



**UTPL**  
*La Universidad Católica de Loja*

**Modalidad Abierta y a Distancia**

# Estadística Inferencial

## Guía didáctica



Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos



Departamento de Ciencias Biológicas

Sección departamental de Ecología y Sistemática

---

## Estadística Inferencial

*Guía didáctica*

Autor:

Griffith Daniel M.



E S T A \_ 3 0 2 1

Asesoría virtual  
[www.utpl.edu.ec](http://www.utpl.edu.ec)

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

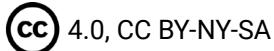
Recursos

## Estadística Inferencial

Guía didáctica

Griffith Daniel M.

Universidad Técnica Particular de Loja



## Diagramación y diseño digital:

Ediloja Cía. Ltda.

Telefax: 593-7-2611418.

San Cayetano Alto s/n.

[www.ediloja.com.ec](http://www.ediloja.com.ec)

[edilojainfo@ediloja.com.ec](mailto:edilojainfo@ediloja.com.ec)

Loja-Ecuador

ISBN digital - 978-9942-25-841-0



La versión digital ha sido acreditada bajo la licencia Creative Commons 4.0, CC BY-NY-SA: Reconocimiento-No comercial-Compartir igual; la cual permite: copiar, distribuir y comunicar públicamente la obra, mientras se reconozca la autoría original, no se utilice con fines comerciales y se permiten obras derivadas, siempre que mantenga la misma licencia al ser divulgada. <https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>

24 de septiembre, 2020

# Índice

<b>1. Datos de información.....</b>	<b>9</b>
1.1. Presentación de la asignatura .....	9
1.2. Competencias Genéricas de la UTPL .....	9
1.3. Competencias específicas de la carrera .....	9
1.4. Problemática que aborda la asignatura .....	10
<b>2. Metodología de aprendizaje.....</b>	<b>10</b>
<b>3. Orientaciones didácticas por resultados de aprendizaje .....</b>	<b>12</b>
 <b>Primer bimestre.....</b>	 <b>12</b>
Resultado de aprendizaje 1 .....	12
Contenidos, recursos y actividades de aprendizaje.....	12
 <b>Semana 1 .....</b>	 <b>12</b>
<b>Unidad 1. Introducción a la estadística inferencial .....</b>	<b>13</b>
1.1. Repaso de la estadística descriptiva.....	13
Actividades de aprendizaje recomendadas .....	14
 <b>Semana 2 .....</b>	 <b>15</b>
1.2. Distribución muestral de un estadístico.....	15
1.3. Estadística inferencial: definiciones y métodos .....	15
Actividades de aprendizaje recomendadas .....	16
Autoevaluación 1 .....	18
 <b>Semana 3 .....</b>	 <b>23</b>
<b>Unidad 2. Intervalo de confianza.....</b>	<b>23</b>
2.1. Estimación puntual.....	23
2.2. Construcción e interpretación de intervalos de confianza .....	23
Actividades de aprendizaje recomendadas .....	24

<b>Índice</b>	
<b>Primer bimestre</b>	
<b>Segundo bimestre</b>	
<b>Solucionario</b>	
<b>Referencias bibliográficas</b>	
<b>Recursos</b>	
<b>Semana 4 .....</b>	<b>25</b>
2.3. Intervalos de confianza de muestras grandes para una media y una proporción poblacional.....	25
2.4. Intervalos de confianza para la diferencia entre dos medias poblacionales y entre dos proporciones binomiales .....	25
Actividades de aprendizaje recomendadas .....	26
Autoevaluación 2 .....	27
Resultado de aprendizaje 2 .....	32
Contenidos, recursos y actividades de aprendizaje.....	32
<b>Semana 5 .....</b>	<b>32</b>
<b>Unidad 3. Prueba de hipótesis .....</b>	<b>33</b>
3.1. Procedimientos de una prueba estadística de hipótesis .....	33
Actividades de aprendizaje recomendadas .....	34
<b>Semana 6 .....</b>	<b>35</b>
3.2. Una prueba de muestra grande acerca de una media poblacional.....	35
3.3. Una prueba de hipótesis para la diferencia entre dos medias poblacionales.....	35
Actividades de aprendizaje recomendadas .....	36
<b>Semana 7 .....</b>	<b>37</b>
3.4. Una prueba de hipótesis para una proporción binomial.	37
3.5. Una prueba de hipótesis para la diferencia entre dos proporciones binomiales .....	37
Actividades de aprendizaje recomendadas .....	38

<b>Índice</b>	
<b>Primer bimestre</b>	
<b>Segundo bimestre</b>	
<b>Solucionario</b>	
<b>Referencias bibliográficas</b>	
<b>Recursos</b>	
<b>Semana 8 .....</b>	<b>38</b>
Actividades de aprendizaje recomendadas .....	39
Autoevaluación 3 .....	40
Actividades finales del bimestre.....	47
<b>Segundo bimestre .....</b>	<b>48</b>
Resultado de aprendizaje 3.....	48
Contenidos, recursos y actividades de aprendizaje.....	48
<b>Semana 9 .....</b>	<b>48</b>
<b>Unidad 4. Prueba t de Student .....</b>	<b>49</b>
4.1. Prueba t para hacer inferencias de muestra pequeña respecto a una media poblacional.....	49
4.2. Prueba t para la diferencia entre dos medias: muestras independientes.....	49
Actividades de aprendizaje recomendadas .....	50
Autoevaluación 4 .....	51
<b>Semana 10 .....</b>	<b>56</b>
<b>Unidad 5. Análisis de varianza (ANOVA).....</b>	<b>56</b>
5.1. Introducción a los modelos lineales .....	56
5.2. ANOVA de un factor .....	56
Actividades de aprendizaje recomendadas .....	57
<b>Semana 11 .....</b>	<b>58</b>
5.3. Pruebas de comparaciones múltiples (post hoc) .....	58
Actividades de aprendizaje recomendadas .....	59
Autoevaluación 5 .....	60
Resultado de aprendizaje 4.....	66

Contenidos, recursos y actividades de aprendizaje.....	66
<b>Semana 12 .....</b>	<b>66</b>
<b>    Unidad 6. Regresión línea simple y correlación .....</b>	<b>66</b>
6.1. Regresión lineal.....	66
Actividades de aprendizaje recomendadas .....	68
<b>    Semana 13 .....</b>	<b>69</b>
6.2. Correlación bivariada .....	69
Actividades de aprendizaje recomendadas .....	70
Autoevaluación 6 .....	71
Resultado de aprendizaje 5 .....	75
Contenidos, recursos y actividades de aprendizaje.....	75
<b>    Semana 14 .....</b>	<b>75</b>
<b>    Unidad 7. Pruebas no paramétricas .....</b>	<b>76</b>
7.1. Tablas de contingencias: una clasificación de dos vías	76
7.2. Prueba chi cuadrado: evaluación de la asociación entre dos variables categóricas .....	76
Actividades de aprendizaje recomendadas .....	77
<b>    Semana 15 .....</b>	<b>78</b>
7.3. Prueba de suma de rango de Wilcoxon (muestras independientes) y de rango con signo de Wilcoxon (un experimento pareado).....	78
Actividades de aprendizaje recomendadas .....	79
<b>    Semana 16 .....</b>	<b>79</b>
7.4. Prueba de análisis de varianza de Kruskal-Wallis.....	79
Actividades de aprendizaje recomendadas .....	80
Autoevaluación 7 .....	81

Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos

Actividades finales del bimestre .....	86
<b>4. Solucionario .....</b>	<b>87</b>
<b>5. Referencias bibliográficas .....</b>	<b>99</b>
<b>6. Recursos .....</b>	<b>100</b>

Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos



## 1. Datos de información

### 1.1. Presentación de la asignatura



### 1.2. Competencias Genéricas de la UTPL

Pensamiento crítico y reflexivo

### 1.3. Competencias específicas de la carrera

Sustenta técnica y científicamente propuestas pertinentes y factibles para el manejo y conservación de los recursos ambientales.

## 1.4. Problemática que aborda la asignatura

Débil sustento-técnico científico a las propuestas de manejo y conservación de los recursos ambientales.



## 2. Metodología de aprendizaje

A lo largo del desarrollo de esta asignatura, se utilizarán tres metodologías de aprendizaje:

1. **Autoaprendizaje:** es un proceso donde el/la estudiante orientado por el/la docente desarrolla de forma autónoma e independiente competencias que le permitan adquirir los diferentes resultados de aprendizaje. Para cada uno de las 7 Unidades propuestas en esta asignatura, cada estudiante tendrá que resumir y realizar esquemas para entender y estudiar los contenidos propuestos orientados siempre por el/la docente.
2. **Aprendizaje de indagación:** esta metodología de aprendizaje se basa en la investigación y reflexión que siguen los/as estudiantes para llegar a una solución ante un problema planteado por el/la docente. Específicamente y con el objetivo de desarrollar un pensamiento crítico en los/as estudiantes se expondrán situaciones y problemáticas actuales relacionadas con la Gestión Ambiental a través de los Foros y de actividades práctico experimentales, donde estudiantes y docente interactuarán constantemente a través del EVA para buscar soluciones.

3. **Aula invertida:** consiste en un enfoque pedagógico para el proceso de enseñanza donde los roles del docente y del estudiante se invierten. Específicamente, se diseñarán actividades en los dos bimestres en las cuáles los/as estudiantes tendrán que asumir el rol de docente y crear contenidos didácticos en grupo y posteriormente compartir entre el resto de participantes, siempre con la supervisión y apoyo del docente.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos



---

### 3. Orientaciones didácticas por resultados de aprendizaje

---



#### Primer bimestre

##### Resultado de aprendizaje 1

Realiza estimaciones de parámetros poblacionales

#### Contenidos, recursos y actividades de aprendizaje

El primer resultado de aprendizaje se trabajará durante las cuatro primeras semanas del primer bimestre. A través de este resultado de aprendizaje se repasarán algunas de las herramientas básicas de la estadística descriptiva con el objetivo de aprender a organizar, explorar e interpretar un conjunto de datos, para posteriormente iniciar con algunos conceptos generales sobre la estimación de parámetros poblacionales, principalmente estimación puntual e intervalos de confianza.



##### Semana 1



## Unidad 1. Introducción a la estadística inferencial

### 1.1. Repaso de la estadística descriptiva

Bienvenidos/as a la primera semana de Estadística Inferencial. En esta primera semana vamos a definir el concepto de estadística y a repasar algunos de los conceptos más importantes de la estadística descriptiva (Unidad 1 apartado 1.1). Primero repasaremos las diferentes variables estadísticas que podemos encontrar en un estudio o investigación y sus escalas de medición (Mendenhall et al. 2015, Capítulo 1, apartados 1.1 y 1.2). Despues repasaremos los principales estadísticos descriptivos (medidas de centro, de variación y de posición) que nos permiten comparar entre muestras y dar una idea rápida de cómo se distribuyen los datos (Mendenhall et al. 2015, Capítulo 2, apartados 2.1-2.3 y 2.6). Por último repasaremos la distribución normal de probabilidad que es considerada como la más importante y útil en el campo de la probabilidad y la estadística (Mendenhall et al. 2015, Capítulo 6, apartado 6.2).

Esta semana no tendremos ninguna actividad calificada, pero sí les propongo realizar una pequeña presentación personal y académica en el Foro Académico e interactuar en el horario de tutorías para resolver cualquier duda que pueda surgir sobre el contenido impartido.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar Introducción y apartados 1.1-1.2, 2.1-2.3, 2.6, 6.2.

### Recurso Educativo Abierto (REA) nº 1

Por último, les recomiendo visualizar el siguiente video introductorio titulado “[Tutorial 1: Introducción a la inferencia estadística, distribución en el muestreo](#)”

¿Qué les ha parecido esta primera semana? ¿Les han sido familiares los conceptos impartidos? Recuerda que estos conceptos de la Unidad 1 son la base para poder entender bien las siguientes unidades de la materia de Estadística Inferencial.



### Actividades de aprendizaje recomendadas

- Realizar una corta presentación personal y académica en el Foro Académico. Comente qué utilidad le dará la estadística inferencial en su vida profesional.
- Lectura del texto básico recomendado
- Revisar el Recurso Educativo Abierto (REA) nº 1.
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.



## Semana 2

### 1.2. Distribución muestral de un estadístico

### 1.3. Estadística inferencial: definiciones y métodos

En esta segunda semana de clases vamos a continuar con la Unidad 1 de Introducción a la estadística inferencial, pero vamos a profundizar en el apartado 1.2 sobre distribución muestral de un estadístico repasando el teorema del límite central y también recordaremos cómo calcular la distribución muestral de la media y de una proporción (Mendenhall et al. 2015, Capítulo 7, apartados 7.3 y 7.6).

Por último y para terminar esta unidad de repaso definiremos qué es la estadística inferencial y sus métodos (Mendenhall et al. 2015, Capítulo 8, apartados 8.1 y 8.2). Recordemos que al estudiar las probabilidades para el estudio de fenómenos aleatorios seguimos un razonamiento deductivo: establecíamos ciertas hipótesis (modelo de distribución) y a partir de ahí se deducían las propiedades sobre el fenómeno en cuestión (media, varianza, probabilidad, etc.). Por el contrario mediante la inferencia estadística realizamos el proceso inverso: a partir de un conjunto de datos experimentales, se infiere o estima las propiedades del fenómeno bajo estudio. En otras palabras, la estadística inferencial nos permite extraer conclusiones y generalizaciones sobre una población basándonos en la información suministrada por la muestra.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Presentación unidad 1: Distribuciones muestrales y el Teorema del Límite Central

Ir a recursos

### Recursos de aprendizaje

#### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar el Capítulo 7, apartados 7.3-7.6 y Capítulo 8, apartados 8.1-8.2.

#### Recurso Educativo Abierto (REA) nº 2

Les recomiendo revisar el siguiente artículo del REA nº2 para preparar su participación en el foro académico 1: [Cambio climático - amenaza en cascada a la biodiversidad](#) (scidev.net)

Finalmente, no se olviden de resumir los contenidos referentes a distribuciones muestrales, al Teorema del Límite Central y a la definición de estadística inferencial.



#### Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Realizar alguna de las actividades propuestas en el texto básico recomendado

Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos

- Elaborar un resumen con los contenidos referentes a distribuciones muestrales, al Teorema del Límite Central y a la definición de estadística inferencial del texto básico.
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.
- Realizar la autoevaluación 1 para repasar los conocimientos adquiridos en esta Unidad 1.



## Autoevaluación 1

A. En los ítems del 1 al 4, dentro del paréntesis, escriba V si la afirmación es correcta y F si es falsa.

1. ( ) La etapa de exploración de los datos está vinculada exclusivamente con las variables categóricas.
2. ( ) La estadística inferencial busca extraer conclusiones para la población a partir de la muestra.
3. ( ) El teorema del límite central afirma que, a mayor número de muestras, la distribución se aleja de la normalidad.
4. ( ) La distribución muestral de un estadístico es la distribución de probabilidad para los posibles valores del estadístico, que resulta cuando muestras aleatorias de tamaño  $n$  se sacan repetidamente de la población.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

**B. Encierre el literal que corresponde a la opción correcta.**

5. El propósito de la Estadística es:

- a. Resumir la información contenida en un conjunto de datos.
- b. Demostrar el grado en que un hipótesis tiene fundamento.
- c. Extraer conclusiones sobre datos empíricos que pueden ser aplicadas a otros casos.
- d. Todas las respuestas son correctas
- e. Ninguna respuesta es correcta

6. Una característica que puede cambiar entre los elementos de una población o de una muestra se denomina:

- a. Una muestra
- b. Una variable
- c. Una probabilidad
- d. Un parámetro
- e. Estadística descriptiva

7. En la planificación y diseño de un estudio ambiental, ¿cuál de los siguientes objetivos es correcto?

- a. Para optimizar recursos y tiempo, se estudiará toda la población.
- b. Para optimizar recursos y tiempo, se estudiará una muestra representativa de la población.
- c. Para mayor precisión, se estudiará la muestra en vez de la población.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

8. En base a los fundamentos de estadística cuál de las siguientes afirmaciones es correcta:
  - a. Una variable se calcula en función de los parámetros.
  - b. El parámetro se calcula en función de la muestra.
  - c. Un estadístico se determina en función de la muestra.
9. La tasa anual de secuestro de carbono por superficie (tonelada/ha), es una variable de tipo:
  - a. Binomial
  - b. Categórica nominal
  - c. Categórica ordinal
  - d. Numérica discreta
  - e. Numérica continua
10. La gráfica que permite identificar la variación y el valor central de una muestra es un:
  - a. Histograma
  - b. Diagrama de dispersión
  - c. Diagrama de frecuencias
  - d. Diagrama de cajas
11. Para conocer qué dap (diámetro a la altura del pecho) en una plantación forestal es superado por el 5% de los individuos, usted usaría:
  - a. El percentil 95
  - b. El percentil 5
  - c. El cuartil 3
  - d. El decil 9

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

12. Al comparar la gráfica de la distribución de la variable original vs la distribución muestral de la media, se observa que:
  - a. El rango de valores en el eje X es menor para la distribución original
  - b. El rango de valores en el eje X es mayor en la distribución original
  - c. Alejamiento de la media muestral vs la media de la variable original
13. El error estándar de la media muestral se define como el cociente entre:
  - a. La varianza y el tamaño de la muestra
  - b. La desviación estándar y la raíz cuadrada del tamaño de muestra
  - c. El tamaño de muestra y la desviación estándar
  - d. El valor z y el tamaño de la muestra
14. La distribución de la proporción muestral se aproxima a una distribución normal cuando:
  - a. El tamaño de muestra es grande
  - b. El tamaño de muestra es pequeño
  - c. El tamaño de muestra  $n < 30$
  - d. Para cualquier tamaño de muestra
15. La estadística inferencial busca dar respuesta a dos problemas básicos:
  - a. Calcular tamaños de muestras y estimar la media poblacional
  - b. Plantear hipótesis y hacer gráficas
  - c. Estimar parámetros y probar hipótesis

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

Ir al solucionario

Luego de haber respondido la autoevaluación, le recomiendo comparar sus respuestas con el solucionario. En caso de haber discordancia en alguna respuesta, se le sugiere identificar el error y volver a revisar los contenidos respectivos.

[Índice](#)[Primer bimestre](#)[Segundo bimestre](#)[Solucionario](#)[Referencias bibliográficas](#)[Recursos](#)

## Semana 3



### Unidad 2. Intervalo de confianza

#### 2.1. Estimación puntual

#### 2.2. Construcción e interpretación de intervalos de confianza

Bienvenidos/as a la semana 3 de Estadística Inferencial. Esta semana vamos a comenzar con la Unidad 2 sobre el intervalo de confianza, específicamente vamos a trabajar la Unidad 2 apartados 2.1 y 2.2 sobre estimación puntual y las propiedades básicas de los intervalos de confianza. La estimación puntual consiste en dar un valor real a partir de un estadístico, es decir, una función de la muestra que no depende de ningún parámetro desconocido. Por ejemplo, supongamos que queremos estudiar el tamaño (peso en kg) de los osos de anteojos del Parque Nacional Podocarpus que se puede representar mediante una variable aleatoria  $X$  de la que se conoce el tipo de distribución (i.e., normal) pero no los parámetros que la determinan. El objetivo de esta unidad es que usted sea capaz de obtener un valor que pueda asignarse a cada uno de los parámetros desconocidos. A ese rango de valores en los que se

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

confía que contenga el parámetro poblacional, con un determinado nivel de confianza se le denomina intervalo de confianza que se estudiará en el apartado 2.2.

Presentación unidad 2: Intervalo de confianza para una media poblacional

Ir a recursos

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar el Capítulo 8, apartados 8.3-8.5.

### Recurso Educativo Abierto (REA) nº 1

Les recomiendo visualizar el siguiente video titulado “[Intervalos de confianza en Excel](#)”

¿Recordaban algunos conceptos del intervalo de confianza impartidos el año pasado? No se olviden de participar en el foro académico esta semana.



### Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Realizar alguno de los ejercicios prácticos resueltos en el texto básico recomendado.

- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.



## Semana 4

### 2.3. Intervalos de confianza de muestras grandes para una media y una proporción poblacional

### 2.4. Intervalos de confianza para la diferencia entre dos medias poblacionales y entre dos proporciones binomiales

Esta semana 4 vamos a finalizar la Unidad 2. Primero estudiaremos los intervalos de confianza en muestras grandes para una media y una proporción poblacional (Unidad 2 apartado 2.3). Luego estudiaremos intervalos de confianza para la diferencia entre dos medias poblacionales y entre dos proporciones binomiales (apartado 2.4). Por lo general, con muestras grandes ( $n > 30$ ) la estimación de estos parámetros se hace por aproximación a la normal.

#### Recursos de aprendizaje

##### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar el Capítulo 8, apartados 8.6-8.7.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Recurso Educativo Abierto (REA) nº 1

Les recomiendo visualizar el siguiente video titulado “[Intervalos contrastes de proporciones Excel](#)”

Para repasar contenidos recuerda realizar el ejercicio práctico sobre intervalos de confianza



### Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Realización de alguna de las actividades propuestas en el texto básico recomendado.
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.
- Realizar la autoevaluación 2 para repasar los conocimientos adquiridos en esta Unidad 2.



## Autoevaluación 2

- Seleccione y marque el literal que corresponde a la opción correcta.**
1. ¿Cuál es el mejor estimador puntual de la diferencia entre dos medias poblacionales  $\mu_1 - \mu_2$ ?
    - a. La desviación estándar de la muestra:  $s$
    - b. La puntuación:  $Z_a$
    - c. La diferencia entre dos medias muestrales:  $x_1 - x_2$
    - d. La diferencia entre dos proporciones muestrales:  $p^1 - p_2$
    - e. Ninguna respuesta es correcta
  2. El punto central de un intervalo de confianza para una media corresponde a:
    - a. El estadístico muestral
    - b. La media poblacional
    - c. Al tamaño de muestra
  3. Suponga que, al estimar la proporción poblacional, se obtiene un intervalo dado por:  $[0.10; 0.25]$  al 90% de confianza. Esto nos dice que:
    - a. Se espera que el 90% de las medias muestrales estén contenidas en el intervalo.
    - b. El intervalo es 100% que contenga a la proporción buscada
    - c. Se espera que el 90% de las proporciones muestrales estimadas estén contenidas en el intervalo.
    - d. Todas las respuestas son correctas

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

4. Escoger un nivel de confianza alto para construir un intervalo, implica:
  - a. Obtener un intervalo estrecho
  - b. Mayor amplitud en el intervalo
  - c. Que el intervalo sea más preciso
  - d. Ninguna respuesta es correcta
5. Para obtener un intervalo de confianza estrecho, es necesario:
  - a. Un tamaño de muestra pequeño
  - b. Con cualquier tamaño de muestra siempre será estrecho
  - c. Un tamaño de muestra grande
  - d. Ninguna respuesta es correcta
6. Sea  $[0.24; 0.57]$  un intervalo de confianza para la diferencia de proporciones. Se puede deducir que:
  - a. La proporción de la segunda población es igual al 57%
  - b. Las proporciones poblacionales son diferentes
  - c. La diferencia de proporciones es igual a:  $0.57 - 0.24 = 0.33$
  - d. Las proporciones poblacionales son iguales
  - e. Respuestas a y c son correctas

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

7. Una médica está realizando un estudio sobre el efecto del consumo de grasa no saturada en la presión sistólica de un grupo de pacientes con diabetes tipo II. El promedio de la presión obtenido de  $n = 40$  pacientes tomados aleatoriamente fue 167 mmHg con una desviación estándar de 28,3 mmHg. Use esta información para estimar el promedio de presión sistólica de todos los pacientes con diabetes II con un margen de error de 95%. Exprese su respuesta con 3 cifras decimales y use el símbolo ‘ $\pm$ ’ en su respuesta.
- a.  $28,3 \pm 51,754$
  - b.  $1,96 \pm 1255,692$
  - c.  $167 \pm 1,649$
  - d.  $167 \pm 8,770$
  - e. Ninguna respuesta es correcta

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

8. Se realizó un experimento para probar el efecto de un nuevo medicamento para el tratamiento del virus del Zika. La infección fue inducida en 100 ratones y éstos se dividieron al azar en dos grupos de 50. El primer grupo, el grupo de control, no recibió tratamiento para el Zika. El segundo grupo recibió el medicamento. Después de un periodo de 90 días, las proporciones de sobrevivientes en el grupo de control y en el grupo de medicamento se encontraron de 0,69 y 0,80, respectivamente.

La pregunta es: ¿Hay evidencia suficiente para indicar que el medicamento es efectivo para tratar el Zika? En vez de usar el método comparando el estadístico de prueba con el valor crítico, utilice el método de intervalo de confianza. Use un intervalo de confianza de **95%**.

¿Cuál es la hipótesis nula? Use el formato que usamos en clase (y Mendenhall et al. 2015).

- a.  $H_0: \mu_1 = \mu_2$
- b.  $H_0: p_1 = p_2$
- c.  $H_0: p_1 < p_2$
- d.  $H_0: (p_1 - p_2) = 0$
- e. Respuestas c y d son correctas
- f. Respuestas b y d son correctas

9. Usando el ejemplo de Pregunta 8, ¿cuál es la hipótesis alternativa? Use el formato que usamos en clase.

- a.  $H_a: \mu_1 = \mu_2$
- b.  $H_a: p_1 = p_2$
- c.  $H_a: p_1 < p_2$
- d.  $H_a: (p_1 - p_2) = 0$
- e. Respuestas c y d son correctas
- f. Respuestas b y d son correctas

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

10. Usando el ejemplo de Pregunta 8, resulta que los límites inferior y superior del intervalo de confianza son -0.279 y 0.059, respectivamente. ¿Hay evidencia suficiente para indicar que el medicamento es efectivo para tratar el Zika?
- a. Sí
  - b. No
  - c. No se puede contestar la pregunta basada en la información dada

Luego de haber respondido la autoevaluación, compare con el solucionario que se encuentra al final de la guía didáctica, y realice los correctivos si es necesario.

[Ir al solucionario](#)

## Resultado de aprendizaje 2

Plantea hipótesis para una o más muestras acerca de las características poblacionales y realiza su comprobación.

### Contenidos, recursos y actividades de aprendizaje

Esta semana comenzamos el segundo resultado de aprendizaje que se trabajará hasta el final del primer bimestre. A través de este resultado de aprendizaje se estudiará con la Unidad 3 sobre contrastes de hipótesis. El objetivo de este resultado de aprendizaje es que al final del primer bimestre cada estudiante sea capaz de: 1) formular una hipótesis nula  $H_0$  y su correspondiente hipótesis alternativa  $H_a$ , 2) contrastar la  $H_0$  con los resultados de la muestra, 3) calcular la probabilidad de encontrar el resultado (valor P) y 4) interpretar los resultados decidiendo si se rechaza o no la hipótesis nula.



#### Semana 5



## Unidad 3. Prueba de hipótesis

### 3.1. Procedimientos de una prueba estadística de hipótesis

Esta semana 5 vamos a iniciar la Unidad 3 sobre Prueba de hipótesis, específicamente con el apartado 3.1 sobre procedimientos de una prueba estadística de hipótesis. A modo de ejemplo, podríamos decir que el razonamiento empleado en una prueba estadística de hipótesis es similar al proceso de un tribunal:

#### Paralelismo entre contraste de hipótesis y un juicio a una persona

Observe que en este ejemplo, en el juicio no se demuestra que el demandado es inocente, sino sólo que no hubo evidencia suficiente para concluir que el demandado era culpable.

