Universidad del Valle
Programa Académico de Estadística
Estadística Aplicada III
Cali, marzo de 2019

Diseño de bloques completos al azar y diseño factorial

Kevin Steven García[a], Alejandro Vargas[b], Alejandro Soto[c]

1. Punto 1

Uno de los objetivos de la industria metalúrgica nacional es determinar cuál de los tres elementos: níquel, hierro o cobre es el mejor agente soldante. Se sueldan una serie de lingotes de acero utilizando cada uno de los posibles agentes soldantes (níquel, hierro o cobre). Existen diferencias entre los lingotes de acero utilizados en este proceso, se utiliza en total 7 lingotes (todos diferentes), se mide la fuerza (expresada en 1000 libras por pulgada cuadrada) necesaria para soldar los lingotes. Los resultados fueron:

	AGENTE	SOLDANTE	
LINGOTE	NIQUEL	HIERRO	COBRE
1	67.0	71.9	72.2
2	67.5	68.8	66.4
3	76.0	82.6	74.5
4	72.7	78.1	67.3
5	73.1	74.2	73.2
6	65.8	70.8	68.7
7	75.6	84.9	69.0

a. El diseño de bloques completos al azar (DBCA) es el adecuado para este problema, ya que, se quieren comparar 3 tratamientos contenidos en 7 bloques (lingotes) relativamente homogéneos, se consideran los lingotes como bloques, ya que nos dicen que los 7 son diferentes y teniendo en cuenta esto, se podría decir que es un factor perturbador el cual tiene un efecto en la variable de respuesta pero no es de nuestro interés, entonces se desea controlar esta variabilidad que surge de este factor (lingote) que podría afectar en los resultados.

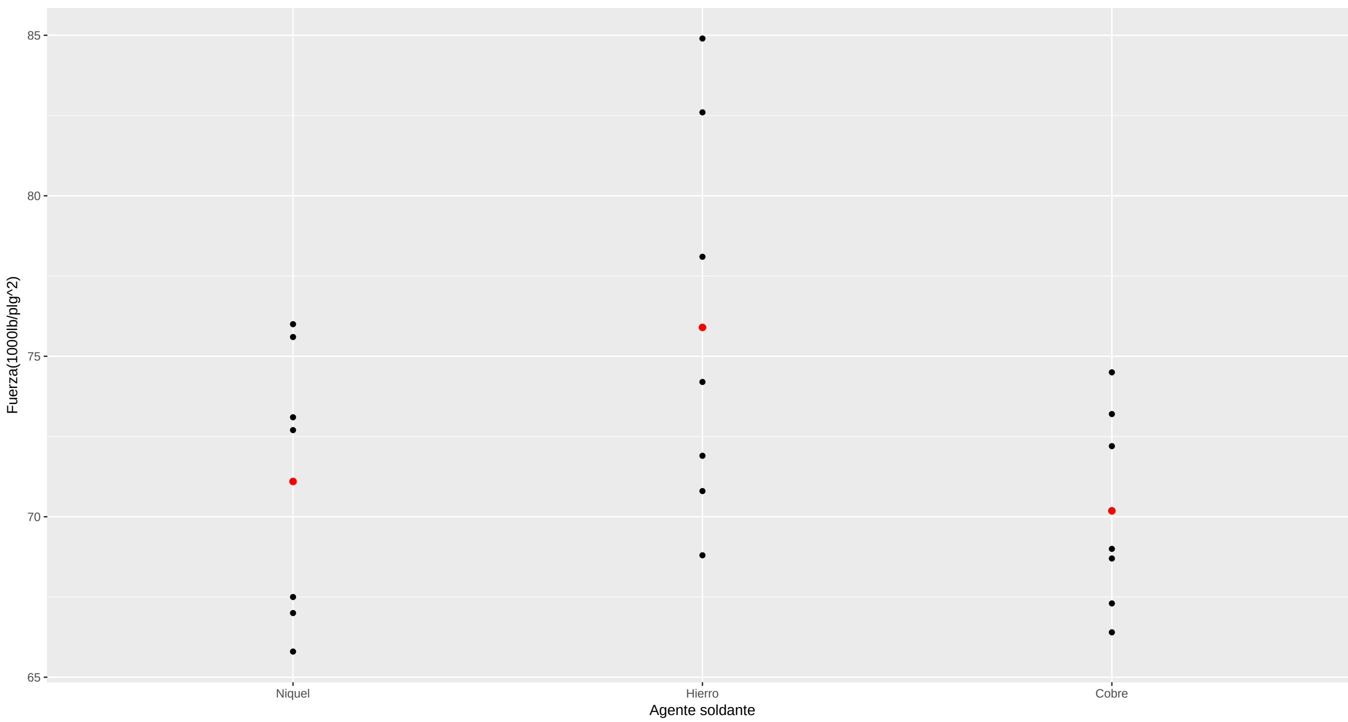
b.
- Unidad experimental: Lingote
- Factor de tratamiento: Agente Soldante
- Niveles: Niquel, Hierro, Cobre.
- Factor de control: Tipo de lingote
- Niveles: 1,2,3,4,5,6 y 7.
- Tratamientos: Niquel, Hierro y Cobre.
- Variable de respuesta: Fuerza necesaria para soldar los lingotes ($1000lb/plg^2$)

