



UTPL
La Universidad Católica de Loja

Modalidad Abierta y a Distancia

Estadística

Guía didáctica

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas



Facultad de Ciencias Económicas y Empresariales

Departamento de Economía

Estadística

Guía didáctica

Carrera	PAO Nivel
▪ Logística y Transporte	IV

Autor:

Griffith Daniel Michael



Asesoría virtual
www.utpl.edu.ec

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

Universidad Técnica Particular de Loja

Estadística

Guía didáctica

Griffith Daniel Michael

Diagramación y diseño digital:

Ediloja Cía. Ltda.

Telefax: 593-7-2611418.

San Cayetano Alto s/n.

www.ediloja.com.ec

edilojacialtda@ediloja.com.ec

Loja-Ecuador

ISBN digital - 978-9942-39-060-8



Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0)

Usted acepta y acuerda estar obligado por los términos y condiciones de esta Licencia, por lo que, si existe el incumplimiento de algunas de estas condiciones, no se autoriza el uso de ningún contenido.

Los contenidos de este trabajo están sujetos a una licencia internacional Creative Commons **Reconocimiento-NoComercial-CompartirIgual** 4.0 (CC BY-NC-SA 4.0). Usted es libre de **Compartir** — copiar y redistribuir el material en cualquier medio o formato. **Adaptar** — remezclar, transformar y construir a partir del material citando la fuente, bajo los siguientes términos: **Reconocimiento-** debe dar crédito de manera adecuada, brindar un enlace a la licencia, e indicar si se han realizado cambios. Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante. **No Comercial-** no puede hacer uso del material con propósitos comerciales. **Compartir igual-** Si remezcla, transforma o crea a partir del material, debe distribuir su contribución bajo la misma licencia del original. No puede aplicar términos legales ni medidas tecnológicas que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia. <https://creativecommons.org/licenses/by-nc-sa/4.0/>

8 de marzo, 2021

Índice

1. Datos de información.....	8
1.1. Presentación de la asignatura	8
1.2. Competencias genéricas de la UTPL.....	8
1.3. Competencias específicas de la carrera	8
1.4. Problemática que aborda la asignatura	9
2. Metodología de aprendizaje.....	9
3. Orientaciones didácticas por resultados de aprendizaje	10
Primer bimestre.....	10
Resultado de aprendizaje 1	10
Contenidos, recursos y actividades de aprendizaje.....	10
Semana 1	11
Unidad 1. Introducción a la estadística	11
1.1. ¿Qué es la estadística?	11
Actividades de aprendizaje recomendadas	12
Autoevaluación 1	13
Semana 2	15
Unidad 2. Exploración de datos.....	15
2.1. Tablas de frecuencia.....	15
2.2. Representación gráfica de variables (categóricas y numéricas).....	15
Actividades de aprendizaje recomendadas	17
Semana 3	18
2.3. Medidas de centralización	18
2.4. Medidas de variación.....	18

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

2.5. Medidas de posición relativa	18
Actividades de aprendizaje recomendadas	19
Autoevaluación 2	21
Resultado de aprendizaje 2	23
Contenidos, recursos y actividades de aprendizaje.....	23
Semana 4	23
Unidad 3. Introducción a la probabilidad	23
3.1. Conceptos básicos	23
3.2. Reglas fundamentales de la probabilidad.....	24
Actividades de aprendizaje recomendadas	24
Autoevaluación 3	25
Semana 5	28
Unidad 4. Distribuciones de variables	28
4.1. Distribución binomial de la probabilidad.....	28
4.2. Distribución de la probabilidad de Poisson.....	28
Semana 6	29
4.3. Distribución normal de la probabilidad.....	29
Actividades de aprendizaje recomendadas	30
Autoevaluación 4	32
Semana 7	35
Unidad 5. Introducción a la estadística inferencial	35
5.1. Estadística inferencial: definiciones y métodos	35
5.2. Distribución muestral	35
5.3. Control estadístico de procesos	35
Actividades de aprendizaje recomendadas	36
Autoevaluación 5	38

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

Semana 8	41
Actividades de aprendizaje recomendadas	41
Actividades finales del bimestre	42
Segundo bimestre	43
Resultado de aprendizaje 3	43
Contenidos, recursos y actividades de aprendizaje.....	43
Semana 9	43
Unidad 6. Intervalo de confianza	44
6.1. Estimación puntual.....	44
6.2. Construcción e interpretación de un intervalo de confianza	44
Semana 10	45
6.3. Intervalo de confianza de muestra grande para una media y una proporción poblacional.....	45
6.4. Intervalo de confianza para la diferencia entre dos medias poblacionales y entre dos proporciones binomiales	45
Actividades de aprendizaje recomendadas	46
Autoevaluación 6	47
Semana 11	50
Unidad 7. Prueba de hipótesis	50
7.1. Procedimientos de una prueba estadística de hipótesis	50
7.2. Una prueba de muestra grande acerca de una media poblacional.....	50

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

Semana 12	51
7.3. Una prueba de hipótesis para la diferencia entre dos medias poblacionales.....	51
7.4. Una prueba de hipótesis para una proporción binomial.	51
7.5. Una prueba de hipótesis para la diferencia entre dos proporciones binomiales	51
Actividades de aprendizaje recomendadas	52
Semana 13	52
7.6. Prueba t de student	52
Actividades de aprendizaje recomendadas	53
Autoevaluación 7	54
Semana 14	60
Unidad 8. Regresión lineal simple y correlación	60
8.1. Regresión lineal simple (calidad del modelo, evaluación de suposiciones).....	60
Actividades de aprendizaje recomendadas	61
Semana 15	61
8.2. Correlación bivariada (Pearson, Spearman).....	61
Actividades de aprendizaje recomendadas	62
Autoevaluación 8	63
Semana 16	66
Actividades finales del bimestre	66
4. Solucionario	67
5. Referencias bibliográficas	78

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

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

1. Datos de información

1.1. Presentación de la asignatura



1.2. Competencias genéricas de la UTPL

- Identifica problemas de logística y transporte.
- Resuelve problemas de ingeniería en logística y transporte.
- Asume pensamiento crítico y reflexivo.

1.3. Competencias específicas de la carrera

- Orientación a la innovación y a la investigación.
- Pensamiento crítico y reflexivo.

1.4. Problemática que aborda la asignatura

- El componente de estadística aporta al núcleo problema 2 transporte y se vincula con el sector industrial, empresarial y afines.



2. Metodología de aprendizaje

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

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 una de las 8 unidades propuestas en esta asignatura, cada estudiante tendrá que resumir y realizar esquemas para entender y estudiar los contenidos propuestos.

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 los/as estudiantes expondrán propuestas de movilidad sostenible y los objetivos de desarrollo sostenible a través de los foros y de las resoluciones de problemas.



3. Orientaciones didácticas por resultados de aprendizaje



Primer bimestre

Resultado de aprendizaje 1

Comprende los diferentes conceptos básicos de la estadística como base instrumental para el trabajo comunitario e investigativo del área de la ingeniería de transporte.

El primer resultado de aprendizaje se trabajará durante las 3 primeras semanas del primer bimestre, a través de la unidad 1.- Introducción a la estadística y la unidad 2.- Exploración de datos. A través de diferentes ejemplos se identificarán los diferentes tipos de variables estadísticas y aprenderemos a organizar e interpretar un conjunto de datos.

Contenidos, recursos y actividades de aprendizaje

Índice

Primer bimestre

Segundo bimestre

Solucionario

Referencias bibliográficas



Semana 1



Unidad 1. Introducción a la estadística

1.1. ¿Qué es la estadística?