Antes de avanzar a los siguientes apartados, les recomiendo repasar a través de la siguiente infografía las diferentes partes que componen una prueba estadística de hipótesis:

#### [Prueba de hipótesis](#)

#### [Recursos de aprendizaje](#)

#### **Texto básico:**

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar el Capítulo 9, apartados 9.1 y 9.2.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Recurso Educativo Abierto (REA) nº 1

Les recomiendo visualizar el siguiente video titulado “[Contraste de hipótesis sobre la media de una población](#)”

¿Qué les ha parecido esta unidad? ¿Están entendiendo los contenidos? Solo recordar que esta semana tendremos la primera Evaluación Parcial con el contenido de las Unidades 1 y 2.



### Actividades de aprendizaje recomendadas

- Lectura del texto básico
- Revisar el REA nº1 sobre contraste de hipótesis
- Realización de alguna de las actividades propuestas en el texto básico recomendado
- Repasar todos los contenidos de las Unidades 1 y 2.
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.



## Semana 6

### 3.2. Una prueba de muestra grande acerca de una media poblacional

### 3.3. Una prueba de hipótesis para la diferencia entre dos medias poblacionales

En esta semana 6 continuaremos con las pruebas de hipótesis de una muestra grande acerca de una media poblacional y para la diferencia entre dos medias poblacionales (Unidad 3, apartados 3.2 y 3.3), además de estudiar el valor P (el nivel de significancia observado) y los tipos de errores que podemos cometer (Error Tipo I y Error Tipo II). La siguiente infografía nos ayuda a entender el significado de una muestra grande acerca de una media poblacional:

Ejemplo práctico de una prueba de hipótesis de muestra grande acerca de una media poblacional

[Ir a recursos](#)

No obstante les sugiero revisar el siguiente ejemplo para comprender mejor los conceptos del apartado 3.2:

Ejemplo práctico de una prueba de hipótesis de muestra grande acerca de una media poblacional

[Ir a recursos](#)

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar el Capítulo 9, apartados 9.3 y 9.4.

Para finalizar esta semana, les recuerdo que debéis de subir vuestras preguntas en el segundo Foro del bimestre 1.



### Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Realización de alguna de las actividades propuestas en el texto básico recomendado
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.



## Semana 7

### 3.4. Una prueba de hipótesis para una proporción binomial

### 3.5. Una prueba de hipótesis para la diferencia entre dos proporciones binomiales

Iniciamos la semana 7 del bimestre estudiando los dos últimos apartados de la Unidad 3 sobre prueba de hipótesis para una proporción binomial (apartado 3.4) y para la diferencia entre dos proporciones binomiales (apartado 3.5). Les recomiendo repasar bien los ejemplos propuestos en la siguiente infografía para entender mejor estos apartados:

Presentación unidad 3: Una prueba para la diferencia entre dos proporciones binomiales ( $p_1 - p_2$ )

[Ir a recursos](#)

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar el Capítulo 9, apartados 9.5 y 9.6.

Ya casi estamos al final del bimestre, sería recomendable repasar bien los contenidos impartidos hasta el momento en la Unidad 3 para poder realizar la segunda evaluación parcial del bimestre.



## Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Repasar los contenidos impartidos en la Unidad 3 para poder realizar la segunda evaluación parcial del bimestre
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.



## Semana 8

Estamos en la última semana del bimestre, la semana 8. Durante esta semana le recomendamos estudiar todo el contenido impartido en las Unidades 1, 2 y 3, así como realizar las diferentes autoevaluaciones de cada una de las unidades para repasar antes de la evaluación presencial del primer bimestre.

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar el Capítulo 9, apartado 9.7, además los Capítulos 1, 2, 6, 7, 8 y 9.

Por fin llegamos a la última semana del bimestre. Repase bien los contenidos y las autoevaluaciones de las Unidades 1, 2 y 3. ¡Mucha suerte en la evaluación presencial!

Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos



## Actividades de aprendizaje recomendadas

- Lectura y estudio del texto básico recomendado.
- Realizar las autoevaluaciones de las unidades 1, 2 y 3 para repasar los contenidos.
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente



## Autoevaluación 3

**Seleccione y marque el literal que corresponde a la opción correcta.**

1. En una prueba de hipótesis, el valor  $P$ :
  - a. Se conoce como el error tipo II
  - b. Toma valores en el intervalo  $[-1, +1]$
  - c. Se fija antes de realizar la prueba
  - d. Se compara con el nivel alfa
  - e. Todas las respuestas son correctas
  
2. El nivel de significancia alfa:
  - a. Siempre es mayor que el valor  $P$
  - b. Es la probabilidad de rechazar  $H_0$  cuando  $H_0$  es verdadera
  - c. Valores próximos a 1 son adecuados para realizar una prueba
  - d. Es la probabilidad de aceptar  $H_0$  cuando  $H_0$  es falsa
  - e. Ninguna respuesta es correcta

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

3. Cuando la hipótesis alternativa adopta la forma  $H_a: \mu_1 \neq \mu_2$ , se deberá efectuar una:
  - a. Prueba de una cola
  - b. Prueba de dos colas
  - c. Prueba Z
  - d. Ninguna respuesta es correcta
4. Cuando el valor  $P$  es mayor al valor de  $\alpha=0,05$  para una prueba de hipótesis de tipo  $H_0: \mu_1 = \mu_2$ , la decisión correcta es:
  - a. Rechazar la hipótesis
  - b. No rechazar la hipótesis
  - c. Reservarse en su decisión
  - d. No se debe comparar el valor P con  $\alpha$
5. En una prueba de hipótesis la probabilidad con la cual estamos dispuestos a arriesgar el rechazo de una hipótesis aun cuando esta es verdadera, está dado por:
  - a. El valor de Z
  - b. El valor de t
  - c. El nivel de significancia
  - d. Ninguna respuesta es correcta
6. En un contraste de hipótesis se comete un error de tipo I cuando la declaración respecto a la hipótesis nula es:
  - a. Aceptar una  $H_0$  que es verdadera
  - b. Rechazar una  $H_0$  que es verdadera
  - c. Aceptar una  $H_0$  que es falsa

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

7. En un estudio de médicos y médicas del Ecuador se encuentra que el promedio semanal de ganancias para médicos es \$675. La pregunta es: ¿los hombres de la misma posición tienen ganancias semanales promedio más altas que los de las mujeres? Una muestra aleatoria de  $n = 50$  médicos hombres mostró  $\bar{x} = \$715$  y  $s = \$98$ . Pruebe la hipótesis apropiada usando  $\alpha = 0,01$ , el cual corresponde a un valor crítico  $z = 2,33$ .

¿Cuál es la hipótesis nula? Use el formato que usamos en clase (i.e., Mendenhall *et al.* 2015).

- a.  $H_0: \mu_1 = \mu_2$
  - b.  $H_0: \mu = 715$
  - c.  $H_0: \mu = 675$
  - d.  $H_0: \mu > 675$
  - e. Respuestas b y c son correctas
  - f. Ninguna respuesta es correcta
8. Usando el ejemplo de Pregunta 7, ¿cuál es la hipótesis alternativa? Use el formato que usamos en clase.
- a.  $H_a: \mu_1 = \mu_2$
  - b.  $H_a: \mu = 715$
  - c.  $H_a: \mu = 675$
  - d.  $H_a: \mu > 675$
  - e. Respuestas b y c son correctas
  - f. Ninguna respuesta es correcta
9. Usando el ejemplo de Pregunta 7, ¿cuál es el estadístico de prueba?
- a.  $z = 2,89$
  - b.  $z = 2,33$
  - c.  $\alpha = 2,89$
  - d.  $\alpha = 2,33$
  - e.  $F = 715$

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

10. Usando el ejemplo de Pregunta 7, ¿cuál es su conclusión en términos de la pregunta planteada?
- a. Se rechaza la hipótesis nula y se concluye que las médicas ganan más de los médicos
  - b. No se rechaza la hipótesis nula y se concluye que las ganancias de los médicos y las médicas son estadísticamente equivalentes
  - c. Se rechaza la hipótesis nula y se concluye que las ganancias de los médicos y las médicas son estadísticamente equivalentes
  - d. Se rechaza la hipótesis nula y se concluye que los médicos ganan más de las médicas
  - e. No se puede contestar la pregunta basada en la información dada.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

11. Para determinar si la propiedad de una computadora afecta el rendimiento académico de un/a estudiante, se tomaron dos muestras aleatorias de 100 estudiantes. El promedio de calificaciones para los  $n_1 = 100$  que tienen computadora tuvieron un promedio y variancia igual a  $\bar{x}_1 = 27,0$  y  $s_1^2 = 5,6$ , en tanto que  $\bar{x}_2 = 26,4$  y  $s_2^2 = 4,0$  para los  $n_2 = 100$  que no tienen computadora. ¿Los datos presentan suficiente evidencia para indicar una diferencia en el rendimiento medio entre propietarios/as de computadoras y no propietarios/as? Pruebe usando  $\alpha = 0,05$ .

¿Cuál es la hipótesis nula? Use el formato que usamos en clase (i.e., Mendenhall et al. 2015).

- a.  $H_0: \mu_1 = \mu_2$
- b.  $H_0: p_1 = p_2$
- c.  $H_0: \mu_1 = 27,0$
- d.  $H_0: \mu_1 < \mu_2$
- e.  $H_0: \mu_1 \neq \mu_2$
- f. Ninguna respuesta es correcta

12. Usando el ejemplo de Pregunta 11, ¿cuál es la hipótesis alternativa? Use el formato que usamos en clase.

- a.  $H_a: \mu_1 = \mu_2$
- b.  $H_a: p_1 = p_2$
- c.  $H_a: \mu_1 = 27,0$
- d.  $H_a: \mu_1 < \mu_2$
- e.  $H_a: \mu_1 \neq \mu_2$
- f. Ninguna respuesta es correcta

Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos

13. Usando el ejemplo de Pregunta 11, ¿cuál es el estadístico de prueba apropiado?
- a.  $z = 2,33$
  - b.  $z = 1,94$
  - c.  $z = 1,96$
  - d.  $\alpha = 1,96$
  - e.  $\alpha = 0,05$
14. Usando el ejemplo de Pregunta 11, ¿cuál es su conclusión en términos de la pregunta planteada?
- a. Se rechaza la hipótesis nula y se concluye no hay una diferencia en el promedio de los rendimientos académicos para los estudiantes que tienen computadora y los/as que no la tienen.
  - b. Se rechaza la hipótesis nula y se concluye hay una diferencia en el promedio de los rendimientos académicos para los estudiantes que tienen computadora y los/as que no la tienen.
  - c. No se rechaza la hipótesis nula y se concluye no hay una diferencia en el promedio de los rendimientos académicos para los estudiantes que tienen computadora y los/as que no la tienen.
  - d. No se rechaza la hipótesis nula y se concluye hay una diferencia en el promedio de los rendimientos académicos para los estudiantes que tienen computadora y los/as que no la tienen.

Luego de haber respondido la autoevaluación, se le recomienda revisar el solucionario y comparar con sus respuestas. En las preguntas donde no haya acertado, se le sugiere volver a revisar el contenido respectivo.

[Ir al solucionario](#)

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos



## Actividades finales del bimestre

**Actividad 1:** Dedicar esta semana a terminar de estudiar los temas revisados durante este bimestre. Revise nuevamente sus apuntes, las autoevaluaciones y reforzar si es necesario con lecturas del texto base.

**Actividad 2:** Recuerde asistir a su centro universitario a rendir la prueba bimestral. En ésta, se analizan solamente los contenidos estudiados en el primer bimestre.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos



## Segundo bimestre

**Resultado de aprendizaje 3** | Caracteriza y aplica la técnica del análisis de varianza

### Contenidos, recursos y actividades de aprendizaje

A través de este tercer resultado de aprendizaje vamos a aprender a comparar dos medias de distribuciones independientes y normales. Usaremos la prueba  $t$  de Student cuando la variable independiente presenta dos niveles (Unidad 4) y el análisis de varianza (ANOVA) cuando presenta más de dos niveles (Unidad 5).



### Semana 9

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos



## Unidad 4. Prueba t de Student

**4.1. Prueba t para hacer inferencias de muestra pequeña respecto a una media poblacional**

**4.2. Prueba t para la diferencia entre dos medias: muestras independientes**

Bienvenido a la primera semana del segundo bimestre de Estadística Inferencial. En esta semana vamos a estudiar uno de los test estadísticos más utilizados: la prueba *t* de Student. Este estadístico se utiliza para determinar si hay diferencias significativas entre las medias de dos grupos, es decir, cuando queremos comparar dos medias de poblaciones independientes y de distribución normal que presenta como máximo dos niveles. A continuación puedes revisar las principales características de la prueba *t* de Student:

Presentación unidad 4: Pasos para seguir para ajustar un modelo lineal (Prueba *t*, ANOVA, regresión)

[Ir a recursos](#)

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar Capítulo 10, apartados 10.1, 10.2, 10.3 y 10.4.

¡Comenzamos el segundo bimestre! Les recomiendo realizar un resumen con las principales características de la prueba  $t$ .



### Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Realización de alguna de las actividades propuestas en el texto básico recomendado
- Elaborar un resumen con las principales características de la prueba  $t$ .
- Realizar la autoevaluación 4 para repasar los conocimientos adquiridos en esta Unidad 4
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.



