Diseño de Experimentos 3009340, semestre 02 de 2021, equipo de trabajo # 4

INFORME TALLER 1

Karen Andrea Amaya [*] Ana Maria Sanchez [**] Jhonatan Smith Garcia [***] Carolina Vergara Clavijo [****]

Fecha de entrega: $16-11-2021$

Índice

[*] Estudiante Estadística, Universidad Nacional de Colombia – Sede Medellín.
[**] Estudiante Estadística, Universidad Nacional de Colombia – Sede Medellín.
[***] Estudiante Estadística, Universidad Nacional de Colombia – Sede Medellín.
[****] Estudiante Estadística, Universidad Nacional de Colombia – Sede Medellín.

1. Punto 1

Se desea investigar la influencia de la temperatura en el rendimiento de un proceso químico, en particular interesa investigar un rango de temperatura entre 60 a 120° C. Se tienen recursos para realizar 20 corridas experimentales:

(a) Los niveles de temperatura con los que se experimentan son 60, 65, 70 y 120° C, haciendo cinco replicaciones en cada nivel. ¿Considera que es adecuado el diseno experimental usado? Argumente, y de ser necesario proponga alternativas.

Dada la naturaleza del experimento y el interés particular del investigador se aconseja utilizar medidas equidistantes respecto a la temperatura, puesto que, los niveles propuestos tienen una tendencia a acumularse entre 60 y 70, dejando un sesgo de información entre los niveles de temperatura de 70 a 120 pues allí no se recolectan datos. Esto representa un problema puesto que dentro del marco teórico y las corridas experimentales, no se obtendría información relevante de los otros niveles.

Por tanto, los valores recomendados para los nuevos niveles son: 60, 80, 100, 120 grados centígrados. Estos valores se distribuyen de manera uniforme entre el rango de interés. Además, permite que al tomar 5 muestras en cada nivel, se siga cumpliendo el requisito de 20 corridas experimentales sumado al hecho de ahora tener información sobre todo el rango de temperatura de interés.

Al ser un experimento de factores fijos dado 4 niveles, se debe asegurar homogeneidad entre las unidades experimentales. Dado que el enunciado no tiene ningún tipo de aseveraciones respecto a la no dependencia de unidades entre si, se asumen independientes y homogéneas. Esto garantiza que fijar la temperatura en niveles diferentes no modificará la respuesta dada una UE particular.

(b) El orden en que se decidieron hacer las corridas experimentales para facilitar el trabajo experimental fue: primero las cinco del nivel bajo de temperatura, luego las cinco del siguiente y así hasta finalizar ¿Es correcto lo que hicieron? Argumente.

El diseño de experimento de un experimento particular debe asegurar aleatoriedad en el proceso de toma de datos, tanto para el análisis de datos por niveles, como para la selección de la muestra, la corrida del experimento e inferencias a realizar. La forma en que el problema plantea la toma de datos va en contra de un principio de aleatoriedad puesto que, claramente se está sesgando la toma de datos al asignarle un orden especifico. La forma correcta de realizar las corridas de los experimentos seria, aleatorizar el proceso, sin importar niveles en el mismo; es decir, si hay 4 niveles y cada nivel tendrá 5 muestras, aleatorizar el proceso de toma de muestras. Puede ser con ayuda de un software, con el paquete *"sample"* R por ejemplo.

(c) Para hacer el análisis estadístico, compararon mediante una prueba t de Student de dos en dos las medias en los niveles de temperatura y obtuvieron conclusiones ¿Es adecuado el análisis? Argumente y en su caso proponga alternativas.

Para una correcta realización del experimento comparando por pares de medias se ha de tener en cuenta que dicha comparacion por medias pareadas esta dada por dos métodos, LSD y HSD. Cada uno de estos métodos asegura la comparación de medias muestrales entre muestras independientes. En este caso, el experimento se asume con muestras independientes entre si.

