

Aplicación experimental

Frecuencias de Saeta

Autores:

Aramayo Jesús Gino.

Burgos Zambrano Santiago.

Curiacus Maximiliano.

García Tomás.

Resumen:

A modo de experimentación para verificar los conocimientos aprendidos en la Cátedra decidimos realizar una experiencia para cumplir este objetivo, la misma se basa en poder tomar mediciones de los tiempos de arribo y salidas entre colectivos Saeta en una parada específica y de las líneas previamente seleccionadas (6B y 8Trasnversal), con los datos obtenidos de la experimentación, los objetivos por orden de prioridad son poder verificar el tipo de distribución que poseen los mismos ya que de forma inductiva se sospecha que los mismos poseen una distribución normal es decir la frecuencia es una sola (media) pero existen valores oscilantes y atípicos alrededor de la misma, una vez que realicemos esto, se pretende realizar comparaciones entre colectivos de la misma línea y posteriormente entre buses de ambas líneas, para poder determinar , primeramente la frecuencia media de cada bus en particular, y luego definir qué línea tiene la mayor o menor frecuencia. A modo de obtener una “sospecha más general” vamos a realizar una encuesta a personas que tomen dichos colectivos, preguntándoles “cual línea piensa que pasa mas rápido”. Una vez realizado esto procederemos a concluir si nuestras sospechas eran ciertas o no.

Introducción:

El medio de transporte más utilizado por las mayorías de las personas en la provincia de Salta es el colectivo, ya sea para realizar viajes de corta o larga distancias (Urbanos e interurbanos).

Una de las preguntas más frecuentes que se hacen las personas que utilizan el colectivo es la frecuencia con la cual pasan los mismos, es decir el intervalo de tiempo que hay entre 2 colectivos de una misma línea al pasar por una misma parada. Esto se debe a que las personas presentan molestias ya que muchas veces los colectivos tardan mas de lo común en llegar a una parada o también porque afirman que colectivos de otra línea pasan con mas frecuencia.

Este trabajo surgió de la inquietud de determinar si los colectivos realmente pasan por las paradas según como especifica la empresa prestadora de servicio de transporte **SAETA** (Sociedad Anónima del Estado de Transporte Automotor), y además de poder responder las preguntas antes mencionadas.

Metodología:

Para realizar este experimento y poner en práctica los conocimientos adquiridos por la materia de estadística experimental, tuvimos que recolectar datos realizando mediciones

cronometrando el tiempo que transcurre entre la llegada de un colectivo y el próximo de la misma línea a la misma parada.

Estas mediciones la realizamos en un horario donde se supone que el tránsito es normal y no hay problemas debido al embotellamiento por horarios picos o cambios de recorridos producidos por manifestaciones, ya que estos factores afectan las frecuencias de llegada de los mismos, produciéndonos alteraciones circunstanciales en las mediciones, el horario elegido fue de 15 a 17 horas.

Para elegir las líneas de colectivos en las cuales vamos a analizar nos basamos en los que pasan por las cercanías del B° Parque Belgrano los cuales son el 6B y 8 Transversal.

Realizamos una encuesta a las personas que se encontraban en la parada para saber su opinión acerca de que línea poseía una mayor frecuencia. La pregunta realizada fue: ¿Cual colectivo pasa más seguido?, teniendo como opciones de respuesta:

- A. El 8 transversal.
- B. El 6B.
- C. Igual.

Las mediciones que realizamos fueron tomando el tiempo que tardan 2 colectivos consecutivos de las líneas 6B y 8 Transversal que llegan y salen del barrio en dichas paradas.

De la página de Saeta adquirimos el tiempo teórico que deberían tardar las líneas en pasar por la parada, basándonos en este dato para realizar la comprobación con las mediciones.

Desarrollo :

Una vez recolectado todos los datos, procedemos a realizar los cálculos, probando si los datos provienen de una distribución normal, prueba de homogeneidad y prueba de hipótesis.

- Datos:

Los datos de la siguiente tabla se encuentran en minutos.

