

Aplicacion de tecnicas de estimacion y prueba de hipotesis

Caso: tendencias socio-economicas de algunas lineas de carrea de Ingenieria de sistemas

Alvarez Bautista Burga Casanova Cuyate

Facultad de Ingenieria Industrial y de Sistemas
Universidad Nacional de Ingenieria

Octubre 2022

Tabla de Contenido

- 1 Problema
- 2 Objetivos
- 3 Importancia de los objetivos
- 4 Resultados
- 5 Conclusiones

Tabla de Contenido

- 1 Problema
- 2 Objetivos
- 3 Importancia de los objetivos
- 4 Resultados
- 5 Conclusiones

- Empiricamente se observa que en el mundo la precarizacion del trabajo se acrecenta cada vez mas, de igual modo con la evolucion de personas casadas y los salarios promedio de los jovenes
- Con motivo de generar informacion util para la prediccion de estas tendencias socio-economicas se ha procedio a realizar un analisis estadistico sobre 9 hipotesis planteadas

Tabla de Contenido

- 1 Problema
- 2 **Objetivos**
- 3 Importancia de los objetivos
- 4 Resultados
- 5 Conclusiones

Objetivos del trabajo

General

Generar informacion relevante para la prediccion de tendencias socio-economicas en el mundo tomando como referencia datos provenientes de distintos paises.

Hipotesis especificas

- La distribucion de ingresos de sigue la ley normal
- Las personas que trabajan una cantidad de horas superior a la media tienen una mejor distribucion de ingresos que aquellas que no lo hacen
- En paises desarrollados existe una mayor cantidad de mujeres con puestos de trabajos relacionados a ingenieria que en paises en via de desarrollo
- Los cientificos de datos poseen un mejor distribucion de ingresos que los ingenieros de datos
- El sector (*publico / privado*) al que pertenece un trabajador es causa de la diferencia de salarios
- El promedio de ingresos de la poblacion mexicana es mayor que la peruana

Hipotesis especificas

- El **promedio de ingresos** de las personas que trabajan una cantidad de horas superior a la mediana es mayor al promedio de ingreso de personas que laburan una cantidad menor de horas que la mediana
- Las personas de mediana edad poseen una mejor distribucion de ingreso que las personas jovenes
- La eleccion de especialidad es causa de la formacion de 2 grupos en la poblacion.

Cada hipotesis tiene asociado el objetivo de comprobar o rechazar la suposicion

Tabla de Contenido

- 1 Problema
- 2 Objetivos
- 3 Importancia de los objetivos**
- 4 Resultados
- 5 Conclusiones

Acerca de la normalidad de la distribución de ingresos

Ever no te olvides del floro

Sobre los demas objetivos:

- 1 analizar causas de la separacion en grupos
- 2 Comparar

Sobre la eleccion de especialidad

Antes de realizar cualquier tecnica de inferencia es necesario conocer la forma de las distribuciones, incluso antes de analizar la varianza

Se desea analizar principalmente:

- 1 Si la eleccion de la especializacion es causa de la formacion de 2 clusters en la poblacion
- 2 Que tan confiable resulta usar **ANOVA** ¹

¹*Analysis of Variance*

Tabla de Contenido

- 1 Problema
- 2 Objetivos
- 3 Importancia de los objetivos
- 4 Resultados**
- 5 Conclusiones

Hipotesis 1

Se aplicó el test de *Jarque-Bera*, para comprobar si la muestra presenta una **curtosis** y **asimetría** correspondientes a una ley normal.