## Autoevaluación 4

**Seleccione y marque el literal que corresponde a la opción correcta.**

1. Se aplica una prueba de  $t$  cuando la variable independiente es:
  - a. Continua
  - b. Categórica y hay sólo dos categorías (i.e., niveles)
  - c. Categórica y hay más de dos categorías (i.e., niveles)
  - d. La variable independiente puede ser categórica o continua
  - e. Ninguna respuesta es correcta
  
2. La forma de la distribución  $t$  depende de:
  - a. El nivel de significancia  $\alpha$
  - b. Los grados de libertad  $df$
  - c. El tamaño muestral  $n$
  - d. Respuestas a y b son correctas
  - e. Respuestas b y c son correctas
  - f. Todas las respuestas son correctas

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

3. Un diseño de muestras independientes para la diferencia entre dos medias poblacionales requiere que:
  - a. Las características de los individuos entre las dos muestras sean iguales.
  - b. Las dos muestras tengan la misma cantidad de observaciones.
  - c. Los individuos que constituye una muestra son completamente distintos de los de la otra muestra.
  - d. Las varianzas de las muestras son iguales
  - e. Respuestas a y b son correctas
  - f. Respuestas b y c son correctas
  - g. Respuestas c y d son correctas
4. Un agricultor selecciona 10 frutales al azar y registra la cantidad de frutos producidos por individuo. Luego les aplica un abono orgánico para probar su efectividad en la producción de frutos y vuelve a medir la cantidad de frutos producidos por cada individuo. Estamos frente a un ejemplo de:
  - a. Muestras pareadas
  - b. Muestras independientes
  - c. Muestras múltiples
  - d. Ninguna respuesta es correcta
5. La prueba  $t$  para dos grupos independientes compara:
  - a. Las varianzas
  - b. Las desviaciones estándar
  - c. Las medias

Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos

6. Los grados de libertad para una prueba  $t$  de muestras independientes es:
- $n_1 + n_2$
  - $n_1 + n_2 + 2$
  - $n_1 + n_2 - 2$
  - $1/n_1 + 1/n_2$
7. ¿Qué suposiciones se hacen cuando se usa la prueba  $t$  para probar una hipótesis respecto a dos medias poblacionales?
- Las muestras deber ser aleatorias e independientes
  - Las poblaciones de la que haga muestreo deben estar normalmente distribuidas
  - Las poblaciones de la que haga muestreo deben tener la misma varianza:  $\sigma_1 = \sigma_2$
  - Respuestas a y c son correctas
  - Todas las respuestas son correctas

[Índice](#)[Primer bimestre](#)[Segundo bimestre](#)[Solucionario](#)[Referencias bibliográficas](#)[Recursos](#)

8. Se realizó un análisis de la concentración de glucosa en plasma, a 10 pacientes diabéticos y a 10 pacientes no diabéticos. Con estos datos ¿se puede asegurar que la concentración media de glucosa es mayor en los pacientes diabéticos que en los no diabéticos?

Análisis										
Diabéticos	95	87	105	90	108	119	99	76	80	98
No diabéticos	78	85	88	82	85	88	93	90	75	85

Probar usando la función **T.TEST()** en Excel con el nivel de significancia  $\alpha = 0.05$ . ¿Cuál es la hipótesis nula? Use el formato que usamos en clase (y Mendenhall *et al.* 2015).

- a.  $H_0: \mu_1 = \mu_2$   
b.  $H_0: \mu_1 > \mu_2$   
c.  $H_0: \mu_1 < \mu_2$   
d.  $H_0: \mu_1 - \mu_2 > 0$   
e. Respuestas b y d son correctas
9. Usando el ejemplo de Pregunta 9, ¿cuál es la hipótesis alternativa? Use el formato que usamos en clase.
- a.  $H_a: \mu_1 = \mu_2$   
b.  $H_a: \mu_1 > \mu_2$   
c.  $H_a: \mu_1 < \mu_2$   
d.  $H_a: \mu_1 - \mu_2 > 0$   
e. Respuestas b y d son correctas

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

10. Usando el ejemplo de Pregunta 9, ¿cuál es estadístico de prueba observado?
  - a.  $t = 0,0164$
  - b.  $t = 0,0328$
  - c.  $t = 0,05$
  - d.  $t = 2,4116$
  - e.  $t = 12,007$
  
11. Usando el ejemplo de Pregunta 9, ¿cuál es el valor  $P$ ?
  - a.  $P = 0,0164$
  - b.  $P = 0,0328$
  - c.  $P = 0,05$
  - d.  $P = 2,4116$
  - e.  $P = 12,007$
  
12. Usando el ejemplo de Pregunta 9, se debe rechazar la hipótesis nula?
  - a. Sí
  - b. No
  - c. No se puede comprobar la hipótesis porque la muestra ( $n$ ) es demasiada pequeña.

Al finalizar la autoevaluación, se le recomienda revisar el solucionario y comparar con sus respuestas. En las preguntas donde no haya acertado, se le sugiere volver a revisar el contenido respectivo.

[Ir al solucionario](#)

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos



## Semana 10



### Unidad 5. Análisis de varianza (ANOVA)

#### 5.1. Introducción a los modelos lineales

#### 5.2. ANOVA de un factor

En esta semana, vamos a iniciar una de las unidades más importantes del semestre sobre análisis de varianza (ANOVA), específicamente los apartados 5.1 y 5.2 de introducción y de ANOVA de un factor. El ANOVA sirve para comparar varios grupos en una variable cuantitativa, en otras palabras, sería como una generalización de la prueba t para dos muestras a diseños con más de dos muestras. A la variable categórica (nominal u ordinal) que define los grupos que deseamos comparar la llamamos variable independiente y a la variable cuantitativa (de intervalo o razón) en la que deseamos comparar los grupos la llamamos dependiente.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar Capítulo 11, apartados 11.1, 11.2, 11.3, 11.4 y 11.5.

### Recurso Educativo Abierto (REA) nº 3 y 4

Para profundizar su conocimiento sobre los modelos lineales y ANOVA, les recomiendo revisar el REA nº 3 (Capítulo 2. Modelos lineales apartados 1, 2 y 4): [enlace web](#)

Para preparar su participación en el foro académico 1, les sugiero que revise este video interesante del REA nº 4 sobre cambio climático y género:

[En el marco de la #Cop25 OIJ organiza: Cambio climático y género.](#)



### Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Realización de alguna de las actividades propuestas en el texto básico
- Elaborar un resumen las características de los modelos lineales y las suposiciones del ANOVA

- Revisar el REA nº 3 sobre modelos lineales y ANOVA (Capítulo 2. Modelos lineales, apartados 1, 2 y 4)
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.

Para facilitar el estudio de esta unidad, les recomiendo realizar un resumen de las características de los modelos lineales y las suposiciones del ANOVA.



## Semana 11

### 5.3. Pruebas de comparaciones múltiples (post hoc)

A lo largo de la semana 11, vamos a finalizar la Unidad 5 (apartado 5.3). Sin embargo, esta vez estudiaremos las pruebas de comparaciones múltiples o también conocidas como pruebas “post hoc”. Estas pruebas se utilizan para comparar entre los niveles de un factor cuando F resulta significativo. Entre las más conocidas están Prueba de Tukey, Scheffé, Duncan y Bonferroni que serán las que estudiemos durante esta semana.

#### Recursos de aprendizaje

##### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar Capítulo 11, apartados 11.1, 11.2, 11.3, 11.4 y 11.5.



## Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Realización de alguna de las actividades propuestas en el texto básico
- Elaborar un resumen sobre las diferentes pruebas de comparaciones múltiples y especifica cuando se deben usar cada una de ellas
- Mida su nivel de conocimientos con la autoevaluación 5 que está al final de la Unidad 5
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.

Para facilitar el estudio de esta unidad, lo animamos a elaborar un resumen sobre las diferentes pruebas de comparaciones múltiples y especifica cuando se deben usar cada una de ellas. También les recuerdo que esta semana deben participar en el primer foro académico calificado del bimestre 2.



## Autoevaluación 5

**Seleccione y marque el literal que corresponde a la opción correcta.**

1. Una suposición necesaria para aplicar un análisis de varianza (ANOVA) es:
  - a. Las varianzas de cada población deben ser iguales
  - b. Las muestras deben ser dependientes entre sí
  - c. Las poblaciones de donde provienen las muestras deben tener una distribución binomial
  - d. Los promedios deben ser iguales
  - e. Todas las respuestas son correctas
  
2. El ANOVA nos permite contrastar si existe o no igualdad entre:
  - a. Las varianzas de múltiples muestras
  - b. Las desviaciones típicas de varias muestras
  - c. Los tamaños muestrales de varias muestras
  - d. Las medias de múltiples muestras
  
3. Para realizar ciertas pruebas de significancia el ANOVA de un factor utiliza para este fin los estadísticos:
  - a.  $z$  obtenida versus  $z$  crítica
  - b.  $t$  obtenida versus  $t$  crítica
  - c.  $\alpha$  obtenida versus  $\alpha$  crítica
  - d.  $F$  obtenida versus  $F$  crítica

4. Un ANOVA de un factor se refiere a:
- Sóla una variable independiente
  - Sóla una variable dependiente
  - La clasificación en una sola dirección
  - Sólo un nivel ( $k=1$ ) o categoría
  - Respuestas a y c son correctas
  - Respuestas b y d son correctas
5. Para la prueba  $F$  los grados de libertad de los tratamientos (i.e., niveles) en el numerador se calculan en base a:
- El número de tratamientos que se desean comparar menos uno:  $df = (k - 1)$
  - La suma de las observaciones de todos los tratamientos menos el número de tratamientos:  $df = (n - k)$
  - El nivel con mayor número de observaciones
  - Ninguna respuesta es correcta
6. Para la prueba  $F$  los grados de libertad del error en el denominador se calculan en base a:
- El número de tratamientos que se desean comparar menos uno:  $df = (k - 1)$
  - La suma de las observaciones de todos los tratamientos menos el número de tratamientos:  $df = (n - k)$
  - El nivel con mayor número de observaciones
  - Ninguna respuesta es correcta

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

7. Una ecóloga selecciona una muestra de 20, 22 y 19 árboles en tres diferentes niveles altitudinales para determinar su área foliar. Para determinar si el área foliar media varía respecto al nivel altitudinal, ¿cuál estadístico de prueba debe utilizar?
- a. Prueba  $z$
  - b. Prueba  $t$
  - c. Prueba Wilcoxon
  - d. Coeficiente de correlación  $r$
  - e. Prueba  $F$
8. Se usa una prueba de comparaciones múltiples (*post hoc*) cuando:
- a. Una vez que haya determinado si hay en realidad una diferencia significativa en las medias de los  $k$  niveles
  - b. No hay diferencia significativa en las medias de los  $k$  niveles
  - c. El valor  $P$  general es menor que el nivel de significancia  $\alpha$
  - d. Las  $k$  muestras no son independientes
  - e. Respuestas b y c son correctas
  - f. Respuestas a y c son correctas
  - g. Respuestas c y d son correctas

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

9. **Neonicotinoides y abejas.** Existe una preocupación mundial ante el riesgo que supone para las abejas el uso de neonicotinoides, una familia de insecticidas que actúan en el sistema nervioso central de los insectos y, con menor toxicidad, en vertebrados (aves y mamíferos). Por ejemplo, hay evidencia de una posible conexión entre neonicotinoides y el desorden del colapso de colonias apícolas, lo cual tiene graves repercusiones para la polinización de muchos cultivos y plantas nativas.

En un estudio sobre el efecto de neonicotinoides en el abejorro (*Bombus terrestris*), se comparó el aumento en el peso (en gramos) de colonias de abejas alimentadas con: (1) polen tratado con una dosis alta de neonicotinoide, (2) polen tratado con una dosis media de neonicotinoide y (3) polen no tratado (control).

Utilizando los resultados indicados en las siguientes preguntas, verifica si la hipótesis de que el peso de la colonia es significativamente diferente entre los tres tratamientos es verdadera. Use  $\alpha = 0,05$ .

¿Qué tipo de variable es la variable independiente?

- a. Categórica
- b. Continua
- c. Discreta
- d. Ordinal
- e. Respuestas a y c son correctas
- f. Respuestas a y d son correctas
- g. Respuestas b y d son correctas

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

10. Usando el ejemplo de Pregunta 9, ¿cuál es la hipótesis nula?
- a.  $H_0: \mu_1 = \mu_2 = \mu_3$
  - b.  $H_0: p_1 = p_2 = p_3$
  - c.  $H_0: \mu_1 = 0,05$
  - d.  $H_0: \mu_1 < \mu_2 < \mu_3$
  - e.  $H_0: \mu_1 = \mu_2$
  - f. Ninguna respuesta es correcta
11. Usando el ejemplo de Pregunta 9, resulta que el valor P de un ANOVA de los datos es  $5.212 \times 10^{-7}$ . ¿Se debe rechazar la hipótesis nula?
- a. Sí
  - b. No
  - c. No se puede comprobar la hipótesis.
12. Usando una prueba de múltiples comparaciones que se llama Bonferroni, el valor P para la diferencia entre el tratamiento de dosis media de neonicotinoide y el control es  $4.8 \times 10^{-6}$ . ¿Qué se concluye de este resultado?
- a. Hay una diferencia significativa entre el tratamiento de dosis alta de neonicotinoide y el control.
  - b. Hay una diferencia significativa entre el tratamiento de dosis media de neonicotinoide y el control.
  - c. No hay una diferencia significativa entre el tratamiento de dosis media de neonicotinoide y el control.
  - d. Hay diferencias significativas entre todos los tratamientos.

Al finalizar la autoevaluación, se le recomienda revisar el solucionario y comparar con sus respuestas. En las preguntas donde no haya acertado, se le sugiere volver a revisar el contenido respectivo.

[Ir al solucionario](#)

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Resultado de aprendizaje 4

Aplica análisis de correlación y regresión simple

### Contenidos, recursos y actividades de aprendizaje

Hasta el momento hemos estudiado las relaciones existentes para una variable continua, ahora a través de este cuarto resultado de aprendizaje vamos a aprender a estimar una forma específica de relación entre dos variables continuas; en particular, el grado en el cual están relacionadas linealmente, mediante la regresión lineal simple y correlación (Unidad 6).



Semana 12



## Unidad 6. Regresión línea simple y correlación

### 6.1. Regresión lineal

Cuando queremos analizar datos, es importante centrarse en dos grandes objetivos: comparar grupos y estudiar relaciones. Por ello, en esta Unidad 6, vamos a estudiar la regresión lineal y

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

la correlación que nos permiten encontrar si dos variables están relacionadas o no, y expresar esta relación a través de una ecuación. Al igual que en unidades anteriores, el objetivo de esta unidad es proporcionar los fundamentos del análisis de regresión, sin entrar en aspectos técnicos o teóricos del análisis. Aprenderemos cuándo y cómo utilizar el análisis de regresión lineal y cómo interpretar los resultados.