6B Norte-Centro	
8,47	3,88
3,52	7,47
6,05	2,92
5,13	2,4
3,18	4,83
6,23	6,2
5,03	4,5
2,15	2,57
9,2	6,12
1,37	4,68
Promedio	4,795
DesvioM	2,12377643
DesvioP	2,07000121

6B Centro-Norte	
0,888333	5,23523
1,0695667	5,4054
2,336333	5,7042
2,9523	5,784233
3,9527	5,8254667
4,174	6,3025
4,5558	6,5381333
4,63383	6,7501
4,83963333	8,08463
4,9572	8,13833
varianzap	3,70531977
desviacion	1,97492699
promedio	4,90639595

8T Norte-Sur

10,4	5,77
10,083	5,63
8,62	5,55
8,28	3,97
7,67	3,23
6,4	2,87
6,08	2,8
6	2,07
5,93	1,82
5,93	0,87

8 T Sur-Norte	
5,35	3,68
5,78	3,63
6,63	3,7
8,68	11,67
2,08	2,37
5,52	3,15
6,18	5,25
11,3	8,42
1,07	1,72
9,83	7,27

Promedio	5,664
DesvioP	3,03786998
DesvioM	3,11678883

Promedio	5,664
DesvioP	3,03786998
DesvioM	3,11678883

- Cálculos realizados:

Bondad de ajuste: normalidad

H°) Los datos provienen de una distribución normal.

Ha) Los datos no provienen de una distribución normal.

6B Norte-Centro

- Shapiro y Wilk (nivel de significancia = 0,05)

I	Aj	Y	yj	Aj*(yn-j - yj)
1	0,4734	9,2	1,37	3,706722
2	0,3211	8,47	2,15	2,029352
3	0,2565	7,47	2,4	1,300455
4	0,2085	6,23	2,57	0,76311
5	0,1686	6,2	2,92	0,553008
6	0,1344	6,12	3,18	0,395136
7	0,1013	6,05	3,52	0,256289
8	0,0711	5,13	3,88	0,088875
9	0,0422	5,03	4,5	0,022366
10	0,014	4,83	4,68	0,0021

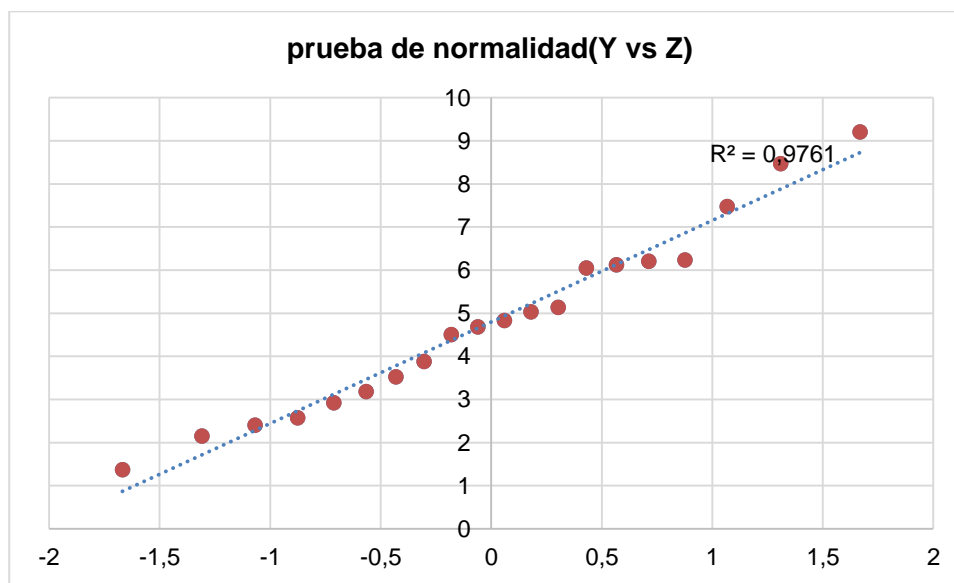
Wcalc = 0,97000073

Wcrit = 0,905

- Prueba de normalidad

	I/21	Z	
1	0,04761905	-1,66839119	1,37
2	0,0952381	-1,30917172	2,15
3	0,14285714	-1,06757052	2,4
4	0,19047619	-0,87614285	2,57
5	0,23809524	-0,71244303	2,92