c. **Estadísticas descriptivas:**
Dado que nuestro interés se centra en los tratamientos que en este caso son los niveles de los agentes soldantes (Niquel, Hierro y Cobre), obtendremos descriptivas por estos niveles y no por los tipos de lingotes(bloques) los cuales no son de nuestro interés:

[a]Código: 1533173. E-mail: kevin.chica@correounivalle.edu.co
[b]Código: 1525953. E-mail: jose.alejandro.vargas@correounivalle.edu.co
[c]Código: 1532457. E-mail: asotomurillo@gmail.com

	Mín	Q1	Mediana	Media	Q3	Máx	Des. Estándar	CV(%)
Niquel	65,8	67,25	72,7	71,1	74,35	76,0	4,25597619	5,985901815
Hierro	68,8	71,35	74,2	75,9	80,35	84,9	6,137860648	8,086772922
Cobre	66,4	68	69,0	70,2	72,7	74,5	3,109892051	4,430947356

En la tabla anterior, se podría decir a priori, que no se aprecian diferencias significativas entre los lotes en cuanto a la cantidad de calcio que presentan. Todas tienen valores muy parecidos en cada uno de los estadísticos obtenidos. Poseen casi la misma media, mediana y desviación.



En ésta gráfica dado que el eje y reduce demasiado su escala, se pueden apreciar diferencias decimales, y se observa que en decimales, el lote 1, 2 y 3 son parecidos, siendo el 3 el que más calcio contiene, y los lotes 4 y 5 son casi iguales, y tienen menor calcio que los otros 3 lotes.

d.

Se plantea el siguiente modelo:

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \qquad i = 1, ..., 3 \qquad j = 1, ..., 7$$

El cuál se construye bajo los siguientes supuestos:

$$\varepsilon_{ij} \approx N(0, \sigma^2), \ \ E[\varepsilon_{ij}] = 0 \ \ \forall_{ij}, \ \ V[\varepsilon_{ij}] = \sigma^2, \ \ cov[\varepsilon_{ij}, \varepsilon_{i'j'}] = 0$$

Donde:
y_{ij} =Fuerza necesaria para soldar el lingote j-ésimo con el agente soldante i-ésimo.
μ =Media general de contenido de calcio por lote sin tener en cuenta el agente soldante ni el tipo de lingote.
τ_i =Efecto del i-ésimo agente soldante sobre la fuerza necesaria para soldar el lingote.
β_j =Efecto del j-ésimo lingote sobre la fuerza necesaria para soldarlo.
ε_{ij} =Error aleatorio debido al i-ésimo agente soldante y a el j-ésimo lingote.

e.

Dado que no nos interesan los bloques(tipo de lingote), ya que a priori nos dijeron que entre estos existen diferencias. Nos centraremos en los efectos de los tres tratamientos o agentes soldantes. Por lo que se plantea la siguiente hipótesis:

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau$$
H_1 :Al menos una de las igualdades no se cumple.

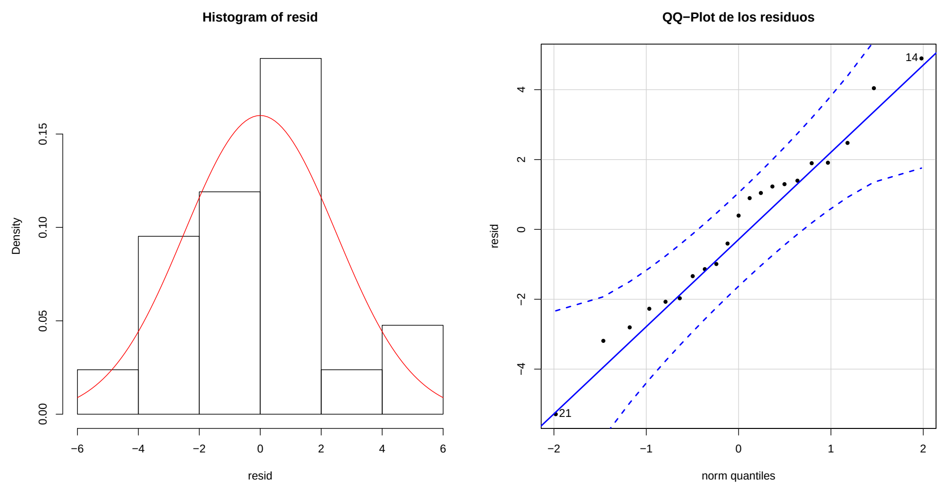
f.

