



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

## Guía de Repaso Examen Estadística II, sección 2

### Ejercicios

#### Parte 1: Recomendación de test a utilizar

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 (si no se cumplen todos los supuestos, explique por qué recomienda utilizarlo a pesar de ello), 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. Unos investigadores están interesados en estimar el porcentaje de desempleados en Chile, a través de la CASEN 2013, y analizar si este alcanza las dos cifras. Se disponen de los siguientes datos de la muestra:

Datos Muestra (n=200.000)	
Valor P en prueba Kolmogorov-Smirnov de normalidad de variable desempleado (1 Si, 0 No)	0,0001

2. Unos investigadores construyeron un índice de tolerancia a los inmigrantes, donde el valor 10 indica total tolerancia y 0 indica nula. Los investigadores quieren analizar si el valor de este índice difiere en la población de Inglaterra y Escocia a través de una muestra probabilística de 100 casos. Se disponen de los siguientes datos de la muestra:

Datos Muestra (n=100)	
Valor P en prueba Kolmogorov-Smirnov de normalidad del índice	0,1
Valor P en prueba Jarque-Vera de simetría del índice	0,5

3. Los mismos investigadores, y a través de la misma muestra, quieren analizar si existe asociación entre votar a favor del *Brexit* (Que el Reino Unido salga de la Unión Europea) y la tolerancia a los inmigrantes. Se disponen de los siguientes datos de la muestra:

Tabla de Frecuencias esperadas en caso de no exista asociación entre las variables		
	Contra el <i>Brexit</i>	A favor del <i>Brexit</i>
Baja Tolerancia a la inmigración	16.7	16.7
Media Tolerancia a la inmigración	16.7	16.7
Alta Tolerancia a la inmigración	16.7	16.7

## Parte 2: Verdadero (V) o Falso (F)

En los siguientes casos, indique si las afirmaciones son verdaderas o falsas. En caso de ser falsas indique porqué lo son, e indique qué sería lo correcto. En caso de ser verdaderas, explique por qué lo son, agregue información que complemente la frase, y/o explique qué quiere decir la afirmación.

4. \_\_\_\_\_. Un buen estimador es aquel que es insesgado y eficiente.
5. \_\_\_\_\_. La distribución muestral de medias siempre distribuye normal estándar.
6. \_\_\_\_\_. Si con una muestra probabilística de profesores del país, se estima un intervalo [10,15] horas de trabajo semanal no lectivas (horas dedicadas a la preparación de clases), con 95% de confianza, esto implica que, con cualquier nivel de significancia no rechazamos la hipótesis nula de que el promedio de horas no lectivas de trabajo semanal de los profesores de Chile es igual a 10.
7. \_\_\_\_\_. Si al realizar un test Z (a partir de una muestra donde se cumplen todos los supuestos del test) donde la hipótesis alternativa corresponde a “el promedio de distancia que recorren los habitantes de la zona nororiente de Santiago para llegar al trabajo es menor que el de los habitantes de otras zonas de Santiago” obtengo un estadístico  $Z= 1,95$ ; esto implica que si quiero poner a prueba la hipótesis alternativa: “el promedio de distancia que recorren los habitantes de la zona nororiente de Santiago para llegar al trabajo es MAYOR que el de los habitantes de otras zonas de Santiago” el valor del estadístico Z será el mismo.
8. \_\_\_\_\_. El test U-Mann-Whitney suele utilizarse cuando no se cumplen los supuestos para realizar un Test T de igualdad de medias.
9. \_\_\_\_\_. Si al realizar un Test de Mc Nemmar (en un caso donde se cumplen sus supuestos), para poner a prueba la hipótesis nula de que la proporción de estudiantes de la Universidad de Chile a favor del paro era igual antes y después de que se hiciera público el proyecto de ley de reforma a la educación superior (con la hipótesis alternativa que plantea que dicha proporción cambió), si  $\chi^2=1,6424$  el Valor  $P=0,2$ .
10. \_\_\_\_\_. El test de Z de proporciones permite comparar las proporciones de dos muestras relacionadas.
11. \_\_\_\_\_. Si se tiene una muestra pequeña, sería correcto utilizar el Test Exacto de Fisher para poner a prueba la hipótesis nula de que existe asociación entre las variables nivel socioeconómico y considerarse lector habitual.
12. \_\_\_\_\_. En una muestra de 30 estudiantes, si al realizar un test (donde se cumplen todos los supuestos del mismo) no rechazamos la hipótesis nula de que la proporción de estudiantes interesados en jugar Pokemon GO es igual a 50%, con un valor  $P=0,1$ ; sería recomendable realizar un estudio con una muestra más grande para estar seguros de que dicha hipótesis es cierta.
13. \_\_\_\_\_. Al realizar un ANOVA de dos vías, en un caso donde se cumplen todos los supuestos y donde la variable dependiente es la cobertura del arancel de referencia respecto al arancel real de una carrera de la Universidad de Chile (arancel de referencia\*100/arancel real), y la variable

independiente es el área de la carrera (Salud/ Ciencias Sociales/ Ciencias Exactas/ Tecnología, etc.), se obtiene un Valor P asociado al estadístico  $F=0,01$ , y un eta cuadrado  $=20\%$ , esto quiere decir que el área de la carrera explica el 20% de la varianza de la cobertura del arancel de referencia respecto al arancel real de las carreras, y con 95% de confianza, la cobertura de estos promedio difiere en todas las áreas de carreras.

14. \_\_\_\_\_. Al analizar la correlación entre la cobertura del arancel de referencia respecto al arancel real y el porcentaje de estudiantes del quintil V en las carreras de la Universidad de Chile, la correlación es de 0,3, y el Valor P del test de hipótesis asociado es 0,01. Esto indica que, no existe relación lineal entre ambas variables.

### Parte 3: Regresión lineal múltiple

Un grupo de investigadores realiza un estudio para predecir el porcentaje de estudiantes del quintil V en las carreras de la Universidad de Chile, a partir de la variable (1) cobertura del arancel de referencia respecto al arancel real, obteniendo los siguientes resultados.

Observaciones: 100

R cuadrado: 0,2

Prob >F: 0,001

Variable	Coefficiente	P> t
Cobertura del arancel de referencia respecto al arancel real (%)	1	0.01

15. Refiérase al ajuste del modelo (interprete el R cuadrado y el test asociado al estadístico F)

16. Interprete estadísticamente los coeficientes de su modelo. Interprete sociológicamente sus resultados.