En esta semana 12 empezamos con el apartado 6.1 sobre regresión lineal simple. El análisis de regresión lineal nos permite estudiar la relación entre una variable llamada dependiente (Y) y una o más variables llamadas independientes o predictoras (X), así como para desarrollar una ecuación lineal con fines predictivos.

Presentación unidad 6: Regresión lineal simple

[Ir a recursos](#)

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar Capítulo 12, apartados 12.1, 12.2, 12.3, 12.4 y 12.5.

### Texto complementario

Martínez Bencardino, C. (2015). *Estadística y Muestreo*. Ecoe Ediciones Ltda. Revisar Capítulo 8.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Recurso Educativo Abierto (REA) nº 3

Para profundizar su conocimiento sobre regresión simple, les recomiendo revisar el REA nº 3 (Capítulo 2. Modelos lineales apartado 3): [enlace web](#)



### Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Lectura del texto complementario recomendado
- Revise el REA nº3 sobre regresión simple (Capítulo 2. Modelos lineales, apartado 3)
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.

Durante esta semana, no se olviden de realizar la primera evaluación parcial (Unidades 4 y 5) y de realizar el ejercicio práctico sobre resolución de problemas 1 para reforzar los contenidos impartidos durante las semanas anteriores.



## Semana 13

### 6.2. Correlación bivariada

En la anterior semana, vimos la regresión lineal y cómo es posible estudiarla a través de un diagrama de dispersión. Sin embargo, a veces, es posible ajustar una línea recta mejor a unas nubes de puntos que a otras. Por ello, y con el objetivo de cuantificar ese grado de ajuste con mayor precisión utilizamos los coeficientes de correlación (apartado 6.2). Estos coeficientes nos permiten cuantificar el grado de relación lineal existente entre dos variables cuantitativas, al mismo tiempo que ayudan a valorar el grado de ajuste de la nube de puntos a una línea recta.

#### Recursos de aprendizaje

##### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar Capítulo 12, apartados 12.6, 12.7 y 12.8.

##### Texto complementario

Martínez Bencardino, C. (2015). *Estadística y Muestreo*. Ecoe Ediciones Ltda. Revisar Capítulo 8.



## Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Lectura del texto complementario recomendado
- Realización de alguna de las actividades propuestas en el texto básico
- Resuma los contenidos sobre las suposiciones de regresión, la predicción usando la recta ajustada y el análisis de correlación
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.
- Realizar la autoevaluación 6 para repasar los conocimientos adquiridos en esta Unidad 6

¿Qué les ha parecido el análisis de regresión lineal? ¿Han entendido bien el contenido? Les animo a realizar un pequeño resumen sobre las suposiciones de regresión, la recta ajustada y el análisis de correlación les va a ayudar mucho a entender mejor esta unidad.



## Autoevaluación 6

**A. Escriba entre los paréntesis V si el enunciado es verdadero o F si es falso.**

1. ( ) La recta de mejor ajuste se refiere a la que reduce al mínimo la suma de cuadrados de las desviaciones de los valores observados desde los pronosticados.
2. ( ) En la fórmula de la recta de mejor ajuste ( $y = \beta_0 + \beta_1 x_1$ ) la pendiente de la recta es  $x_1$ .
3. ( ) El coeficiente de correlación  $r$  es el cuadrado del coeficiente de determinación  $R^2$ .
4. ( ) Cuando  $R^2=0$ , la pendiente  $\beta_1=0$  y no hay relación lineal entre  $x$  e  $y$ .

**B. Encierre el literal que corresponde a la opción correcta**

5. En un conjunto de datos cuando  $y$  disminuye conforme  $x$  aumenta, el valor de coeficiente de correlación  $r$  es:
  - a. Menor que cero
  - b. Mayor que cero
  - c. Igual a cero
  - d. Igual a uno

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

6. ¿En cuál de los siguientes casos es el análisis más apropiado una regresión?
- El sexo y la edad de una población de colibríes en Parque Nacional Podocarpus (PNP)
  - La forma del pico y la longitud corporal de colibríes en PNP
  - La densidad de colibríes con altitud en PNP
  - La concentración sanguínea de corticosteroide en colibríes y altitud en PNP
  - Respuestas b y c son correctas
  - Respuestas c y d son correctas
  - Ninguna respuesta es correcta
7. El modelo de regresión lineal simple  $y = \beta_0 + \beta_1 x_1$  expresa lo siguiente:
- El incremento de  $x$  por cada unidad de incremento de  $y$
  - El incremento de  $y$  por cada unidad de incremento de  $x$
  - El incremento de  $\beta_0$  por cada unidad de incremento de  $\beta_1$
  - El incremento de  $\beta_1$  por cada unidad de incremento de  $x$
  - El valor de  $y$  cuando  $\beta_0$  y  $\beta_1$  valen 1.

8. Luego de ajustar un modelo lineal simple entre el número de automóviles y el contenido de plomo en el aire, se determina que el coeficiente de determinación  $R^2=0,89$ . Esto significa que:
- Existe una relación lineal fuerte entre las dos variables
  - Aparentemente existe una relación lineal entre las variables pero es débil
  - El modelo explica el 0,89% de la variabilidad de los datos
  - El modelo explica el 89% de la variabilidad de los datos
  - No existe relación entre las variables
  - Respuestas a y d son correctas
  - Respuestas b y c son correctas
9. Se determina que la cobertura de suelo y el riesgo de derrumbes posee un coeficiente de correlación de  $r=0,70$ , lo cual indica que:
- El coeficiente de correlación explica el 70% de la varianza
  - El riesgo de derrumbes se incrementa un 70%
  - El riesgo de derrumbes se disminuye un 70%
  - Las variables están muy relacionadas
10. El modelo:  $y = 0.36x + 2.1$ , expresa la relación entre la incidencia de dengue ( $y$ ) y la precipitación mensual ( $x$ ) en mm. ¿Cuál de las siguientes declaraciones es correcta?
- Se observa una relación débil negativa entre  $x$  e  $y$
  - El modelo explica 36% de la variabilidad de los datos
  - Para una precipitación mensual de 100 mm la incidencia de dengue es 38,1
  - Para una precipitación mensual de 100 mm la incidencia de dengue es  $0,36/2,1$
  - Ninguna respuesta es correcta

11. ¿Cuáles son los parámetros de un modelo lineal simple que se pretenden estimar con un análisis de regresión?
- a. La variable dependiente
  - b. La pendiente
  - c. Los residuos
  - d. Los grados de libertad
  - e. La intersección
  - f. Respuestas a y b son correctas
  - g. Respuestas b y e son correctas
  - h. Respuestas d y e son correctas

[Ir al solucionario](#)

Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Resultado de aprendizaje 5

Conoce y utiliza pruebas no paramétricas

### Contenidos, recursos y actividades de aprendizaje

En ocasiones, los datos que vamos a estudiar pueden no seguir una distribución normal y por ello no sería correcto utilizar pruebas paramétricas como el ANOVA o la prueba t. Cuando unos datos cuantitativos, ordinales o nominales presentan una distribución libre (no necesariamente normal) utilizamos los llamados análisis estadísticos no paramétricos (Unidad 7) que estudiaremos a través de este resultado de aprendizaje 5. Este tipo de análisis se suelen utilizar cuando las muestras son pequeñas.



Semana 14



## Unidad 7. Pruebas no paramétricas

### 7.1. Tablas de contingencias: una clasificación de dos vías

### 7.2. Prueba chi cuadrado: evaluación de la asociación entre dos variables categóricas

Durante esta semana vamos a estudiar las tablas de contingencia (apartado 7.1) y la prueba chi cuadrado (apartado 7.2). Es posible que los datos obtenidos en una investigación van a corresponder a dos variables categóricas (por ejemplo, gremios alimenticios de aves en diferentes ecosistemas) para poder explorar estos datos utilizaremos una tabla de contingencia, que no es más que una tabla de dos dimensiones y cada una de estas dimensiones contiene una variable. A su vez es posible que cada variable se subdivida en dos o más categorías. Una vez realizada la tabla de contingencia, podemos utilizar la prueba chi cuadrado para evaluar la hipótesis acerca de la relación entre estas dos variables categóricas.

Presentación unidad 7: Prueba Chi cuadrado

[Ir a recursos](#)

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar Capítulo 14, apartados 14.1, 14.2 y 14.4.

### Recurso Educativo Abierto (REA) nº 1

Les recomiendo visualizar el siguiente video titulado “[Contrastes no paramétricos de independencia](#)”



### Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Revise el REA nº 1
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.

¿Ha entendido bien en qué supuestos se pueden usar las pruebas no paramétricas? Seguro que sí. También, recuerde que se puede realizar el ejercicio práctico de resolución de problemas 2 para repasar el contenido de la Unidad 6.



## Semana 15

### 7.3. Prueba de suma de rango de Wilcoxon (muestras independientes) y de rango con signo de Wilcoxon (un experimento pareado)

En esta semana 15 aprenderemos una de las pruebas no paramétricas más usadas y conocidas como la prueba de suma de rango de Wilcoxon (apartado 7.3). Pero, ¿cuándo debemos utilizar esta prueba? Pues bien, esta prueba se utilizará en lugar de una prueba t, en los casos que los datos no presentan una distribución normal.

Presentación unidad 7: Pruebas paramétricas y su contraparte no paramétrica

[Ir a recursos](#)

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar Capítulo 15, apartados 15.1, 15.2 y 15.5.



## Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Realice alguna de las actividades del texto básico recomendado para esta semana
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.

¿Sabe explicar las diferencias entre una prueba chi cuadrado y un test de Wilcoxon? ¿En qué supuestos se debe utilizar el test de Wilcoxon? Repase bien el contenido teórico impartido hasta el momento y plantee algunas preguntas a sus compañeros en el Foro 2 evaluado.



## Semana 16

### 7.4. Prueba de análisis de varianza de Kruskal-Wallis

Enhorabuena, ya hemos llegado a la última semana del curso, la semana 16 donde estudiaremos la prueba de Kruskal-Wallis (apartado 7.4) que es una valiosa alternativa para un análisis de varianza de una vía (ANOVA) cuando es violada la suposición de normalidad. Además, le animo a repasar el contenido de las Unidades 4, 5, 6 y 7 para preparar mejor la evaluación presencial del segundo bimestre.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Recursos de aprendizaje

### Texto básico:

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la Probabilidad y Estadística*. Bogotá, Colombia: CENGAGE Learning. Revisar Capítulo 15, apartado 15.6.



### Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado
- Realice la autoevaluación 7 para repasar el contenido es esta Unidad 7. También sería recomendable repasar todas las autoevaluaciones de este segundo bimestre para reforzar los contenidos impartidos antes del examen presencial
- Interacción en el Entorno Virtual de Aprendizaje (EVA), chat de tutoría y consulta o teléfono para preguntar sus dudas a la docente.

¡Enhorabuena! Ya han llegado al final del semestre. Le recomiendo repasar bien los contenidos y las autoevaluaciones de las Unidades 4, 5, 6 y 7. ¡Mucho suerte en la evaluación presencial!

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos



## Autoevaluación 7

**Seleccione y marque el literal que corresponde a la opción correcta.**

1. Se usa la prueba chi cuadrado  $\chi^2$  para:
  - a. Probar la independencia de dos variables categóricas
  - b. Probar la asociación entre dos variables categóricas
  - c. Medir la diferencia entre frecuencias observadas y frecuencias esperadas
  - d. Comprobar la  $H_0$  al comparar el estadístico  $\chi^2$  de la muestra con el valor crítico  $\chi^2$
  - e. Todas las respuestas son correctas

2. Se dispone de 300 ratones de laboratorio y se decide tratar a 200 con una vacuna experimental contra la influenza porcina H1N1 y dejar 100 como controles. Luego de tratarlos se expone a los 300 al contagio de la influenza porcina. Después de un período experimental adecuado, el recuento fue:

	Enfermos	Sanos	Total
Tratados	56	144	200
No Tratados	71	29	100
Total	127	173	300

El tratado con la vacuna es una variable de tipo \_\_\_\_\_ y la condición enfermo o sano es una variable de tipo \_\_\_\_\_.