Antes de realizar la ANOVA e interpretar los resultados, se comprobarán los supuestos del modelo para saber si las interpretaciones que obtengamos si van a ser correctas.
Sobre los errrores se tienen los siguientes supuestos:

$$\varepsilon_{ij} \approx N(0, \sigma^2) \rightarrow \quad Normalidad$$

$$E[\varepsilon_{ij}] = 0 \quad \forall_{ij} \rightarrow \quad Correcta\ especificacion$$
$$V[\varepsilon_{ij}] = \sigma^2 \quad \forall_{ij} \rightarrow \quad Homocedasticidad$$
$$cov[\varepsilon_{ij}, \varepsilon_{i'j'}] = 0 \quad \forall_{i \neq i'} \quad \wedge \quad \forall_{j \neq j'} \rightarrow \quad Independencia$$

- Normalidad: Para probar el supuesto de normalidad, se realizarán dos pruebas gráficas (histograma y QQ-plot) y una prueba estadística(Shapiro Wilk).



En estas gráficas se podría pensar que el supuesto no se cumple, ya que la cola derecha de la distribución se sale de los intervalos de confianza de la recta de cuantiles teórica de la distribución normal. Sin embargo, las pruebas gráficas no son muy convincentes ya que su interpretación puede ser distinta de una persona a otra y depende mucho de la experiencia del investigador. Por ello se acude a la prueba formal.

* Test Shapiro-Wilk:

$$H_0 : Los\ datos\ provienen\ de\ una\ distribucion\ normal$$
$$H_a : Los\ datos\ no\ provienen\ de\ una\ distribucion\ normal$$
$$W = 0.98289, p - value = 0.9606$$

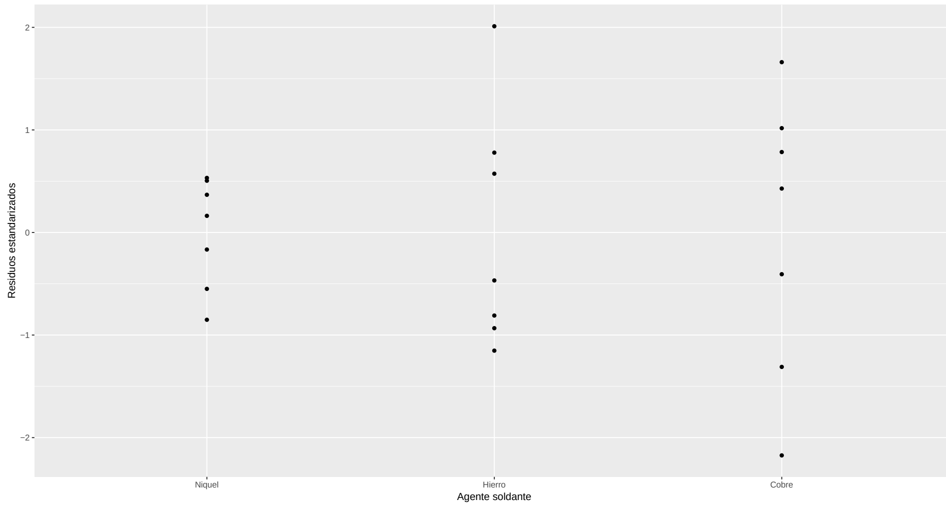
Como el p-valor=0.1192 es superior a nuestro nivel de significancia $\alpha = 0.05$, no se rechaza la hipótesis nula, y concluimos que no hay evidencia estadística suficiente para decir que los datos no provienen de una distribución normal, lo que nos indica que este supuesto de normalidad en los errores no se incumple.

- Correcta Especificación:Sabemos que este supuesto siempre se cumple, ya que el método o la construcción del modelo asegura que la esperanza de los residuales sea igual a cero, sin embargo, se realizó una prueba t para comprobarlo.

* T-test:

$$H_0 : La\ verdadera\ media\ es\ igual\ a\ 0$$
$$H_a : La\ verdadera\ media\ no\ es\ igual\ a\ 0$$
$$t = 1.9109e - 16, df = 24, p - value = 1$$

Como el p-valor=1 es superior a nuestro nivel de significancia $\alpha = 0.05$, no se rechaza la hipótesis nula, y se verifica lo dicho anteriormente.



