



Inferencia Estadística

Ejercicio: Análisis de Varianza

DSLab

noviembre, 2024

Ejercicio 1

Pregunta: Se realizó un experimento para investigar si hay diferencias significativas en el tiempo de reacción ante un estímulo visual entre tres grupos de participantes que fueron sometidos a diferentes niveles de iluminación: normal, baja y alta. Se registraron los tiempos de reacción (en milisegundos) para cada participante en cada grupo. Los datos obtenidos se muestran a continuación:

Grupo Normal: 40, 45, 38, 42, 39

Grupo Baja: 48, 50, 55, 52, 51

Grupo Alta: 35, 37, 32, 40, 36

Usando un nivel de significancia de 0.05, ¿hay diferencias significativas en el tiempo de reacción entre los tres grupos?

Ejercicio 2

Pregunta: Un investigador quiere evaluar el efecto de dos factores en el rendimiento de los estudiantes URJC en un examen final. Los dos factores considerados son el tipo de método de estudio (Método Tradicional y Método Moderno) y el ambiente de estudio (Ambiente Silencioso y Ambiente con Ruido). Para ello, se seleccionan 5 estudiantes aleatoriamente para cada combinación de métodos y ambientes, y se registra su puntuación en el examen final.

Los datos obtenidos se muestran en la siguiente tabla:

Ambiente / Método	Método Tradicional	Método Moderno
Silencioso	85, 90, 88, 92, 87	91, 93, 89, 94, 92
Con Ruido	70, 75, 72, 78, 73	77, 79, 74, 80, 76

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en las puntuaciones del examen debido al tipo de método de estudio, al ambiente de estudio, y a la interacción entre ambos factores.

Ejercicio 3

Pregunta: Un equipo de investigadores quiere evaluar el efecto de dos factores en la satisfacción laboral de los empleados. Los dos factores considerados son el tipo de horario de trabajo (Horario Flexible y Horario Fijo) y el tipo de entorno de trabajo (Oficina Abierta y Oficina Cerrada). Se seleccionan 5 empleados aleatoriamente para cada combinación de horarios y entornos, y se registra su puntuación de satisfacción laboral en una escala de 1 a 100.

Los datos obtenidos se muestran en la siguiente tabla:

Entorno / Horario	Horario Flexible	Horario Fijo
Oficina Abierta	78, 82, 85, 80, 84	60, 65, 58, 63, 61
Oficina Cerrada	88, 90, 87, 85, 89	70, 73, 75, 72, 71

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en la satisfacción laboral debido al tipo de horario de trabajo, al entorno de trabajo, y a la interacción entre ambos factores.

Ejercicio 4

Pregunta: Un investigador quiere evaluar el efecto de dos factores en el rendimiento de los estudiantes en una prueba de matemáticas. Los dos factores considerados son el método de enseñanza (Tradicional y Moderno) y el tipo de material de apoyo (Libro, Video, Software, Sin material). Se seleccionan 3 estudiantes aleatoriamente para cada combinación de métodos y materiales, y se registra su puntuación en la prueba de matemáticas.

Los datos obtenidos se muestran en la siguiente tabla:

Material / Método	Método Tradicional	Método Moderno
Libro	75, 80, 85	90, 92, 88

Material / Método	Método Tradicional	Método Moderno
Video	70, 72, 78	85, 87, 84
Software	65, 68, 70	80, 83, 79
Sin material	60, 62, 65	75, 77, 73

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en las puntuaciones de la prueba de matemáticas debido al método de enseñanza, al tipo de material de apoyo, y a la interacción entre ambos factores.

Ejercicio 5

Pregunta: ¿Es posible simplificar el modelo anterior eliminando la interacción? ¿Es posible eliminar alguno de los efectos individuales? De ser así, construye el nuevo modelo. ¿Dónde aparecen las principales diferencias en la puntuación? Plantea un contraste de hipótesis de igualdad de distribuciones. Plantea un intervalo de confianza para la diferencia de medias.