Bienvenidos/as a la primera semana de estadística. En esta primera semana vamos a definir el concepto de estadística y su aplicabilidad en todas las esferas de la vida cotidiana (Unidad 1, Apartado 1.1 ¿Qué es la estadística?). Me imagino que muchos/as de ustedes se preguntarán: *¿para qué necesito la estadística en la carrera Logística y transporte?* Pues bien, un buen profesional es capaz de entender la información y de usarla eficientemente. Después, en el apartado 1.2 sobre variables estadísticas y escalas de medición, aprenderemos a identificar los diferentes tipos de variables y para terminar esta unidad se presentarán las diferentes escalas de medición.

Recursos de aprendizaje

Entre los recursos de aprendizaje planificados para esta semana deben revisar el capítulo de introducción y los apartados 1.1 y 1.2 de la unidad 1 del texto básico (Mendenhall *et al.* 2015). Esta lectura les ayudará a comprender la importancia de la estadística y servirá

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

como una breve introducción a algunos conceptos básicos que usaremos a lo largo del curso académico.

Una vez revisado el texto, ¿sería capaz usted de diferenciar una variable discreta de una continua? Imagine que quiere estudiar el importe de los impuestos que como empresa de transporte ha estado pagando en los últimos 12 meses, ¿esta variable sería discreta o continua? ¡Correcto! Sería continua porque puede tomar cualquier valor dentro de un intervalo específico.



Actividades de aprendizaje recomendadas

- Realizar una corta presentación personal y académica en el foro académico. Comente si tiene conocimientos previos de estadística y qué utilidad le dará en su vida profesional.
- Realizar la autoevaluación 1 para repasar los conocimientos adquiridos en esta unidad 1.

¿Ha pensado alguna vez en las numerosas aplicaciones de la estadística en nuestra vida cotidiana? ¿Sería capaz de explicar la diferencia entre variables cualitativas y cuantitativas? ¿Qué le ha parecido esta primera semana? ¿Ha cambiado su visión de la estadística? Puede expresar su opinión en el foro de presentación personal o a través de comentarios en el anuncio académico semanal.



Autoevaluación 1

Introducción a la estadística

Lea con atención los enunciados del 1 al 5 y encierre en un círculo el literal que corresponda a la opción correcta.

1. Una de las ramas de la estadística es:
 - a. Una variable.
 - b. La estadística inferencial.
 - c. La escala de medición.
2. Una característica que puede cambiar entre los elementos de una población o de una muestra se denomina:
 - a. Una variable.
 - b. Una muestra.
 - c. Estadística descriptiva.
3. Realizar un censo equivale a:
 - a. Tomar datos de algunos elementos de la población.
 - b. Extraer varias muestras de la población.
 - c. Extraer información de todos los elementos de la población.
4. Los parámetros se relacionan con las:
 - a. Características de la muestra.
 - b. Características de la población.
 - c. Variables numéricas.

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

5. La variable “distancia” entre centros urbanos, es de tipo:

- a. Numérica discreta.
- b. Numérica continua.
- c. Categórica.

En los ítems del 6 al 10, dentro del paréntesis, escriba V si es correcta y F si es falsa.

- 6. () Un ejemplo de escala nominal sería: “El número de pasajeros/as que usan taxis entre las horas de 07h00 y 09h00 los días lunes a viernes en la ciudad de Loja”.
- 7. () La etapa de exploración de los datos está vinculada exclusivamente con las variables categóricas.
- 8. () La estadística inferencial busca extraer conclusiones para la población a partir de la muestra.
- 9. () El muestreo estratificado se caracteriza por dividir la población en grupos, todos de igual tamaño.
- 10. () Una muestra se denomina de conveniencia cuando se eligen los individuos u objetos que van a conformar la muestra, sin aleatorizar.

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](#)



Semana 2



Unidad 2. Exploración de datos

2.1. Tablas de frecuencia

2.2. Representación gráfica de variables (categóricas y numéricas)

Bienvenidos/as a la semana 2 de estadística. Esta semana vamos a comenzar con la unidad 2 sobre la exploración de datos, específicamente vamos a trabajar el apartado 2.1 sobre tablas de frecuencia. Este tipo de tabla es muy importante en cualquier estudio estadístico de una muestra ya que es el primer procedimiento a seguir para ordenar y presentar los datos obtenidos de una forma sencilla. Después trabajaremos el apartado 2.2 sobre tipos de gráficas estadísticas para examinar el lugar y la forma de la distribución de datos de una sola variable o el patrón y la relación entre dos o más variables. Se usa un diagrama circular (i.e., pastel) o diagrama de barras para ver las frecuencias o la cantidad medida de una sola variable categórica. Se usa un histograma o polígono de frecuencias para ver la distribución de una sola variable

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

cuantitativa. En cuanto datos bivariados (i.e., datos que resultan de dos variables medidas en una sola unidad experimental), se puede usar un diagrama de cajas o un diagrama de barras para examinar la relación entre una variable categórica y una cuantitativa. Para ver la relación entre dos variables cuantitativas se usa un diagrama de dispersión.

Las gráficas estadísticas son muy utilizadas, principalmente para visualizar unos resultados.

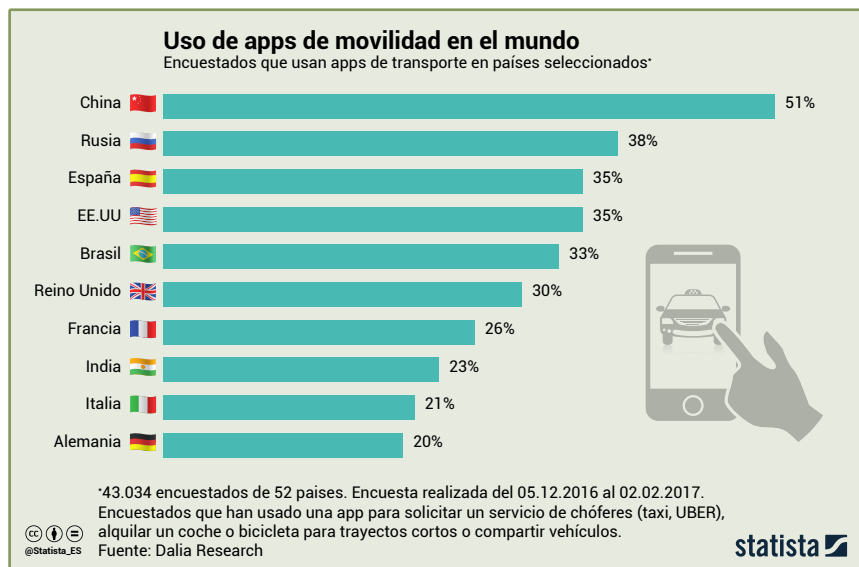
Por ejemplo: Una empresa que se dedica al transporte de materiales pesados y a usted le interesa clasificar el volumen de su venta por provincias, para intensificar la campaña de *marketing* en aquellas provincias que registran menor volumen de ventas: a) ¿Podría utilizar una tabla de frecuencias? b) ¿Qué tipo de gráfica usaría para representar estos datos?

Solución: a) Si. b) Cualquier gráfica bivariada para variables categóricas (diagrama de cajas o diagrama de barras).

Recursos de aprendizaje

Lea detenidamente la unidad 1, apartados 1.3, 1.4 y 1.5 del texto básico (Mendehall *et al.* 2015), con el objetivo de aprender sobre las tablas de frecuencia y sobre algunas de las gráficas estadísticas para datos cualitativos o cuantitativos más comunes. Después de la lectura del texto básico y para finalizar este apartado les animo a practicar con el siguiente ejercicio que les ayudará a interpretar las gráficas estadísticas para datos categóricos.

Gráfico 1.



