

FUEGO, MANTENLO PRENDIDO

Carlos Gabriel Peñaloza Gómez- 201531973 – c.penaloza
Juan David Vega Guzmán – 201520792 – jd.vega11

IIND-2107 Probabilidad y Estadística II
Ingeniería Industrial

Universidad de los Andes
2018-10

12/03/2018

FUEGO, MANTENLO PRENDIDO

1. Introducción

1.1 Contexto del problema

Se desea generar una fogata para una noche de campamento y, para prevenir un escenario en que falten fósforos, se decidió determinar los factores que influyen en la generación de la primera llama de fuego por medio de una lupa, materiales del entorno y ayuda del sol. De esta forma, se espera tener un buen plan de contingencia efectivo que considere los resultados obtenidos en el análisis de los factores y cómo estos se configuran para que la primera llama se genere en el menor tiempo posible.

1.2. Objetivos a desarrollar

Determinar qué factores (o interacciones entre los mismos) influyen en el tiempo de generación de la primera llama de fuego y, a partir de dicho entendimiento, identificar qué combinaciones particulares de dichos factores permiten optimizar el proceso de encender la fogata; minimizando, por tanto, el tiempo de generación mencionado.

1.3. Factores evaluados

A continuación, se detallan los factores involucrados junto con los niveles correspondientes a cada uno:

FACTOR/NIVEL	Nivel 1	Nivel 2	Nivel 3
F1 (ANGULO)	30°	60°	90°
F2 (DISTANCIA)	10 cm	20 cm	30 cm
F3 (MATERIAL)	Papel bond	Cartón	Papel periódico

*Los ángulos se toman con respecto a la vertical.

1.4. Metodología empleada:

Se decidió realizar 3 réplicas por cada combinación de factores (celdas), de forma que, en total, se efectuaron 81 mediciones del tiempo de generación de la primera llama. Además, con el fin reducir la variabilidad asociada a factores externos, fue preciso que la toma de datos se llevara a cabo bajo las siguientes condiciones:

- La hora en que se iban a tomar las mediciones debía ser aquella en la que el sol estuviera más cerca del cenit; es decir, al mediodía. De esta forma, se decidió llevar a cabo el experimento de 12:00 a 1:15 del día 28 de febrero. Lo anterior, con el objetivo de que el ángulo con el que impactan los rayos de sol no aportara un nivel de variabilidad significativa en el estudio.
- El día escogido debía ser soleado con el fin de garantizar que, en efecto, se generara la llama. En la siguiente referencia se puede verificar que el 28 de febrero tuvo dicha característica: <https://www.accuweather.com/es/co/bogota/107487/february-weather/107487>
- En todos los casos se utilizó una lupa de 7 cm de diámetro con un lente de 5x y 37 mm.
- El instrumento de medición, en todos los casos, fue un cronómetro de dos dígitos de precisión.

De esta forma se procuró que la variabilidad del experimento se explicara más en términos de los tres factores evaluados, que por diferencias entre las unidades experimentales causadas por la influencia del

entorno. No obstante, es importante notar que, aunque se procuró ser muy precisos al tomar las mediciones, los cálculos de ángulo y distancia están sujetos a cierto error del experimentador (considerando que no se utilizó ningún soporte para asegurar la posición de la lupa). Por otra parte, es importante notar que la selección de la combinación de los niveles para cada réplica se hizo al azar utilizando una función `Math.random()`, de forma que no hubo un patrón de recolección de datos.

Finalmente, es importante notar que, durante todo el análisis, cuando se habla de la primera llama se hace referencia a la primera muestra de combustión del material, que en general se da cuando se observa humo proveniente del mismo.

1.5. Supuestos utilizados para el desarrollo del experimento:

Para el análisis del presente experimento se tuvieron en cuenta los siguiente dos supuestos:

1. Normalidad sobre el conjunto de datos correspondientes a cada tratamiento.
2. La varianza se mantiene constante a lo largo de todas las observaciones; es decir, se supone el cumplimiento de la homocedasticidad.

En particular, puesto que el segundo supuesto es fundamental para que las pruebas estadísticas a desarrollar tengan validez, en la sección 3 se comprueba el supuesto de homocedasticidad a través de dos etapas:

1. Inspección gráfica
2. Prueba de Brown-Forsythe

2. Descripción de datos

2.1. Estadísticas descriptivas

Con el fin de entender mejor los datos recolectados se realizaron estadísticas descriptivas utilizando la herramienta R

