

# Formulario 7: Diseño Factorial II

Diseño de experimentos

Andrés Felipe Pico Zúñiga

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.colors as mcolors
import seaborn as sns
import statsmodels
import statsmodels.api as sm
from statsmodels.formula.api import ols
from statsmodels.stats.multicomp import pairwise_tukeyhsd
import statsmodels.formula.api as smf
import scipy.stats as stats
```

## Enunciado

Se tiene interés en el rendimiento de un proceso; en particular para ello se consideran tres factores:

- A: Efecto de la Temperatura (100, 120, 140 °C).
- B: La Presión (400, 450).
- C: El Tiempo del lavado del producto en seguida del proceso de enfriamiento (30 y 35 minutos).

## Creación y visualización del dataframe

```
data = {
    'Temperatura': [100]*12 + [120]*12 + [140]*12,
    'Presion': ([400]*6 + [450]*6) * 3,
    'Tiempo': ([30]*3 + [35]*3) * 6,
    "Replica": [1,2,3]*12,
    'Rendimiento': [
        # Grupo de Temperatura 100°C
        34, 33, 33, # Presión: 400, Tiempo: 30
        27, 29, 28, # Presión: 400, Tiempo: 35
        32, 32, 33, # Presión: 450, Tiempo: 30
        28, 28, 27, # Presión: 450, Tiempo: 35
        # Grupo de Temperatura 120°C
        32, 34, 34, # Presión: 400, Tiempo: 30
        26, 28, 29, # Presión: 400, Tiempo: 35
        32, 33, 32, # Presión: 450, Tiempo: 30
        30, 25, 27, # Presión: 450, Tiempo: 35
        # Grupo de Temperatura 140°C
        36, 36, 37, # Presión: 400, Tiempo: 30
```

```

        29, 29, 30, # Presión: 400, Tiempo: 35
        34, 34, 34, # Presión: 450, Tiempo: 30
        27, 28, 29 # Presión: 450, Tiempo: 35
    ]
}

df = pd.DataFrame(data)
print(df)

```

##	Temperatura	Presion	Tiempo	Replica	Rendimiento
## 0	100	400	30	1	34
## 1	100	400	30	2	33
## 2	100	400	30	3	33
## 3	100	400	35	1	27
## 4	100	400	35	2	29
## 5	100	400	35	3	28
## 6	100	450	30	1	32
## 7	100	450	30	2	32
## 8	100	450	30	3	33
## 9	100	450	35	1	28
## 10	100	450	35	2	28
## 11	100	450	35	3	27
## 12	120	400	30	1	32
## 13	120	400	30	2	34
## 14	120	400	30	3	34
## 15	120	400	35	1	26
## 16	120	400	35	2	28
## 17	120	400	35	3	29
## 18	120	450	30	1	32
## 19	120	450	30	2	33
## 20	120	450	30	3	32
## 21	120	450	35	1	30
## 22	120	450	35	2	25
## 23	120	450	35	3	27
## 24	140	400	30	1	36
## 25	140	400	30	2	36
## 26	140	400	30	3	37
## 27	140	400	35	1	29
## 28	140	400	35	2	29
## 29	140	400	35	3	30
## 30	140	450	30	1	34
## 31	140	450	30	2	34
## 32	140	450	30	3	34
## 33	140	450	35	1	27
## 34	140	450	35	2	28
## 35	140	450	35	3	29

## Tabla ANOVA

```

model = ols(
    'Rendimiento ~ C(Temperatura) + C(Presion) + C(Tiempo) + \
    C(Temperatura):C(Presion) + C(Temperatura):C(Tiempo) + \
    C(Presion):C(Tiempo) + C(Temperatura):C(Presion):C(Tiempo)',

```

```

data=df
).fit()

# Perform the ANOVA analysis
anova_table = sm.stats.anova_lm(model, typ=2) # typ=2 for Type II ANOVA

# Display the ANOVA table
print(anova_table)

```

	sum_sq	df	F	PR(>F)
C(Temperatura)	22.388889	2.0	9.595238	8.667203e-04
C(Presion)	10.027778	1.0	8.595238	7.293760e-03
C(Tiempo)	283.361111	1.0	242.880952	4.708271e-14
C(Temperatura):C(Presion)	2.722222	2.0	1.166667	3.284467e-01
C(Temperatura):C(Tiempo)	3.722222	2.0	1.595238	2.236316e-01
C(Presion):C(Tiempo)	1.361111	1.0	1.166667	2.908225e-01
C(Temperatura):C(Presion):C(Tiempo)	0.055556	2.0	0.023810	9.764947e-01
Residual	28.000000	24.0	NaN	NaN

## Prueba HSD

```

tukey_temp = pairwise_tukeyhsd(endog=df['Rendimiento'],
                                groups=df['Temperatura'],
                                alpha=0.05)
print("--- Prueba HSD de Tukey para Temperatura ---")

```

```

## --- Prueba HSD de Tukey para Temperatura ---
print(tukey_temp)

```

```

## Multiple Comparison of Means - Tukey HSD, FWER=0.05
## =====
## group1 group2 meandiff p-adj lower upper reject
## -----
## 100 120 -0.1667 0.9908 -3.3309 2.9976 False
## 100 140 1.5833 0.4456 -1.5809 4.7476 False
## 120 140 1.75 0.3747 -1.4142 4.9142 False
## -----

```