<https://cdn.statcdn.com/Infographic/images/normal/8714.jpeg>



Actividades de aprendizaje recomendadas

En la gráfica anterior se muestra los 10 países con mayor uso de apps de movilidad y transporte en el mundo. Según los datos de la gráfica, ¿podrían contestar a las siguientes preguntas?

- ¿Qué país presenta el mayor porcentaje de uso de aplicaciones de transporte?
- ¿Cuántos países presentan al menos un 35% del uso de aplicaciones de transporte?

Solución: **a)** Con más del 51% China presenta el mayor porcentaje de encuestados que usan aplicaciones de transporte **b)** Solo 4 países (China, Rusia, España y EEUU).

A través de estos recursos aprenderán qué es y cómo se realiza una tabla de frecuencia, ¿se animan a realizar algún ejemplo ya resuelto de los que han podido encontrar en los recursos de aprendizaje? A través de estos recursos veremos algunos ejemplos de diagramas circulares y de barras y sus aplicaciones. Seguro que les suenan algunas de estas gráficas, pero, ¿serían capaz de realizar alguna de ellas en Excel con un conjunto de datos? Intenten realizar algún ejemplo del apartado 1.3 del texto básico en Excel para practicar.



Semana 3

2.3. Medidas de centralización

2.4. Medidas de variación

2.5. Medidas de posición relativa

En los anteriores apartados se estudió el uso de gráficas para resumir y presentar datos estadísticos; sin embargo, en esta unidad vamos a aprender a sintetizar la información a través de indicadores numéricos. A estos indicadores comúnmente se los conoce como estadísticos descriptivos y los estudiaremos a lo largo de esta semana. Estos estadísticos descriptivos permiten comparar entre muestras y dar una idea rápida de cómo se distribuyen los datos.

Primeramente estudiaremos las medidas de centralización (apartado 2.3) que suelen ser las más utilizadas y nos indican el valor promedio de los datos, o en torno a qué valor se distribuyen

estos. Para comprobar la representatividad de las medidas de centralización usaremos las medidas de variación o de dispersión (apartado 2.4) y las medidas de posición relativa (percentil y cuartil; apartado 2.5).

Recursos de aprendizaje

Como recursos de aprendizaje les recomiendo revisar la unidad 2 del texto básico recomendado (Mendenhall *et al.* 2015). Los apartados 2.1 y 2.2 para medidas de centralización y los apartados 2.3 y 2.6 para medidas de variación y de posición, respectivamente. Para comprobar si han entendido bien el concepto de estadísticos descriptivos, les animo a practicar con el siguiente ejercicio.



Actividades de aprendizaje recomendadas

1. El Ministerio de Transporte y Obras Públicas está interesado en estimar un precio medio del boleto de autobuses en trayectos urbanos en las principales ciudades del país. A continuación se muestran los precios medios de venta en dólares para 10 ciudades del Ecuador:

1.08	0.98	1.09	1.24	1.33	1.14	1.55	1.08
1.22	1.05						

- a. Calcule la media del precio del boleto.
- b. Calcule la mediana del precio del boleto.
- c. ¿Cuál sería el rango de precios de los boletos?

Recuerde: para calcular la mediana, se debe ordenar los precios en orden ascendente. En el caso de que el número de datos es par, la mediana se calcula como el promedio de los dos datos centrales, en

este caso los datos de posiciones 5 y 6. **Solución: a) 1.176; b) 1.115 y d) 0.98-1.55**

- Revisar el contenido en el texto básico recomendado para las unidades 1 y 2.
- Revisa el siguiente recurso educativo abierto ([REA nº 1: Sustainable Urban Mobility Congress](#)) para poder preparar el primer foro académico calificado del primer bimestre.
- Las tablas de frecuencia y las gráficas nos ayudan a visualizar los datos y según el tipo de datos tendremos que usar un tipo de gráfica. ¿Sería capaz de realizar un esquema con los diferentes tipos de gráficas y para qué tipo de datos se utilizan? ¿Podría hacer lo mismo para los principales estadísticos descriptivos?
- Realizar la autoevaluación 2 para repasar los conocimientos adquiridos en esta unidad 2.

¿Qué le ha parecido? ¿Ha entendido bien los conceptos de la tabla de frecuencia? ¿Ha escuchado antes de alguno de los estadísticos descriptivos que estamos estudiando? ¿Cuál le resulta más conocido? Puede expresar su opinión a través de comentarios en el anuncio académico semanal.



Autoevaluación 2

Exploración de datos

- A. Seleccione y marque el literal que corresponde a la opción correcta.**
- La frecuencia relativa es el cociente entre:
 - El total de los datos y la frecuencia absoluta.
 - La frecuencia acumulada y la absoluta.
 - La frecuencia absoluta y el número total de datos.
 - Un diagrama de barras sirve para representar:
 - Una variable categórica.
 - Dos variables numéricas.
 - Una variable numérica y una categórica.
 - La desviación estándar es:
 - El cuadrado de la varianza.
 - La raíz cuadrada de la varianza.
 - El cuadrado de la amplitud.
 - Cuando la distribución de una variable numérica es simétrica, entonces:
 - La media y la varianza coinciden.
 - La media, mediana y moda coinciden.
 - La mediana es menor que la media aritmética.

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

5. El percentil 90:

- a. Deja a su izquierda el 90% de las observaciones.
- b. Está por debajo el 90% de las observaciones.
- c. Deja 10 observaciones a su derecha.

B. Escriba en el paréntesis, V si el enunciado es correcto y F si es falso:

- 6. () La amplitud es una medida de centralización.
- 7. () Se conoce como rango intercuartil a la diferencia entre el valor máximo y el mínimo.
- 8. () La medida central que es insensible ante valores extremos se denomina mediana.
- 9. () El valor más alto en la curva de densidad corresponde a la varianza.
- 10. () El coeficiente de variación es una medida de dispersión que se obtiene dividiendo la desviación estándar para la media aritmética.

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](#)

Resultado de aprendizaje 2

Representa en términos probabilísticos las situaciones generadas en un contexto creado de acuerdo a una problemática.

El resultado de aprendizaje 2 se trabajará hasta el final del primer bimestre (semanas 4, 5, 6, 7 y 8) a través de las unidades 3, 4 y 5. Mediante la lectura del texto básico y algunos ejercicios prácticos aprenderemos sobre la probabilidad y las distribuciones de variables.

Contenidos, recursos y actividades de aprendizaje



Semana 4



Unidad 3. Introducción a la probabilidad

3.1. Conceptos básicos

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

3.2. Reglas fundamentales de la probabilidad

En esta cuarta semana de clases vamos a definir el concepto de la probabilidad y sus reglas fundamentales. Es muy común relacionar términos como probabilidad y estadística, sin embargo, hay que tener claro que la probabilidad es una herramienta usada en la estadística para evaluar la confiabilidad de unas conclusiones acerca de una población cuando tenga solo información muestral (Mendehall *et al.* 2015).

Recursos de aprendizaje

Lea comprensivamente la unidad 4 y para practicar realice los ejemplos 4.5 (p. 128) y 4.17 (p. 142) del texto básico (Mendenhall *et al.* 2015). ¿Ha entendido bien los principios básicos de la probabilidad? Estos principios son la base de la inferencia estadística que estaremos trabajando a lo largo del segundo bimestre, así que le recomiendo que los repase muy bien para facilitar la comprensión de las siguientes unidades temáticas.



Actividades de aprendizaje recomendadas

- Revisar el contenido de la unidad 4 del texto básico recomendado.
- Realizar la autoevaluación 4 para repasar los conocimientos adquiridos en esta unidad 4.