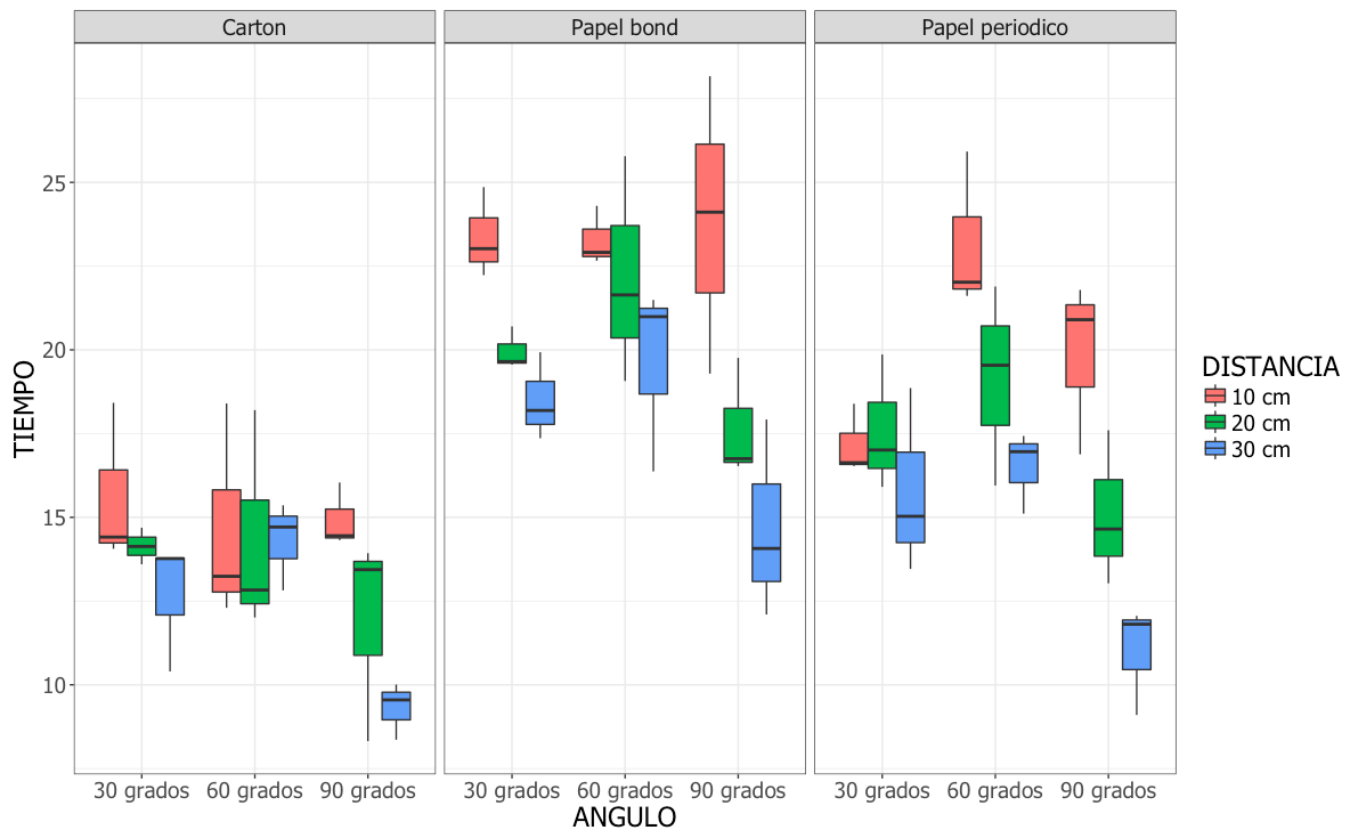
Celda	Ángulo	Distancia	Material	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis
1	30 grados	10 cm	Carton	1	3	15.63	2.42	14.41	15.63	0.52	14.06	18.42	4.36	0.38	-2.33
2	60 grados	10 cm	Carton	1	3	14.65	3.28	13.24	14.65	1.39	12.30	18.40	6.10	0.35	-2.33
3	90 grados	10 cm	Carton	1	3	14.94	0.96	14.45	14.94	0.19	14.32	16.04	1.72	0.38	-2.33
4	30 grados	20 cm	Carton	1	3	14.14	0.55	14.13	14.14	0.79	13.60	14.69	1.09	0.02	-2.33
5	60 grados	20 cm	Carton	1	3	14.35	3.36	12.83	14.35	1.22	12.01	18.20	6.19	0.36	-2.33
6	90 grados	20 cm	Carton	1	3	11.90	3.11	13.44	11.90	0.73	8.32	13.93	5.61	-0.37	-2.33
7	30 grados	30 cm	Carton	1	3	12.65	1.95	13.77	12.65	0.03	10.40	13.79	3.39	-0.38	-2.33
8	60 grados	30 cm	Carton	1	3	14.30	1.32	14.71	14.30	0.96	12.82	15.36	2.54	-0.28	-2.33
9	90 grados	30 cm	Carton	1	3	9.31	0.85	9.55	9.31	0.68	8.36	10.01	1.65	-0.26	-2.33
10	30 grados	10 cm	Bond	1	3	23.37	1.35	23.02	23.37	1.17	22.23	24.86	2.63	0.24	-2.33
11	60 grados	10 cm	Bond	1	3	23.29	0.88	22.91	23.29	0.37	22.66	24.30	1.64	0.35	-2.33
12	90 grados	10 cm	Bond	1	3	23.86	4.45	24.11	23.86	6.02	19.29	28.17	8.88	-0.06	-2.33
13	30 grados	20 cm	Bond	1	3	19.97	0.63	19.65	19.97	0.13	19.56	20.70	1.14	0.38	-2.33
14	60 grados	20 cm	Bond	1	3	22.16	3.39	21.64	22.16	3.81	19.07	25.78	6.71	0.15	-2.33
15	90 grados	20 cm	Bond	1	3	17.68	1.80	16.75	17.68	0.33	16.53	19.76	3.23	0.38	-2.33
16	30 grados	30 cm	Bond	1	3	18.49	1.31	18.19	18.49	1.23	17.36	19.93	2.57	0.22	-2.33
17	60 grados	30 cm	Bond	1	3	19.62	2.82	20.99	19.62	0.74	16.37	21.49	5.12	-0.37	-2.33
18	90 grados	30 cm	Bond	1	3	14.70	2.96	14.07	14.70	2.92	12.10	17.92	5.82	0.20	-2.33
19	30 grados	10 cm	Periódico	1	3	17.18	1.05	16.63	17.18	0.15	16.53	18.39	1.86	0.38	-2.33
20	60 grados	10 cm	Periódico	1	3	23.18	2.38	22.02	23.18	0.61	21.61	25.92	4.31	0.37	-2.33
21	90 grados	10 cm	Periódico	1	3	19.86	2.62	20.90	19.86	1.32	16.88	21.79	4.91	-0.34	-2.33
22	30 grados	20 cm	Periódico	1	3	17.59	2.04	17.01	17.59	1.63	15.91	19.86	3.95	0.26	-2.33
23	60 grados	20 cm	Periódico	1	3	19.13	2.99	19.54	19.13	3.48	15.95	21.89	5.94	-0.14	-2.33
24	90 grados	20 cm	Periódico	1	3	15.09	2.32	14.65	15.09	2.40	13.03	17.60	4.57	0.18	-2.33
25	30 grados	30 cm	Periódico	1	3	15.78	2.78	15.03	15.78	2.33	13.46	18.86	5.40	0.25	-2.33
26	60 grados	30 cm	Periódico	1	3	16.50	1.23	16.96	16.50	0.70	15.11	17.43	2.32	-0.32	-2.33
27	90 grados	30 cm	Periódico	1	3	10.99	1.64	11.81	10.99	0.37	9.10	12.06	2.96	-0.37	-2.33

