



Universidad de Chile
Facultad de Ciencias Sociales
Profesora: Catalina Canals
Ayudantes: Daniel Cifuentes, Camila Lizama,
Daniela Olivares, Sofía Villalobos

Guía de Ejercicios Modulo 5 Estadística II, sección 2

Parte I: Identificación de Hipótesis y elección del Test

En los siguientes casos: (1) identifique la hipótesis de investigación, (2) indique si tiene toda la información necesaria sobre los datos para decidir qué test utilizar, (3.a) si tiene toda la información necesaria sobre los datos para decidir qué test utilizar: sugiera qué test de hipótesis utilizar y por qué, indique si se cumple cada uno de los supuestos del test sugerido, y plantee las hipótesis nula y alternativa del test a realizar, (3.b) si NO tiene toda la información necesaria sobre los datos para decidir qué test utilizar, indique qué información requeriría para saber qué test recomendar, sugiera cómo podría obtenerla, e indique qué test recomendaría –y cuáles serían la hipótesis nula y la alternativa-en cada caso posible.

1. Un grupo de investigadores está realizando un estudio sobre las desigualdades educativas entre la Población que reside en la Región Metropolitana y la Población que reside en otras las regiones. Particularmente, tienen la sospecha de que la población de la Región Metropolitana tienden a tener más escolaridad, para lo cual están interesados en utilizar una muestra probabilística de 2000 chilenos, para hacer dos comparaciones:
 - a. Comparar los años de escolaridad de la población de la Región Metropolitana y de la que reside en otras las regiones.
 - b. Comparar el porcentaje de personas con educación superior completa en ambas poblaciones.

Los antecedentes de la muestra que tienen los investigadores son los siguientes.

	Región Metropolitana	Otras regiones
Tamaño muestral	300	1700
Valor P en Test Kolmogorov Simirnov de normalidad de la variable escolaridad (años)	0.1	0.04
Valor P en Test de Levene de igualdad de varianzas de la variable escolaridad (años)	0.2	

2. Los investigadores quieren repetir el análisis pero distinguiendo ahora en las regiones de la zona Norte, del Centro y del Sur del país. Particularmente, quieren comparar como es la escolaridad en dichas 3 zonas, usando una muestra probabilística de 3000 casos. En este caso, los investigadores esperan encontrar que la escolaridad sea diferente en las distintas zonas. Los antecedentes que disponen de la muestra son los siguientes:

	Norte	Sur	Centro
Tamaño muestral	750	750	500
Valor P en Test Kolmogorov Simirnov de normalidad de la	0.1	0.04	0.09

variable escolaridad (años)			
Valor P en Test de Levene de igualdad de varianzas de la variable escolaridad (años)	0.2		

3. Un grupo de investigadores está realizando un estudio sobre las desigualdades educativas entre la Población que reside en Puerto Montt y Puerto Varas. Particularmente, tienen la sospecha de que la población de Puerto Varas tienden a tener mejores resultados educativos. Para estudiar este tema, cuentan con una muestra de 100 estudiantes que rindieron el SIMCE de II Medio en el 2013, a través de la cual quisieran hacer dos comparaciones:
- Comparar los puntajes SIMCE de los dos sectores.
 - Comparar la proporción de alumnos con nivel de aprendizaje adecuado en el SIMCE¹ de los dos sectores.

Los antecedentes de la muestra que tienen los investigadores son los siguientes.

	Puerto Montt	Puerto Varas
Tamaño muestral	50	50
Valor P en Test Kolmogorov de igualdad de distribuciones de la variable Puntaje SIMCE	0.12	

Parte II: Resolución de test de hipótesis y estimación de indicadores de tamaño del efecto

