

Teorema del Límite Central

Bryan Cutipa Carcasi

Actividad N° 4

Ejemplos Prácticos de Tests Estadísticos en R

1 Tests para Tendencia Central

1.1 Tests de Una Muestra

1.1.1 One-sample t-test

Variables: Puntuaciones de satisfacción laboral (escala 1-10)

Caso: Una empresa quiere saber si la satisfacción promedio de sus empleados difiere significativamente del estándar nacional ($\mu = 7,2$).

```
1 satisfaccion <- c(6.5, 7.8, 8.1, 6.9, 7.3, 8.0, 6.7, 7.5, 8.2, 7.1)
2 t.test(satisfaccion, mu = 7.2)
```

1.1.2 One-sample z-test

Variables: Pesos de productos manufacturados (gramos)

Caso: Control de calidad: verificar si el peso promedio de cereales empacados cumple con el estándar (500g), conociendo que $\sigma = 3g$.

```
1 library(BSDA)
2 pesos <- c(502, 498, 505, 495, 501, 499, 503, 497, 500, 504)
3 z.test(pesos, mu = 500, sigma.x = 3)
```

1.1.3 Wilcoxon signed-rank test

Variables: Tiempo de reacción en milisegundos

Caso: Evaluar si un entrenamiento cognitivo reduce el tiempo de reacción comparado con el valor basal (250ms), sin asumir normalidad.

```
1 tiempo_reaccion <- c(245, 240, 255, 235, 248, 252, 238, 242, 250, 246)
2 wilcox.test(tiempo_reaccion, mu = 250)
```

1.1.4 Sign test

Variables: Cambio en síntomas de dolor (mejor/peor)

Caso: Evaluar si un tratamiento mejora los síntomas en más del 50% de los pacientes.

```
1 mejoras <- c(1, 1, 0, 1, 1, 1, 0, 1, 1, 0)
2 binom.test(sum(mejoras), length(mejoras), p = 0.5)
```

2 Tests Comparando Dos Grupos

2.1 Grupos Independientes

2.1.1 Independent samples t-test

Variables: Puntuaciones de ansiedad (escala 0-30)

Caso: Comparar niveles de ansiedad entre grupo control y grupo experimental tras una intervención terapéutica.

```
1 control <- c(18, 20, 16, 22, 19, 17, 21, 18, 20, 19)
2 experimental <- c(14, 12, 15, 13, 11, 16, 12, 14, 13, 15)
3 t.test(control, experimental, var.equal = TRUE)
```

2.1.2 Welch's t-test

Variables: Ingresos mensuales (en miles)

Caso: Comparar ingresos entre graduados de dos universidades diferentes, sin asumir varianzas iguales.

```
1 universidad_A <- c(3.2, 4.1, 3.8, 4.5, 3.9, 4.2, 3.7, 4.0)
2 universidad_B <- c(2.8, 3.1, 2.9, 3.3, 2.7, 3.0, 2.9, 3.2, 2.8)
3 t.test(universidad_A, universidad_B, var.equal = FALSE)
```

2.1.3 Mann-Whitney U test

Variables: Calificaciones de satisfacción del cliente (ordinal 1-5)

Caso: Comparar satisfacción entre clientes de dos sucursales diferentes, datos no normales.

```
1 sucursal_norte <- c(4, 5, 3, 4, 5, 4, 3, 5, 4)
2 sucursal_sur <- c(3, 2, 4, 3, 3, 2, 4, 3, 2, 3)
3 wilcox.test(sucursal_norte, sucursal_sur)
```

2.2 Grupos Relacionados/Pareados

2.2.1 Paired t-test

Variables: Presión arterial sistólica (mmHg)

Caso: Evaluar el efecto de un medicamento comparando la presión antes y después del tratamiento en los mismos pacientes.

```
1 antes <- c(140, 135, 142, 138, 145, 139, 141, 136, 143, 137)
2 despues <- c(132, 128, 135, 130, 138, 131, 134, 129, 136, 130)
3 t.test(antes, despues, paired = TRUE)
```

2.2.2 Wilcoxon signed-rank (pareado)

Variables: Puntuaciones de dolor (escala 0-10, no normal)

Caso: Comparar niveles de dolor antes y después de fisioterapia en los mismos pacientes.

```
1 dolor_antes <- c(8, 7, 9, 6, 8, 7, 9, 8, 7, 6)
2 dolor_despues <- c(5, 4, 6, 3, 5, 4, 7, 5, 4, 3)
3 wilcox.test(dolor_antes, dolor_despues, paired = TRUE)
```

3 Tests Comparando Múltiples Grupos

3.1 ANOVA y Equivalentes No Paramétricos

3.1.1 One-way ANOVA

Variables: Rendimiento académico (puntuaciones 0-100)

Caso: Comparar el rendimiento entre estudiantes de tres métodos de enseñanza diferentes.

```

1 metodo_A <- c(85, 88, 82, 87, 84, 86, 83, 89, 85, 87)
2 metodo_B <- c(78, 80, 76, 82, 79, 77, 81, 78, 80, 79)
3 metodo_C <- c(92, 94, 90, 93, 91, 95, 89, 92, 94, 91)
4
5 datos <- data.frame(
6   rendimiento = c(metodo_A, metodo_B, metodo_C),
7   metodo = factor(rep(c("A", "B", "C"), each = 10))
8 )
9 resultado_anova <- aov(rendimiento ~ metodo, data = datos)
10 summary(resultado_anova)
11 TukeyHSD(resultado_anova)