De dichas estadísticas es importante notar que:

- La media de todos los tratamientos se encuentra en el rango de 9.31 seg a 23.86 seg
- La máxima desviación estándar fue de 4.45 seg; pero en general se encuentra por debajo de 3.5 seg
- El tiempo máximo de generación de la primera llama fue de 28.17 seg y sucedió cuando la lupa se ubicó a 90° con respecto a la vertical, 10 cm sobre el suelo y utilizando papel bond.
- El tiempo mínimo de generación de la primera llama fue de 8.32 seg y sucedió cuando la lupa se ubicó a 90° con respecto a la vertical, 20 cm sobre el suelo y utilizando cartón.

2.2. Gráficas de interés

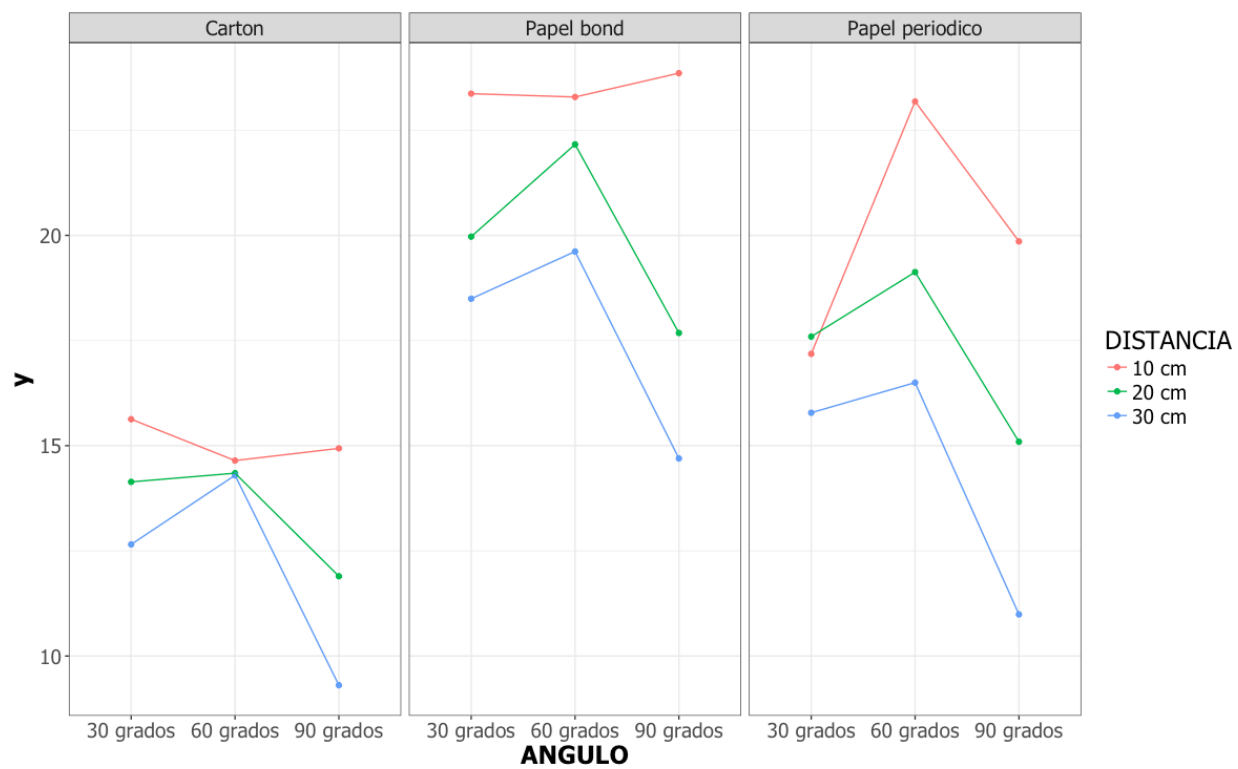
Diagrama de caja



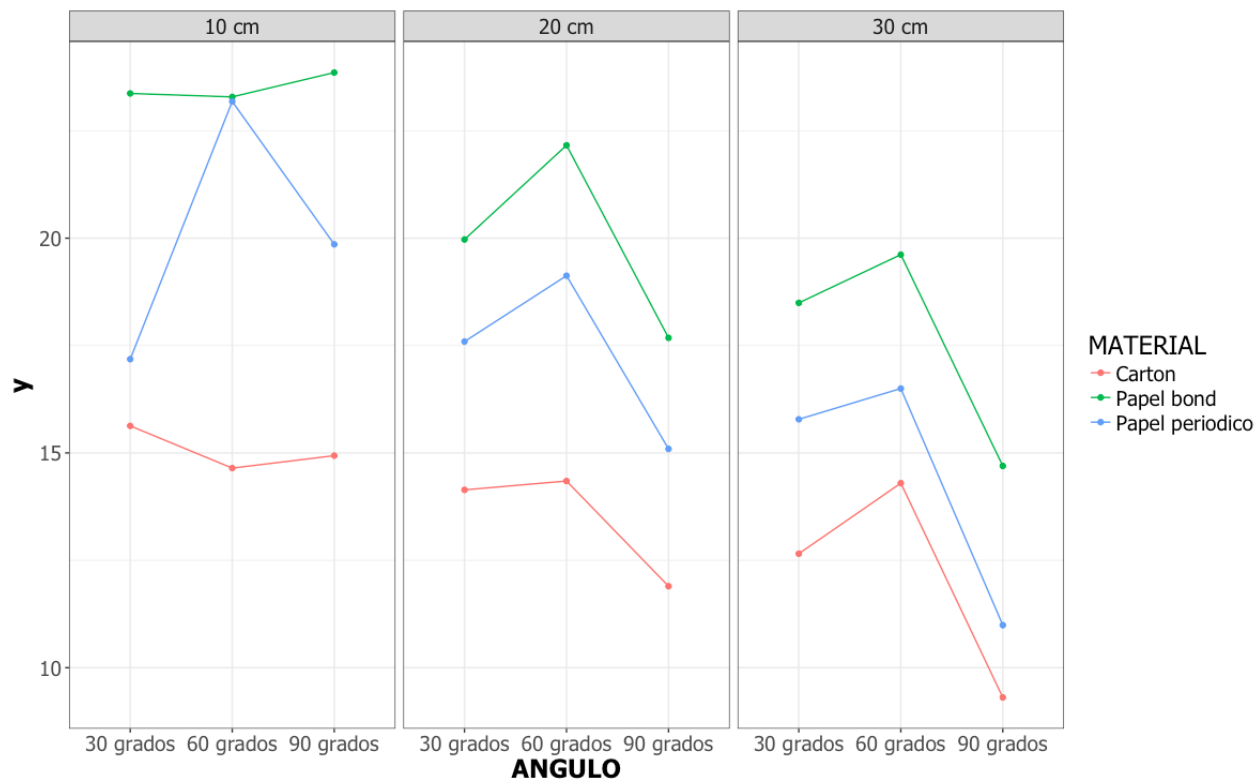
El diagrama de caja permite estudiar la variabilidad de los datos gráficamente. Este permite entender mejor cómo se comportan los datos en sus diferentes niveles, pero no resulta tan intuitivo evidenciar la significancia de factores en este caso.

Interacción triple

A continuación, se grafica cómo interactúan los diferentes factores entre sí. Este análisis gráfico permite evidenciar que la interacción triple no parece ser significativa; sin embargo, esto se estudiará de manera detallada durante la ANOVA.



También se puede graficar como:



2.3. Transformaciones de datos

Para este proyecto no es necesario realizar transformaciones a los datos pues estos cumplen con el supuesto de homocedasticidad requerido para realizar el análisis ANOVA. La prueba de este supuesto se incluye más adelante en el documento.

3. Métodos

Nota: Para todas las pruebas estadísticas se utilizó un nivel de significancia del 5%.

3.1. ANOVA

Modelo ANOVA con interacción triple:

Puesto que se desea analizar la influencia de los 3 factores, sus interacciones dobles y la interacción triple de estos se decidió realizar la siguiente prueba de hipótesis:

Sea $Y = \text{"Tiempo de generación de la primera llama de fuego"}$

