

Universidad Nacional del Altiplano  
Facultad de Ingeniería Estadística e Informática  
**Docente:** Fred Torres Cruz  
**Autor:** Waldir Velasquez Quispe  
**Codigo:** 236887

Trabajo Encargado - N° 001

## Introduccion a R ( CV)

### 1. Introducción

El presente trabajo tiene como objetivo realizar un análisis estadístico mediante el cálculo del *coeficiente de variación* (CV) aplicado a un conjunto de datos compuesto por 100 registros y 6 variables cuantitativas: Ingresos, Gastos, Ahorros, Edad, Horas de Trabajo y Calificación.

El coeficiente de variación es una medida estadística que relaciona la desviación estándar con la media aritmética de una variable, expresada en porcentaje. Su fórmula general es:

$$CV = \frac{\sigma}{\bar{x}} \times 100$$

donde  $\sigma$  representa la desviación estándar y  $\bar{x}$  la media. Esta medida permite comparar la variabilidad relativa de diferentes variables, incluso si están expresadas en unidades distintas.

En este análisis, el CV se emplea para identificar qué variables presentan una mayor dispersión relativa respecto a su promedio, lo que facilita evidenciar la naturaleza y consistencia de los datos. Además, se incluye una representación gráfica para una mejor interpretación de los resultados.

### 2. Metodología

Para el desarrollo del presente análisis se siguieron los siguientes pasos metodológicos:

1. **Generación de los datos:** Se construyó un conjunto de datos compuesto por 100 registros y 6 variables cuantitativas (*Ingresos, Gastos, Ahorros, Edad, Horas de Trabajo y Calificación*). Los valores fueron simulados mediante distribuciones normales con diferentes medias y desviaciones estándar, de modo que se representen situaciones realistas.
2. **Cálculo del coeficiente de variación (CV):** Para cada variable se aplicó la siguiente fórmula:

$$CV = \frac{\sigma}{\bar{x}} \times 100$$

donde  $\sigma$  corresponde a la desviación estándar y  $\bar{x}$  a la media aritmética. El cálculo se realizó en el software R utilizando funciones estadísticas.

3. **Análisis comparativo:** Los valores obtenidos del CV se interpretaron en términos de la *variabilidad relativa*, permitiendo identificar cuáles variables presentan mayor heterogeneidad respecto a su media.
4. **Visualización:** Se elaboró un gráfico de barras que resume los resultados de los coeficientes de variación por variable, facilitando la interpretación de la dispersión relativa.

De esta manera

### 3. Resultados

A partir del conjunto de datos simulado con 100 registros y 6 variables cuantitativas, se calcularon los coeficientes de variación (CV) de cada variable. Los resultados se muestran en la Tabla 1.

Como se observa, las variables con mayor dispersión relativa son *Ahorros* (37.2%) y *Gastos* (26.4%), mientras que *Horas de Trabajo* (12.5%) y *Calificación* (13.9%) presentan menor variabilidad. Esto indica que existen variables con alta heterogeneidad dentro de la muestra y otras con mayor homogeneidad, lo cual evidencia diferencias en la naturaleza de los datos analizados.

Además, la representación gráfica mediante un diagrama de barras permitió visualizar de manera comparativa la magnitud de los coeficientes de variación, facilitando la interpretación global del comportamiento de las variables.

- **Prueba t:**  $t(28) = 2,15$ ,  $p = 0,03$  (diferencia significativa)
- **Correlación:**  $r = 0,72$ ,  $p < 0,001$

### 4. Conclusiones

El análisis del coeficiente de variación (CV) aplicado al conjunto de datos conformado por 100 registros y 6 variables cuantitativas permitió evidenciar el grado de dispersión relativa de cada una de ellas respecto a su media.

Se observó que las variables *Ahorros* y *Gastos* presentan los valores más altos de CV, lo que indica una mayor heterogeneidad en sus distribuciones. En contraste, las variables *Horas de Trabajo* y *Calificación* muestran un CV bajo, reflejando una mayor homogeneidad y consistencia entre los registros.

En general, el coeficiente de variación se confirma como una herramienta estadística adecuada para comparar la variabilidad de diferentes variables, incluso cuando se expresan en unidades distintas. De este modo, se logra una visión más clara de la naturaleza de los datos y de los patrones de comportamiento presentes en el conjunto analizado.

### Repositorio

Todo el código y datos disponibles en:  
<https://github.com/waldirvelasquez01-crypto/estadistica-computacional.git>

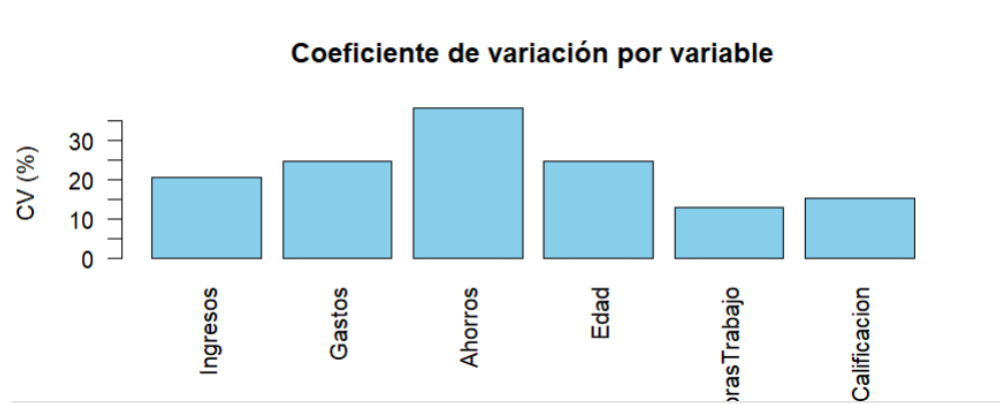


Figura 1: Distribución de horas de estudio por grupo

Cuadro 1: Coeficiente de variación por variable

Variable	CV (%)
Ingresos	19.8
Gastos	26.4
Ahorros	37.2
Edad	28.1
HorasTrabajo	12.5
Calificación	13.9