Ejercicio 6

Pregunta: Una empresa de manufactura quiere evaluar el efecto de dos factores en la durabilidad de un producto. Los dos factores considerados son el tipo de material utilizado (Material A y Material B) y la temperatura de fabricación (Baja, Media y Alta). Se seleccionan 4 muestras aleatoriamente para cada combinación de materiales y temperaturas, y se mide la durabilidad en días. Los datos obtenidos se muestran en la siguiente tabla:

Temperatura / Material	Material A	Material B
Baja	15, 18, 16, 17	14, 17, 16, 15
Media	20, 22, 21, 19	19, 18, 20, 21
Alta	25, 27, 26, 24	23, 22, 24, 25

Se realiza un ANOVA de dos factores y se aplica la prueba de Tukey HSD.

```
# Datos de durabilidad
```

```
durabilidad <- c(15, 18, 16, 17, 14, 17, 16, 15,
                 20, 22, 21, 19, 19, 18, 20, 21,
                 25, 27, 26, 24, 23, 22, 24, 25)
```

```
# Factor de material
```

```
material <- factor(rep(c("A", "B"), each = 12))

# Factor de temperatura
temperatura <- factor(rep(c("Baja", "Media", "Alta"), each = 4, times = 2))

# Crear un data frame con los datos
datos <- data.frame(durabilidad, material, temperatura)

# Realizar ANOVA de dos factores
modelo <- aov(durabilidad ~ material * temperatura, data = datos)

# Resumen del modelo
summary(modelo)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## material          1 170.67   170.67    102.4 7.44e-09 ***
## temperatura       2   65.33    32.67     19.6 3.03e-05 ***
## material:temperatura 2   65.33    32.67     19.6 3.03e-05 ***
## Residuals        18   30.00     1.67
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

En esta salida:

- El factor “material” no es significativo ($p > 0.05$).
- El factor “temperatura” es significativo ($p < 0.05$).
- La interacción “material:temperatura” no es significativa ($p > 0.05$).

Realizamos la prueba post-hoc de Tukey HSD de comparación de medias:

```
# Realizar la prueba de Tukey HSD
tukey <- TukeyHSD(modelo)

# Mostrar los resultados
print(tukey)

##      Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = durabilidad ~ material * temperatura, data = datos)
##
## $material
##      diff      lwr      upr p adj
```