6	0,28571429	-0,56594882	3,18
7	0,33333333	-0,4307273	3,52
8	0,38095238	-0,30298045	3,88
9	0,42857143	-0,18001237	4,5
10	0,47619048	-0,0597171	4,68
11	0,52380952	0,0597171	4,83
12	0,57142857	0,18001237	5,03
13	0,61904762	0,30298045	5,13
14	0,66666667	0,4307273	6,05
15	0,71428571	0,56594882	6,12
16	0,76190476	0,71244303	6,2
17	0,80952381	0,87614285	6,23
18	0,85714286	1,06757052	7,47
19	0,9047619	1,30917172	8,47
20	0,95238095	1,66839119	9,2



6B Centro-Norte

- Shapiro y Wilk(nivel de significancia = 0,05)

i	Aj20	Y(n-j+1)	Yj	Aj20*(Y(n-j+1)-Yj)
1	0,4734	8,13833	0,888333	3,43214858
2	0,3211	8,08463	1,0695667	2,25253683
3	0,2565	6,7501	2,336333	1,13213124
4	0,2085	6,5381333	2,9523	0,74764624
5	0,1686	6,3025	3,9527	0,39617628
6	0,1344	5,8254667	4,174	0,22195712
7	0,1013	5,784233	4,5558	0,12444026
8	0,0711	5,7042	4,63383	0,07610331
9	0,0422	5,4054	4,83963333	0,02387535

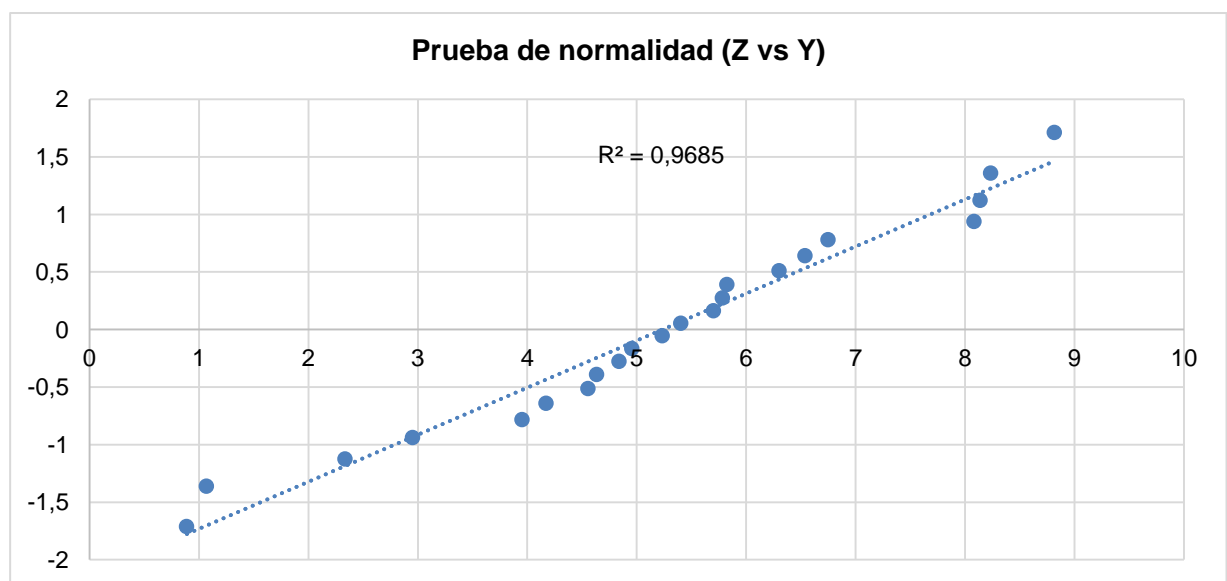
10	0,014	5,23523	4,9572	0,00389242
----	-------	---------	--------	------------