```

3.1.2 Kruskal-Wallis test

Variables: Nivel de estrés percibido (escala ordinal 1-7)

Caso: Comparar niveles de estrés entre trabajadores de cuatro departamentos diferentes.

```

1 ventas <- c(5, 6, 4, 5, 6, 5, 4, 6, 5)
2 marketing <- c(3, 4, 3, 4, 3, 4, 3, 4)
3 finanzas <- c(6, 7, 6, 7, 6, 7, 6, 7, 6)
4 rrhh <- c(4, 5, 4, 5, 4, 5, 4, 5, 4, 5)
5
6 datos_estres <- data.frame(
7   estres = c(ventas, marketing, finanzas, rrhh),
8   departamento = factor(rep(c("Ventas", "Marketing", "Finanzas", "RRHH"),
9                               ,
10                               c(length(ventas), length(marketing),
11                               length(finanzas), length(rrhh))))
12 )
13 kruskal.test(estres ~ departamento, data = datos_estres)

```

3.1.3 Repeated measures ANOVA

Variables: Tiempo de concentración (minutos)

Caso: Evaluar el efecto de tres tipos de música en la concentración de los mismos participantes.

```

1 participante <- factor(rep(1:8, 3))
2 musica <- factor(rep(c("Clasica", "Ambiental", "Silencio"), each = 8))
3 concentracion <- c(
4   c(25, 28, 22, 30, 26, 29, 24, 27),
5   c(30, 32, 28, 35, 31, 33, 29, 31),
6   c(20, 22, 18, 24, 21, 23, 19, 22)
7 )
8
9 datos_rm <- data.frame(participante, musica, concentracion)
10 resultado_rm <- aov(concentracion ~ musica + Error(participante/musica),
11                    data = datos_rm)
12 summary(resultado_rm)

```

3.1.4 Friedman test

Variables: Preferencia de sabor (ranking 1-5)

Caso: Comparar preferencias de cinco sabores de helado evaluados por los mismos catadores.

```
1 preferencias <- matrix(c(
2   4, 5, 2, 3, 1,
3   5, 4, 1, 3, 2,
4   3, 5, 2, 4, 1,
5   4, 5, 1, 2, 3,
6   5, 4, 2, 3, 1,
7   4, 5, 1, 3, 2
8 ), nrow = 6, byrow = TRUE)
9
10 friedman.test(preferencias)
```

4 Tests para Varianza

4.0.1 F-test para igualdad de varianzas

Variables: Variabilidad en tiempos de producción (minutos)

Caso: Comparar si dos máquinas tienen la misma variabilidad en sus tiempos de producción.

```
1 maquina_1 <- c(12.5, 12.8, 12.3, 12.7, 12.4, 12.6, 12.5, 12.9)
2 maquina_2 <- c(12.1, 13.2, 11.8, 13.5, 12.0, 13.1, 11.9, 13.3)
3 var.test(maquina_1, maquina_2)
```

4.0.2 Levene's test

Variables: Variabilidad en ventas mensuales (miles de dólares)

Caso: Evaluar si tres regiones tienen la misma variabilidad en sus ventas mensuales.

```
1 library(car)
2 region_norte <- c(150, 145, 155, 148, 152, 149, 151)
3 region_centro <- c(120, 135, 115, 140, 125, 130, 118)
4 region_sur <- c(180, 165, 190, 175, 185, 170, 195)
5
6 datos_ventas <- data.frame(
7   ventas = c(region_norte, region_centro, region_sur),
8   region = factor(rep(c("Norte", "Centro", "Sur"), each = 7))
9 )
10 leveneTest(ventas ~ region, data = datos_ventas)
```

4.0.3 Bartlett's test

Variables: Puntuaciones de creatividad (escala continua)

Caso: Verificar homogeneidad de varianzas en puntuaciones de creatividad entre cuatro grupos de edad.

```
1 grupo_20_30 <- c(75, 78, 72, 76, 74, 77, 73, 79)
2 grupo_31_40 <- c(68, 70, 65, 72, 69, 71, 67, 73)
3 grupo_41_50 <- c(62, 65, 60, 67, 63, 66, 61, 68)
4 grupo_51_60 <- c(58, 60, 55, 62, 59, 61, 57, 63)
5
6 datos_creatividad <- data.frame(
7   creatividad = c(grupo_20_30, grupo_31_40, grupo_41_50, grupo_51_60),
8   edad = factor(rep(c("20-30", "31-40", "41-50", "51-60"), each = 8))
```

```
9 )  
10 bartlett.test(creatividad ~ edad, data = datos_creatividad)
```

Notas Importantes

- **Verificación de paquetes:** Los tests que requieren paquetes específicos (BSDA, car) necesitan instalación previa
- **Presupuestos:** Cada test tiene sus propios presupuestos que deben verificarse antes de su aplicación
- **Interpretación:** Los resultados deben interpretarse en el contexto del problema de investigación
- **Nivel de significancia:** Se recomienda establecer $\alpha = 0,05$ como nivel de significancia estándar