H_0 : Los factores ni sus interacciones (de cualquier tipo) influyen sobre Y

$$H_0: \mu_{111} = \mu_{112} = \dots = \mu_{abc}$$

donde a, b, c son el número de niveles correspondientes a cada factor

H_1 : Al menos uno de los factores o sus interacciones SI influyen sobre Y

H_1 : Existe al menos un par de medias $\mu_{ijk} \neq \mu_{lmn}$ para cualquier $i, j, k \neq l, m, n$

Con base en lo anterior, a continuación, se muestra el resultado final de la ANOVA con triple interacción:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
MATERIAL	2	627.7	313.85	58.432	3.11e-14	***
ANGULO	2	139.7	69.87	13.008	2.45e-05	***
DISTANCIA	2	318.1	159.04	29.610	2.08e-09	***
MATERIAL:ANGULO	4	15.7	3.92	0.731	0.5750	
MATERIAL:DISTANCIA	4	24.9	6.22	1.157	0.3399	
ANGULO:DISTANCIA	4	65.0	16.24	3.024	0.0253	*
MATERIAL:ANGULO:DISTANCIA	8	34.5	4.32	0.804	0.6022	
Residuals	54	290.0	5.37			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						

En particular, puesto que se desea saber si la interacción triple influye sobre Y con el fin de verificar si es posible combinar su efecto con el error (para estimar la varianza σ^2 con más grados de libertad); entonces, se realiza la siguiente prueba estadística para verificar la existencia de interacción triple:

H_0 : No hay interacción triple entre los factores

H_1 : Existe interacción triple entre los factores

Se utiliza el siguiente estadístico de prueba:

$$F^{ABC} = \frac{MS_{ABC}}{MSE} \sim F(8, 54)$$
$$F^{ABC} = 0.804$$

Se calcula la F crítica con un nivel de significancia del 5%

$$F_{critico} = F(1 - 0.05, 8, 54) = 2.115$$

Ahora, puesto que $F^{ABC} < F_{critico}$, entonces, hay suficiente evidencia estadística para afirmar que no hay interacción triple entre los factores; es decir, no se rechaza H_0 . Lo anterior también se puede comprobar notando que el $PValue(ABC) > \alpha$

De esta forma, se puede concluir que si es posible combinar el efecto de la interacción tripe con el del error, de forma que σ^2 se estime mucho mejor.

