupo está formado por 31 alumnos, de los cuales 15 participan en el Proyecto, quedando 16 alumnos como no participantes. Esto nos permite realizar un estudio comparativo dentro del Grupo 2, en relación a la participación del alumnado en el Proyecto.

La variable que se va a someter al estudio estadístico es la *calificación* obtenida por los alumnos en la asignatura de Técnicas Estadísticas para el Turismo (*TET*), en la convocatoria oficial de junio de 2012, por ser ésta la que recibe la aplicación de los resultados obtenidos por los alumnos participantes en el Proyecto.

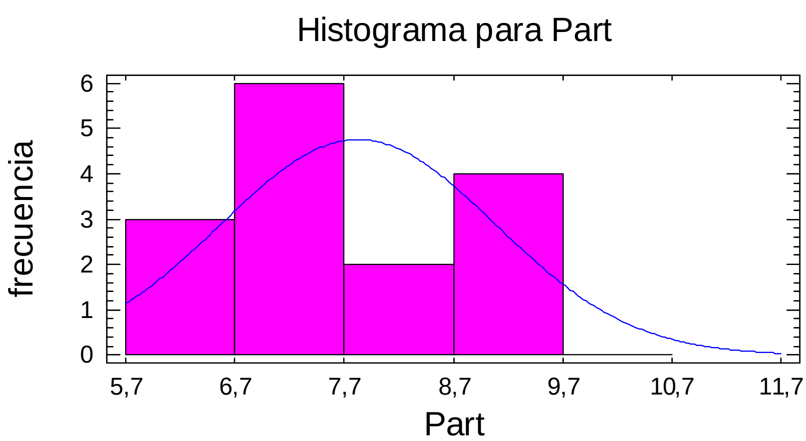
Los valores de la variable *calificaciónTET* se introducen en un fichero de datos de la aplicación estadística Statgraphics.


STATGRAPHICS Plus - Comparacion en TET.sgp - [ComparacionTET.sf3]

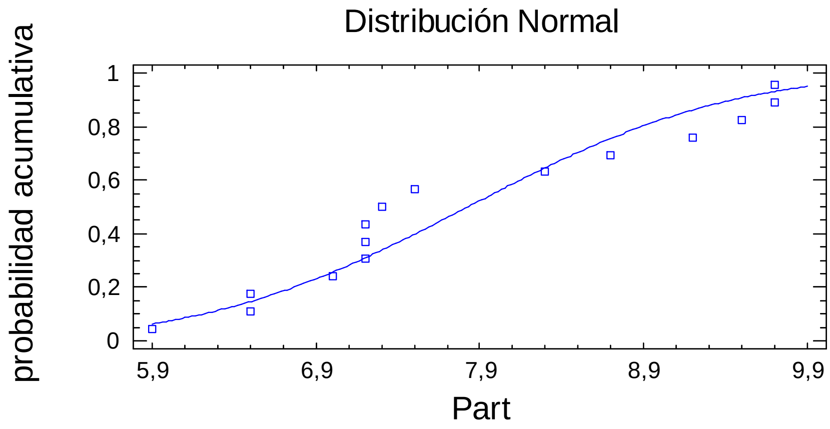
Archivo Edición Gráficos Descripción Comparación Dependencia Avanzado SnapStats!! Ver Ventana Ayuda

	calificacionTET	Participacion	No Part	Part	Col_5	Col_6	Col_7	Col_8	Col_9
1	7,1	n	7,1	7,5					
2	4,3	n	4,3	7					
3	8	n	8	8,7					
4	7	n	7	7,2					
5	7,9	n	7,9	9,2					
6	8,7	n	8,7	9,5					
7	3,8	n	3,8	7,2					
8	9,7	n	9,7	9,7					
9	8,8	n	8,8	6,5					
10	7,5	n	7,5	7,2					
11	8,7	n	8,7	9,7					
12	9,3	n	9,3	5,9					
13	9,7	n	9,7	7,3					
14	9,2	n	9,2	6,5					
15	6,7	n	6,7	8,3					
16	6,9	n	6,9						
17	7,5	p							
18	7	p							
19	8,7	p							
20	7,2	p							
21	9,2	p							

Los datos se han clasificados según la participación o no de los alumnos en el Proyecto. Observemos el comportamiento de los datos según esta clasificación. Comenzando por los alumnos participantes (*Part*). Mediante Statgraphics, obtenemos el siguiente histograma como representación gráfica de los datos:



Y la gráfica de normalidad:

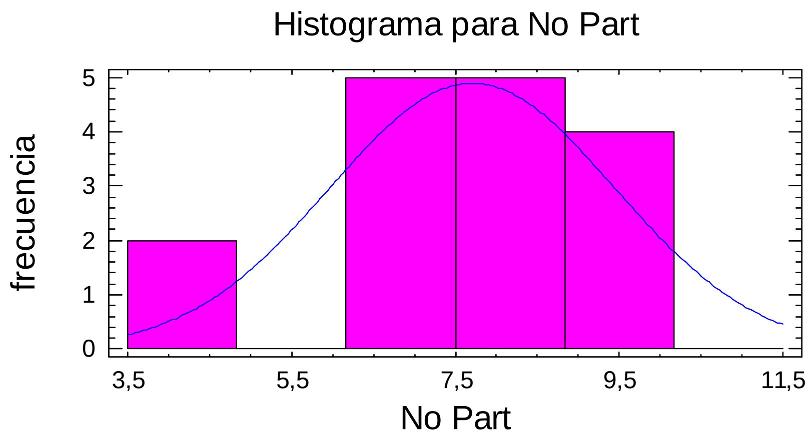


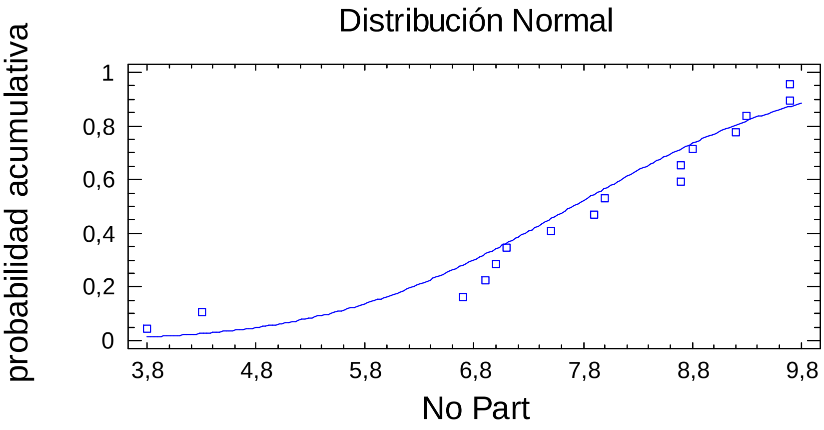
Esta distribución apunta a que puede ajustarse una normal. Para determinar si asumimos tal distribución de los datos, sometemos la muestra de tamaño 15 a contrastes de bondad de ajuste. Al disponer de un tamaño inferior a 50, aplicamos el test de Shapiro-Wilks. Los resultados son:

Estadístico W de Shapiro-Wilks = 0,911338
P-valor = 0,142533

El P-valor es inferior a 0.05, que podemos tomar como nivel de significación. Por tanto, no tenemos evidencias para rechazar la normalidad de la distribución de la que proceden los datos para la variable *Part*, y en consecuencia asumimos que la variable se distribuye normalmente.

Repetimos este análisis para la variable *No Part*, a partir de la muestra de tamaño 16 de la que disponemos. Las representaciones gráficas correspondientes al histograma y al gráfico de normalidad son los siguientes:





Al igual que la variable *Part* , la distribución de *No Part* apunta a que puede ajustarse una normal. En este caso disponemos de una muestra de tamaño 15, inferior a 50, por lo que nuevamente aplicamos el test de Shapiro-Wilks. Los resultados son:

Estadístico W de Shapiro-Wilks = 0,888791
P-valor = 0,0538427

El P-valor es inferior a 0.05, que podemos tomar como nivel de significación. Por tanto, no tenemos evidencias para rechazar la normalidad de la distribución de la que proceden los datos para la variable *No Part*, y en consecuencia asumimos que la variable se distribuye normalmente.

Pasamos a comparar las distribuciones de las variables *Part* y *No Part*. Las técnicas, en principio, pueden ser paramétricas, dado que hemos asumido que ambas variables se distribuyen normalmente.

