

PD.1.2

Estrella Guerra, Danilo David

2024-09-05

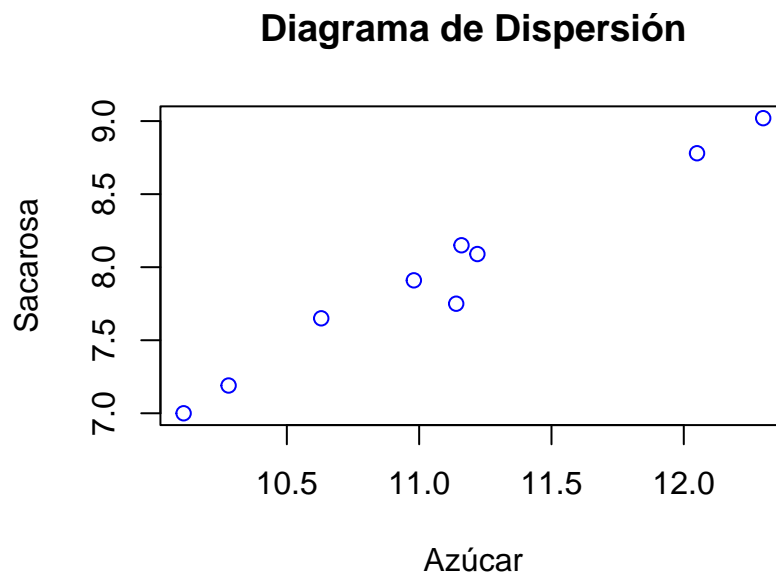
La empresa azucarera del Norte llevo a cabo un estudio con la finalidad de analizar la influencia de algunos factores sobre el rendimiento de azúcar. Se desea determinar si existe relación entre el porcentaje de azúcar por caña (Y) con el porcentaje de sacarosa por caña (X). Los resultados obtenidos se presentan en la siguiente tabla:

Análisis de Correlación

- **Prueba t para correlación:** La prueba t evalúa si un coeficiente de correlación es significativamente diferente de cero. Es una prueba específica que se utiliza cuando se está examinando la relación lineal entre dos variables continuas en un análisis de correlación simple.

1) Elabore el gráfico de dispersión

```
plot(PD2$X,PD2$Y,main="Diagrama de Dispersión",  
     col="blue", xlab = "Azúcar", ylab = "Sacarosa")
```



Se puede apreciar una relación directa entre el azúcar y la sacarosa.

a) Estime el coeficiente de correlación de Pearson

```
cor(PD2$X,PD2$Y,method="p") # p: pearson
```

```
## [1] 0.9852942
```

El coeficiente de correlación de Pearson determina si hay una asociación entre las variables azúcar y sacarosa. Debido a que fluctúa en el intervalo de -1 a 1, al ejecutar el código, el valor del coeficiente de correlación de Pearson 0.9852942 nos indica que hay una fuerte asociación lineal directa entre el azúcar y la sacarosa.

b) Pruebe a un nivel de significación de 0.03 si el coeficiente de correlación es significativamente diferente de cero

1)

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

2)

$$\alpha = 0.03$$

3)

```
cor.test(PD2$X,PD2$Y,method="p",alternative = "t") # t: student
```

```
##
## Pearson's product-moment correlation
##
## data: PD2$X and PD2$Y
## t = 15.257, df = 7, p-value = 0.000001252
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.9291982 0.9970143
## sample estimates:
## cor
## 0.9852942
```

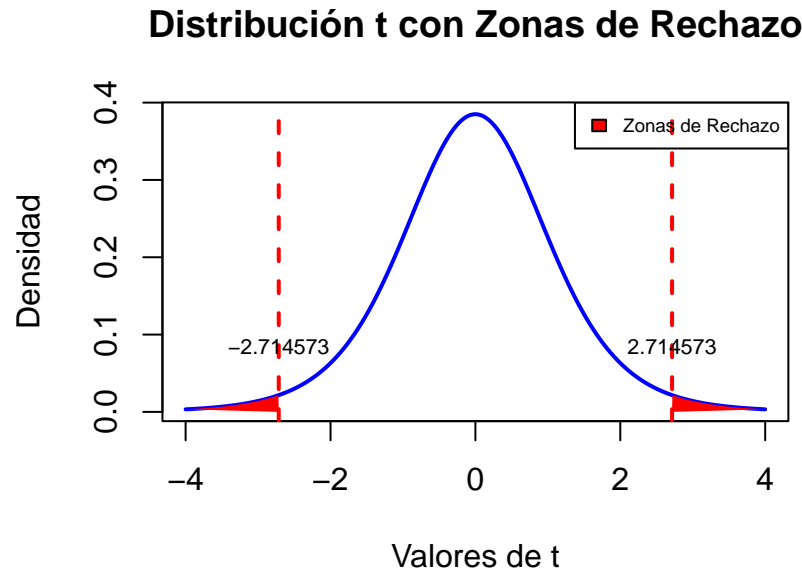
del anterior código hemos obtenido el valor de:

$$t = 15.257, p - value = 0.000001252 \neq 0.000 \rightarrow \alpha \text{ se rechaza } H_0$$