¿Le ha gustado este tema de probabilidad?
¿Considera que la probabilidad está presente en las actividades de nuestra vida diaria?
Puede contestar a estas preguntas a través de comentarios en el anuncio académico de la semana 4.



Autoevaluación 3

Introducción a la probabilidad

Una vez que ha culminado la revisión de los fundamentos teóricos y ejemplos de aplicación de las probabilidades, le recomiendo responder la siguiente autoevaluación conforme se indica en cada enunciado.

En cada uno de los enunciados siguientes, complete con el término adecuado de manera que la afirmación sea verdadera.

1. La probabilidad _____, depende de la repetición del experimento.
2. Los estudios que se efectúan sin modificar las condiciones del entorno se denominan _____
3. La probabilidad del espacio muestral es igual a la unidad siempre que los sucesos sean _____
4. La probabilidad de un suceso A condicionada a un suceso B, será nula cuando _____
5. Dos sucesos son independientes cuando la probabilidad de la intersección se expresa como _____ de las probabilidades de cada suceso.

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

En los siguientes ítems, seleccione y encierre el literal que corresponde a la respuesta correcta.

6. La expresión divide y vencerás se relaciona con.
 - a. El valor de una probabilidad.
 - b. La probabilidad de la unión de sucesos.
 - c. El teorema de la probabilidad total.
7. Por sus propiedades, las probabilidades se relacionan con.
 - a. La teoría de conjuntos.
 - b. La física.
 - c. La geometría.
8. Si la probabilidad de que una torre eléctrica resista el viento hasta una velocidad de 200 km/hora es igual a X , entonces la probabilidad de no resistir el viento a una velocidad mayor de 200 km/hora será igual a.
 - a. 1.
 - b. $1-X$.
 - c. $X-1$.
9. ¿Cuándo dos sucesos M y N son mutuamente excluyentes?
 - a. $P(M \cup N) = P(M) - P(N)$.
 - b. $P(M \cup N) = P(M) + P(N)$.
 - c. $P(M \cup N) = 1 - (P(M) + P(N))$.

10. Un conjunto está formado de 4 elementos, ¿cuántos arreglos de dos en dos, sin importar el orden serían?

- a. 6.
- b. 8.
- c. 12.

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](#)



Semana 5



Unidad 4. Distribuciones de variables

4.1. Distribución binomial de la probabilidad

4.2. Distribución de la probabilidad de Poisson

Esta semana vamos a iniciar uno de los temas más importantes del semestre sobre las distribuciones de variables. Durante las próximas dos semanas aprenderemos sobre tres de las distribuciones más importantes y más usadas en la estadística: la distribución binomial, la distribución de Poisson y la distribución normal.

Por ejemplo: si una empresa de logística ecuatoriana dedicada al envío de paquetería internacional afirma que el 90% de sus paquetes llegan a tiempo, la agencia de protección al consumidor quizás verifique si es cierta esa afirmación con una muestra de 10 paquetes. Si la afirmación es cierta, es improbable que un paquete llegue con retraso a su destino y es muy probable que 9 de cada 10 paquetes sean entregados en el tiempo estimado. Esto sería un

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

ejemplo de una variable aleatoria discreta que se podría modelar con la distribución de Poisson.

Finalmente, aprenderemos a identificar y calcular las probabilidades de una distribución de Poisson, basada en variables aleatorias discretas, que en el ámbito de las empresas se suele utilizar como modelo para describir la distribución de errores en una entrada de datos, el número de imperfecciones en un producto finalizado, el número de partes defectuosas en envíos, el número de clientes que se esperan tener en un periodo de tiempo, etc.

Recursos de aprendizaje

Para empezar y a modo de introducción general de este capítulo les recomiendo revisar los aspectos relacionados con las variables aleatorias discretas y sus distribuciones de probabilidad (Mendenhall *et al.* 2015, unidad 4, apartado 4.8) para luego comenzar a estudiar las principales distribuciones de variables discretas comenzando con la distribución *binomial* (unidad 5, apartado 5.2) y de *Poisson* (apartado 5.3) del texto básico (Mendenhall *et al.* 2015).



Semana 6

4.3. Distribución normal de la probabilidad

En los dos primeros apartados de la unidad 4 hemos estudiado las distribuciones de probabilidad de variables aleatorias discretas, pero en este apartado 4.3 estudiaremos las distribuciones de las variables aleatorias continuas, concretamente la distribución normal.

La distribución normal es con diferencia la distribución de variables más utilizada en la estadística. Específicamente aprenderemos: i) cómo estandarizar una variable continua, ii) que el área bajo la curva normal siempre equivale a 1 y iii) a calcular áreas bajo la curva normal estándar que nos permite asignar valores de probabilidad a una variable aleatoria continua X .

Recursos de aprendizaje

Esta semana como recurso de aprendizaje deben revisar la unidad 6, apartados 6.1, 6.2 y 6.3 del texto básico recomendado (Mendenhall *et al.* 2015) para estudiar la distribución normal de probabilidad.

No se olvide de que una distribución de probabilidad normal es simétrica, el área bajo la curva a la izquierda de la media es 0.5 y el área bajo la curva a la derecha de la media es también 0.5. Por ejemplo, suponiendo que el precio de los boletos de buses en Quito sigue una distribución normal y su media es de \$0.35, la probabilidad de que un boleto tenga un costo mayor o menor a \$0.35 es de 0.5, respectivamente. Pero si queremos calcular la probabilidad acumulada de que un boleto cueste entre \$0.25 y \$0.50, debemos utilizar la distribución de probabilidad normal estándar que presenta una media de 0 y una desviación estándar de 1.



Actividades de aprendizaje recomendadas

- Para repasar la distribución normal, les recomiendo revisar el siguiente recurso educativo abierto ([REA nº 2: Distribución de probabilidad normal](#)) donde se explica con detalle cómo calcular en la práctica con Excel una distribución normal.

- Realizar la autoevaluación 4 para repasar los conocimientos adquiridos en esta unidad 4.

¿Ha entendido bien el contenido de esta unidad 4? Después de leer los recursos de aprendizaje propuestos, ¿sería capaz de plantear algún ejemplo de una distribución normal, de una distribución binomial y de una distribución de Poisson?

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas



Autoevaluación 4

Distribuciones de variables

Finalmente, para repasar los contenidos de esta unidad, le animo a realizar la siguiente autoevaluación que se presenta a continuación.

A. Lea con atención los enunciados del 1 al 5 y marque la opción correcta:

1. Un fenómeno se dice aleatorio cuando:
 - a. Es posible predecir sus resultados.
 - b. No tiene posibilidad de ocurrencia.
 - c. No existe certeza de los resultados que ocurrirán.
2. Entre los siguientes ejemplos, identifique aquel que corresponde a una variable aleatoria binomial:
 - a. Encuestar a un grupo de personas para identificar si conocen o no la normativa ambiental.
 - b. Seleccionar una persona que puede conocer o no las formas de reciclar los residuos sólidos.
 - c. Muestrear 10 puentes para identificar si están o no en riesgo de colapsar.
 - d. Todas las respuestas son correctas.
 - e. Ninguna respuesta es correcta.

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

3. La distribución normal:
- a. Es única y posee los parámetros media y tamaño de muestra.
 - b. Está compuesta por una familia de distribuciones normales cada una con sus parámetros media y varianza (o desviación estándar).
 - c. Es apta para calcular probabilidades relacionadas con variables numéricas discretas.
4. La curva normal estándar se caracteriza por:
- a. Posee parámetros desconocidos.
 - b. La desviación estándar es igual a 0 y la media es igual a 1.
 - c. La media es igual a 0 y la desviación estándar es igual a 1.
5. Los valores de la variable Z:
- a. Se usan para caracterizar una distribución binomial.
 - b. Pueden tomar valores positivos y negativos.
 - c. Siempre serán positivos, ya que para obtenerlos se divide por un valor positivo.
- B. Escriba en el paréntesis, V si el enunciado es correcto y F si es falso:**
6. () Simbólicamente la función de distribución acumulada de probabilidad se representa por $P(X \geq x)$.
7. () Los parámetros que definen la distribución binomial son: la probabilidad de éxito p y el número de ensayos n .