Comencemos con la comparación de las varianzas de las dos distribuciones. Aplicamos el test de hipótesis F para comparar las dos varianzas mediante Statgraphics:

	Part	No Part
Desviación Típica	1,74068	1,25838
Varianza	3,02996	1,58352
GL	15	14

Cociente de varianzas = 1,91343

95,0% Intervalos de Confianza
Desviación Típica No Part: [1,28585;2,69403]
Desviación Típica de Part: [0,921294;1,98459]
Cociente de varianzas: [0,648769;5,53264]

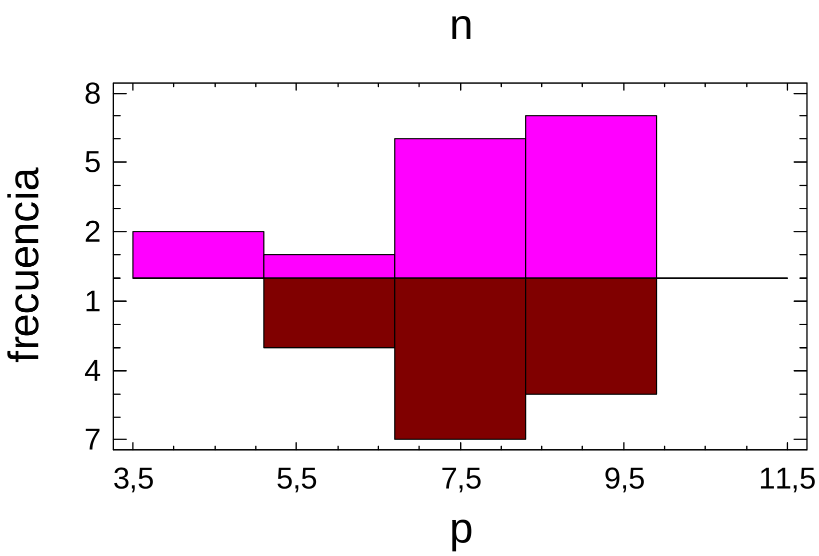
Contrastes F para comparar varianzas

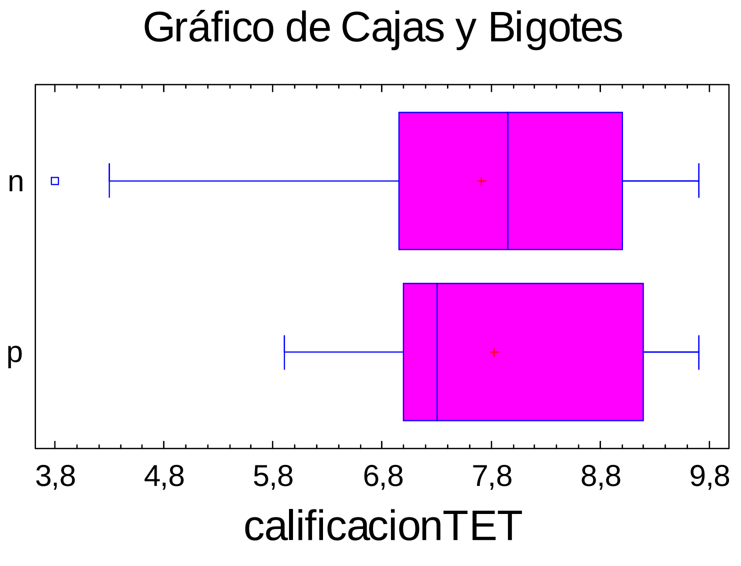
Hipótesis nula: sigma1 = sigma2
(1) Hipótesis alt.: sigma1 <> sigma2
F = 1,91343   P-Valor = 0,232872

Puesto que el p-valor calculado no es inferior a 0,05, no podemos rechazar la hipótesis nula, y en consecuencia asumimos la igualdad de varianzas.

Pasamos a la comparación de las medias, partiendo de igualdad de varianzas, que damos como válido a raíz del contraste anterior.

Los siguientes gráficos nos dan una idea del comportamiento comparativo de ambas poblaciones:





95,0% intervalo de confianza para la media de No Part [6,77871,8,63379]
95,0% intervalo de confianza para la media de Part [7,1298,8,52354]
95,0% intervalos de confianza para la diferencia de medias:
suponiendo varianzas iguales: [-1,24283,1,002]

El contrastes t de comparación de medias:

Hipótesis nula: media1 = media2
Hipótesis alt.: media1 <> media2
suponiendo varianzas iguales: t = -0,21942 P-Valor = 0,82786

Por tanto, no tenemos evidencias para rechazar la igualdad de las medias a un nivel de significación de 0.05. De hecho, el p-valor=0,82786 es alto. **En consecuencia, no podemos establecer diferencias significativas entre las distribuciones de *Part* y *No Part*.**

CONCLUSIONES

1. No podemos establecer diferencias significativas entre las distribuciones de *Part* y *No Part*.
2. No obstante, a partir de los diagramas de cajas y bigotes, observamos que todos los alumnos participante tienen calificaciones por encima de 5, cosa que no ocurre con los alumnos no participantes.
3. observamos más homogeneidad en las calificaciones de los alumnos participantes que en los no participantes.