	Método LSD	Método de Tukey
Estadístico de prueba	$D_{ij} = \left\|\bar{Y}_{i\bullet} - \bar{Y}_{j\bullet}\right\|$	$D_{ij} = \left\|\bar{Y}_{i\bullet} - \bar{Y}_{j\bullet}\right\|$
Valor crít. para D_{ij}	$LSD_{ij} = t_{\gamma/2,N-a}\sqrt{\text{MSE}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$	$HSD_{ij} = \frac{1}{\sqrt{2}}q_\gamma(a,N-a)\sqrt{MSE\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$
Rechazo de H_0	si $D_{ij} > LSD_{ij}$	si $D_{ij} > HSD_{ij}$

$$H_0 : \mu_1 = \mu_2 \Longleftrightarrow \mu_d = 0 \text{ vs } H_1 : \mu_1 \neq \mu_2 \Longleftrightarrow \mu_d \neq 0$$

Bajo los supuestos $d_j \overset{\text{iid}}{\sim} N\left(\mu_d, \sigma_d^2\right)$, con σ_d^2 desconocido, el estadístico de la prueba, su distribución bajo H_0 y el criterio de decisión, es:

$$T_0 = \frac{d}{S_d/\sqrt{n}} \sim t_{n-1}, \text{ con,}$$
$$\bar{d} = \frac{1}{n}\sum_{j=1}^n d_j, \quad s_d^2 = \frac{1}{n-1}\sum_{j=1}^n \left(d_j - d\right)^2$$

rechazar H_0 si $P\left(|t_{n-1}| > |T_0|\right)$ es pequeño. Ahora, en el experimento la muestra no es aleatoria y esto no asegura el principio de independencia entre si. Por tanto el estadístico de prueba t no tiene validez estadística. Se recomienda nuevamente aleatorizar el proceso como ya se mencionó en el inciso anterior.

2. Punto 5

Se realizó un experimento para determinar cuál tipo de batería no recargable para linternas era la más económica. En particular el experimentador estaba interesado en comparar los tiempos de vida por unidad de costo de la marca particular que él a menudo compraba con la marca del almacén donde usualmente compra. También deseaba conocer si valía la pena pagar dinero extra por baterías alcalinas respecto a las de larga duración. Dado que la vida de la bombilla de una linterna es considerablemente mayor a la de cualquier batería, se decide usar la misma bombilla para probar las baterías del experimento. Se mantuvieron aproximadamente constantes variables de ruido como la temperatura ambiental y las baterías fueron compradas con distintas fechas de fabricación y en distinto lugares con el fin de dar una representación amplia de las fechas de manufactura, así la variabilidad causada por tal factor sería medida como parte de la variabilidad natural. Las condiciones de corrida se mantuvieron constantes y cada batería fue usada en forma continua hasta agotarse, el tiempo fue registrado en minutos. El diseño se realizo en forma completamente aleatorizado y se tomaron como tratamientos los siguientes:

Nivel	**Tratamiento**
1	alcalina, marca A, costo=$0.985 dolares/unidad
2	alcalina, marca del almacen, costo=$0.935 dolares/unidad
3	de larga duracion, marca A, costo=$0.520 dolares/unidad
4	de larga duracion, marca del almacen, costo=$0.495 dolares/unidad

Los espacios de tiempo o turnos de prueba fueron asignados aleatoriamente a los tipos de bateria para asi determinar el orden en el cual las baterías iban a ser observadas, y se hicieron cuatro replicaciones de cada nivel. Se monto un circuito que conectaba una bateria a una bombilla de linterna y se coloco un pequeno reloj en el circuito el cual paraba cuando se agotaba la bateria. El tiempo transcurrido en el reloj fue tomado como la medicion de la vida de la bateria (minutos). Los datos segun el orden de observacion se muestran a continuacion:

Tipo	Vida
1	602.00
2	863.00
1	529.00
4	235.00
1	534.00
1	585.00
2	743.00
3	232.00
4	282.00
2	773.00
2	840.00
3	255.00
4	238.00
3	200.00
4	228.00
3	215.00

(a) Postule el modelo ANOVA, haga la tabla ANOVA y pruebe la significancia del modelo. Determine si el tamano de muestra usado era suficiente para detectar diferencias entre las medias de 250 minutos/$ con un nivel de significancia de 0.05.

(b) Valide los supuestos del modelo. Tenga en cuenta que como se conoce el orden de las corridas, debe probar primero incorrelacion, use para ello test Ljung-Box (consulte sobre esta prueba para que pueda formularla y aplicarla).