- a. nominal, continua
  - b. categórica, categórica
  - c. continua, ordinal
  - d. cuantitativa, continua
  - e. dependiente, independiente
3. Usando el ejemplo de Pregunta 2, ¿qué tipo de contraste se puede realizar?
- a. Prueba *t* de Student
  - b. Regresión simple
  - c. Prueba de suma de rangos de Wilcoxon
  - d. Prueba chi cuadrado
  - e. Prueba Kruskal-Wallis

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

4. Usando el ejemplo de Pregunta 2, ¿cuántos grados de libertad hay?
  - a. 0
  - b. 1
  - c. 2
  - d. 3
  - e. 4
  - f. 9
  
5. Aplicando la función **CHISQ.TEST()** en Excel al ejemplo de Pregunta 2, ¿cuál es el valor  $X^2$  para la prueba?
  - a. 0,05
  - b. 3,841
  - c. 5,991
  - d. 7,814
  - e. 48,748
  
6. Según el nivel de alfa dado y los grados de libertad del ejemplo de Pregunta 2, ¿cuál es el valor crítico de  $X^2$  que se debería comparar con el estadístico de prueba de  $X^2$  (usando el método que usábamos antes de tener smartphones y comíamos carne de dinosaurios)? Pista: Ver Tabla 5 de Apéndice I de Mendenhall et al. 2015.
  - a. 0,05
  - b. 3,841
  - c. 5,991
  - d. 7,814
  - e. 48,748

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

7. Aplicando la función **CHISQ.TEST()** en Excel al ejemplo de Pregunta 2, ¿cuál es el valor P para la prueba?
- a. 0,05
  - b. 3,841
  - c. 5,991
  - d.  $2,911 \times 10^{-12}$
  - e. 2,911
8. Usando el ejemplo de Pregunta 2, ¿los datos dan suficiente evidencia para indicar que la vacuna experimental es efectiva para tratar la influenza porcina? En otra palabra, hay una diferencia significativa en el patrón del número de ratones enfermos y sanos, dependiendo del tratado de la vacuna? Pruebe usando  $\alpha = 0.05$ .
- a. Sí
  - b. No
  - c. No se puede comprobar esta hipótesis.
9. Si el/la investigador/a no está seguro de que las suposiciones requeridas para una prueba t de dos muestras sea satisfecha, una alternativa es sustituir los valores de las observaciones por sus rangos y proceder como si éstos fueran las observaciones reales. Una prueba no paramétrica que usa un estadístico de prueba basado en estos rangos de muestra se llama:
- a. Prueba de suma de rango de Wilcoxon
  - b. Prueba chi cuadrado
  - c. Prueba de t
  - d. Prueba U de Mann-Whitney
  - e. Respuestas a y b son correctas
  - f. Respuestas a y d son correctas

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

10. **Neonicotinoides y abejas** Usando el ejemplo de Pregunta 9 de la Autoevaluación 5, ¿cuál sería el análisis más apropiado en el caso que no se cumpliera la suposición de normalidad de los residuos?
- a. Prueba t de Student
  - b. Regresión simple
  - c. Prueba de suma de rangos de Wilcoxon
  - d. Prueba chi cuadrado
  - e. Prueba Kruskal-Wallis
11. En referencia al ejemplo de Pregunta 9 de la Autoevaluación 5, si en vez de comparar el peso de colonias de abejas entre los tres tratamientos de neonicotinoides, se relacionará el peso de colonias con distintos valores de dosis de neonicotinoides, ¿cuál sería el análisis más apropiado?
- a. Prueba *t* de Student
  - b. Regresión simple
  - c. Prueba de suma de rangos de Wilcoxon
  - d. Prueba chi cuadrado
  - e. Prueba Kruskal-Wallis

Al finalizar la autoevaluación, se le recomienda revisar el solucionario y comparar con sus respuestas. En las preguntas donde no haya acertado, se le sugiere volver a revisar el contenido respectivo.

[Ir al solucionario](#)

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos



## Actividades finales del bimestre

**Actividad 1:** Dedique esta semana a estudiar y a repasar los temas revisados durante este bimestre.

**Actividad 2:** Si aún no lo ha hecho, realice todas las autoevaluaciones para repasar los contenidos impartidos este segundo bimestre.

**Actividad 3:** Recuerde asistir a su centro universitario a rendir la prueba bimestral. En ésta, se analizan solamente los contenidos estudiados en el segundo bimestre



## 4. Solucionario

A continuación, usted dispone de la solución para cada una de las autoevaluaciones planteadas a lo largo de esta guía didáctica.

Autoevaluación 1		
Pregunta	Respuesta	Retroalimentación
1	f	Cuando se realiza exploración de datos, se toma en cuenta los diversos tipos de variables, no exclusivamente las categóricas.
2	v	Inferir se relaciona con el método inductivo, es decir, partir de lo particular hacia lo general. En el caso de la estadística sería partir de la muestra hacia la población.
3	f	La distribución de datos numéricos tiende más hacia la normalidad mientras el tamaño de muestra incrementa.
4	v	Esta es precisamente la definición de la distribución muestral de un estadístico. La distribución muestral es el resultado de un muestreo aleatorio de la población.
5	d	La Estadística tiene todos los propósitos de la lista. La <b>estadística descriptiva</b> está formada por procedimientos empleados para resumir y describir las características importantes de un conjunto de mediciones. La <b>estadística inferencial</b> está formada por procedimientos empleados para hacer inferencias acerca de características poblacionales (i.e., parámetros), a partir de información contenida en una muestra sacada de esta población.
6	b	Variable por definición es aquella que cambia entre individuos (o elementos) de un conjunto, muestra o población de estudio.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

Autoevaluación 1		
Pregunta	Respuesta	Retroalimentación
7	b	Para el diseño de un estudio científico de cualquier disciplina, es más eficiente estudiar una muestra (un subconjunto de datos) siempre y cuando es representativa de la población.
8	c	Un estadístico es una medida numérica descriptiva (p.e., la media, la mediana, la varianza, etc.) que se calculen de la muestra.
9	e	La tasa anual de secuestro de carbono por superficie (tonelada/ha) puede expresarse en decimales (fracciones), por tanto es continua.
10	d	El diagrama de cajas nos proporciona dos tipos básicos de información visual: variación y tendencia central de los datos.
11	a	El <b>pésimo percentil</b> es el valor de $x$ que es mayor a $p\%$ de las mediciones y es menor que el restante ( $100 - p\%$ ), cuando las mediciones están ordenadas de su valor mínimo a su valor máximo. En este caso, el percentil 95 representa el dap que es mayor al 95% y menor del 5% de las mediciones.
12	b	La distribución de la media muestral construida a partir de datos re-muestreados, tienden hacia el valor central, de manera que los promedios de las muestras van a generar un rango más estrecho que los datos originales.
13	b	El error estándar de la media muestral, que es un estadístico de variación más ajustado cuando se trabaja con muestras, se construye dividiendo la desviación estándar para la raíz cuadrada de $n$ .
14	a	La distribución binomial (aquella que se ajusta a proporciones) se aproxima a una normal para muestras grandes.
15	c	La estadística inferencial está formada por procedimientos empleados para hacer inferencias acerca de características poblacionales (i.e., parámetros) y contrastar hipótesis.

Ir a la  
autoevaluación

[Índice](#)[Primer bimestre](#)[Segundo bimestre](#)[Solucionario](#)[Referencias bibliográficas](#)[Recursos](#)

Autoevaluación 2		
Pregunta	Respuesta	Retroalimentación
1	c	El mejor estimador puntual de la diferencia entre dos medias poblacionales es la diferencia entre dos medias muestrales obtenidas de sus respectivas poblaciones.
2	b	Para construir un intervalo de confianza para la media poblacional, se debe disponer de la media muestral, que se ubica en el centro del intervalo.
3	c	Un nivel de confianza (por ejemplo, el $k\%$ ) indica que, de 100 muestras aleatorias, se espera que $k$ arrojen un valor estimado del parámetro, dentro de dicho intervalo.
4	b	La relación entre el nivel de confianza del intervalo y el ancho del intervalo es directa.
5	c	La relación entre el tamaño de muestra y la amplitud del intervalo de confianza es inversa. Por ejemplo, la formula del intervalo de confianza para una media poblacional $\mu$ es: $x \pm z \alpha_{z/2} (s / \sqrt{n})$
6	b	Cuando el intervalo de confianza no contiene el valor 0, se concluye que las dos medias o dos proporciones son diferentes.
7	d	Hay que usar la formula del intervalo de confianza para una media poblacional $\mu$ (cuando $n \geq 30$ ): $x \pm z \alpha_{z/2} (s / \sqrt{n})$
8	b	La hipótesis nula es que no hay diferencia entre las proporciones de sobrevivientes en el grupo de control y en el grupo de medicamento.
9	c	La hipótesis alternativa es que la proporción de sobrevivientes en el grupo de medicamento será mayor que la proporción de sobrevivientes en el grupo de control.
10	b	Debido a que uno de los posibles valores dentro del intervalo de confianza es 0, no se debe rechazar la hipótesis nula $H_0: p_1 = p_2$ . Por lo tanto, no hay evidencia suficiente para indicar que el medicamento es efectivo para tratar el Zika.

[Ir a la autoevaluación](#)

Autoevaluación 3		
Pregunta	respuesta	Retroalimentación
1	d	El valor $P$ es la probabilidad de observar el estadístico de muestra cuando $H_0$ sea correcta. Se compara el valor $P$ con el nivel de significancia $\alpha$ , lo cual es el máximo riesgo tolerable elegido por el/la investigador/a de rechazar incorrectamente $H_0$ .
2	b	El nivel de significancia alfa se define como la probabilidad de rechazar $H_0$ cuando $H_0$ es verdadera y es elegido por el/la investigador/a.
3	b	Cuando la hipótesis alternativa postula que las medias poblacionales son diferentes pero no especifica cuál es mayor o menor, se denomina una prueba de dos colas. Las colas se refieren a la cola superior y la cola inferior de la curva normal.
4	a	Cuando el valor $P$ – definido como la probabilidad de observar el estadístico de muestra cuando $H_0$ sea correcta – es mayor que $\alpha$ (i.e., alta), no se rechaza $H_0$ .
5	c	El nivel de significancia $\alpha$ se define como el máximo riesgo tolerable elegido por el/la investigador/a de rechazar incorrectamente $H_0$ .
6	b	Un error de tipo I se define como la decisión de rechazar la hipótesis nula cuando es verdadera.
7	c	La hipótesis nula $H_0$ es que el sueldo de hombres médicos <i>no es diferente</i> que el sueldo de médicos en general.
8	d	La hipótesis alternativa $H_a$ es que el sueldo de hombres médicos es <i>mayor</i> que el sueldo de médicos en general.
9	a	Dado que el tamaño muestral $n \geq 30$ , se puede asumir que la distribución muestral de $x$ es normal o normalmente aproximada. Por lo tanto, se puede utilizar el estadístico de prueba $z$ de la curva normal estándar para contrastar esta hipótesis: $z = (x - \mu_0) / (s / \sqrt{n})$
10	d	El estadístico de prueba $z = 3,41$ es mayor que el valor crítico $z = 2,33$ , entonces cae en la región de rechazo y se rechaza $H_0$ . Se concluye que la ganancia para hombres médicos es significativamente mayor que la de mujeres.

[Índice](#)[Primer bimestre](#)[Segundo bimestre](#)[Solucionario](#)[Referencias bibliográficas](#)[Recursos](#)[Ir a la autoevaluación](#)

### Autoevaluación 3

Pregunta	respuesta	Retroalimentación
11	a	La hipótesis nula es que no hay diferencia en el rendimiento académico medio entre estudiantes que tienen computadora y estudiantes que no tienen computadora.
12	e	La hipótesis alternativa es que hay una diferencia en el rendimiento medio entre propietarios/as de computadoras y no propietarios/as. No se especifica cuál grupo tendrá el mayor rendimiento.
13	b	Dado que el tamaño muestral de cada grupo $\geq 30$ , se puede asumir que la distribución muestral de $x$ es normal. Por lo tanto, se puede utilizar el estadístico de prueba $z$ para contrastar esta hipótesis. Ver la formula de $z$ para la diferencia entre dos medias poblaciones en apartado 9.4 de Mendenhall <i>et al.</i> 2015.
14	c	El estadístico de prueba $z = 1,94$ es menor que el valor crítico $z = 1,96$ que corresponde a una prueba de dos colas con $\alpha=0,05$ . Entonces no se rechaza $H_0$ .

**Autoevaluación 4**

Pregunta	Respuesta	Retroalimentación
1	b	La prueba $t$ es un caso particular del análisis de varianza (ANOVA) en el que sólo hay dos categorías o niveles de la variable independiente.
2	e	La forma de la distribución $t$ depende del tamaño muestral $n$ . Los grados de libertad (degrees of freedom en inglés) para una muestra, en cambio, depende de $n$ : $df = n - 1$ .
3	g	Dos suposiciones para la prueba $t$ de dos muestras independientes son (i) independencia de las unidades muestrales y (ii) igualdad de varianzas.
4	a	Debido a que la agricultora mide la producción de frutos de los mismos frutales antes y después de aplicar abono, no se puede considerar que las muestras son independientes sino pareadas.
5	c	El propósito de la prueba $t$ es probar si hay diferencia estadística entre las medias de dos poblaciones.
6	c	Como cada muestra tiene $(n - 1)$ grados de libertad, el número total es la suma: $(n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$
7	e	Todas son suposiciones para la prueba $t$ y para los modelos lineales en general, por ejemplo ANOVA y regresión.
8	a	La hipótesis nula es que no hay diferencia en la concentración de glucosa en plasma entre pacientes diabéticos y pacientes no diabéticos.
9	e	La hipótesis alternativa es que la concentración media de glucosa es mayor en los pacientes diabéticos que en los no diabéticos. Respuestas b y d son idénticas.
10	d	Dado que el tamaño muestral de cada grupo es $n_1 = n_2 = 10$ , no se puede utilizar el estadístico de prueba $z$ para contrastar esta hipótesis. La fórmula para $t$ es distinta (ver apartado 10.4 de Mendenhall <i>et al.</i> 2015). No obstante, T.TEST() en Excel nos da una opción más sencilla para realizar la prueba.
11	a	T.TEST() en Excel da el valor $P = 0,0164$ .

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

#### Autoevaluación 4

Pregunta	Respuesta	Retroalimentación
12	a	El valor $P$ es menor que el nivel de significancia $\alpha=0,05$ . Entonces se rechaza $H_0$ .

Ir a la  
autoevaluación

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

### Autoevaluación 5

Pregunta	Respuesta	Retroalimentación
1	a	Las suposiciones para los modelos lineales (i.e., prueba <i>t</i> , ANOVA y regresión) son: (i) muestras aleatorias independientes, (ii) poblaciones con una distribución normal y (iii) homogeneidad de varianza.
2	d	En un ANOVA, se divide la variación total de las mediciones de respuesta en partes que pueden ser atribuidas a varios factores (i.e., variables independientes) de interés para el experimentador. Sin embargo, el propósito del ANOVA es hacer comparaciones de más de dos medias poblacionales.
3	d	El ANOVA se apoya en estadísticos con distribuciones muestrales que son modeladas por la distribución F.
4	e	El diseño de ANOVA de un factor comprende sólo un factor o variable independiente. En Mendenhall <i>et al.</i> 2015 se usa la designación como una clasificación en una dirección.
5	a	Los grados de libertad de los tratamientos que corresponden a la suma de cuadrados para tratamientos (SST), el componente que mide la variación <i>entre</i> los <i>k</i> tratamientos, se calculan en base a ( <i>k</i> – 1).
6	b	Los grados de libertad del error que corresponden a la suma de cuadrados para el error (SSE), el componente que mide la variación agrupada <i>dentro</i> de los <i>k</i> tratamientos, se calculan en base a ( <i>n</i> – <i>k</i> ).
7	e	Dado que la variable independiente es categórica <i>y hay más de dos categorías, un ANOVA es apropiado en este caso. El estadístico de prueba para ANOVA es F.</i>
8	f	Solo se aplica una prueba de comparaciones múltiples si el valor <i>P</i> del ANOVA en general es menor que <b>a</b> . <i>En otra palabra, existe por lo menos una diferencia significativa entre las medias de los <i>k</i> niveles del factor.</i>
9	f	La variable independiente es categórica <i>y, específicamente, ordinal porque se puede ordenar los niveles desde el control hasta la dosis alta.</i>

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

### Autoevaluación 5

Pregunta	Respuesta	Retroalimentación
10	a	La hipótesis nula es que no hay diferencia en el peso medio de la colonia entre los tres tratamientos.
11	a	Dado que el valor $P < 0,05$ , se rechaza la hipótesis nula y se concluye que hay por lo menos una diferencia en el peso medio de la colonia entre los tres tratamientos.
12	b	Dado que el valor $P < 0,05$ , se concluye que hay diferencia significativa entre los tratamientos mencionados.