$W_{calc} = 0,95461892$

$W_{crit} = 0,905$

• Prueba de normalidad

i	Yi	i/21	Zi
1	0,888333	0,04761905	-1,66839119
2	1,0695667	0,0952381	-1,30917172
3	2,336333	0,14285714	-1,06757052
4	2,9523	0,19047619	-0,87614285
5	3,9527	0,23809524	-0,71244303
6	4,174	0,28571429	-0,56594882
7	4,5558	0,33333333	-0,4307273
8	4,63383	0,38095238	-0,30298045
9	4,83963333	0,42857143	-0,18001237
10	4,9572	0,47619048	-0,0597171
11	5,23523	0,52380952	0,0597171
12	5,4054	0,57142857	0,18001237
13	5,7042	0,61904762	0,30298045
14	5,784233	0,66666667	0,4307273
15	5,8254667	0,71428571	0,56594882
16	6,3025	0,76190476	0,71244303
17	6,5381333	0,80952381	0,87614285
18	6,7501	0,85714286	1,06757052
19	8,08463	0,9047619	1,30917172
20	8,13833	0,95238095	1,66839119



8 T Sur-Norte

- Shapiro y Wilk(nivel de significancia = 0,05)

i	Aj20	Y(n-j+1)	Yj	Aj20*(Y(n-j+1)-Yj)
1	0,4734	11,67	1,07	5,01804
2	0,3211	11,3	1,72	3,076138
3	0,2565	9,83	2,08	1,987875
4	0,2085	8,68	2,37	1,315635
5	0,1686	8,42	3,15	0,888522
6	0,1344	7,27	3,63	0,489216
7	0,1013	6,63	3,68	0,298835
8	0,0711	6,18	3,7	0,176328
9	0,0422	5,78	5,25	0,022366
10	0,014	5,52	5,35	0,00238

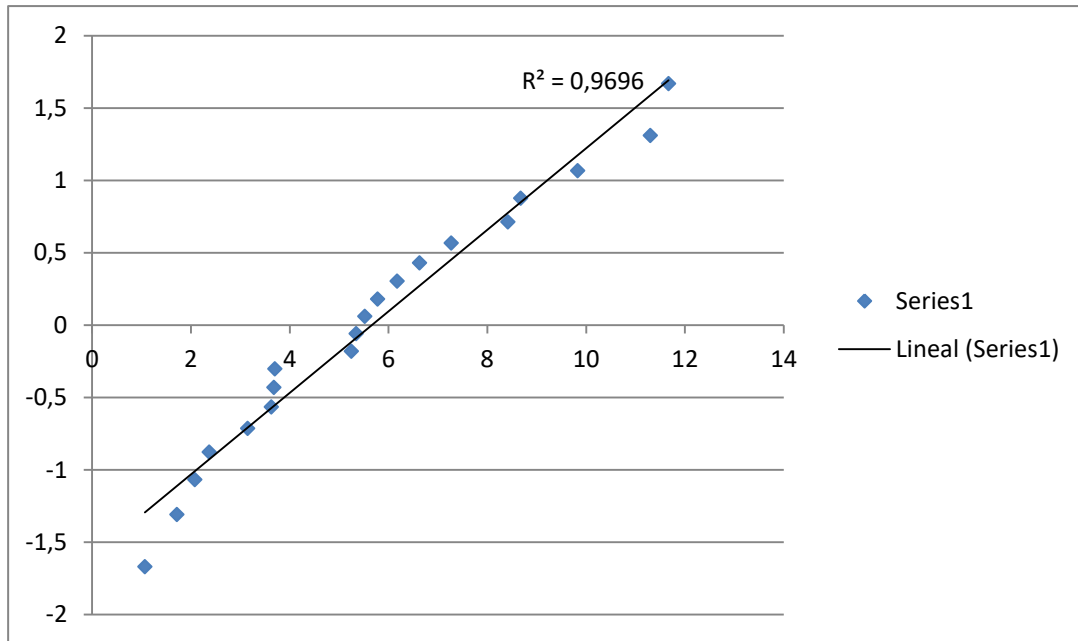
Wcalc = 0,95482244

Wcrit = 0,905

- Prueba de normalidad

i	Yi	Zi	i/21
1	1,07	-1,66839119	0,04761905
2	1,72	-1,30917172	0,0952381
3	2,08	-1,06757052	0,14285714
4	2,37	-0,87614285	0,19047619
5	3,15	-0,71244303	0,23809524
6	3,63	-0,56594882	0,28571429
7	3,68	-0,4307273	0,33333333
8	3,7	-0,30298045	0,38095238
9	5,25	-0,18001237	0,42857143
10	5,35	-0,0597171	0,47619048
11	5,52	0,0597171	0,52380952
12	5,78	0,18001237	0,57142857
13	6,18	0,30298045	0,61904762
14	6,63	0,4307273	0,66666667
15	7,27	0,56594882	0,71428571
16	8,42	0,71244303	0,76190476
17	8,68	0,87614285	0,80952381
18	9,83	1,06757052	0,85714286
19	11,3	1,30917172	0,9047619
20	11,67	1,66839119	0,95238095