(c) Determine si ¿las baterías alcalinas son más económicas que las de larga duración? ¿Existe diferencias entre la vida media por unidad de costo de las baterías de la marca A vs.las de marca del almacén?En este último¿cuáles son más económicas?

(d) Verifique que para todas las comparaciones por pares, Tukey da I.C más cortos que el método de Scheffé (consulte sobre este metodo y vea manual de la librería agricolae) y haga el agrupamiento de medias por cada uno de estos dos metodos.

(e) Considere ahora como respuesta la vida de las baterías sin importar su costo. Suponiendo un modelo ANOVA de efectos fijos de un solo factor en un DCA, pruebe la normalidad y la igualdad de las varianzas. ¿Sera que la transformacion raiz cuadrada estabiliza varianza y nos lleva a las mismas conclusiones sobre cual es la mejor bateria cuando el analisis se hace en la escala original de los tiempos de vida?

2.1. Análisis descriptivo

La variable respuesta a estudiar en este casi resulta ser el tiempo de vida por unidad de costo. La unidad experimental son las baterías, y se tiene un diseño de experimento DCA con un solo factor. A continuación se puede observar el comportamiento de la vida de las baterías, de acuerdo al tipo y costo de las mismas, siendo las baterías tipo 2, alcalinas con marca del almacén, aquellas con una mayor duración vida, aunque del mismo modo parce ser aquella con la varianza más alta.

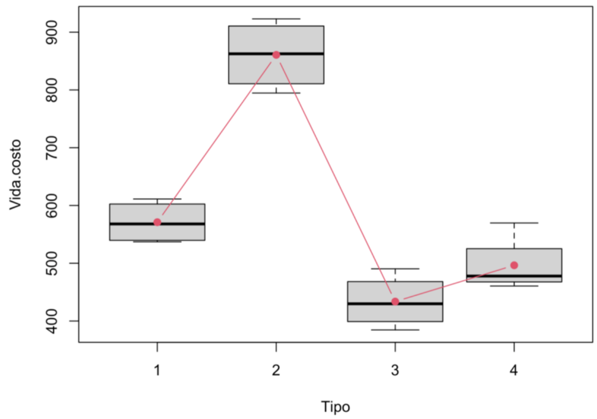


Figura 1: Diagrama de cajas Vida/costo
según el tipo de baterías

En el gráfico 1 se observa que el tipo de batería nivel 2 (alcalina, marca de almacén) tiene la relación vida-costo más alta de todas. Contrario al nivel 3 (larga duración, marca A) y al nivel 4 (larga duración, marca de almacén) tienen la relación vida-costo más baja.

2.2. (a) Modelo y test ANOVA

El modelo corresponde a un modelo ANOVA de un factor de efectos fijos en una estructura de diseño DCA:

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad \text{con} \quad \varepsilon_{ij} \sim \text{IID}\, N\left(0, \sigma^2\right), \quad \sum_{i=1}^{4} \alpha_i = 0 \tag{1}$$

Donde: Y_{ij} Es la respuesta vida/costo para los tipos de batería i en la batería j. μ corresponde a la media global de la relación vida/costo y ε_{ij} el error de la j-esima batería en el i-ésimo tipo de batería. Con fines prácticos se reorganiza la tabla de la siguiente manera:

Cuadro 1: Observaciones experimentales

NIVEL	Vida/Costo				$\bar{y}_i$	S_i^2
1	611.167	537.056	542.132	593.909	571.066	1374.58
2	922.995	794.652	826.738	898.396	860.69525	3605.80
3	446.154	490.385	384.615	413.462	433.654	2062.40
4	474.747	569.697	480.808	460.606	496.4645	2455.20
					$y_{..}$ = 590.4699375	S^2 = 873287.52

Se plantea la siguiente prueba de hipótesis:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_1 : \text{ para al menos un par, } \mu_i \neq \mu_j$$

Donde μ_i corresponde a la media de cada tipo de batería(niveles) con $i = \{1, 2, 3, 4\}$. Por otro lado, a continuación se presentan los resultados de la tabla Anova:

Cuadro 2: Tabla Anova

Fuente	gl.	SC	CM	F	P
Tipo	a-1 = 3	427305.972	142435.324	59.99	1.01e-10
Error	N-a = 12	28493.964	2374.497	-	-