El estadístico de *Jarque Bera* es asintóticamente un pivote de una *Chi-Cuadrado* y toma como hipótesis nula que los datos de la muestra siguen la ley normal

Test de Jarque-Bera

$JB = \frac{n}{6}(S^2 + \frac{1}{4}(K - 3)^2)$, Siendo n los grados de libertad

Estimadores de momentos centrales

- Tercer Momento Central

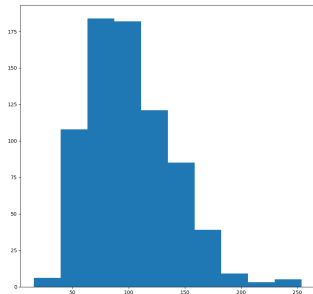
$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3}$$

- Cuarto Momento Central

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4}$$

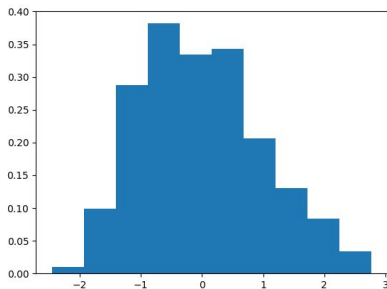
Adicionalmente, se usara el test de *Kolmogorov-Smirnov*, donde se plantea que la distribucion de ingresos en la poblacion de ciencia de datos no sigue la ley normal y se comparará con la funcion acumulada teoria de esta

Figura: Data sin estandarizar



Hipotesis 1

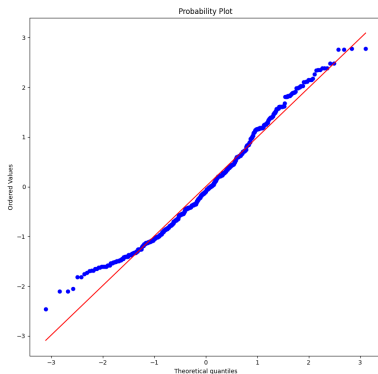
Figura: Data estandarizada y sin outliers



Puede parecer una distribucion Normal

Hipotesis 1

Figura: Grafica Q-Q



Resultados hipotesis 1

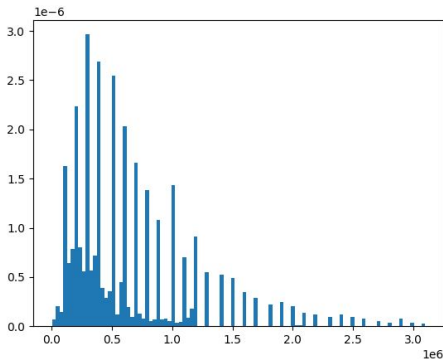
Ever no te olvides del floro

Distribución de software

Aplicacion del test **Kolmogórov-Smirnov**

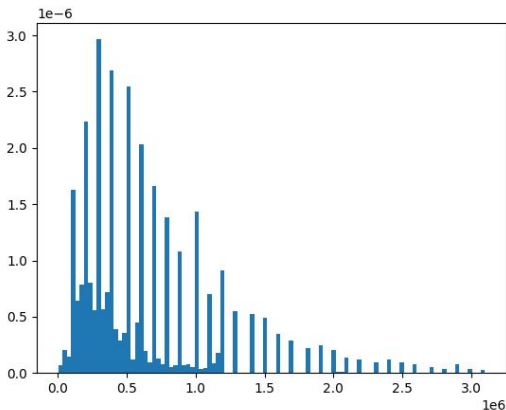
En este caso se va a comprar la funcion de distribucion acumulada observada con la de la distribucion teoria de la normal y la de pareto

Figura: Distribución de ingresos de ingenieros de software en la India



Hipotesis 2

Figura: Distribución de ingresos de ingenieros de software en la India



se puede notar como existen *2 grupos en la población*

Tabla de Contenido

- 1 Problema
- 2 Objetivos
- 3 Importancia de los objetivos
- 4 Resultados
- 5 Conclusiones**

Conclusiones

In this slide, some important text will be highlighted because it's important. Please, don't abuse it.

Remark

Sample text

Important theorem

Sample text in red box

Examples

Sample text in green box. The title of the block is "Examples".

This is a text in first column.

$$E = mc^2$$

- First item
- Second item

This text will be in the second column and on a second thought this is a nice looking layout in some cases.