Ir a la  
autoevaluación

[Índice](#)[Primer bimestre](#)[Segundo bimestre](#)[Solucionario](#)[Referencias bibliográficas](#)[Recursos](#)

Autoevaluación 6		
Pregunta	Respuesta	Retroalimentación
1	v	La recta de mejor ajuste es aquella que minimiza las distancias de los puntos observados a la recta ajustada por el modelo a través del método de mínimos cuadrados.
2	f	En la fórmula de la recta de mejor ajuste la pendiente de la recta es el coeficiente de $x$ , en este caso $\beta_1$ .
3	f	Es al contrario: el coeficiente de determinación $R^2$ es el cuadrado del coeficiente de correlación $r$ .
4	v	Cuando $R^2=0$ , la pendiente $\beta_1=0$ y no hay relación lineal entre $x$ e $y$ .
5	a	Cuando $y$ disminuye conforme $x$ aumenta, $r$ es menor que cero y puede que exista una relación lineal negativa.
6	f	El análisis de regresión solo se puede aplicar con dos variables continuas (i.e., cuantitativas).
7	b	La variable dependiente $y$ responde en función de la variable independiente $x$ . El modelo expresa que $y$ incrementa por cada unidad de incremento de $x$ . $\beta_0$ y $\beta_1$ son constantes desconocidos y no cambian.
8	f	$R^2$ representa el porcentaje de variabilidad explicado por el modelo de regresión. Valores de $R^2$ cercanos a 1, nos indica un buen ajuste del modelo y fuerte relación entre $x$ e $y$ .
9	d	Cuando los valores del coeficiente de correlación $r$ son cercanos a 1, nos indica una mayor relación entre las variables.
10	c	Para calcular $y$ para valores concretos de $x$ , se sustituye el valor $x = 100$ en el modelo dado: $y = 0.36(100) + 2.1$ .
11	g	Los parámetros que un análisis de regresión pretende estimar son la pendiente ( $\beta_1$ ) y la intersección ( $\beta_0$ ).

[Ir a la autoevaluación](#)

[Índice](#)[Primer bimestre](#)[Segundo bimestre](#)[Solucionario](#)[Referencias bibliográficas](#)[Recursos](#)

### Autoevaluación 7

Pregunta	Respuesta	Retroalimentación
1	e	El objetivo de la prueba de chi cuadrado es determinar si un método de clasificación de una variable categórica es contingente o dependiente del método de clasificación de otra variable categórica. Se logra esto al comparar el estadístico $X^2$ de la muestra (basado en las diferencias entre frecuencias observadas y frecuencias esperadas) con el valor crítico $X^2$ , lo cual está basado en $\alpha$ y los grados de libertad.
2	b	Ambas variables son categóricas ya que no se puede hacer operaciones algebraicas con ellas.
3	d	Para probar si la frecuencia de ratones enfermos y sanos depende de la vacuna, es adecuado usar la prueba chi cuadrado.
4	b	Los grados de libertad de una tabla de contingencia se calculan con la fórmula: $df = (\text{columnas} - 1) * (\text{filas} - 1) = (2-1)*(2-1) = 1$ .
5	e	La función <b>CHISQ.TEST()</b> nos da un valor de $X^2 = 48,748$ .
6	b	Según Tabla 5 “Valores críticos de ji cuadrada” de Apéndice I de Mendenhall <i>et al.</i> 2015, se busca el valor crítico en la columna $X^2_{.050}$ y la fila $df=1$ .
7	d	La función <b>CHISQ.TEST()</b> nos da un valor de $P = 2,911 \times 10^{-12}$
8	a	Dado que el valor $P < 0,05$ , se concluye que hay una diferencia significativa en el patrón del número de ratones enfermos y sanos, dependiendo del tratado de la vacuna. La vacuna es efectiva para tratar la influenza porcina.
9	f	Dos pruebas no paramétricas que usan un estadístico de prueba basado en los rangos de valores de las observaciones son la prueba de suma de rango de Wilcoxon y la prueba U de Mann-Whitney.
10	e	La prueba <i>H</i> de Kruskal-Wallis es la alternativa no paramétrica al ANOVA de un factor. Se usa para detectar diferencias en ubicaciones entre más de dos distribuciones poblacionales basadas en muestreo aleatorio independiente.

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

### Autoevaluación 7

Pregunta	Respuesta	Retroalimentación
11	b	Si en vez de una variable categórica la variable independiente es continua con valores de dosis de neonicotinoides distintos, el análisis apropiado sería una regresión simple.

Ir a la  
autoevaluación

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos



## 5. Referencias bibliográficas

### Texto básico

Mendenhall W., Beaver R.J. y Beaver B.M. (2015). **Introducción a la Probabilidad y Estadística** (14a edición). México D.F.: Cengage Learning.

### Texto complementario

Martínez Bencardino, C. (2015). **Estadística y Muestreo**. Ecoe Ediciones Ltda.

### Recursos Educativos Abiertos

Título del REA	Enlace
REA nº1. Estadística Online (YouTube)	<a href="https://www.youtube.com/channel/UCqkM4FQL-p9G-liqitrzgyg/videos">https://www.youtube.com/channel/UCqkM4FQL-p9G-liqitrzgyg/videos</a>
REA nº2. Cambio climático: amenaza en cascada a la biodiversidad (scidev.net)	<a href="https://www.scidev.net/america-latina/medio-ambiente/noticias/cambio-climatico-amenaza-en-cascada-a-la-biodiversidad.html">https://www.scidev.net/america-latina/medio-ambiente/noticias/cambio-climatico-amenaza-en-cascada-a-la-biodiversidad.html</a>
REA nº3. Luis Cayuela - Curso de análisis de datos ecológicos en R	<a href="http://luiscayuela.blogspot.com/2012/07/material-cursos-de-r-para-modelos.html">http://luiscayuela.blogspot.com/2012/07/material-cursos-de-r-para-modelos.html</a>
REA nº4. Video sobre cambio climático y género (COP 25)	<a href="https://www.youtube.com/watch?v=pONITMuXFnw">https://www.youtube.com/watch?v=pONITMuXFnw</a>

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos



## 6. Recursos

### Documento 1. Distribuciones muestrales y el Teorema del Límite Central

The image shows the cover of a document titled "ESTADÍSTICA INFERENCIAL". At the top left is the UTPL logo with the text "UTPL La Universidad Católica de Loja". The title "ESTADÍSTICA INFERENCIAL" is centered above a yellow rectangular box containing the text "Distribuciones muestrales y el Teorema del Límite Central". Below this box, the text "Elaborado: Daniel Griffith, Ph.D." is visible. In the bottom right corner, there is a yellow square with the text "soy+ utpl". The background of the cover features a dark blue color with a pattern of light blue circles and dots.

## Distribuciones muestrales de $\bar{x}$

### 1) Muestreo a partir de poblaciones que siguen una distribución normal

En este caso, la distribución de la media de la muestra tiene las siguientes propiedades:

1. La distribución de  $\bar{x}$  será normal.
2. La media de la distribución de  $\bar{x}$  será igual a la media de la población de la cual se seleccionaron las muestras.
3. La varianza  $s_x^2$  de la distribución de  $\bar{x}$  será igual a la varianza de la población dividida por el tamaño de la muestra.

$$s_x^2 = \frac{\sigma^2}{n}$$



## Error estándar de la media

La varianza  $s^2$  de la distribución de  $\bar{x}$  es igual a la varianza de la población dividida por el tamaño de la muestra.

$$s^2 = \frac{\sigma^2}{n}$$

**Error estándar de la media:**

$$s = \frac{\sigma}{\sqrt{n}}$$

Como consecuencia, la dispersión de la distribución de medias muestrales es considerablemente menor que la dispersión de la población muestreada.



**2) Muestreo a partir de poblaciones que no siguen una distribución normal**

En este caso, se utiliza un teorema matemático conocido como:

**Teorema del límite central**

Dada una población de cualquier forma funcional no normal con una media  $\mu$  y varianza finita  $\sigma^2$ , la distribución muestral de  $\bar{x}$ , calculada a partir de muestras de tamaño  $n$  de dicha población, será casi normal con media  $\mu$  y varianza  $\frac{\sigma^2}{n}$  cuando la muestra es muy grande.



Entonces, en el caso de la media de la muestra, se tiene la seguridad de que la distribución muestral está distribuida en forma al menos aproximadamente normal con tres condiciones:

1. Cuando se hace el muestreo a partir de una población con distribución normal.
2. Cuando se hace el muestreo a partir de una población que no exhibe una distribución normal y la muestra es grande.
3. Cuando se hace el muestreo a partir de una población cuya forma funcional se desconoce,  
+ siempre que el tamaño de la muestra sea grande.

¿Qué tan grande debe ser la muestra?  **$n > 30$**

**Ir al contenido**

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

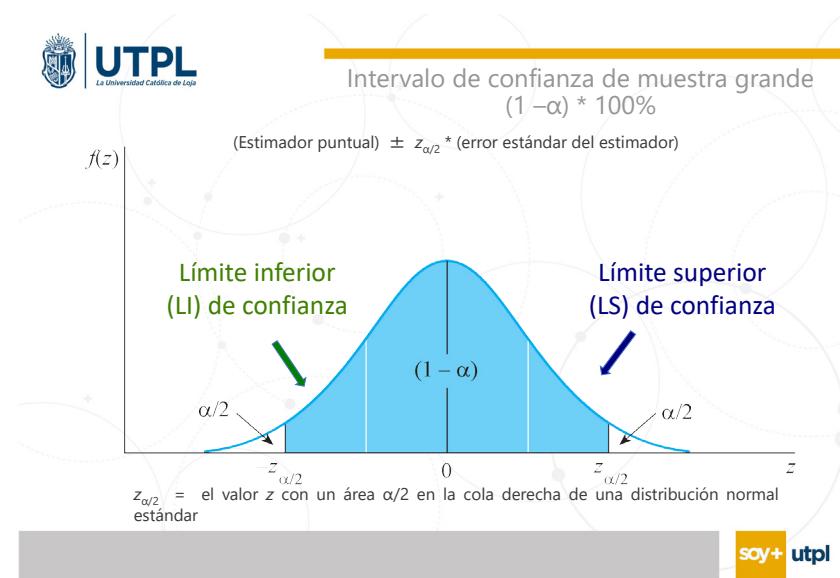
## Documento 2. Intervalo de confianza para una media poblacional

**ESTADÍSTICA INFERENCIAL**

Intervalo de confianza para una media poblacional

Elaborado: Daniel Griffith, Ph.D.

soy+ utpl



## Un intervalo de confianza $(1 - \alpha)^*$ 100% para una media poblacional $\mu$

Para calcular el intervalo de confianza  $(1 - \alpha)$  de muestra grande ( $n > 30$ ) para una media poblacional  $\mu$  usemos la siguiente formula:

(Estimador puntual)  $\pm Z\alpha/2 * (\text{error estándar del estimador})$

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

$\bar{x}$  = media muestral

$z_{\alpha/2}$  = valor z de la distribución normal estándar (=1.96 para un intervalo de confianza de 95%)

$s$  = desviación estándar de la muestra

$n$  = tamaño muestral



[Ir al contenido](#)

Índice

Primer  
bimestre

Segundo  
bimestre

Solucionario

Referencias  
bibliográficas

Recursos

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Documento 3. Ejemplo práctico de una prueba de hipótesis de muestra grande acerca de una media poblacional

The cover features the UTPL logo and the text 'ESTADÍSTICA INFERENCIAL'. A yellow box highlights the title 'Ejemplo práctico de una prueba de hipótesis de muestra grande acerca de una media poblacional'. Below it, the author is listed as 'Elaborado: Daniel Griffith, Ph.D.' and the publisher as 'soy+ utpl'.



Una prueba de una cola acerca de una media poblacional  $\mu$

- **Ejemplo:** El promedio semanal de ganancias para médicos es \$670. ¿Los hombres de la misma posición tienen ganancias semanales promedio más altas que los de las mujeres? Una muestra aleatoria de  $n = 40$  médicos hombres mostró  $\bar{x} = \$725$  y  $s = \$102$ . Pruebe la hipótesis apropiada usando  $\alpha = 0.01$



1-2) **Hipótesis:**  $H_0: \mu = 670$  contra  $H_a: \mu > 670$

3) **Estadístico de prueba:**  $z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{725 - 670}{102/\sqrt{40}} = 3.41$

4) **Región de rechazo:** valor crítico es  $\alpha = 0.01 \rightarrow z = 2.33$

5) **Conclusión:** estadístico de prueba  $z = 3.41 >$  valor crítico  $z = 2.33$ , entonces cae en la región de rechazo.

Rechazamos  $H_0$  y concluimos que la ganancia para médicos hombres es significativamente mayor que la de mujeres.