8. () El valor esperado de una variable aleatoria discreta se define como el valor que puede asumir la variable con mayor probabilidad de ocurrencia.
9. () La curva normal estándar es simétrica respecto a todas las medidas de tendencia central.
10. () Una característica de la distribución normal estándar es que el área total bajo la curva es igual a 2.

Luego de haber respondido la autoevaluación, le recomiendo revisar el solucionario, en caso de no haber acertado en algún ítem, se sugiere volver a revisar los contenidos.

[Ir al solucionario](#)

[Índice](#)

[Primer
bimestre](#)

[Segundo
bimestre](#)

[Solucionario](#)

[Referencias
bibliográficas](#)



Semana 7



Unidad 5. Introducción a la estadística inferencial

5.1. Estadística inferencial: definiciones y métodos

5.2. Distribución muestral

5.3. Control estadístico de procesos

Durante esta semana 7 vamos a iniciar la unidad 5 sobre la estadística inferencial. Primeramente veremos el apartado 5.1 con los diferentes procedimientos de muestreo para continuar con el apartado 5.2 sobre distribuciones muestrales, el Teorema del Límite Central y cálculos de la distribución muestral de la media, utilizando el cálculo del error estándar. Finalmente se realizará una breve introducción al control estadístico de procesos en el apartado 5.3.

Recordemos que al estudiar las probabilidades para el estudio de fenómenos aleatorios seguíamos un razonamiento deductivo: establecíamos ciertas hipótesis con un modelo de distribución y

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

a partir de ahí se deducían las propiedades sobre el fenómeno en cuestión (p.e., media, varianza, probabilidad, etc.). Por el contrario, mediante la inferencia estadística realizamos el proceso inverso a partir de un conjunto de datos experimentales, se infieren o se estiman 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.

Recursos de aprendizaje

Como recursos de aprendizaje en esta semana deben revisar la unidad 7, apartados 7.3-7.7 del texto básico recomendado (Mendenhall *et al.* 2015).

Este contenido es muy importante ya que identificar y realizar bien un tipo de muestreo nos permitirá obtener mejores resultados en su análisis estadístico. Dependiendo del objetivo de nuestra investigación, tendremos que decidir cuál tipo de muestreo realizar, por ello, es importante que tengamos muy claro las diferencias y los usos de cada uno de estos muestreos. ¿Se animan a proponer un ejemplo de cada uno de los muestreos?



Actividades de aprendizaje recomendadas

- Realización de alguna de las actividades propuestas en el apartado 7.5 sobre el cálculo del error estándar del texto básico recomendado.
- Realizar la autoevaluación 5 para repasar los conocimientos adquiridos en esta Unidad 5.

¿Ha entendido bien los diferentes tipos de muestreo? Sería recomendable realizar un resumen con las principales características de los diferentes tipos de muestreo y con las principales características del Teorema del Límite Central para asumir la forma aproximada de la distribución muestral de la media.

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



Autoevaluación 5

Introducción a la estadística inferencial

Luego de haber culminado con la revisión de las distribuciones muestrales, tanto para la media como para la proporción, su construcción y aplicaciones le recomiendo responder la siguiente evaluación.

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

1. () El muestreo estratificado se caracteriza por seleccionar aleatoriamente algunos estratos para realizar el muestreo.
2. () El muestro por conglomerados busca construir grupos homogéneos a partir de una población homogénea.
3. () En la selección del tamaño de una muestra aleatoria, a mayor margen de error muestral (E) se obtendrá una muestra más pequeña.
4. () El valor de un estadístico, se calcula a partir de los datos de la muestra.
5. () El teorema del límite central dice que, a mayor número de muestras, la distribución se aleja de la normalidad.

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

B. Encierre el literal que corresponde a la opción correcta

6. 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 mayor para la distribución original.
 - b. El rango de valores en el eje X es menor en la distribución original.
 - c. Alejamiento de la media muestral vs la media de la variable original.
7. 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 del tamaño de muestra.
 - c. El tamaño de muestra y la desviación estándar.
8. La distribución de la proporción muestral se aproxima a una normal cuando:
 - a. Para cualquier tamaño de muestra.
 - b. El tamaño de muestra es pequeño.
 - c. El tamaño de muestra n es grande.
9. El procedimiento de muestreo sistemático consiste en seleccionar:
 - a. El primer elemento de la muestra al azar.
 - b. Todos los elementos de la muestra al azar.
 - c. Elementos equidistantes todos al azar.

10. En el muestreo estratificado, los estratos son:

- a. Cualquier grupo de la población.
- b. Cada elemento de la población.
- c. Grupos homogéneos que se forman conforme algún criterio.

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.

[Ir al solucionario](#)

[Índice](#)

[Primer
bimestre](#)

[Segundo
bimestre](#)

[Solucionario](#)

[Referencias
bibliográficas](#)



Semana 8

Repaso del primer bimestre

Por fin llegamos a la última semana del semestre, ¡¡¡buen trabajo!!!
Para esta semana está planificado repasar todo el contenido impartido en las últimas 7 semanas antes de la primera evaluación presencial.



Actividades de aprendizaje recomendadas

- Revisar el contenido en el texto básico y complementario recomendado para las unidades 1, 2, 3, 4 y 5 como repaso antes de la evaluación.

¿Le ha resultado complicado este bimestre?
Después de repasar el contenido impartido, ¿sería capaz de interpretar una gráfica estadística?
¿Ha entendido bien los diferentes estadísticos descriptivos y su aplicación? ¿Podría poner un ejemplo de distribución binomial y de distribución de Poisson? Le animo a revisar los esquemas y resúmenes que hemos realizado a lo largo del semestre para repasar los contenidos antes de la evaluación presencial. ¡Mucha suerte, nos vemos en el segundo bimestre!

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



Actividades finales del bimestre

Actividad 1: dedique esta semana a terminar de estudiar y repasar las unidades 1, 2, 3, 4 y 5 a través de las lecturas recomendadas en el texto básico.

Actividad 2: si aún no lo ha hecho realice las diferentes autoevaluaciones para repasar los contenidos impartidos este primer bimestre.

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

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas



Segundo bimestre

Resultado de aprendizaje 3

Aplica los conceptos adecuados para almacenar y procesar datos.

Este resultado de aprendizaje se trabajará a través de las unidades 6, 7 y 8 sobre intervalo de confianza, prueba de hipótesis y regresión lineal, durante las 8 semanas de este segundo bimestre. El objetivo de este resultado de aprendizaje es que al final del segundo semestre cada estudiante será capaz de reconocer y diseñar un muestreo y de extraer conclusiones y generalizaciones sobre una población basándose en la información suministrada por la muestra.

Contenidos, recursos y actividades de aprendizaje



Semana 9



Unidad 6. Intervalo de confianza

6.1. Estimación puntual

6.2. Construcción e interpretación de un intervalo de confianza

Esta semana vamos a comenzar con la unidad 6 sobre el intervalo de confianza, específicamente vamos a trabajar los apartados 6.1 y 6.2 sobre estimación puntual y la construcción e interpretación de un intervalo 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 paquetes que se envían de Quito a Loja 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 confía que contenga el parámetro poblacional, con un determinado nivel de confianza se le denomina intervalo de confianza que se estudiará en el apartado 6.2.