4. Un grupo de investigadores desarrolla un índice que mide el nivel de aprobación del gobierno, siendo 0 la total desaprobación y 10 la total aprobación. Para evaluar si hay diferencias entre la aprobación del gobierno en los trabajadores dependientes o independientes en comparación con los empleadores, a partir de una muestra probabilística de 1000 casos se desarrolló un test Z para evaluar diferencia de medias en muestras independientes. A partir de su estimación, y siendo \bar{X}_t el promedio de aprobación del gobierno en la muestra de trabajadores, y \bar{X}_e el de los empresarios, el estadístico Z observado en su muestra fue igual el siguiente:

$$Z_{obs} = \frac{\bar{X}_t - \bar{X}_e}{SE(\bar{X}_t - \bar{X}_e)} = 2$$

- Si su hipótesis nula fuese que $\mu_t = \mu_e$, y su hipótesis alternativa fuese $\mu_t \neq \mu_e$,
 - ¿Cuál sería el Valor P? Interprete el valor P y muestre gráficamente a qué probabilidad corresponde.
 - Con 5% de significancia, ¿Rechazaría la hipótesis nula?
- Si su hipótesis nula fuese que $\mu_t = \mu_e$, y su hipótesis alternativa fuese $\mu_t > \mu_e$,
 - ¿Cuál sería el Valor P? Interprete el valor P y muestre gráficamente a qué probabilidad corresponde.
 - Con 5% de significancia, ¿Rechazaría la hipótesis nula?

¹ A partir de los resultados SIMCE, el nivel de logro de cada alumno es categorizado como insuficiente, elemental o adecuado, siendo este último el que indica un mayor nivel de logro.

- c. Interprete sociológicamente sus resultados.
5. Analice si existe asociación o no entre estar estudiando en educación superior y tomar la píldora anticonceptiva para las chilenas menores de 30 años (con 5% de significancia), y en caso de haberla estime la magnitud de la asociación. Repita el análisis controlando por nivel socioeconómico. interprete sus resultados.

Para realizar las estimaciones utilice los siguientes datos, suponga que provienen de una muestra probabilística.

Mujeres de NSE Medio o Alto		
	Estudia	No estudia
Toma píldora	40	30
No toma píldora	10	25
Mujeres de NSE Bajo		
	Estudia	No estudia
Toma píldora	30	20
No toma píldora	10	25

Parte III: Identificación de errores

Lea el siguiente análisis realizado, identifique los errores que se comenten, indique porqué es incorrecto y cuál sería la corrección a tal error.

6. Para comparar la proporción de fumadores en Chile antes y después de la ley que restringe el consumo de tabaco en lugares públicos cerrados, se utilizó una muestra de 1000 individuos, que fueron encuestados antes y después de la vigencia de dicha ley. Para comparar si la proporción de fumadores disminuyó (hipótesis de investigación) se realizó un utilizar un test Z para proporciones, planteando la hipótesis nula de que la proporción de fumadores se mantuvo constante y la hipótesis alternativa de que la proporción de fumadores disminuyó. Al realizar la estimación en SPSS se obtuvieron los siguientes resultados. Al ser el Valor P mayor que la significancia (5%), se rechaza la hipótesis nula y se concluye que la proporción de fumadores aumentó.

Test T para proporciones						
	t	gl	Sig. (bilateral)	Diferencia de medias	95% de intervalo de confianza de la diferencia	
					inferior	superior
Fumador	36,165	2530	,100	0,36224	0,1799	0,5445

7. Para comparar el promedio PSU de los estudiantes de los estudiantes de Psicología, Sociología y Antropología de la U. de Chile, considerando la información registrada de todos los estudiantes se realiza un ANOVA de una vía, dado que se cumplen todos los supuestos de dicho test. El valor P asociado al test es =0,001, por lo cual se rechaza la hipótesis nula y se concluye que el promedio PSU es diferente en las tres carreras. Para confirmar esta información, se realizan tres test T para muestras relacionadas: (1) para comparar el promedio PSU de Psicología y Sociología, (2) para comparar el promedio PSU de Antropología y Sociología, y (3) para comparar el promedio PSU de Psicología y

Antropología. Por otra parte, el eta cuadrado estimado corresponde a un 20%, lo que quiere decir que la carrera con mayor puntaje PSU promedio supera en un 20% al promedio de aquella con menor puntaje PSU.