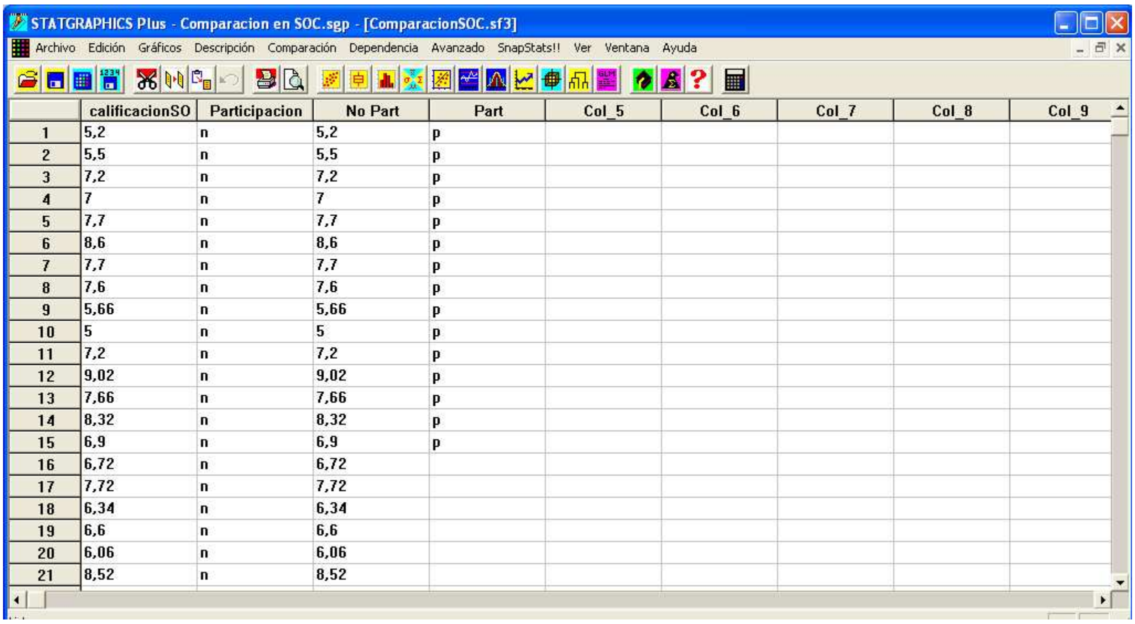
Estudio estadístico de la asignatura

SOCIOLOGÍA TURÍSTICA Y MEDIOAMBIENTAL

La asignatura se estructura en dos grupos, en los que el alumnado matriculado se distribuye de una forma aleatoria. Dentro de esta asignatura, se selecciona el Grupo 2 para la aplicación del contenido del Proyecto. Este Grupo está formado por 43 alumnos, de los cuales 15 participan en el Proyecto, quedando 28 alumnos como no participantes, de los cuales uno no asiste a clase ni se presenta a la convocatoria de examen, por lo que se ha optado por excluirlo de este estudio al carecer de calificación. Esta informacón nos permite realizar un estudio comparativo dentro del Grupo 2, en relación a la participación del alumnado en el Proyecto.