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

Recursos de aprendizaje

Leer comprensivamente la unidad 8, apartados 8.1, 8.2, 8.3, 8.4 y 8.5 del texto básico (Mendenhall *et al.* 2015).



Semana 10

6.3. Intervalo de confianza de muestra grande para una media y una proporción poblacional

6.4. Intervalo de confianza para la diferencia entre dos medias poblacionales y entre dos proporciones binomiales

Esta semana 10 vamos a finalizar la unidad 6. Primero estudiaremos los intervalos de confianza de muestra grande para una media y una proporción poblacional (apartado 6.3). Luego estudiaremos los intervalos de confianza de muestra grande para la diferencia entre dos medias poblacionales y entre dos proporciones binomiales (apartado 6.4). Por lo general, con muestras grandes ($n > 30$) la estimación de estos parámetros se hace por aproximación a la normal.

Recursos de aprendizaje

Lea comprensivamente la unidad 8, apartados 8.6 y 8.7 del texto básico (Mendenhall *et al.* 2015). Después de leer los recursos de aprendizaje, ¿sería capaz de estimar una media o una proporción poblacional? A veces, para entender mejor el contenido es necesario realizar algún ejercicio práctico. Por ello, le animo a realizar los

ejemplos 8.4 y 8.5 ya resueltos del texto básico (Mendenhall *et al.* 2015) que se encuentran en la unidad 8, pp. 287-288.



Actividades de aprendizaje recomendadas

- Lectura del texto básico recomendado.
- Realización de alguna de las actividades propuestas en los recursos de aprendizaje.
- Para repasar el intervalo de confianza le recomiendo revisar el siguiente recurso educativo abierto ([REA nº 3: Intervalos de confianza](#)) dónde se explica con detalle como calcular en la práctica con Excel un intervalo de confianza.
- Realizar la autoevaluación 6 para repasar los conocimientos adquiridos en esta unidad 6.

Después de la lectura del texto básico recomendado, ¿ha entendido bien el concepto de intervalo de confianza? ¿Podría pensar en algún ejemplo donde se pueda aplicar el cálculo del intervalo de confianza para dos medias poblacionales?

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas



Autoevaluación 6

Intervalo de confianza

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

1. La estadística inferencial busca dar respuesta a dos problemas básicos:
 - a. Calcular tamaños de muestras y estimar la media poblacional.
 - b. Estimar parámetros y probar hipótesis.
 - c. Plantear hipótesis y hacer gráficas.
2. El punto central de un intervalo de confianza 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.

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

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.
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.

En los literales del 6 al 10, escriba entre paréntesis V si el enunciado es verdadero o F si es falso.

6. () El error estándar de la media muestral se relaciona inversamente con la desviación estándar muestral.
7. () El error estándar de la media muestral se define como el cociente entre la varianza y el tamaño de la muestra.
8. () El intervalo de confianza para la proporción poblacional se emplea cuando se trata con variables binomiales o cualitativas.
9. () Buscando estimar la edad media de los estudiantes que ingresan a la universidad, al 99% se obtuvo el rango $[17.5 - 19.5]$, entonces el estimador puntual de la media fue 18.5.

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

10. () Cuando no se conoce la varianza poblacional para estimar la media utilizamos la puntuación normal estándar Z .

Luego de haber respondido la autoevaluación, compare con el solucionario y realice los correctivos si fuera necesario.

[Ir al solucionario](#)

[Índice](#)

[Primer bimestre](#)

[Segundo bimestre](#)

[Solucionario](#)

[Referencias bibliográficas](#)



Semana 11



Unidad 7. Prueba de hipótesis

7.1. Procedimientos de una prueba estadística de hipótesis

7.2. Una prueba de muestra grande acerca de una media poblacional

Esta semana vamos a estudiar la unidad 7 sobre una prueba de hipótesis. Al final de esta unidad usted debe ser capaz de: i) formular una hipótesis nula H_0 y su correspondiente hipótesis alternativa H_a , ii) contrastar la H_0 con los resultados de la muestra, iii) calcular la probabilidad de encontrar el resultado (valor P) y iv) interpretar los resultados decidiendo si se rechaza o no la hipótesis nula.

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](#)

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

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

Recursos de aprendizaje

Para estudiar el contenido de esta semana es necesario leer comprensivamente la unidad 9, apartados 9.1, 9.2 y 9.3 del texto básico recomendado (Mendenhall *et al.* 2015).

¿Ha entendido bien el significado del contraste de hipótesis? Demuestre sus conocimientos a través del siguiente juego, donde usted tendrá que relacionar las diferentes partes que componen una prueba estadística de hipótesis con su definición. ¡Suerte!

Prueba de hipótesis



Semana 12

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

7.4. Una prueba de hipótesis para una proporción binomial

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

En esta semana 12 continuaremos con las pruebas de hipótesis para la diferencia entre dos medias poblacionales (unidad 7, apartado 7.3), 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). Finalmente acabaremos la semana estudiando los dos últimos apartados de la unidad 7 sobre una prueba de hipótesis para una proporción binomial (apartado 7.4) y para la diferencia entre dos proporciones binomiales (apartado 7.5).

Recursos de aprendizaje

Esta semana como recurso de aprendizaje deben revisar la unidad 9, apartados 9.4, 9.5 y 9.6 del texto básico recomendado (Mendenhall *et al.* 2015).



Actividades de aprendizaje recomendadas

- Para repasar una prueba de hipótesis, le recomiendo revisar el siguiente recurso educativo abierto ([REA nº 3: Contraste de hipótesis sobre la media de una población](#)) **dónde se explica con detalle** dos procedimientos para realizar en la práctica con Excel una prueba de hipótesis acerca de una media.



Semana 13

7.6. Prueba t de student

En esta semana vamos a terminar la unidad 7 estudiando 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.

Recursos de aprendizaje

Como recursos de aprendizaje en esta semana deben revisar la unidad 10, apartados 10.1, 10.2, 10.3 y 10.4 del texto básico recomendado (Mendenhall *et al.* 2015). Para reforzar los contenidos 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 7 para repasar los conocimientos adquiridos en esta unidad 7.

¿Qué le ha parecido esta unidad de contraste de hipótesis? ¿Ha entendido bien los conceptos? Para reforzar el contenido le recomiendo realizar un resumen de los principales procedimientos para una prueba de hipótesis estadística.



Autoevaluación 7

Prueba de hipótesis

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.
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.

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

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:
- Rechazar la hipótesis.
 - No rechazar la hipótesis.
 - Reservarse en su decisión.
 - No se debe comparar el valor P con α .
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:
- El valor de Z .
 - El valor de t .
 - El nivel de significancia.
 - 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:
- Aceptar una H_0 que es verdadera.
 - Rechazar una H_0 que es verdadera.
 - Aceptar una H_0 que es falsa.

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 $z = 2,33$. (i.e., Mendenhall *et al.* 2015).

¿Cuál es la hipótesis nula? Use el formato que usamos en clase

- 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, 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 las 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 $z = 2,33$. (i.e., Mendenhall *et al.* 2015).

¿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. $H_a: \mu > 675$.

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. $F = 715$.
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 que 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 que de las médicas.
 - e. No se puede contestar la pregunta basada en la información dada.
11. Se aplica una prueba de t cuando la variable independiente es:
- a. Continua.
 - b. Categórica y hay solo 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.

