

**Práctica 1**

***Alumno****: Francisco Javier Piqueras Martínez*

***Asignatura****: Modelado Estadístico de Datos*

***Fecha******de******entrega****: 16 de diciembre de 2019*

Índice

[1. Descripción del documento 3](#_Toc27423538)

[2. Ejercicios 3](#_Toc27423539)

[2.1. Ejercicio 1 3](#_Toc27423540)

[2.2. Ejercicio 2 4](#_Toc27423541)

[2.3. Ejercicio 3 5](#_Toc27423542)

[2.4. Ejercicio 4 5](#_Toc27423543)

[2.5. Ejercicio 5 8](#_Toc27423544)

# Descripción del documento

Este documento consiste en la realización de la Práctica 1 de la asignatura de MED.

# Ejercicios

## Ejercicio 1

**Enunciado**: *Se ha realizado un estudio para ver si influye la metodología docente a la hora de aprobar. Para ello* ***75*** *estudiantes han recibido la metodología 1 y* ***25*** *la metodología 2. De cada estudiante se ha registrado si al final aprobaban (1) o no (2). Los datos experimentales se dan en la tabla siguiente, donde el número de individuos con perfil aprobar = 1 y metodología = 1 es 35, con perfil aprobar = 1 y metodología = 2 es 15, con perfil aprobar = 2 y metodología = 1 es 40 y con perfil aprobar = 2 y metodología = 2 es 10. ¿Hay diferencias estadísticamente significativas entre las dos metodologías?*

*Imagen que contiene dispositivo

Descripción generada automáticamente*

**Resolución:**

|  |
| --- |
| Para saber si hay diferencias estadísticamente significativas entre las personas que han aprobado o no para cada una de las metodologías, se va a realizar un estudio de diferencia de proporciones poblacionales entre la metodología 1 y la metodología 2.  Puesto que la variable respuesta “aprobar” solo puede tomar dos valores, así como la variable explicativa “metodología”, y como el tamaño muestral es grande (>5), vamos a aplicar el **test z de diferencia de proporciones**, enmarcado en un esquema D🡨D cuyo parámetro poblacional *θ* de interés es la diferencia de proporciones poblacionales entre y (). Para ello se considera el estadístico dado por la v.a. (es decir, ).  En primer lugar, se muestran los cálculos detallados para el intervalo de confianza al 95% dado por:  Y para el contraste de hipótesis dado por:  Con lo que al ser se acepta la hipótesis nula o de igualdad y se concluye en que no existe diferencias significativas entre las poblaciones. Es decir, **no hay diferencia estadísticamente significativa entre ambas metodologías**. |

## Ejercicio 2

**Enunciado**: *En el modelo de regresión lineal, se define la matriz H (matriz “hat”) como aquella matriz que pone el sombrero a la y, es decir que , entonces se verifica que H es simétrica e idempotente.*

1. *Verdadero*
2. *Falso*

**Resolución:**

|  |
| --- |
| **Verdadero**. H es simétrica e idempotente.  En primer lugar, se definen los conceptos simétrica e idempotente aplicado a matrices. Se dice que una matriz es simétrica cuando es igual a su traspuesta. Se dice que una matriz es idempotente cuando es igual a el producto por sí misma.  Por el método mínimos cuadrados, sabemos:  Por lo tanto:  Demostración de que H es simétrica:  Demostración de que H es idempotente: |

## Ejercicio 3

**Enunciado**: *En el modelo de regresión lineal, se define la matriz H (matriz “hat”) como aquella matriz que pone el sombrero a la y, es decir que , entonces se verifica que los elementos de la diagonal de H, vienen dados por , siendo*

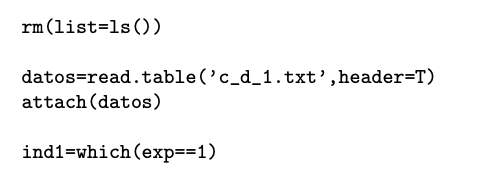
1. *Verdadero*
2. *Falso*

**Resolución:**

|  |
| --- |
| **Verdadero**.  Asumiendo, tal y como hemos calculado en el ejercicio anterior:  Vamos a asumir que .  Por lo tanto, los elementos de la matriz H que pertenecen a la diagonal vienen dados por la siguiente fórmula: |

## Ejercicio 4

**Enunciado**: *Dado el código en R, se pide rellenar el mayo número de ‘xxx’ posible:*

**

*Imagen que contiene captura de pantalla, texto

Descripción generada automáticamente*

**Resolución:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | Estimate | Std. Error | T value | Pr(> |t|) | | (Intercept) | **25.85714** | **3.5057** | **7.376** | **0.000000321** | | Exp2 | **0.3429** | **4.572** | **0.075** | **0.9412** |   Residual standard error: xxx on **15** degrees of freedom Multiple R-squared: xxx, Adjusted R-squared: xxx F-statistic: **0.005626** on **1** and **15**, p-value**: 0.9412**  Intercept 🡪 (X=0), por lo tanto, exp1.  Estimate  La media de exp1 es 25.85714  exp2:  Asumiendo homocedasticidad ():  Std. Error:  Intercept🡪  exp2 🡪  t-Student:  exp2 🡪  Intercept 🡪  p-value:  exp2 🡪  Intercept 🡪  Grados de libertad:  f-statistic: |

## Ejercicio 5

**Enunciado:** *Dado el código en R, se pide rellenar el mayo número de ‘xxx’ posible:*

*Imagen que contiene animal

Descripción generada automáticamente*

*Imagen que contiene texto

Descripción generada automáticamente*

**Resolución***:*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | |  | Estimate | Std. Error | T value | Pr(> |t|) | | (Intercept) | **25.85714** | **3.4746** | **7.442** | **0.0000000482** | | Exp2 | **0.3429** | **4.5303** | **0.076** | **0.940** | | Exp3 | **3.2571** | **5.3828** | **-0.605** | **0.552** |   Residual standard error: xxx on **19** degrees of freedom Multiple R-squared: xxx, Adjusted R-squared: xxx F-statistic: xxx on **2** and **19**, p-value: xxx  Intercept 🡪 (X=0), por lo tanto, exp1.  Estimate  La media de exp1 es 25.85714  exp2:  exp3:  Asumiendo homocedasticidad ():  Std. Error:  Intercept🡪  exp2 🡪  exp3 🡪  t-Student:  exp3 🡪  exp2 🡪  Intercept 🡪  p-value:  exp3 🡪  exp2 🡪  Intercept 🡪  Grados de libertad: |