Donde,

$$SSA = \sum_{i=1}^{a} n_i \bar{y}_{i}^2 - N\bar{y}_{..}^2 = 4 \times \left(\sum_{i=1}^{4} \bar{y}_i^2 \right) - 16 \times (590.4699)^2 = 427305.971$$

$$SSE = \sum_{i=1}^{a} (n_i - 1) S_i^2 = 3 \times \left(\sum_{i=1}^{4} S_i^2 \right) = 28493.964$$

Así entonces,

$$MSA = \frac{SSA}{a-1} = \frac{427305.971}{4-1} = 142435.324$$

$$MSE = S_p^2 = \frac{\sum_{i=1}^{a} (n_i - 1) S_i^2}{N-a} = 2374.497$$

Por lo cual, el estadístico de prueba resulta ser $F_0 = \frac{MSA}{MSE} = 59.985 \sim f_{a-1=3,n-a=14}$. Por lo tanto, $P(f_{3,12} > 59.985) = 1.703674e-07$, es decir que, se rechaza la hipótesis nula a favor de la alterna dado que el valor p es menor que el nivel de significancia $0.01 < \alpha < 0.1$. Esto implica que al menos dos tipos de pila poseen una respuesta media estadísticamente diferente, luego el tipo de pila tiene efectos significativos sobre la vida/costo de las baterías. En el software R se obtienen los siguientes resultados:

```
            Df Sum Sq Mean Sq F value   Pr(>F)    
Tipo         3 427306  142435   59.99 1.7e-07 ***
Residuals   12  28494    2374                    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

2.3. Análisis del tamaño de muestra

Para determinar el tamaño de muestra por tratamiento se sigue el siguiente procedimiento, considerando iguales tamaños de muestra por tratamientos $n_i = 4 \forall i$, con una amplitud del I.C L= 250 minutos/$

$$HSD_{ij} = \frac{1}{\sqrt{2}} q_{\gamma}(a, g.l) \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} = \frac{1}{\sqrt{2}} q_{0.05}(4, 4(n-1)) \sqrt{2374.497 \times \frac{2}{n}} \leq 125$$

Variando el valor de n se obtienen los siguientes resultados:

Cuadro 3: Tamaño de muestra variando n

n	$MSEq_{0.05}^2(4, 4(n-1)) - 125^2 n$	Acción
2	47449.67	Incrementar n
3	1826.21	Incrementar n
4	-20640.59	No incrementar n

De esta manera se concluye que el tamaño de muestra usado por el experimentador si era suficiente para detectar diferencias entre las medias de 250 minutos/$ con un nivel de significancia de 0.05.

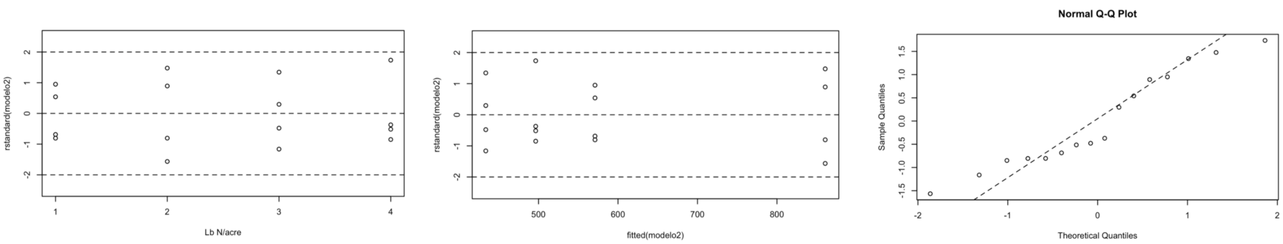
2.4. (b) Análisis de residuos y validación de supuestos

En la figura 2.a se percibe que no hay carencia de ajuste, debido a que se distribuyen de manera aleatoria al rededor del cero. En este mismo gráfico, no se observan outliers (valores atípicos por encima de $\pm 2\sigma$).

En la figura 2.b, se observa por cada uno de los niveles y de manera vertical en su respectivo nivel que para los niveles 1,2 y 4 en un principio, parecen tener una varianza similar, debido a que la dispersión de los datos es similar en el eje vertical. Por el contrario, para el nivel 3, los datos en apariencia se agrupan ligeramente más que en los niveles 1, 2 y 4. Esto podría afectar el supuesto de varianza constante y se requieren pruebas para verificarlo.

Finalmente, para el análisis de normalidad en los residuos (figura 2.c), se observan dos valores que podrían considerarse atípicos (en las colas del gráfico de normalidad). Además, los datos aparentan no ajustarse suficiente a la recta de normalidad. Por todo esto, se podría concluir que los supuestos, al menos en un primer acercamiento descriptivo, parecen no cumplirse; sin embargo, se debe tener en cuenta que todos y cada uno de los supuestos son sensibles a tamaños muestrales pequeños, como lo es este experimento.

Teniendo en cuenta lo anterior, puede que de manera descriptiva/gráfica los supuestos no se cumplan, pero se debe tener en cuenta que el tamaño de la muestra total es de 16 y de 4 para cada nivel, estos valores son considerados muy pequeños para obtener la confiabilidad de las pruebas.



(a) Gráfico de caja humedad **(b)** Gráfico de caja marca **(c)** Gráfico de caja humedad

Figura 2: Gráficos de residuales

Se validan a continuación los supuestos de normalidad, varianza constante e independencia respectivamente de manera analítica:

- **Independencia**

  Se plantea la siguiente prueba de hipótesis:

  $$H_0 : \operatorname{corr}(\varepsilon_{ij}, \varepsilon_{i'j'}) = 0 \forall (i,j) \neq (i',j')$$
  $$H_1 : \operatorname{corr}(\varepsilon_{ij}, \varepsilon_{i'j'}) \neq 0 \text{ para algún}$$
  $$(i,j) \neq (i',j')$$

  Haciendo uso del test de Ljung-Box Se obtienen los siguientes resultados:

  ```
  Box-Ljung test
  data:  Vida.costo
  X-squared = 0.83862, df = 1, p-value = 0.3598
  ```

  Dado que el valor p = 0.2127 por la prueba de Box-Ljung se concluye que los datos se distribuyen de forma independiente, es decir que los residuales estudentizados son incorrelacionados, con un nivel de significancia $\alpha < 0.05$. Se prosigue entonces, con los análisis de los demás supuestos.

- **Normalidad**

  Se plantea la siguiente prueba de hipótesis:

  $$H_0 : \varepsilon_{ij} \sim \text{ Normal vs..}$$
  $$H_1 : \varepsilon_{ij} \text{no son Normales}$$

  Haciendo uso del test de Shapiro Wilk se obtiene:

  ```
  Shapiro-Wilk normality test
  data:  rstandard(modelo2)
  W = 0.92628, p-value = 0.2127
  ```

  En un principio de manera descriptiva a través del gráfico de normalidad de los residuos se observaba un desajuste respecto a la recta de normalidad. Esto en primera instancia se perciben algunos *outliers* alejados de los datos sumado a que, los datos çentrales”del gráfico no se ajustan lo suficiente a la recta como para asegurar normalidad. Al corroborar con el test de *Shapiro Wilk*, se tiene una conclusión contraria respecto al primer análisis descriptivo, es decir que NO hay evidencia suficiente para rechazar normalidad con un nivel de significancia $0.01 < \alpha < 0.1$, esto puede ser causado a que las pruebas analíticas planteadas de aquí en adelante, son sensibles a tamaños muéstrales pequeños.

- **Varianza Constante**

  Se plantea la siguiente prueba de hipótesis:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2$$

$$H_1 : \text{ algún par} \sigma_i^2 \neq \sigma_j^2$$

  Siendo $\sigma_i^2 = var(\epsilon_{ij})$ la varianza de los errores en el i-ésimo Tipo de bateria, con $\epsilon_{ij} \sim iidN(0, \sigma_i^2)$. Haciendo uso el test de Bartlett obtenemos lo siguiente:

$$\chi_0^2 = \frac{M}{C} \sim \chi_{a-1}^2 = \chi_4^2$$

  Donde:

$$M = (N-a)\log\left(S_p^2\right) - \sum_{i=1}^{a}(n_i-1)\log\left(S_i^2\right) M = 12 \times \log(MSE) - 4 \times \sum_{i=1}^{4}\log\left(S_i^2\right) \approx -13.0916$$

$$C = 1 + \frac{1}{3(a-1)}\left[\left(\sum_{i=1}^{a}\frac{1}{n_i-1}\right) - \frac{1}{N-a}\right] = 1 + \frac{1}{3\times 3}\left[4\times\frac{1}{3} - \frac{1}{12}\right] \approx 1.1388$$

En el software R se obtienen lo siguientes resultados:

```
Bartlett test of homogeneity of variances
data:  Vida.costo by Tipo
Bartlett's K-squared = 0.62261, df = 3, p-value = 0.8912
```

Acorde al valor p del estadístico de Barlett, teniendo en cuenta que la hipótesis nula iba a favor de la igualdad de las varianzas, NO se rechaza la hipótesis nula y se cumple (de manera analítica) el supuesto de varianza constante. Sin embargo se ha de tener en cuenta que, esta prueba no es consistente con el análisis descriptivo anterior. Nuevamente, se atribuye esto a la sensibilidad de la prueba a tamaños muéstrales pequeños, aun así, es favorable que los supuestos del modelo se cumplan.

En resumen, no se rechaza la hipótesis nula y se cumple el supuesto de varianza constante.

2.5. (c)¿Qué baterías son más económicas?

A continuación se puede observar la tabla de medias μ_i para cada tipo de batería $i = 1, 2, 3, 4$ con su respectivo error cuadrático e intervalos de confianza:

Cuadro 4: Medias Vida/Costo por tipo de batería

Tipo	Media	SE	df	Inferior IC	Superior IC
1	571.0660	24.3643	12	517.9807	624.1513
2	860.6952	24.3643	12	807.6099	913.7805
3	433.6538	24.3643	12	380.5685	486.7392
4	496.4646	24.3643	12	443.3793	549.5500

Primero se procede a analizar si si las baterías alcalinas son más económicas que las de larga duración. Dado:

$$W = \frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{2}(\mu_3 + \mu_4)$$

Se plantea la siguiente la siguiente prueba de hipótesis:

$$H_0 : W = 0 \quad \text{vs.} \quad H_1 : W > 0$$

Así entonces $\widehat{W} = 250.8214$. Observe que $c_i = \frac{1}{2}$ para i = 1,2 y $c_j = \frac{1}{2}$ para i = 3,4. Bajo los supuestos del DCA, el estadístico de prueba para los test relativos a contrastes de medias de tratamientos, es:

$$T_0 = \frac{\widehat{W} - W_0}{\sqrt{MSE \sum_{i=1}^{a} \frac{c_i^2}{n_i}}} \sim t_{N-a}$$

$$T_0 = \frac{250.8214}{\sqrt{\frac{2374.497}{4}}} = 10.29458 \sim t_{\gamma=0.05,16-4=12}$$

Dado que $T_0 = 0.29458$ No es mayor que $t_{0.05,12} = 1.782288$ Se concluye que No hay evidencia suficiente para rechazar la hipótesis nula. Es decir que, las baterías alcalinas No resultan ser más económicas que las de larga duración, con un nivel de confianza de $\alpha = 0.05$.

¿Existe diferencias entre la vida media por unidad de costo de las baterías de la marca A vs.las de marca del almacén?

Dado:

$$W = \frac{1}{2}(\mu_1 + \mu_3) - \frac{1}{2}(\mu_2 + \mu_4)$$

Se plantea la siguiente la siguiente prueba de hipótesis:

$$H_0 : W = 0 \quad \text{vs.} \quad H_1 : W \neq 0$$

Así entonces $\widehat{W} = -176.22$. Observe a continuación que $c_i = \frac{1}{2}$ para i = 1,3 y $c_j = \frac{1}{2}$ para i = 2,4. Bajo los supuestos del DCA, el estadístico de prueba para los test relativos a contrastes de medias de tratamientos, es:

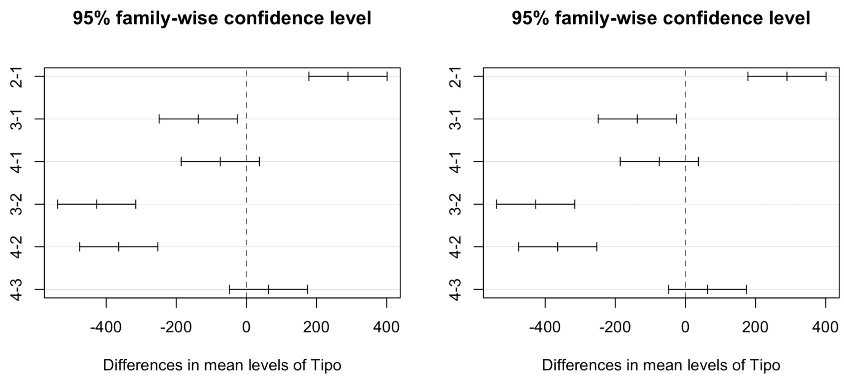
$$T_0 = \frac{\widehat{W} - W_0}{\sqrt{MSE \sum_{i=1}^{a} \frac{c_i^2}{n_i}}} \sim t_{N-a}$$