```
## B-A 5.333333 4.22605 6.440616      0
##
## $temperatura
##           diff           lwr           upr           p adj
## Baja-Alta  -4.0 -5.6474142 -2.3525858 0.0000216
## Media-Alta -1.5 -3.1474142  0.1474142 0.0779292
## Media-Baja  2.5  0.8525858  4.1474142 0.0030417
##
## $`material:temperatura`
##           diff           lwr           upr           p adj
## B:Alta-A:Alta      3  0.09886432  5.901136 0.0403048
## A:Baja-A:Alta     -4 -6.90113568 -1.098864 0.0041136
## B:Baja-A:Alta     -1 -3.90113568  1.901136 0.8768907
## A:Media-A:Alta    -5 -7.90113568 -2.098864 0.0004115
## B:Media-A:Alta     5  2.09886432  7.901136 0.0004115
## A:Baja-B:Alta     -7 -9.90113568 -4.098864 0.0000058
## B:Baja-B:Alta     -4 -6.90113568 -1.098864 0.0041136
## A:Media-B:Alta    -8 -10.90113568 -5.098864 0.0000009
## B:Media-B:Alta     2 -0.90113568  4.901136 0.2889127
## B:Baja-A:Baja      3  0.09886432  5.901136 0.0403048
## A:Media-A:Baja    -1 -3.90113568  1.901136 0.8768907
## B:Media-A:Baja     9  6.09886432 11.901136 0.0000001
## A:Media-B:Baja    -4 -6.90113568 -1.098864 0.0041136
## B:Media-B:Baja     6  3.09886432  8.901136 0.0000456
## B:Media-A:Media   10  7.09886432 12.901136 0.0000000
```

A la vista de los resultados: ¿Qué contrastes de hipótesis tiene sentido plantear para realizar un experimento en el futuro?

Ejercicio 7

Pregunta: Un investigador quiere evaluar el efecto de dos factores en el rendimiento de los empleados en una tarea de productividad. Los dos factores considerados son el tipo de capacitación (Capacitación A y Capacitación B) y el horario de trabajo (Turno de Mañana, Turno de Tarde y Turno de Noche). Para ello, se seleccionan 3 empleados aleatoriamente para cada combinación de capacitación y horario, y se registra su rendimiento en la tarea (en puntos).

Los datos obtenidos se muestran en la siguiente tabla:

Horario / Capacitación	Capacitación A	Capacitación B
Mañana	78, 82, 81	85, 87, 88
Tarde	72, 75, 74	80, 82, 83
Noche	65, 68, 66	60, 62, 61

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en el rendimiento de los empleados debido al tipo de capacitación, al horario de trabajo, y a la interacción entre ambos factores.

Ejercicio 8

Pregunta: Aplica un método de comparación de medias al ejercicio anterior para estudiar dónde se encuentran las diferencias estadísticamente significativas.

Ejercicio 9

Pregunta: Un investigador quiere evaluar el efecto de tres diferentes dietas en el aumento de peso de ratones. Para ello, selecciona al azar 15 ratones y los divide en tres grupos de 5 ratones cada uno. Cada grupo es alimentado con una dieta diferente durante un período de 8 semanas. Al final del experimento, se mide el aumento de peso (en gramos) de cada ratón. Los datos obtenidos se muestran a continuación:

Dieta A: 20, 21, 19, 23, 22

Dieta B: 30, 32, 29, 31, 33

Dieta C: 25, 27, 26, 24, 28

Utilizando un nivel de significancia de 0.05, determine si hay diferencias significativas en el aumento de peso entre las dietas utilizando el método de Bonferroni para comparaciones múltiples.

Ejercicio 10

Pregunta: Un investigador quiere evaluar el efecto de tres factores en el rendimiento de un motor: el tipo de combustible (Gasolina, Diésel, Eléctrico), la temperatura (Baja, Media, Alta) y la presión (Baja, Media, Alta). Para ello, se mide el rendimiento del motor (en kilovatios) bajo todas las combinaciones posibles de estos factores. Los datos obtenidos se muestran a continuación:

Combustible	Temperatura	Presión	Rendimiento
Gasolina	Baja	Baja	150

Combustible	Temperatura	Presión	Rendimiento
Gasolina	Baja	Media	160
Gasolina	Baja	Alta	155
Gasolina	Media	Baja	165
Gasolina	Media	Media	170
Gasolina	Media	Alta	168
Gasolina	Alta	Baja	160
Gasolina	Alta	Media	158
Gasolina	Alta	Alta	162
Diésel	Baja	Baja	140
Diésel	Baja	Media	145
Diésel	Baja	Alta	148
Diésel	Media	Baja	150
Diésel	Media	Media	155
Diésel	Media	Alta	152
Diésel	Alta	Baja	148
Diésel	Alta	Media	150
Diésel	Alta	Alta	149
Eléctrico	Baja	Baja	130
Eléctrico	Baja	Media	135
Eléctrico	Baja	Alta	138
Eléctrico	Media	Baja	140
Eléctrico	Media	Media	145
Eléctrico	Media	Alta	142
Eléctrico	Alta	Baja	138
Eléctrico	Alta	Media	140
Eléctrico	Alta	Alta	139

Estudia el efecto de cada uno de los factores en el rendimiento, así como el efecto de las posibles interacciones.