Z vs Y



8T Norte-Sur

- Shapiro y Wilk(nivel de significancia = 0,05)

i	Aj20	Y(n-j+1)	Yj	Aj20*(Y(n-j+1)-Yj)
1	0,4734	10,4	0,87	4,511502
2	0,3211	10,083	1,82	2,6532493
3	0,2565	8,62	2,07	1,680075
4	0,2085	8,28	2,8	1,14258
5	0,1686	7,67	2,87	0,80928
6	0,1344	6,4	3,23	0,426048
7	0,1013	6,08	3,97	0,213743
8	0,0711	6	5,55	0,031995
9	0,0422	5,93	5,63	0,01266
10	0,014	5,93	5,77	0,00224

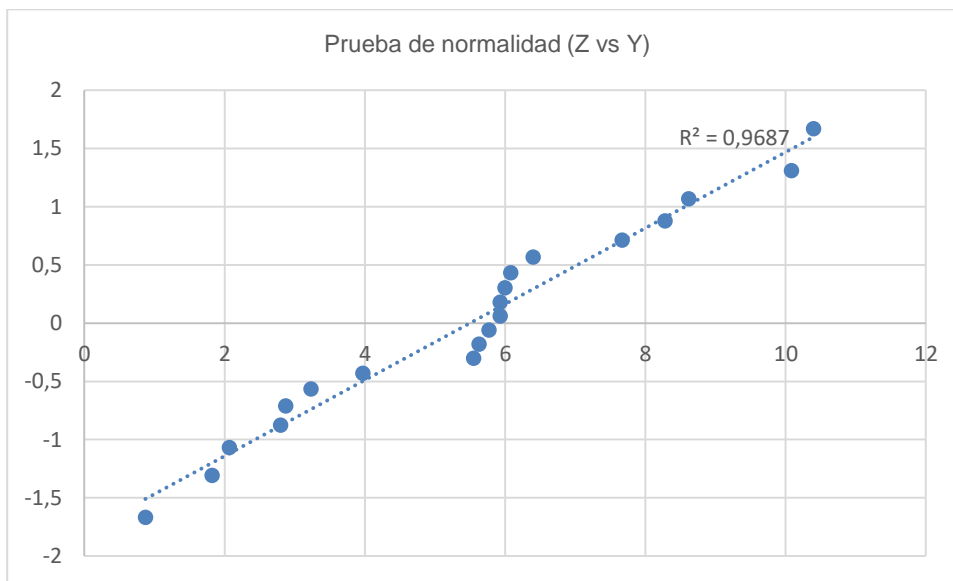
$W_{calc} = 0,95911589$

$W_{crit} = 0,905$

- Prueba de normalidad

i	Y	i/21	Zi
1	0,87	0,047619048	-1,66839119
2	1,82	0,095238095	-1,30917172
3	2,07	0,142857143	-1,06757052
4	2,8	0,19047619	-0,87614285
5	2,87	0,238095238	-0,71244303
6	3,23	0,285714286	-0,56594882
7	3,97	0,333333333	-0,4307273
8	5,55	0,380952381	-0,30298045
9	5,63	0,428571429	-0,18001237
10	5,77	0,476190476	-0,0597171

11	5,93	0,523809524	0,0597171
12	5,93	0,571428571	0,18001237
13	6	0,619047619	0,30298045
14	6,08	0,666666667	0,4307273
15	6,4	0,714285714	0,56594882
16	7,67	0,761904762	0,71244303
17	8,28	0,80952381	0,87614285
18	8,62	0,857142857	1,06757052
19	10,083	0,904761905	1,30917172
20	10,4	0,952380952	1,66839119



Se puede observar que en todos los contraste realizados por Shapiro y Wilk y prueba de normalidad nos da como resultado que los datos provienen de una distribución normal debido que en todos los casos el W calculado es mayor que el W crítico y el R de la grafica es cercano a 1, es decir se puede decir que hay una relación aproximadamente lineal.