$$T_0 = \frac{-176.22}{\sqrt{\frac{2374.497}{4}}} = -7.232682 \sim t_{\gamma=0.05,16-4=12}$$

Ya que $|T_0| = 7.232682 > t_{0.025,12} = 2.178813$ se concluye que se Rechaza la hipótesis Nula. Es decir que efectivamente existen diferencias entre la vida media por unidad de costo de las baterías de la marca A vs.las de marca del almacén, con un nivel de confianza de $\alpha = 0.05$. **¿cuáles son más económicas?** Observe que $\frac{1}{2}(\mu_1 + \mu_3) \leq \frac{1}{2}(\mu_2 + \mu_4)$ Luego las baterías de marca de almacén son más económicas y tienen una mayor duración de Vida que las de marca A.

2.6. (d) Comparaciones por pares

En el Cuadro 4 se puede verificar que para todas las comparaciones por pares, Tukey da I.C más cortos con una longitud L = 204.6, en comparación con el método de Scheffé con una longitud $L^{'}$ = 223. Donde teniendo un factor con 4 niveles, el número de comparaciones distintas son: $\binom{4}{2} = \frac{4 \bullet 3}{2} = 6$.



(a) Tukey **(b)** Scheffé

Figura 3: Representación de los I.C de las diferencias de medias del 95 %, tipo Tukey y Scheffé

Gráficamente, Figura 3, no se evidencia una diferencia muy notoria entre los intervalos de confianza aplicando los 2 métodos, a pesar de tener una longitud menor con el método de Tukey. Se procede entonces a hacer el agrupamiento de medias por cada uno de estos dos métodos.

- **Agrupamiento de medias con Tukey**

  Se plantea la siguiente prueba de hipótesis:

  $$H_0 : \mu_i = \mu_j \text{ vs. } H_1 : \mu_i \neq \mu_j$$

  Se procede a calcular la Diferencia "Honestamente"significativa de la siguiente manera:

  $$HSD_{ij} = \frac{1}{\sqrt{2}} q_\gamma(a, N-a) \sqrt{\text{MSE}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = \frac{1}{\sqrt{2}} q_\gamma(4, 12) \sqrt{\text{MSE}\left(\frac{1}{2}\right)} = 102.2979$$

  Donde se rechaza H_0 si $|D_{ij}| > HSD_{ij}$, con $D_{ij} = \bar{Y}_{i\bullet} - \bar{Y}_{j\bullet}$. Cuyos resultados se pueden observar en el Cuadro 4. Concluyendo de esta manera que los grupos según Tukey son:

  - Grupo 1 : $\mu_1 = \mu_4$
  - Grupo 2 : $\mu_3 = \mu_4$

  No se asignan en un mismo grupo ya que $\mu_1 - \mu_3$ son estadísticamente diferentes.

- **Agrupamiento de medias con Scheffé**

  Gracias a los intervalos de confianza de Scheffé se puede concluir que se obtienen los mismos resultados que con el método de Tukey, puesto que el IC $\mu_1 - \mu_4$ y $\mu_3 - \mu_4$ incluyen al cero, es decir que no son estadísticamente diferentes. Dichos resultados se pueden observar el el Cuadro 4.

Cuadro 5: Comparaciones por pares e IC simultáneos 1(1-α)100 %

		Tukey				Scheffé			
Dif. Poblacional	$D_{ij} = Y_{i\bullet} - Y_{j\bullet}$	Inferior	Superior	Desición	Sup-Inf	Inferior	Superior	Sup-Inf	Desición
$\mu_2 - \mu_1$	289.63	187.33	391.93	Diferentes	204.6	178.13	401.13	223	Diferentes