4) Intervalo de confianza (IC) de la prueba t

```
qt(c(0.03/2, 1- 0.03/2), 7)
```

```
## [1] -2.714573  2.714573
```



Al haber obtenido un $t_{cal} = 15.257$ podemos observar que cae en la zona de rechazo, por lo tanto se acepta la hipótesis alterna H_0 .

Conclusión : a un $\alpha = 0.03$ se puede afirmar que existe correlación significativa entre el porcentaje de azúcar y porcentaje de la sacarosa. Los resultados obtenidos son representativos y válidos por ende se recomienda continuar con el proceso.

c) Pruebe a un nivel de significación de 0.03 si el coeficiente de correlación es menor a 0.8

1)

$$H_0 : \rho \geq 0.8$$

$$H_1 : \rho < 0.8$$

2)

$$\alpha = 0.03$$

3)

```
# Nivel de significancia
```

```
alpha <- 0.03
```

```
# Calcular el coeficiente de correlación
```

```
cor_test <- cor.test(PD2$X, PD2$Y, method = "pearson")
```

```
r_observado <- cor_test$estimate # Valor del coeficiente de correlación observado
```

```

# Valor de correlación de referencia
r_referencia <- 0.8

# Transformación de Fisher Z
z_observado <- 0.5 * log((1 + r_observado) / (1 - r_observado))
z_referencia <- 0.5 * log((1 + r_referencia) / (1 - r_referencia))

# Tamaño de la muestra
n <- length(PD2$X)

# Estadístico Z para la prueba de hipótesis
z_prueba <- (z_observado - z_referencia) / sqrt(1 / (n - 3))

# Valor crítico para un nivel de significancia del 0.03 en una prueba unilateral (cola izquierda)
z_critico <- qnorm(alpha)

# Decisión de la prueba
if (z_prueba < z_critico) {
  cat("Rechazamos la hipótesis nula: El coeficiente de correlación es significativamente menor que 0.8.")
} else {
  cat("No se rechaza la hipótesis nula: No hay evidencia suficiente para afirmar que el coeficiente de correlación es menor que 0.8.")
}

```

```
## No se rechaza la hipótesis nula: No hay evidencia suficiente para afirmar que el coeficiente de correlación es menor que 0.8.
```

```

# Mostrar el valor z y el valor crítico
cat("Valor Z calculado:", z_prueba, "\n")

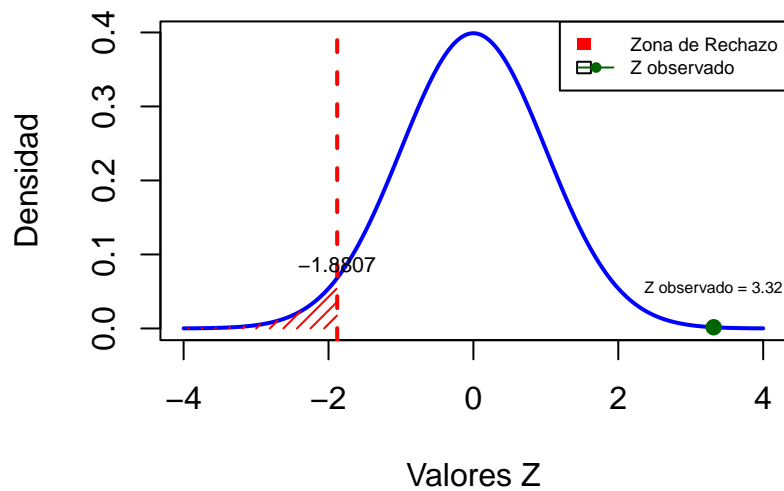
```

```
## Valor Z calculado: 3.316675
```

```
cat("Valor Z crítico:", z_critico, "\n")
```

```
## Valor Z crítico: -1.880794
```

Distribución Z con Zona de Rechazo



d) Halle un intervalo del 95% de confianza para ρ

```
cor.test(PD2$X, PD2$Y, method = "p", alternative = "t", conf.level = 0.95)$conf
```

```
## [1] 0.9291982 0.9970143
## attr(,"conf.level")
## [1] 0.95
```

$IC(\rho) = [0.9291982, 0.9970143]$

El intervalo que va de 0.929 a 0.997 brinda un 95% de confianza de contener a la correlación entre el porcentaje de azúcar y el porcentaje de sacarosa.