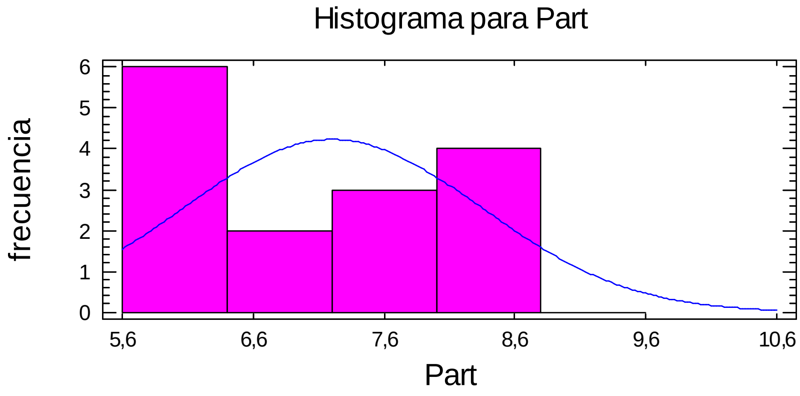
La variable que se va a someter al estudio estadístico es la *calificación* obtenida por los alumnos en la asignatura de Sociología Turística y Medioambiental (*SOC*), en la convocatoria oficial de junio de 2012, por ser ésta la que recibe la aplicación de los resultados obtenidos por los alumnos participantes en el Proyecto.

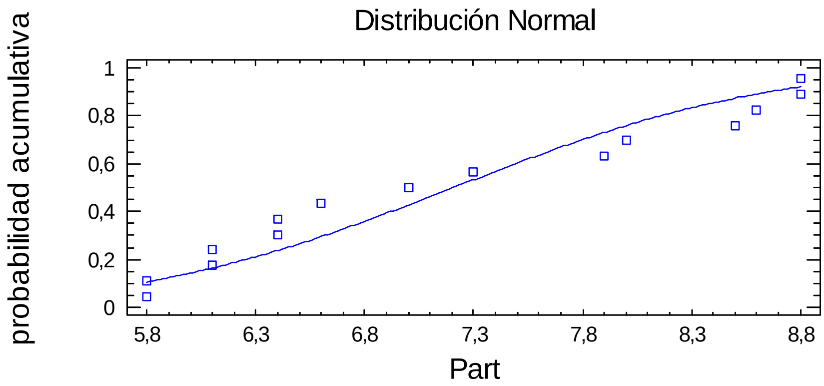
Los valores de la variable *calificaciónSOC* se introducen en un fichero de datos de la aplicación estadística Statgraphics.



	calificacionSO	Participacion	No Part	Part	Col_5	Col_6	Col_7	Col_8	Col_9
1	5,2	n	5,2	p					
2	5,5	n	5,5	p					
3	7,2	n	7,2	p					
4	7	n	7	p					
5	7,7	n	7,7	p					
6	8,6	n	8,6	p					
7	7,7	n	7,7	p					
8	7,6	n	7,6	p					
9	5,66	n	5,66	p					
10	5	n	5	p					
11	7,2	n	7,2	p					
12	9,02	n	9,02	p					
13	7,66	n	7,66	p					
14	8,32	n	8,32	p					
15	6,9	n	6,9	p					
16	6,72	n	6,72						
17	7,72	n	7,72						
18	6,34	n	6,34						
19	6,6	n	6,6						
20	6,06	n	6,06						
21	8,52	n	8,52						

Los datos se han clasificados según la participación o no de los alumnos en el Proyecto. Observemos el comportamiento de los datos según esta clasificación. Comenzando por los alumnos participantes (*Part*). Mediante Statgraphics, obtenemos el siguiente histograma como representación gráfica de los datos:



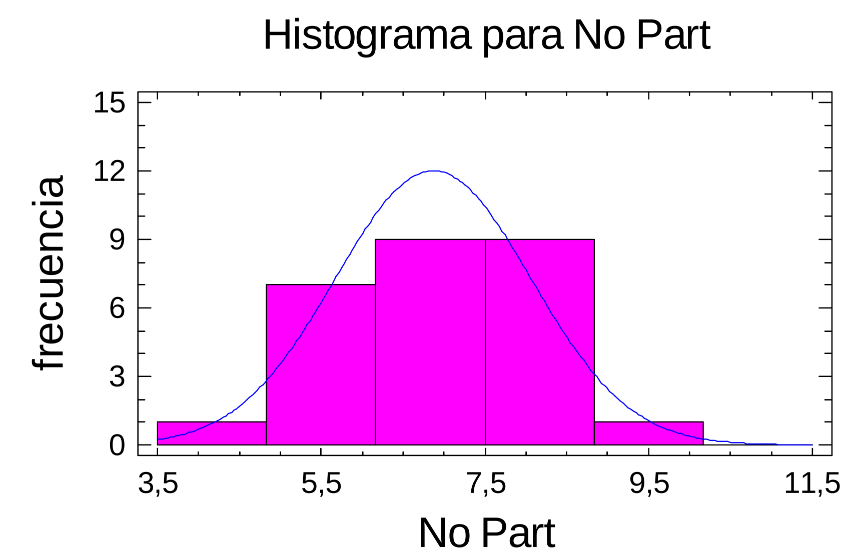


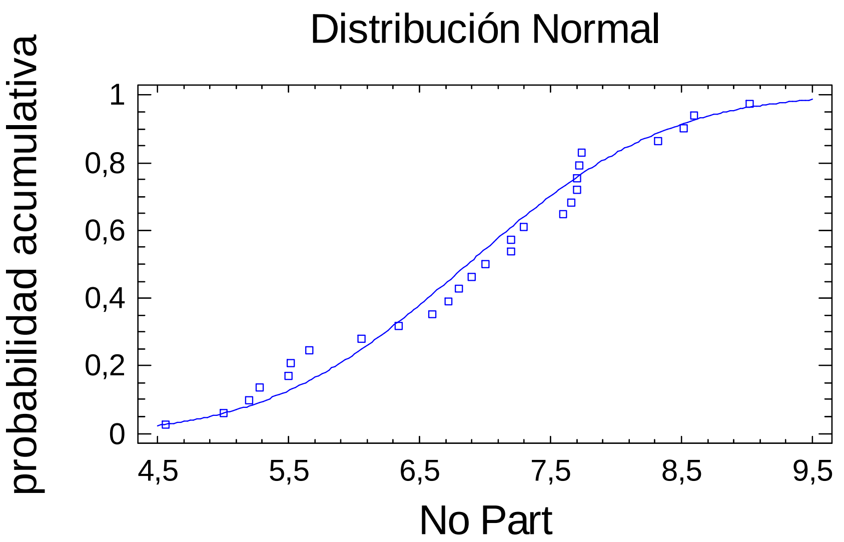
Esta distribución apunta a que puede ajustarse una normal. Para determinar si asumimos tal distribución de los datos, sometemos la muestra de tamaño 15 a contrastes de bondad de ajuste. Al disponer de un tamaño inferior a 50, aplicamos el test de Shapiro-Wilks. Los resultados son:

Estadístico W de Shapiro-Wilks = 0,891383
P-valor = 0,0710663

En este caso tenemos un P-valor inferior a 0.05, que mantenemos como nivel de significación. Con este criterio no tenemos evidencias para rechazar la normalidad de la distribución de la que proceden los datos para la variable *Part*, y en consecuencia asumimos que la variable se distribuye normalmente.

Repetimos este análisis para la variable *No Part*, a partir de la muestra de tamaño 28 de la que disponemos. Las representaciones gráficas correspondientes al histograma y al gráfico de normalidad son los siguientes:





Al igual que la variable *Part* , la distribución de *No Part* apunta a que puede ajustarse a una normal. En este caso disponemos de una muestra de tamaño 27, inferior a 50, por lo que nuevamente aplicamos el test de Shapiro-Wilks. Los resultados son:

Estadístico W de Shapiro-Wilks = 0,964671
P-valor = 0,494729

El P-valor es inferior a 0.05, que mantenemos como nivel de significación. Por tanto, no tenemos evidencias para rechazar la normalidad de la distribución de la que proceden los datos para la variable *No Part*, y en consecuencia asumimos que la variable se distribuye normalmente.

Pasamos a comparar las distribuciones de las variables *Part* y *No Part*. Las técnicas, en principio, pueden ser paramétricas, dado que hemos asumido, mediante la aplicación de test de bondad de ajuste para normalidad, que ambas variables se distribuyen normalmente.

Comencemos con la comparación de las varianzas de las dos distribuciones. Aplicamos el test de hipótesis F para comparar las dos varianzas mediante Statgraphics:

	Part	No Part
Desviación Típica	1,19656	1,13167
Varianza	1,43176	1,28067
GL	26	14

Cociente de varianzas = 1,11798

95,0% Intervalos de Confianza
Desviación Típica No Part: [0,942311;1,63981]
Desviación Típica Part: [0,828522;1,78475]
Cociente de varianzas: [0,403997;2,70224]