Modelo ANOVA con interacción doble

Ahora, con base en lo anterior, se realiza la siguiente prueba estadística:

H_0 : Los factores ni sus interacciones dobles influyen sobre Y

$$H_0: \mu_{111} = \mu_{112} = \dots = \mu_{abc}$$

donde a, b, c son el número de niveles correspondientes a cada factor

H_1 : Al menos uno de los factores o sus interacciones dobles SI influyen sobre Y

H_1 : Existe al menos un par de medias $\mu_{ijk} \neq \mu_{lmn}$ para cualquier $i, j, k \neq l, m, n$

En la siguiente tabla se muestra el resultado de la ANOVA con interacciones dobles:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
MATERIAL	2	627.7	313.85	59.952	3.23e-15	***
ANGULO	2	139.7	69.87	13.347	1.51e-05	***
DISTANCIA	2	318.1	159.04	30.380	6.36e-10	***
MATERIAL:ANGULO	4	15.7	3.92	0.750	0.5620	
MATERIAL:DISTANCIA	4	24.9	6.22	1.187	0.3252	
ANGULO:DISTANCIA	4	65.0	16.24	3.103	0.0216	*
Residuals	62	324.6	5.24			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						

De allí se observa que los PValues asociados a las interacciones dobles Material:Ángulo y Material:Distancia no cambian significativamente con respecto a los obtenidos en la ANOVA con triple interacción. Esto permite concluir que combinar el efecto de la interacción triple con el error no es de gran ayuda a la hora de mejorar la estimación de σ^2 ; de hecho, se observa que las conclusiones sobre la influencia de los factores y sus interacciones dobles no cambia.

Teniendo en cuenta que el primer modelo ofrece una buena estimación de σ^2 y que, como se verá más adelante, se cumple el supuesto de homocedasticidad; entonces, con un nivel de significancia del 5% se puede concluir que los tres factores (material, ángulo y distancia) junto con la interacción doble entre ángulo y distancia influyen en la generación de la primera llama de fuego.

Ahora bien, como ejercicio académico, se realizó un último análisis de varianza combinando el efecto de las interacciones Material:Ángulo y Material:Distancia con el error. De allí, como era de esperarse, se obtuvo el siguiente resultado cuyas conclusiones son iguales a las previamente descritas:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
MATERIAL	2	627.7	313.85	60.168	6.24e-16	***
ANGULO	2	139.7	69.87	13.395	1.19e-05	***
DISTANCIA	2	318.1	159.04	30.490	2.99e-10	***
ANGULO:DISTANCIA	4	65.0	16.24	3.114	0.0204	*
Residuals	70	365.1	5.22			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Puesto que este último modelo confirma los resultados previamente descritos y, además, combina el efecto de múltiples factores no influyentes con el error; se decidió utilizarlo en los análisis posteriores de selección de tratamientos.

La información previamente descrita resulta valiosa para cumplir el objetivo del proyecto; sin embargo, falta validar que se cumplen los supuestos requeridos para realizar el ANOVA y encontrar la configuración óptima para generar la llama en el menor tiempo posible.

3.2. Validación de los modelos y procedimientos estadísticos utilizados

Validación del supuesto de homocedasticidad

Prueba Brown-Forsythe

a) Transformación de los datos para probar con el factor distancia

Se debe realizar la prueba por factor pues no se tienen suficientes datos para que el análisis por celda sea significativo. Para este motivo se eligió el factor distancia. Se encuentran las medianas de cada uno de los niveles del factor:

```
> m1 = median(datos$TIEMPO..Seg.[datos$DISTANCIA=='10 cm'])
> m1
[1] 19.29
> m2 = median(datos$TIEMPO..Seg.[datos$DISTANCIA=='20 cm'])
> m2
[1] 16.75
> m3 = median(datos$TIEMPO..Seg.[datos$DISTANCIA=='30 cm'])
> m3
[1] 14.71
```

Luego, se transforman los datos iniciales convirtiéndolos en el valor absoluto de la resta entre el valor original y la mediana del nivel correspondiente.

b) La prueba

Para validar el supuesto se realizó la siguiente prueba estadística:

$$Y_{ijk} \rightarrow Z_{ijk} = |Y_{ijk} - \widetilde{Y}_{..k}|$$

donde el primer subíndice está asociado al material, el segundo al ángulo y el tercero a la distancia.

$$H_0: \sigma_{111}^2 = \sigma_{112}^2 = \dots = \sigma_{abc}^2$$

donde a, b, c son el número de niveles correspondientes a cada factor

H_1 : Existe al menos un par de varianzas $\sigma_{ijk}^2 \neq \sigma_{lmn}^2$ para cualquier $i, j, k \neq l, m, n$

El resultado de realizar la prueba se observa a continuación:

```

              Df Sum Sq Mean Sq F value Pr(>F)
datos$DISTANCIA  2   38.2   19.109    2.909 0.0605 .
Residuals      78   512.3    6.568
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Luego de realizar el ANOVA con los datos transformados, se puede establecer que, como el valor de F no se encuentra en la región crítica, no se rechaza la hipótesis nula y por lo tanto con una significancia del 5% se concluye que los datos cumplen con el supuesto de homocedasticidad requerido para realizar el ANOVA.

Ahora, puesto que se encontró que el PValue asociado es muy cercano al nivel de significancia propuesto (0.05); entonces, se decidió repetir la prueba de Brown Forsythe considerando las medianas asociadas a los factores material y ángulo. Para ello se utilizó la misma prueba estadística que se indicó previamente y se obtuvieron los siguientes resultados:

1. Para el caso de las medianas de los niveles del material:

```

> # Brown-Forsythe test for material
> # 1. Data transformations to test distance
> m1m = median(datos$TIEMPO..Seg.[datos$MATERIAL=='Papel bond'])
> m1m
[1] 19.93
> m2m = median(datos$TIEMPO..Seg.[datos$MATERIAL=='Carton'])
> m2m
[1] 13.79
> m3m = median(datos$TIEMPO..Seg.[datos$MATERIAL=='Papel periodico'])
> m3m
[1] 16.96
> mediansm<-c(rep(m1m,27),rep(m2m,27),rep(m3m,27))
> datos$z=abs(datos$TIEMPO..Seg.-mediansm)
> # 2. The test
> bf=aov(z~datos$MATERIAL,data=datos)
> summary(bf)
              Df Sum Sq Mean Sq F value Pr(>F)
datos$MATERIAL  2     1.0    0.502    0.051 0.951
Residuals     78   773.9    9.922

```

Se observa que el PValue correspondiente es mayor al nivel de significancia (0.05) y por lo tanto hay evidencia estadística suficiente para no rechazar la hipótesis nula; es decir, las varianzas son iguales.

2. Para el caso de las medianas de los niveles del ángulo:

```
> # Brown-Forsythe test for angle
> # 1. Data transformations to test distance
> m1g = median(datos$TIEMPO..Seg.[datos$ANGULO=='30 grados'])
> m1g
[1] 17.01
> m2g = median(datos$TIEMPO..Seg.[datos$ANGULO=='60 grados'])
> m2g
[1] 18.4
> m3g = median(datos$TIEMPO..Seg.[datos$ANGULO=='90 grados'])
> m3g
[1] 14.45
> mediansg<-c(rep(m1g,27),rep(m2g,27),rep(m3g,27))
> datos$z=abs(datos$TIEMPO..Seg.-mediansg)
> # 2. The test
> bf=aov(z~datos$ANGULO,data=datos)
> summary(bf)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
datos\$ANGULO	2	14.2	7.118	1.165	0.317
Residuals	78	476.6	6.110		

Al igual que en el caso anterior, se observa que el PValue es mayor que el nivel de significancia. Por ende, se confirma que, para todos los factores, la prueba de Brown-Forsythe permite concluir que se cumple el supuesto de homocedasticidad.

Selección de tratamientos

Dado que ya se identificó cuáles factores e interacciones influyen, se realizará una prueba de Tukey para determinar cuáles tratamientos permiten minimizar el tiempo de generación de la primera llama de fuego. Para ello se utilizó la siguiente prueba de hipótesis:

$$H_o^{ij}: \mu_i = \mu_j$$

$$H_1^{ij}: \mu_i \neq \mu_j \forall i \neq j \text{ con } i, j \in \{1, 2, \dots, k\}$$

donde k corresponde al número de niveles del factor correspondiente

Dicha prueba se realizó 3 veces (1 por cada factor), con el fin de determinar cuál (es) niveles de cada factor son más convenientes para minimizar el tiempo de generación de la primera llama. Con este propósito, se utilizó el siguiente estadístico de prueba:

$$EP = \frac{\bar{Y}_i - \bar{Y}_j}{\sqrt{\frac{MSE}{n}}} \sim q(k, g, L Error)$$

Donde el MSE y los grados de libertad del error corresponden a los obtenidos en la última ANOVA calculada.

Con base en lo anterior, al efectuar la prueba con R, se obtuvieron los siguientes resultados:

- En primer lugar, se muestran los resultados de la prueba de Tukey para el factor material:

```
> TukeyHSD(fit3, "MATERIAL")
Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = TIEMPO ~ MATERIAL + ANGULO + DISTANCIA + ANGULO * DISTANCIA, data = datos)

$MATERIAL
              diff      lwr      upr    p adj
Papel bond-Carton    6.809259  5.320799  8.297719 0.00e+00
Papel periodico-Carton 3.717407  2.228948  5.205867 3.00e-07
Papel periodico-Papel bond -3.091852 -4.580312 -1.603392 1.33e-05
```

Como los p-values para las tres comparaciones son menores al nivel de significancia de 5%, se rechazan las hipótesis nulas que plantean que las medias por pares son iguales. Esto significa que las estadísticamente, las medias del tiempo que toma en crearse la primera llama son distintas con cada material. Sin embargo, con la comparación de medias de Tukey se puede establecer un orden. De la primera prueba se obtiene que el tiempo es mayor con papel bond que con cartón. De la segunda prueba se obtiene que el papel periódico toma mayor tiempo en prenderse que el cartón. De la última prueba se obtiene que el papel periódico toma menos tiempo que el papel bond.

Con esto se pueden organizar de la siguiente forma: $\mu \text{ cartón} < \mu \text{ papel periódico} < \mu \text{ papel bond}$. En el contexto del problema esto significa que si se desea minimizar el tiempo que toma en prenderse la primera llama debería utilizarse el cartón.

- En segundo lugar, se muestran los resultados de la prueba Tukey para el factor distancia:

```
> TukeyHSD(fit3, "DISTANCIA")
Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = TIEMPO ~ MATERIAL + ANGULO + DISTANCIA + ANGULO * DISTANCIA, data = datos)

$DISTANCIA
              diff      lwr      upr    p adj
20 cm-10 cm -2.660370 -4.148830 -1.1719105 0.0001708
30 cm-10 cm -4.846296 -6.334756 -3.3578364 0.0000000
30 cm-20 cm -2.185926 -3.674386 -0.6974661 0.0022020
```

Como los p-values para las tres comparaciones son menores al nivel de significancia de 5%, se rechazan las hipótesis nulas que plantean que las medias por pares son iguales. Esto significa que,

estadísticamente, las medias del tiempo que toma en crearse la primera llama son distintas para cada distancia. Sin embargo, con la comparación de medias de Tukey se puede establecer un orden. De la primera prueba se obtiene que el tiempo es mayor con 10cm que con 20cm. De la segunda prueba se obtiene que el tiempo es menor con 30cm que con 20cm. De la última prueba se obtiene que con 30cm el tiempo es menor que con 10cm.