Ir al contenido

[Índice](#)[Primer bimestre](#)[Segundo bimestre](#)[Solucionario](#)[Referencias bibliográficas](#)[Recursos](#)

## Documento 4. Ejemplo práctico de una prueba de hipótesis para la diferencia entre dos medias poblacionales

The cover features the UTPL logo and the text 'ESTADÍSTICA INFERENCIAL'. It also includes the author's name, 'Elaborado: Daniel Griffith, Ph.D.', and the 'soy+ utpl' logo.



Una prueba para la diferencia entre dos medias poblacionales ( $\mu_1 - \mu_2$ )

**Ejemplo:** Para determinar si la propiedad de una computadora afecta el rendimiento académico de un estudiante, se tomaron dos muestras aleatorias de 100 estudiantes. El promedio de calificaciones para los  $n_1 = 100$  que no tenían computadora tuvieron un promedio y variancia igual a  $\bar{x}_1 = 270$  y  $s_1^2 = 0.36$ , en tanto que  $\bar{x}_2 = 2.54$  y  $s_2^2 = 0.40$  para los  $n_2 = 100$  que tenían computadora.

¿Los datos presentan suficiente evidencia para indicar una diferencia en el rendimiento medio entre propietarios de computadoras y no propietarios? Pruebe usando  $\alpha = 0.05$

**Solución:**  $H_0: (\mu_1 - \mu_2) = D_0 = 0$       contra       $H_a: (\mu_1 - \mu_2) \neq 0$

$$z \approx \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{2.70 - 2.54}{\sqrt{\frac{0.36}{100} + \frac{0.40}{100}}} = 1.84$$

**Conclusión:**  $z = 1.84 < 1.96$   
entonces no rechaza  $H_0$

No hay suficiente evidencia que hay una diferencia en el promedio de los rendimientos académicos para los dos grupos.

[Ir al contenido](#)

[Índice](#)[Primer bimestre](#)[Segundo bimestre](#)[Solucionario](#)[Referencias bibliográficas](#)[Recursos](#)

## Documento 5. Una prueba para la diferencia entre dos proporciones binomiales ( $p_1 - p_2$ )

ESTADÍSTICA INFERENCIAL

Una prueba para la diferencia entre dos proporciones binomiales ( $p_1 - p_2$ )

Elaborado: Daniel Griffith, Ph.D.

soy+ utpl



Una prueba para la diferencia entre dos proporciones binomiales ( $p_1 - p_2$ )

Los registros de un hospital indican que 52 hombres de una muestra de 1000 contra 23 mujeres de una muestra de 1000 fueron ingresados por enfermedad del corazón. ¿Estos datos presentan suficiente evidencia para indicar un porcentaje más alto de enfermedades del corazón entre hombres ingresados al hospital? Use  $\alpha = 0.05$

**Solución:**  $H_0: p_1 = p_2$        $(p_1 - p_2) = 0$       cont  $H_a: p_1 > p_2$

$$\hat{p} = \frac{\hat{X}_1 + \hat{X}_2}{n_1 + n_2} = \frac{52 + 23}{1000 + 1000} = 0.0375$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.052 - 0.023}{\sqrt{(0.0375)(0.9625)\left(\frac{1}{1000} + \frac{1}{1000}\right)}} = 3.41$$

**Conclusión:**  $z = 3.41 > 1.65$  entonces rechaza  $H_0$

Los datos presentan suficiente evidencia para indicar que el porcentaje de hombres que ingresan al hospital por enfermedad del corazón es más alto que el de mujeres.



[Ir al contenido](#)

Índice

Primer bimestre

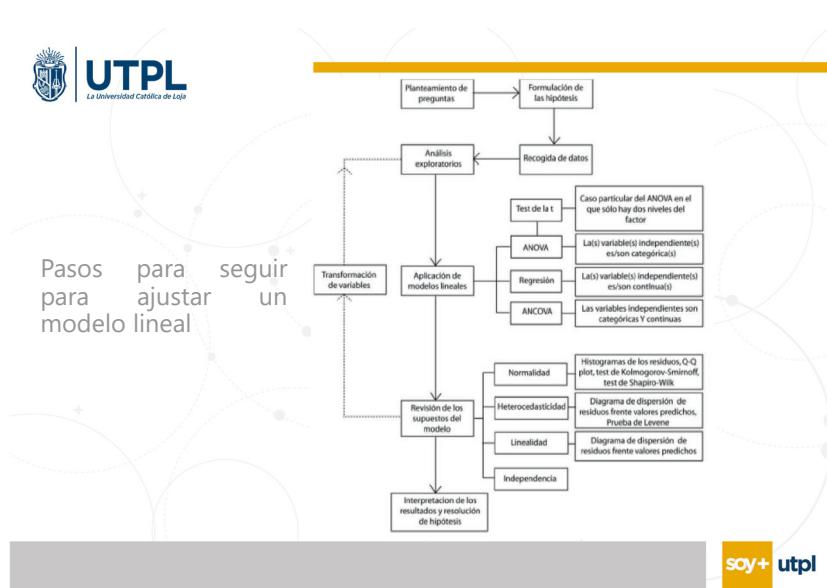
Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Documento 6. Pasos para seguir para ajustar un modelo lineal (Prueba t, ANOVA, regresión)



Ir al contenido

Índice

Primer bimestre

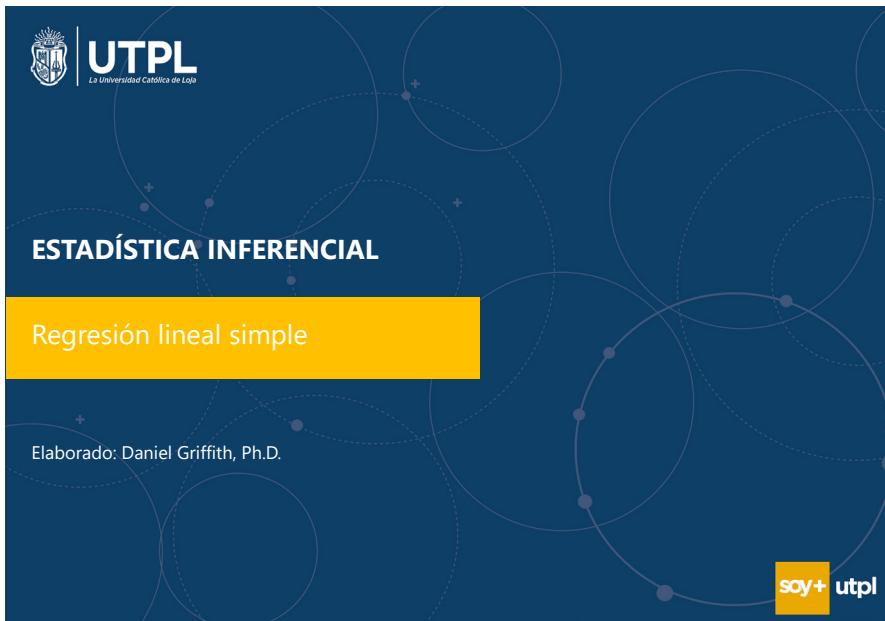
Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Documento 7. Regresión lineal simple



### Regresión lineal simple

Es un tipo de modelo probabilístico que supone que el valor medio de  $Y$  para un valor dado de  $X$  se grafica como una línea recta y que los puntos se desvían de esta *línea de medias* en una cantidad aleatoria (positiva o negativa) igual a  $\varepsilon$ , es decir:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

$\beta_0, \beta_1$  = los parámetros del modelo

$\varepsilon$  = error aleatorio



## Bondad de un ajuste $R^2$

La bondad de un ajuste de un modelo de regresión se mide usando el coeficiente de determinación,  $R^2$

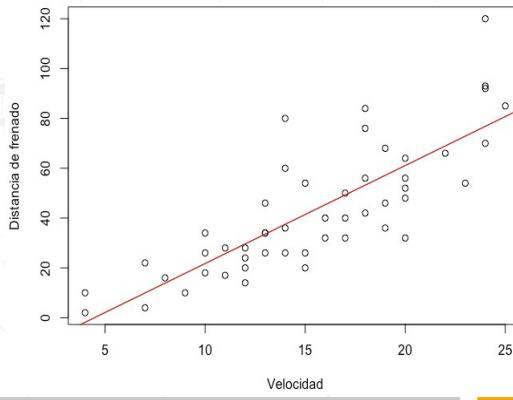
- $R^2$  es una cantidad **adimensional** que sólo puede tomar valores entre **[0, 1]**
- Cuando un **ajuste es bueno**  $R^2$  será cercano a **uno**.
- Cuando un **ajuste es malo**  $R^2$  será cercano a **cero**.
- A  $R^2$  también se le denomina **el porcentaje de variabilidad explicado** por el modelo de regresión.
- $R^2$  puede ser pesado de calcular en modelos de regresión general, pero en el modelo lineal simple, la expresión es de lo más sencilla:  $R^2 = r^2$

## Ejemplo: Regresión lineal simple

¿La distancia que se requiere para frenar un auto depende de su velocidad?

$$H_0: \beta_1 = 0$$

$$H_a: \beta_1 > 0$$



## Ejemplo: Regresión lineal simple

Calcular Y para valores concretos de X

$$Y = 3.93X - 17.58$$

- ¿Cuál distancia se requiere para frenar un automóvil que va a 20 millas / hora?

$$3.93 * 20 - 17.58$$

[1] 61.02

- ¿Cuál distancia se requiere para frenar el automóvil que va a 50 millas / hora?

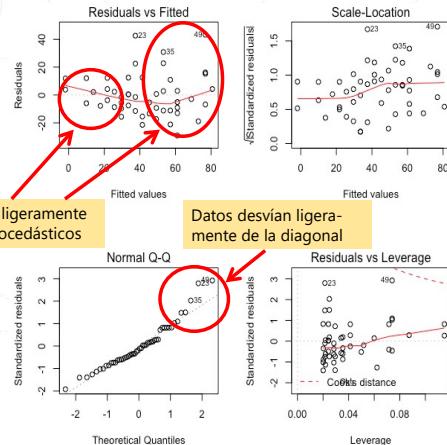
$$3.93 * 50 - 17.58$$

[1] 178.92

**Ojo! Cuidado con la suposición que la relación es lineal mas allá de los valores de X medidos. No se debe extrapolar el modelo mucho mas allá del rango de X medido (Velocidad entre 0 y 30 en este ejemplo).**

## Evaluación de las suposiciones del modelo: Exploración gráfica de los residuos

Hay que comprobar las suposiciones de normalidad y homocedasticidad (= homogeneidad de varianzas)



Datos ligeramente heterocedásticos

Datos desvían ligeramente de la diagonal

## Documento 8. Prueba Chi cuadrado

**UTPL** La Universidad Católica de Loja

**ESTADÍSTICA INFERENCIAL**

**Prueba Chi cuadrado**

Elaborado: Daniel Griffith, Ph.D.

**soy+ utpl**

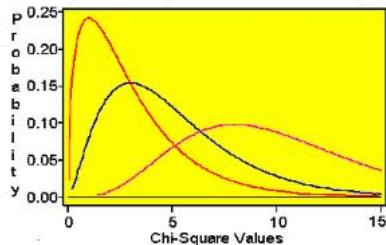


### Características de la distribución de chi cuadrado $\chi^2$

- La distribución  $\chi^2$  se lee con **grados de libertad** en tablas de contingencia:
- $df = (\# \text{ filas} - 1) * (\# \text{ columnas} - 1)$**
- No tiene valores negativos. El valor mínimo es 0.
  - Todas las curvas son asimétricas.
  - Cuando aumentan los grados de libertad las curvas son menos elevadas y más extendidas a la derecha.
  - Sé usa una corrección cuando al menos una Frec. Esperada < 5.
  - Las fórmulas son:

$$\chi^2_c = \sum \left[ \frac{(O_i - E_i)^2}{E_i} \right]$$

$$F. \text{ Esperada} = \frac{\text{Total de Fila} \times \text{Total de Columna}}{\text{Total General}}$$



## Tablas de contingencia: una clasificación de dos vías

Tabla de contingencia

$x \setminus y$	$d_1$	...	$d_k$	...	$d_s$	total
$c_1$	$n_{11}$	...	$n_{1k}$	...	$n_{1s}$	$n_{1\bullet}$
:	:		:		:	:
$c_h$	$n_{h1}$	...	$n_{hk}$	...	$n_{hs}$	$n_{h\bullet}$
:	:		:		:	:
$c_r$	$n_{r1}$	...	$n_{rk}$	...	$n_{rs}$	$n_{r\bullet}$
total	$n_{\bullet 1}$	...	$n_{\bullet k}$	...	$n_{\bullet s}$	$n$

Cuando se registran dos variables categóricas, se puede resumir la información en una **tabla de contingencia** al contar el número observado de unidades que caen en cada una de las diversas intersecciones de niveles de categoría.

El objetivo es determinar si un método de clasificación es o no es **contingente** o **dependiente** del otro método de clasificación:

**$H_0$ :** Los dos métodos de clasificación son independientes

**$H_a$ :** Los dos métodos de clasificación son dependientes

[Ir al contenido](#)

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas

Recursos

## Documento 9. Pruebas paramétricas y su contraparte no paramétrica

ESTADÍSTICA INFERENCIAL

Pruebas paramétricas y su contraparte no paramétrica

Elaborado: Daniel Griffith, Ph.D.

soy+ utpl

### Pruebas paramétricas y su contraparte no paramétrica

Situación	Método no paramétrico	Método paramétrico	Tamaño muestral
1 media	Prueba de rangos señalados de Wilcoxon	Prueba de Z (prueba de t)	$n > 30$ $n < 30$
2 medias independientes	Suma de rangos señalados de Wilcoxon	Prueba de Z para dos muestras independientes (Prueba t)	$n > 30$ $n < 30$
2 medias pareadas	Prueba de rangos señalados de Wilcoxon	Prueba de Z pareada (Prueba t pareada)	$n > 30$ $n < 30$
2 variables continuas	Coeficiente de correlación de Spearman	Coeficiente de correlación de Pearson/Regresión	
Variable categórica con más de 2 niveles	Kruskall-Wallis	ANOVA (prueba F)	

soy+ utpl

Ir al contenido