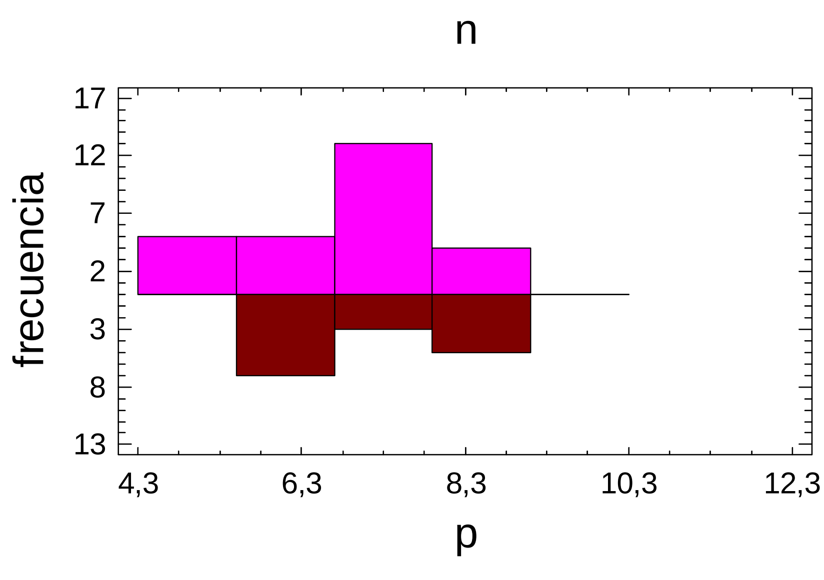
Contrastes F para comparar varianzas

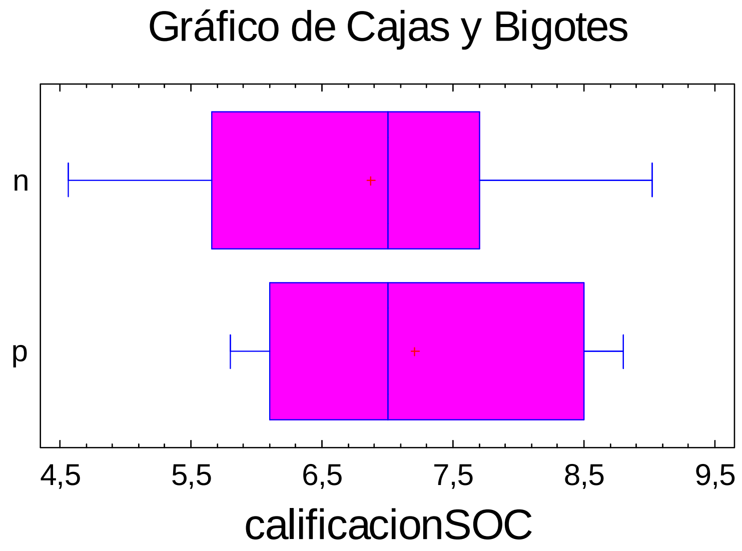
Hipótesis nula: sigma1 = sigma2
(1) Hipótesis alt.: sigma1 <> sigma2
F = 1,11798   P-Valor = 0,852263

Puesto que el p-valor calculado no es inferior a 0,05, no podemos rechazar la hipótesis nula, y en consecuencia asumimos la igualdad de varianzas.

Pasamos a la comparación de las medias, partiendo de igualdad de varianzas, que damos como válido a raíz del contraste anterior.

Los siguientes gráficos nos dan una idea del comportamiento comparativo de ambas poblaciones:





95,0% intervalo de confianza para la media de No Part: [6,39406,7,34075]
95,0% intervalo de confianza para la media de Part: [6,57997,7,83336]
95,0% intervalos de confianza para la diferencia de medias:
suponiendo varianzas iguales: [-1,10352,0,425004]

El contrastes t de comparación de medias:

Hipótesis nula: media1 = media2
Hipótesis alt.: media1 <> media2
suponiendo varianzas iguales: t = -0,897164   P-Valor = 0,375

Por tanto, no tenemos evidencias para rechazar la igualdad de las medias a un nivel de significación de 0.05. De hecho, el p-valor=0,375 es alto. **En consecuencia, no podemos establecer diferencias significativas entre las distribuciones de *Part* y *No Part*.**

CONCLUSIONES

1. No podemos establecer diferencias significativas entre las distribuciones de *Part* y *No Part*.
2. No obstante, a partir de los diagramas de cajas y bigotes, observamos que todos los alumnos participantes tienen calificaciones por encima de 5, cosa que no ocurre con los alumnos no participantes.
3. observamos más homogeneidad en las calificaciones de los alumnos participantes que en los no participantes.