

Análisis del coeficiente de variación en un conjunto de datos de no menos de 100 registros

Autor: Bryan Cutipa Carcasi
Docente: Fred Torres Cruz

Introducción

Este análisis utiliza un conjunto de datos simulado de 100 individuos, con variables relacionadas con características laborales y personales: Edad, Ingresos, Horas de Trabajo, Experiencia, Satisfacción y Productividad. El objetivo es calcular, para cada variable, su media, desviación estándar y coeficiente de variación (CV), con el fin de medir la dispersión relativa y comparar en qué variables existe mayor variabilidad. Al final, se presenta un resumen ordenado de los coeficientes de variación de mayor a menor.

Datos

Se generó un conjunto de datos ficticios con las siguientes variables para 100 individuos:

Edad: edad en años.

Ingresos: ingresos mensuales en unidades monetarias.

Horas de Trabajo: horas semanales trabajadas.

Experiencia: años de experiencia laboral.

Satisfacción: nivel de satisfacción laboral (escala de 1 a 10).

Productividad: índice de productividad (escala aproximada).

Los datos se generaron mediante distribuciones normales (o uniforme en el caso de Satisfacción) con parámetros realistas, asegurando una muestra de tamaño $n = 100$.

Métodos

Se realizaron los siguientes pasos en el análisis:

1. Generación de un conjunto de datos simulado con 100 observaciones y 6 variables.
2. Cálculo de medidas estadísticas descriptivas: media, desviación estándar, mínimo, máximo y coeficiente de variación (CV).
3. Comparación de la variabilidad relativa entre variables mediante el CV.

4. Elaboración de un resumen ordenado de los coeficientes de variación de mayor a menor.

El coeficiente de variación se calculó como:

$$CV = \left(\frac{\sigma}{\mu} \right) \times 100 \%$$

donde σ es la desviación estándar y μ la media de la variable.

Resultados

A continuación se presenta el resumen estadístico del conjunto de datos:

Variable	Media	Desv. Est.	CV (%)	Mínimo	Máximo
Edad	35.08	7.96	22.69	18	53
Ingresos	2487.23	808.58	32.51	852	4225
Horas de Trabajo	39.96	5.02	12.56	28	52
Experiencia	10.04	4.02	40.04	2	19
Satisfacción	5.56	2.88	51.80	1	10
Productividad	84.85	12.09	14.25	56	113

Cuadro 1: Medidas descriptivas y coeficientes de variación por variable

El orden de dispersión relativa (de mayor a menor CV) es:

1. Satisfacción: 51.80 %
2. Experiencia: 40.04 %
3. Ingresos: 32.51 %
4. Edad: 22.69 %
5. Productividad: 14.25 %
6. Horas de Trabajo: 12.56 %

La variable con mayor variabilidad relativa fue **Satisfacción** ($CV = 51.80 \%$), mientras que la de menor variabilidad fue **Horas de Trabajo** ($CV = 12.56 \%$).

Conclusión

El análisis descriptivo permitió identificar diferencias significativas en la variabilidad relativa entre las variables estudiadas. La **Satisfacción** mostró la mayor dispersión relativa, lo que sugiere una percepción muy heterogénea entre los individuos. Por otro lado, las **Horas de Trabajo** presentaron la menor variabilidad, indicando que la mayoría de los individuos trabaja un número de horas similar. En general, el coeficiente de variación resultó ser una medida útil para comparar la estabilidad relativa entre variables con diferentes unidades o escalas.

Código R utilizado

```
set.seed(123)

n <- 100
datos <- data.frame(
  Edad = round(rnorm(n, mean = 35, sd = 8)),
  Ingresos = round(rnorm(n, mean = 2500, sd = 800)),
  HorasdeTrabajo = round(rnorm(n, mean = 40, sd = 5)),
  Experiencia = round(rnorm(n, mean = 10, sd = 4)),
  Satisfaccion = round(runif(n, min = 1, max = 10)),
  Productividad = round(rnorm(n, mean = 85, sd = 12))
)

calcular_cv <- function(x) {
  x <- na.omit(x)
  cv <- (sd(x) / mean(x)) * 100
  return(round(cv, 2))
}

cv_resultados <- sapply(datos, calcular_cv)

for (var in names(datos)) {
  cat("\n==", var, "==" "\n")
  cat("Media:", round(mean(datos[[var]], na.rm = TRUE), 2), "\n")
  cat("Desviación estándar:", round(sd(datos[[var]], na.rm = TRUE), 2), "\n")
  cat("Coeficiente de variación:", calcular_cv(datos[[var]]), "%\n")
  cat("Mínimo:", min(datos[[var]], na.rm = TRUE), "\n")
  cat("Máximo:", max(datos[[var]], na.rm = TRUE), "\n")
}
```