Una vez realizado estas pruebas, pasamos a realizar una prueba de homogeneidad de varianzas.

Para realizar esta prueba realizamos la distribución F, donde las hipótesis son:

H^o) $\sigma_1 / \sigma_2 = 1$

H_a) $\sigma_1 / \sigma_2 > 1$

Prueba de varianzas para el 6B (Norte-Centro y Centro-Norte):

Datos			
S1 =	2,12377643	GI1 =	19
S2 =	1,97492699	GI2 =	19

V1=	Norte-Centro
V2=	Centro - Norte

Estadístico de prueba F: $F=(S1/S2)^2$

Fo = 1,156419761

Fc = 2,168251601

Prueba de varianzas para el 8 Transversal (Norte-Centro y Centro-Norte):

Datos			
S1 =	3,12	GI1 =	19
S2 =	2,69	GI2 =	19

V1=	varianza Sur-Norte
V2=	varianza Norte-Sur

Estadístico de prueba F: $F=(S1/S2)^2$

Fo = 1,34

Fc = 2,17

Prueba de varianzas para el 8 Transversal y 6B (Centro-Norte)

Datos			
S1 =	3,11678883	GI1 =	19
S2 =	1,97492699	GI2 =	19

V1=	8 transversal
V2=	6B

Estadístico de prueba F: $F=(S1/S2)^2$

Fo = 2,490649817

Fc = 2,168251601

Prueba de varianzas para el 8 Transversal y 6B (Norte-Centro)

Datos			
S1 =	2,69002986	GI1 =	19
S2 =	2,12377643	GI2 =	19

V1=	8 transversal
V2=	6B

Estadístico de prueba F: $F=(S1/S2)^2$

Fo = 1,604340733

Fc = 2,168251601

Como se puede observar, en todos los casos el F(observado) es menor que el F(critico), por lo tanto no hay pruebas suficientes para rechazar la hipótesis nula, por lo tanto se concluye que las varianzas son iguales.

Pasaremos a realizar la comparación entre las frecuencias del 8 transversal y del 6B, las comparaciones serán realizadas entre frecuencias de ida y vuelta de colectivos de la misma línea y frecuencia de ida (Centro-Norte) entre colectivos 6B y 8 Transversal, como también la frecuencia de vuelta (Norte-Centro) entre los mismos.

Para realizar dichas comparaciones utilizaremos la distribución "T-student", debido a que los datos provienen de una distribución normal, no conocemos la varianza poblacional, y el N es menor que 30 no podemos aproximar la misma utilizando el teorema del limite central.

Estadístico de Prueba "T-student"

- 8 Transversal.

U1= 8 T Sur-Norte

U2= 8 T Norte-Sur

Ho) U1-U2=0

Ha) U1-U2≠0

Y1=	5,664	S1 =	3,11678883	GL=19
Y2=	5,49865	S2 =	2,69002986	""

to= 0,17960823

tc= 2,02439416

Intervalo de confianza:

1,19 < U1-U2 < 1,53

- 6B

U1= Centro-Norte

U2= Norte-Centro

Ho) U1-U2=0

Ha) U1-U2≠0

Y1=	4,90639595	S1 =	1,97492699	GL=19
Y2=	4,795	S2 =	2,12377643	""

to= 0,17177767

tc= 2,02439416

Intervalo de confianza:

$$-0.85 < U1-U2 < 1,07$$

- 6B- 8 transversal (Centro-Norte)

U1= 8 T Sur-Norte
U2= 6B Centro-Norte

Ho) U1-U2=0

Ha) U1-U2≠0

Y1=	5,664	S1 =	3,11678883	GL=19
Y2=	4,90639595	S2 =	1,97492699	""

to= 0,91823343

tc= 2,02439416

Intervalo de confianza:

$$-0.46 < U1-U2 < 1,98$$

- 6B- 8 transversal (Norte-Centro)