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

12. La forma de la distribución t depende de
- El nivel de significancia α .
 - Los grados de libertad df .
 - El tamaño muestral n .
 - Respuestas a y b son correctas.
 - Respuestas b y c son correctas.
 - Todas las respuestas son correctas.
13. Un diseño de muestras independientes para la diferencia entre dos medias poblacionales requiere que:
- Las características de los individuos entre las dos muestras sean iguales.
 - Las dos muestras tengan la misma cantidad de observaciones.
 - Los individuos que constituye una muestra son completamente distintos de los de la otra muestra.
 - Las varianzas de las muestras son iguales.
 - Respuestas a y b son correctas.
14. ¿Qué suposiciones se hacen cuando se usa la prueba t para probar una hipótesis respecto a dos medias poblacionales?
- Las muestras deben ser aleatorias e independientes.
 - Las poblaciones en las que se haga el muestreo deben estar normalmente distribuidas.
 - Las poblaciones en las que se haga el muestreo deben tener la misma varianza: $\sigma_1 = \sigma_2$
 - Respuestas a y c son correctas.
 - Todas las respuestas son correctas.

15. La prueba t para dos grupos independientes compara:

- a. Las varianzas.
- b. Las desviaciones estándar.
- c. Las medias.

[Ir al solucionario](#)

[Índice](#)

[Primer
bimestre](#)

[Segundo
bimestre](#)

[Solucionario](#)

[Referencias
bibliográficas](#)



Semana 14



Unidad 8. Regresión lineal simple y correlación

8.1. Regresión lineal simple (calidad del modelo, evaluación de suposiciones)

Esta semana vamos a estudiar la unidad 8 que será la última unidad del semestre sobre regresión lineal y correlación. Cuando queremos analizar datos es importante centrarse en dos grandes objetivos: comparar grupos y estudiar relaciones. Por ello, en esta unidad 8 vamos a estudiar la regresión lineal y 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 demasiados técnicos o teóricos del análisis. Enfocaremos en cuándo y cómo utilizar el análisis de regresión lineal y cómo interpretar los resultados.

En esta semana 14 empezamos con el apartado 8.1 sobre regresión lineal simple. El análisis de regresión lineal nos permite estudiar la relación entre una variable llamada dependiente (Y) y una o más

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

variables llamadas independientes o predictoras (X), así como para desarrollar una ecuación lineal con fines predictivos.

Recursos de aprendizaje

Leer comprensivamente la unidad 12, apartados 12.1, 12.2, 12.3, 12.4 y 12.5. del texto básico (Mendenhall *et al.* 2015). Después de leer los recursos de aprendizaje, ¿qué le pareció el análisis de regresión lineal? ¿Entendió bien el contenido? Sería recomendable realizar un breve resumen sobre las principales características del análisis de regresión lineal.



Actividades de aprendizaje recomendadas

- Para profundizar su conocimiento sobre regresión simple, les recomiendo revisar el [REA nº 3 \(Capítulo 2. Modelos Lineales, apartados 1, 2 y 3.1\)](#).



Semana 15

8.2. Correlación bivariada (Pearson, Spearman)

En la semana anterior, vimos la regresión lineal y cómo es posible estudiarla a través de un diagrama de dispersión. Pero 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 8.2). 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

Para una mejor comprensión de esta unidad 8, les recomiendo leer comprensivamente la unidad 12, apartados 12.6, 12.7 y 12.8 del texto básico (Mendenhall *et al.* 2015). Como última actividad del bimestre, los ánimo a realizar un pequeño resumen sobre las suposiciones de regresión, la recta ajustada y el análisis de correlación que les ayudará mucho a entender mejor esta unidad.



Actividades de aprendizaje recomendadas

- Realizar la autoevaluación 8 para repasar los conocimientos adquiridos en esta unidad 8.

¿Qué le pareció el análisis de regresión lineal?
¿Entendió bien el contenido de correlación?
Intente realizar un resumen con las principales características de este análisis.

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas



Autoevaluación 8

Regresión lineal simple y correlación

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

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 β_0 por cada unidad de incremento de β_1 .
 - El incremento de β_1 por cada unidad de incremento de x.
 - El valor de y cuando β_0 y β_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?
- La intersección y los residuos.
 - La variable dependiente y la variable dependiente.
 - La pendiente y la intersección.
 - Los residuos y los grados de libertad.
 - Los grados de libertad y la pendiente.

[Ir al solucionario](#)



Semana 16

Repaso del segundo bimestre

¡Enhorabuena! Ya estamos en la última semana del semestre. Durante 16 semanas hemos revisado las bases de la estadística y espero haber aportado numerosas herramientas que luego van a poder aplicar tanto en el resto de la carrera como en su vida profesional.

En esta última semana del semestre la dedicaremos a repasar los contenidos impartidos durante este segundo bimestre. Para ello, hay que repasar el contenido del texto básico indicado para estudiar las unidades 6, 7 y 8. Deberán estudiar bien todos los resúmenes y esquemas que debieron ir elaborando por cada una de las unidades temáticas y por último realizar todas las autoevaluaciones de nuevo para comprobar el nivel de conocimientos. Recuerden que pueden preguntar cualquier duda en el horario de tutorías.



Actividades finales del bimestre

Actividad 1: dedique esta semana a terminar de estudiar y repasar las unidades 6, 7 y 8 a través de las lecturas recomendadas en el texto básico.

Actividad 2: si aún no lo ha hecho, realice las diferentes autoevaluaciones para repasar los contenidos impartidos este primer bimestre.

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



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	b	Generalmente se describen dos ramas de la estadística: estadística descriptiva e inferencial.
2	a	Variable por definición es aquella que cambia entre individuos (o elementos) de un conjunto, muestra o población de estudio.
3	c	Censar significa levantar información de todos los elementos que conforman la población.
4	b	Los parámetros tienen relación con la población, y los estadísticos con la muestra.
5	b	El área basal de un árbol puede expresarse también en decimales (fracciones), por tanto es continua.
6	F	Al hablar de cantidad de hectáreas, se hace referencia a una variable numérica.
7	F	Cuando se realiza exploración de datos, se toma en cuenta los diversos tipos de variables, no exclusivamente las categóricas.
8	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.
9	F	Estratificar significa separar la población en varios estratos o grupos de acuerdo a cierto criterio, eso implica que puede haber a la vez grupos grandes y grupos pequeños.

Autoevaluación 1		
Pregunta	Respuesta	Retroalimentación
10	V	Seleccionar la muestra sin criterio aleatorio, sino de forma direccionada, se denomina muestra de conveniencia.

[Ir a la autoevaluación](#)

[Índice](#)

[Primer bimestre](#)

[Segundo bimestre](#)

[Solucionario](#)

[Referencias bibliográficas](#)

Autoevaluación 2		
Pregunta	Respuesta	Retroalimentación
1	c	Obtener la frecuencia relativa, implica dividir la absoluta para el número total de individuos.
2	a	Las variables categóricas generalmente se representan mediante diagramas de barras, aunque también mediante diagrama de sectores.
3	b	La varianza es el cuadrado de la desviación estándar.
4	b	Las medidas de tendencia central coinciden en distribuciones simétricas.
5	a	Un valor percentil de orden K, deja tras de sí (a la izquierda) el K% de las observaciones.
6	F	La amplitud sirve para cuantificar la variación de un conjunto de datos.
7	F	El rango inter-cuartil se obtiene de la diferencia entre el primer y tercer cuartil.
8	V	A la mediana no le afecta si en un conjunto de datos se incluye un valor extremadamente alto o bajo, porque para su cálculo solo toma en cuenta los datos centrales.
9	F	Por definición, el valor más alto en una gráfica de distribución se corresponde con la moda.
10	V	Coefficiente de variación o dispersión relativa se obtiene de dividir la desviación estándar para la media aritmética, y ese resultado puede expresarse en porcentaje.