```

print("\n" + "="*50 + "\n")

```

```

##
## =====
# --- Para el factor Presión ---
tukey_pres = pairwise_tukeyhsd(endog=df['Rendimiento'],
                                groups=df['Presion'],
                                alpha=0.05)
print("--- Prueba HSD de Tukey para Presión ---")

```

```

## --- Prueba HSD de Tukey para Presión ---
print(tukey_pres)

```

```

## Multiple Comparison of Means - Tukey HSD, FWER=0.05
## =====

```

```
## group1 group2 meandiff p-adj lower upper reject
## -----
##      400      450 -1.0556 0.3248 -3.2028 1.0917 False
## -----

print("\n" + "="*50 + "\n")

##
## =====
# --- Para el factor Tiempo ---
tukey_tiem = pairwise_tukeyhsd(endog=df['Rendimiento'],
                              groups=df['Tiempo'],
                              alpha=0.05)
print("--- Prueba HSD de Tukey para Tiempo ---")

## --- Prueba HSD de Tukey para Tiempo ---
print(tukey_tiem)

## Multiple Comparison of Means - Tukey HSD, FWER=0.05
## =====
## group1 group2 meandiff p-adj lower upper reject
## -----
##      30      35 -5.6111 0.0 -6.5711 -4.6511 True
## -----
```

## Prueba LSD

```
MSE = 1.1667
df_error = 24
alpha = 0.05

# --- 1. Calcular el valor crítico de t ---
t_critico = stats.t.ppf(1 - alpha / 2, df_error)
print(f"Valor crítico de t para alpha=0.05 y {df_error} gl: {t_critico:.4f}\n")

## Valor crítico de t para alpha=0.05 y 24 gl: 2.0639

# --- 2. Calcular el valor LSD para cada factor ---

# Para Temperatura (n=12 por nivel) y Presión (n=18 por nivel)
# Como las n son diferentes, calcularemos LSD para cada caso.

# a) LSD para Temperatura (12 réplicas por nivel)
n_temp = 12
LSD_temp = t_critico * np.sqrt(MSE * (1/n_temp + 1/n_temp))
print(f"El valor LSD para comparar niveles de Temperatura es: {LSD_temp:.4f}")

## El valor LSD para comparar niveles de Temperatura es: 0.9101

# b) LSD para Presión (18 réplicas por nivel)
n_pres = 18
LSD_pres = t_critico * np.sqrt(MSE * (1/n_pres + 1/n_pres))
print(f"El valor LSD para comparar niveles de Presión es: {LSD_pres:.4f}")

## El valor LSD para comparar niveles de Presión es: 0.7431
```

```

# c) LSD para Tiempo (18 réplicas por nivel)
n_tiem = 18
LSD_tiem = t_critico * np.sqrt(MSE * (1/n_tiem + 1/n_tiem))
print(f"El valor LSD para comparar niveles de Tiempo es: {LSD_tiem:.4f}")

## El valor LSD para comparar niveles de Tiempo es: 0.7431
print("\n" + "="*50 + "\n")

##
## =====

# --- 3. Aplicar la prueba LSD ---
print("--- Aplicación de la Prueba LSD ---")

## --- Aplicación de la Prueba LSD ---

# a) Comparaciones para Temperatura
medias_temp = df.groupby('Temperatura')['Rendimiento'].mean()
print("\nMedias para Temperatura:\n", medias_temp)

##
## Medias para Temperatura:
## Temperatura
## 100    30.333333
## 120    30.166667
## 140    31.916667
## Name: Rendimiento, dtype: float64
print(f"\nComparando diferencias con LSD = {LSD_temp:.4f}:")

##
## Comparando diferencias con LSD = 0.9101:
print(f"|120 - 100| = {abs(medias_temp[120] - medias_temp[100]):.4f} -> No significativo")

## |120 - 100| = 0.1667 -> No significativo
print(f"|140 - 100| = {abs(medias_temp[140] - medias_temp[100]):.4f} -> No significativo")

## |140 - 100| = 1.5833 -> No significativo
print(f"|140 - 120| = {abs(medias_temp[140] - medias_temp[120]):.4f} -> Significativo")

## |140 - 120| = 1.7500 -> Significativo

# b) Comparaciones para Presión
medias_pres = df.groupby('Presion')['Rendimiento'].mean()
print("\nMedias para Presión:\n", medias_pres)

##
## Medias para Presión:
## Presion
## 400    31.333333
## 450    30.277778
## Name: Rendimiento, dtype: float64
print(f"\nComparando diferencia con LSD = {LSD_pres:.4f}:")

##

```

```

## Comparando diferencia con LSD = 0.7431:
print(f"|450 - 400| = {abs(medias_pres[450] - medias_pres[400]):.4f} -> Significativo")

## |450 - 400| = 1.0556 -> Significativo

# c) Comparaciones para Tiempo
medias_tiem = df.groupby('Tiempo')['Rendimiento'].mean()
print("\nMedias para Tiempo:\n", medias_tiem)

##
## Medias para Tiempo:
## Tiempo
## 30    33.611111
## 35    28.000000
## Name: Rendimiento, dtype: float64
print(f"\nComparando diferencia con LSD = {LSD_tiem:.4f}:")

##
## Comparando diferencia con LSD = 0.7431:
print(f"|35 - 30| = {abs(medias_tiem[35] - medias_tiem[30]):.4f} -> Significativo")

## |35 - 30| = 5.6111 -> Significativo

```