- Homocedasticidad:Para analizar la homocedasticidad se realizó una gráfica(Residuos estandarizados por agente soldante) y una prueba estadísticas(Test de Barlett)
  En esta gráfica se confirma o dicho en las estadísticas descriptivas, las varianzas entre los residuos de cada factor son muy parecidas. Claramente se ve que se cumple el supuesto de homocedasticidad, sin embargo se realiza la prueba formal, la cual se espera que tenga un p valor bastante alto, por la evidencia que se tiene.

* Prueba de Barlett:

$$H_0 : Las\ varianzas\ de\ todos\ los\ grupos\ son\ iguales$$

$$H_a : Al\ menos\ una\ delas\ varianzas\ es\ diferente$$

$$Bartlett's\ K - squared = 4.3295, df = 2, p - value = 0.1148$$

  Como el p-valor=0.9998 es superior a nuestro nivel de significancia $\alpha = 0.05$, no se rechaza la hipótesis nula, y concluimos que no hay evidencia estadística suficiente para decir que al menos entre dos grupos la varianza es diferente, lo que nos indica que este supuesto de homocedasticidad en los errores no se incumple.

- Independencia:Para probar el supuesto de independencia en los errores, no se debe realizar un correlograma o una prueba sobre la autocorrelación(prueba de Durbin-Watson) ya que estas dependen del orden en el cuál se obtuvieron los datos, y esa información no se tiene, por lo tanto, solo se realizará una prueba de rachas.

* Test de rachas:

$$H_0 : La\ muestra\ es\ aleatoria\ (las\ observaciones\ son\ independientes)$$

$$H_a : La\ muestra\ no\ es\ aleatoria\ (las\ observaciones\ no\ son\ independientes)$$

$$Standard\ Normal = 0.68395, p - value = 0.494$$

  Como el p valor=0.62212 es mayor que nuestro nivel de significancia $\alpha = 0.05$, no rechazamos la hipótesis nula y concluimos que no hay evidencia estadística suficiente para decir que los residuales no están distribuidos de manera aleatoria, en otras palabras, no podemos decir que no son independientes. Teniendo en cuenta esto, podemos decir que este supuesto de independencia en los errores parece no incumplirse.

Ya que según las pruebas aplicadas, ninguno de los supuestos se incumple, procedemos a realizar e interpretar la ANOVA:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Agente	2	131.9	65.95	6.359	0.0131*
Lingote	6	268.3	44.71	4.311	0.0151*
Residuals	12	124.5	10.37		

Podemos ver que el p valor=0.00363 es inferior a nuestro nivel de significancia $\alpha = 0.05$, por lo cual se rechaza H_0 planteada en el item e, y concluimos que al menos uno de los tratamientos(lote) tiene un efecto significativo sobre la cantidad de calcio.

g. Se realizará la prueba postanova de comparaciones múltiples de Tukey. Los resultados resumidos para esta prueba son:

Agente	lsmean	SE	df	lower.CL	upper.CL	.group
3Cobre	70.2	1.22	12	67.5	72.8	a
1Niquel	71.1	1.22	12	68.4	73.8	a
2Hierro	75.9	1.22	12	73.2	78.6	b

Los tratamientos con la misma letra no son significativamente diferentes. Es decir, los lotes 3 y 2 no son significativamente diferentes, siendo estos, los que más contenido de calcio tienen. El lote 1 no es significativamente diferentes a los dos lotes mencionados anteriormente, sin embargo, tampoco es diferente al lote 4, esto se podría interpretar como que este lote 1 tiene una cantidad de calcio promedio, lo cual se observa en la gráfica del item c. El lote 4 es diferente a los lotes 3 y 2(presenta una cantidad de calcio inferior), y es estadísticamente igual a los lotes 1 y 5. Finalmente el lote 5 es diferente a los lotes 3,2 y 1, y semejante al lote 4, siendo este, el que menor cantidad de calcio presenta.

h. Las conclusión principal que se obtiene es que la suposición del fabricante sobre la existencia de diferencias en el contenido de calcio en los lotes de materia prima resulto ser cierta, en la ANOVA se detectó que al menos uno de los lotes difiere en cuanto al contenido de calcio. Posteriormente en la prueba postanova de Fisher, se observó que los lotes 3 y 2 son los que más contenido de calcio presentan, el lote 1 presenta un contenido de calcio promedio, y los lotes 4 y 5 son los que menor contenido de calcio tienen. Si nos vamos al contexto de los datos, los resultados nos están diciendo que existen diferencias estadisticamente significativas, sin embargo, estas diferencias pueden no ser grandes para el fabricante, ya que se detectaron diferencias decimales.Entonces queda a decisión del fabricante si realmente acepta la decisión o si por el contrario, las diferencias no le parecen significativas dentro del contexto.

2. Punto 2

3. Punto 3

4. Punto 4