U1= 8 T Norte-Sur
U2= 6B Norte-Centro

Ho) U1-U2=0

Ha) U1-U2≠0

Y1=	5,49865	S1 =	2,69002986	GL=19
Y2=	4,795	S2 =	2,12377643	""

to= 0,91815075

tc= 2,02439416

Intervalo de confianza:

$$-0.43 < U1-U2 < 1,84$$

- Frecuencia teorica VS Experimental 6B (Centro-Norte)

Ho) U=5

Ha) U>5

Y=	4,90639595
S=	1,97492699
n=	20

to= -

$$tc = 0,21196228$$

$$tc = 1,72913281$$

Intervalo de confianza:

$$3.98 < U1-U2 < 5.83$$

- Frecuencia teorica VS Experimental 6B (Norte-Centro)

$$H_0) \quad U=5$$

$$H_a) \quad U>5$$

Y=	4,795
S=	2,12377643
n=	20

$$t_0 = 0,43167815$$

$$tc = 1,72913281$$

Intervalo de confianza:

$$3.8 < U1-U2 < 5.79$$

- Frecuencia teorica VS Experimental 8 Transversal (Sur-Norte)

$$H_0) \quad U=8$$

$$H_a) \quad U>8$$

Y=	5,664
S=	3,11678883
n=	20

$$t_0 = 3,35181822$$

$$tc = 1,72913281$$

Intervalo de confianza:

$$4.2 < U1-U2 < 7.12$$

- Frecuencia teorica VS Experimental 8 Transversal (Norte-Sur)

$$H_0) \quad U=8$$

$$H_a) \quad U>8$$

Y=	5,49865
S=	2,69002986

n=	20
----	----

to= -4,1584584

tc= 1,72913281

Intervalo de confianza:

5.24 < U1-U2 < 6.76

Conclusión

Los datos de las medias poblacionales del tiempo entre un colectivo y otro fueron proporcionados por el presidente de prensa del servicio de ómnibus "Saeta" (fuente confiable), a continuación describiremos las observaciones obtenidas a partir de la inferencia (análisis) estadísticas aprendidas en el curso:

Para el caso de la línea 6B los valores obtenidos por los intervalos de confianza (tanto para el colectivo que va hacia el centro como para el que viene del mismo), se puede observar que en ambos casos contienen al 5 (minutos), quiere decir de que el dato proporcionado por SAETA es verdadero; se puede llegar a la misma conclusión observando los valores "t" arrojados por la prueba de hipótesis la cual indica que la media poblacional es igual a la estipulada previamente.

Al analizar los datos de la línea 8 Transversal, si bien mediante la prueba de hipótesis se llega a la conclusión de que la media poblacional cumple con el dato previsto por SAETA (8 minutos); el intervalo nos indica que la verdadera media (de frecuencia del ómnibus) es menor a 8, ya que el mismo no contiene este valor. Lo cual resulta bastante beneficioso para los ciudadanos que ocupan esta línea.

La presunción de la gente respecto a la frecuencia de la línea 8T y 6B fue que la línea 6B tiene mayor frecuencia que la del 8T

Encuesta	¿Qué colectivo cree que pasa mas seguido?		
	6B	8T	Indiferentes
Mayor Frecuencia	33	6	11

Al realizar la prueba de hipótesis de la diferencia de frecuencias entre el 8 transversal y el 6B, llegamos a la conclusión de que en ambos casos (Norte-Centro, Centro-Norte) no hay diferencias entre las frecuencias de los colectivos debido a que los intervalos de confianza contienen al cero y a su vez el estadístico de prueba cae en la zona de aceptación. Con estos resultados podemos ver que la gente se equivocaba al afirmar que el 6B pasa mas seguido que el 8. Cabe destacar que durante la realización de la encuesta varias personas afirmaron que el Transversal recientemente aumento su frecuencia, por lo cual también podemos concluir que el tiempo estimado para esperar el Transversal es menor al que nos dieron en SAETA, pensamos que la frecuencia del mismo recientemente fue modificada.

Bibliografía:

Principios estadísticos de diseño y análisis de investigación. Segunda edición. Robert O. Kuehl.
Introducción a la probabilidad y estadística. Décima tercera edición. Mendenhall.
FUNDAMENTOS DE ESTADISTICA del autor DANIEL PEÑA SANCHEZ DE RIVERA.