Ir a la
autoevaluación

Autoevaluación 3		
Pregunta	Respuesta	Retroalimentación
1	Frecuentista	Efectuar un experimento en repetidas ocasiones y a partir de ello estimar la probabilidad de un suceso específico, se denomina proceso frecuentista.
2	Observacionales	Generalmente los estudios de investigación son de tipo experimentales y observacionales, en los primeros no se cambia ninguna condición y en los segundos se modifica o controla uno o varios factores.
3	Excluyentes	Sucesos excluyentes son aquellos que no comparten elementos o valores de la variable aleatoria, por ejemplo, en el lanzamiento de la moneda los dos sucesos (cara y sello) son excluyentes porque no hay un resultado que sea cara y sello a la vez y la probabilidad del espacio muestral es la suma de las probabilidades. En este caso: $0.50+0.50=1$.
4	Los sucesos son disjuntos	La probabilidad condicionada es el cociente entre la probabilidad de la intersección y la probabilidad del suceso condicionante, por tanto, un cociente será cero solo cuando el denominador sea cero, es decir cuando los sucesos no compartan elementos.
5	Producto	Sucesos independientes es equivalente a sucesos excluyentes y su intersección no involucra elementos. Por la regla del producto: $P(A \text{ y } B) = P(A)*P(B)$.
6	c	El teorema de la probabilidad total tiene como objetivo dividir la región (o probabilidad) objetivo en varias subregiones.
7	a	Las leyes de la teoría de conjuntos son equivalentes a las leyes de las probabilidades.
8	b	Por ley del complemento, dado un suceso A, entonces: $P(A)= 1 - P(A^c)$.
9	b	Por la regla de la suma habría que sumar las dos probabilidades y restar la probabilidad de la intersección, pero al ser excluyentes, dicha probabilidad de intersección es nula.

Autoevaluación 3		
Pregunta	Respuesta	Retroalimentación
10	a	El número de formas sería: $4! / (2! * (4-2)!)$, donde el símbolo “!” representa al factorial.

Ir a la
autoevaluación

Índice

Primer
bimestre

Segundo
bimestre

Solucionario

Referencias
bibliográficas

Autoevaluación 4		
Pregunta	Respuesta	Retroalimentación
1	c	La aleatoriedad se relaciona con la incertidumbre, que es un factor que está presente en los fenómenos que se estudian con la estadística.
2	b	Una variable aleatoria binomial tiene solo dos sucesos o posibilidades. Por falta de un mejor nombre, el resultado uno se llama éxito y el otro fracaso. Por ejemplo, lanzar la moneda una vez, el resultado solo puede ser cara o sello. Cada uno de estos casos en un ejemplo semejante en el cual hay solo dos opciones (sí/no).
3	b	La distribución normal no es única, sino, una familia de distribuciones.
4	c	La distribución normal estándar posee parámetros fijos, por ello se simboliza como $N(0,1)$.
5	a	La operación para obtener la variable Z implica restar la media de la variable numérica X, por tanto, el resultado puede ser positivo o negativo.
6	F	La distribución de probabilidad acumulada hace referencia al concepto de percentil, aquella región que está a la izquierda de cierto valor ($\leq x$).
7	V	La ley binomial se caracteriza por dos parámetros: el número de ensayos y la probabilidad de éxito del suceso en estudio.
8	V	Gráficamente el valor esperado se relaciona con la barra (o línea) más alta, es decir, que representa la mayor probabilidad.
9	V	La curva normal estándar es simétrica respecto a todas las medidas de tendencia central.
10	F	El área total bajo la curva en una distribución normal estándar es igual a 1.

Ir a la
autoevaluación

Autoevaluación 5		
Pregunta	Respuesta	Retroalimentación
1	F	Cuando se realiza un proceso de muestreo estratificado, se toma en cuenta todos los estratos y se extrae una submuestra de cada estrato.
2	F	El muestreo por conglomerados se basa en subdividir la población en varios grupos heterogéneos.
3	V	El margen de error tiene relación inversa con el tamaño de una muestra aleatoria.
4	V	Los estadísticos proceden de la información de la muestra y sirven para estimar los parámetros de la población.
5	F	La distribución de datos numéricos tiende más hacia la normalidad, mientras el tamaño de muestra incrementa.
6	a	La distribución de la media muestral construida a partir de datos remuestreados, tienden hacia el valor central, de manera que los promedios de las muestras van a generar un rango más estrecho que los datos originales.
7	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 .
8	c	La distribución binomial (aquella que se ajusta a proporciones) se aproxima a una normal para muestras grandes.
9	a	El muestreo sistemático es un tipo de muestreo aleatorio, con la característica que únicamente el primer elemento de la muestra se escoge al azar.
10	c	Para aplicar muestreo estratificado se debe disponer de un criterio de estratificación. Por ejemplo, se divide la población de ecuatorianos en habitantes por regiones: costa, sierra, amazonía e insular.

Ir a la
autoevaluación

Autoevaluación 6		
Pregunta	Respuesta	Retroalimentación
1	b	La estimación de parámetros y la prueba de hipótesis, son dos problemas básicos que se abordan con las técnicas de estadística inferencial.
2	a	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.
6	F	El error estándar se obtiene del cociente entre la desviación estándar y la raíz cuadrada del tamaño de muestra, por tanto, la relación con la desviación estándar es directa.
7	F	El cociente entre la desviación estándar, no la varianza.
8	V	A partir de variables cualitativas, es posible estimar proporciones.
9	V	El estimador puntual se ubica en el centro del intervalo de confianza.
10	F	Se emplea la distribución normal estándar (Z) para realizar estimaciones de la media, cuando se conoce la varianza poblacional; caso contrario se emplea la distribución t de <i>student</i> .

Ir a la
autoevaluación

Autoevaluación 7		
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 α , 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	b	Cuando el valor P – definido como la probabilidad de observar el estadístico de muestra cuando H_0 sea correcta – es mayor que α (i.e., alta), no se rechaza H_0 .
5	c	El nivel de significancia α 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 $n > 30$, 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: $z = (x - \mu_0) / (s / \sqrt{n})$

Autoevaluación 7		
Pregunta	Respuesta	Retroalimentación
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.
11	b	La prueba t es un caso particular del análisis de varianza (ANOVA) en el que solo hay dos categorías o niveles de la variable independiente.
12	e	La forma de la distribución t depende del tamaño muestral n . Los grados de libertad (<i>degrees of freedom</i> en inglés) para una muestra, en cambio, depende de n : $df = n - 1$.
13	g	Dos suposiciones para la prueba t de dos muestras independientes son (i) independencia de las unidades muestrales y (ii) igualdad de varianzas.
14	e	Todas son suposiciones para la prueba t y para los modelos lineales en general, por ejemplo ANOVA y regresión.
15	c	El propósito de la prueba t es probar si hay diferencia estadística entre las medias de dos poblaciones.

Ir a la
autoevaluación

Autoevaluación 8		
Pregunta	Respuesta	Retroalimentación
1	V	La recta de mejor ajuste es aquella que minimiza las distancias de los puntos observados a la recta ajustado 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 b_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 . β_0 y β_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	c	Los parámetros que un análisis de regresión pretende estimar son la pendiente (β_1) y la intersección (β_0).

Ir a la
autoevaluación



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-González M.A., Sánchez-Villegas A. y Faulin Fajardo J. (2014). **Bioestadística Amigable** (3ª edición). Barcelona, España: Elsevier.

Recursos educativos abiertos (REAs)

Título del REA	Enlace
REA nº 1. Sustainable Urban Mobility Congress (SUMBILBAO`19)	enlace web
REA nº 2. Distribución de probabilidad normal	enlace web
REA nº 3. Estadística <i>online</i> (YouTube)	enlace web
REA nº 4. Curso de análisis de datos ecológicos en R – Luis Cayuela	enlace web