Con esto se pueden organizar de la siguiente forma: $\mu_{30cm} < \mu_{20cm} < \mu_{10cm}$. En el contexto del problema esto significa que si se desea minimizar el tiempo que toma en prenderse la primera llama debería ponerse la lupa a 30cm de distancia.

```
> TukeyHSD(fit3, "ANGULO")
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = TIEMPO ~ MATERIAL + ANGULO + DISTANCIA + ANGULO * DISTANCIA, data = datos)

$ANGULO
              diff       lwr       upr    p adj
60 grados-30 grados  1.372593 -0.1158673  2.8610524 0.0767713
90 grados-30 grados -1.833704 -3.3221636 -0.3452439 0.0118765
90 grados-60 grados -3.206296 -4.6947561 -1.7178364 0.0000066
```

Como el p-value de la primera prueba es mayor al nivel de significancia, no se rechaza la hipótesis nula y por lo tanto existe evidencia estadística de que las medias de tiempo para 60 ° y 30 ° son iguales. Por otra parte, puesto que el p-value de la segunda prueba también es mayor al nivel de significancia, no se rechaza la hipótesis nula y, nuevamente, se concluye que existe evidencia estadística para afirmar que las medias de tiempo para 90 ° y 30 ° son iguales.

Finalmente, puesto que en la última comparación el p-value es inferior al nivel de significancia, se concluye que las medias para 90 ° y 60 ° son diferentes. Además, dado que el resultado de la diferencia es negativo, se puede afirmar que la media del tiempo es mayor con 60 ° que con 90 °.

De lo anterior se obtiene el siguiente orden: $\mu_{60^\circ} > \mu_{90^\circ} = \mu_{30^\circ}$; por ello, si se desea minimizar el tiempo de generación de la primera llama la lupa se debería ubicar a 90° o a 30° con respecto a la vertical.

4. Resultados y conclusiones

Con base en los resultados obtenidos a través de las distintas pruebas estadísticas se puede concluir que las configuraciones más acertadas para minimizar el tiempo de generación de la primera llama son:

Material	Ángulo	Distancia
Cartón	90°	30 cm
Cartón	30°	30 cm

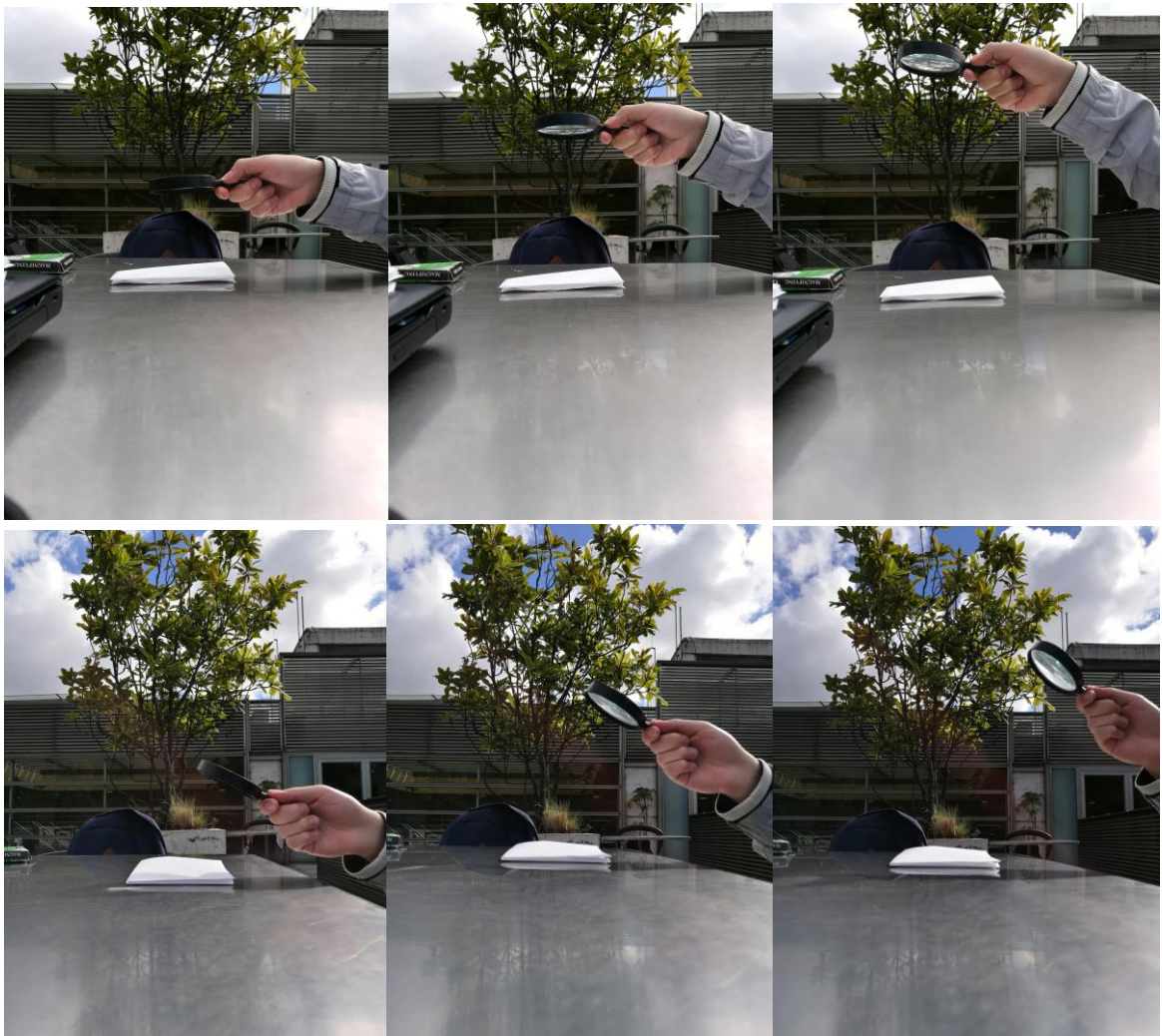
De esta forma, se observa que el material a utilizar para optimizar el proceso de encendido de la fogata es el cartón. Asimismo, la lupa debe ubicarse a 90° o a 30° con respecto a la horizontal y la distancia entre el suelo y la lupa debe ser de 30cm.

Por otra parte, es importante recalcar que las pruebas estadísticas que dieron lugar a tales configuraciones, solo tienen validez cuando se tienen las condiciones descritas en la introducción del presente texto. Es decir, cuando el día es soleado y el sol se encuentra muy próximo al cenit. De igual forma, la calidad de la lupa también es determinante en que dichas combinaciones de niveles sean precisamente las que minimizan el tiempo de generación de la primera llama.

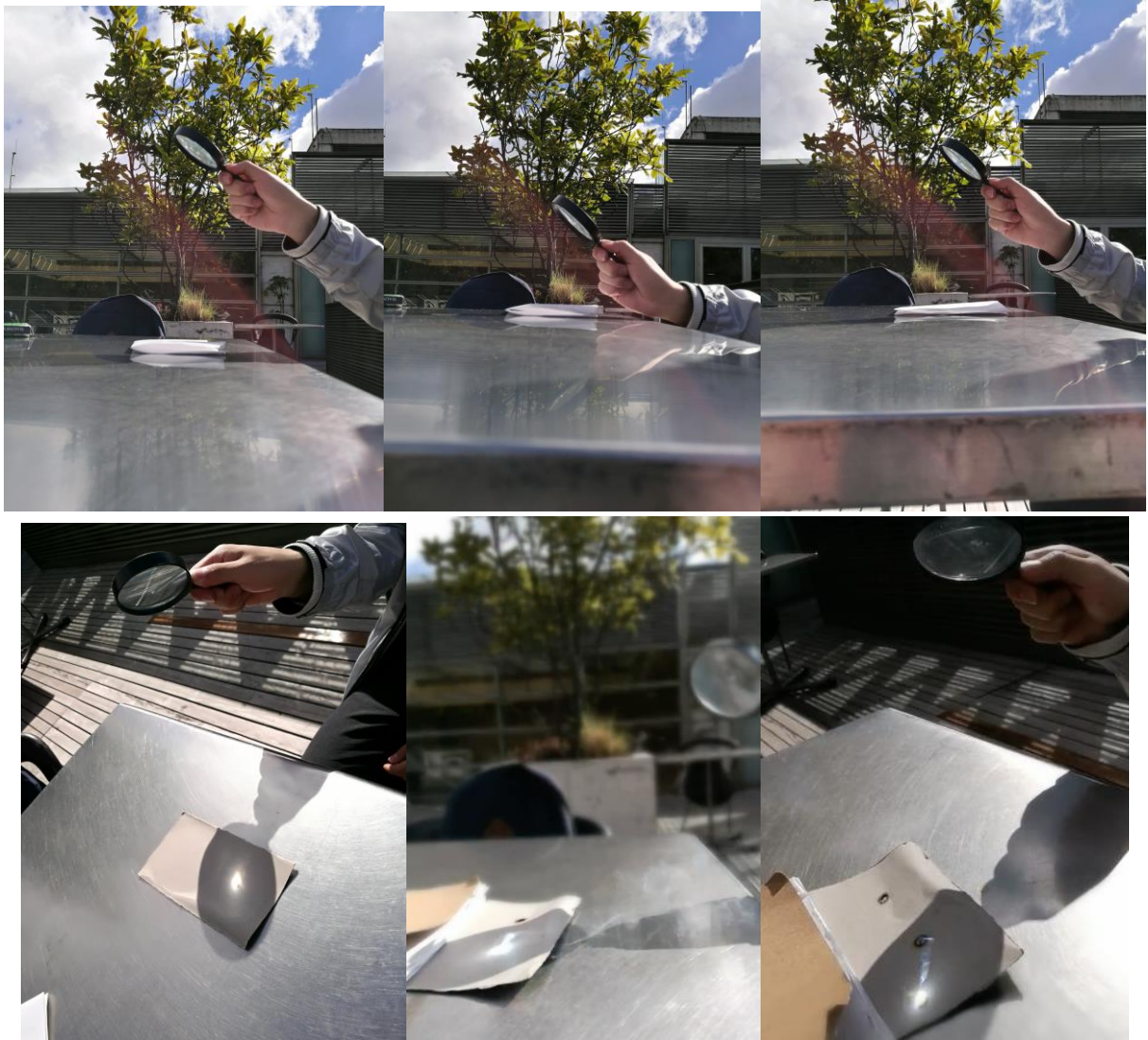
Finalmente, con base en el análisis estadístico realizado, se le recomienda que los exploradores empiecen a encender la fogata al mediodía, utilicen una lupa de 5x y 37 mm a 30 cm del suelo, con un ángulo de 90° (o 30° según vean más conveniente) respecto a la vertical y traten de llevar un cartón ligero dentro de sus implementos del campamento para iniciar el fuego rápidamente.

5. Anexos

5.1. Soporte fotográfico







5.2. Código utilizado

Todas las gráficas y modelos realizados fueron generados utilizando R Studio. El código utilizado durante el proyecto puede ser encontrado en <https://github.com/cpenalozag/ANOVAProject>