$\mu_3 - \mu_1$	-137.41	-239.71	-35.11	Diferentes	204.6	-248.91	-25.92	223	iguales
$\mu_4 - \mu_1$	-74.6	-176.9	27.7	Iguales	204.6	-186.1	36.9	223	iguales
$\mu_3 - \mu_2$	-427.04	-529.34	-324.74	Diferentes	204.6	-538.54	-315.54	223	Diferentes
$\mu_4 - \mu_2$	-364.23	-466.53	-261.93	Diferentes	204.6	-475.73	-252.73	223	Diferentes
$\mu_4 - \mu_3$	62.81	-39.49	165.11	Iguales	204.6	-48.69	174.31	223	iguales

2.7. (e)Respuesta la vida de las baterías sin importar su costo

El modelo corresponde a un modelo ANOVA de un factor de efectos fijos en una estructura de diseño DCA:

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad \text{con} \quad \varepsilon_{ij} \sim \text{IID } N\left(0, \sigma^2\right), \quad \sum_{i=1}^{4} \alpha_i = 0 \tag{2}$$

Donde: Y_{ij} Es la respuesta vida de las baterías sin importar su costo. μ corresponde a la media global de la relación vida/costo y ε_{ij} el error de la j-esima batería en el i-ésimo tipo de batería. Con fines prácticos se reorganiza la tabla de la siguiente manera:

Cuadro 6: Observaciones experimentales

NIVEL	VIDA				$\bar{y}_i$	S^2
1	602	529	534	585	562.5	1333.67
2	863	743	773	840	804.75	3152.25
3	232	255	200	215	225.5	557.67
4	235	282	238	228	245.75	601.58
					459.625	1473754.93

Se plantea la siguiente prueba de hipótesis:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_1 : \text{ para al menos un par, } \mu_i \neq \mu_j$$

Donde μ_i corresponde a la media de cada tipo de batería(niveles) con $i = \{1, 2, 3, 4\}$. Por otro lado, a continuación se presentan los resultados de la tabla Anova:

Cuadro 7: Tabla Anova

Fuente	gl.	SC	CM	F	P
Tipo	a-1 = 3	921006	307002	217.5	1.01e-10
Error	N-a = 12	16935	1411	-	-

Donde,

$$SSA = \sum_{i=1}^{a} n_i \bar{y}_i^2 - N\bar{y}_{..}^2 = 4 \times \left(\sum_{i=1}^{4} \bar{y}_i^2\right) - 16 \times (459.625)^2 = 921006.25$$

$$SSE = \sum_{i=1}^{a} (n_i - 1) S_i^2 = 3 \times \left(\sum_{i=1}^{4} S_i^2\right) = 16935.5$$

Así entonces,

$$MSA = \frac{SSA}{a-1} = \frac{921006.25}{4-1} = 307002.08$$

$$MSE = S_p^2 = \frac{\sum_{i=1}^{a} (n_i - 1) S_i^2}{N-a} = \frac{16935.5}{16-4} = 1411.29$$

Por lo cual, el estadístico de prueba resulta ser $F_0 = \frac{MSA}{MSE} = 217.53 \sim f_{a-1=3, n-a=12}$. Por lo tanto, $P\left(f_{3,12} > 217.53\right) = 1.01e-10$, es decir que, se rechaza la hipótesis nula a favor de la alterna dado que el valor p es menor que el nivel de significancia $0.01 < \alpha < 0.1$. Esto implica que al menos dos tipos de pila poseen una respuesta media estadísticamente diferente sin tener en cuenta el costo de las baterías.

En el software R se obtiene los siguientes resultados:

```
            Df Sum Sq Mean Sq F value   Pr(>F)    
Tipo         3 921006  307002   217.5 1.01e-10 ***
Residuals   12  16935    1411                     
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

2.8. Análisis de residuos y validación de supuestos

- **Normalidad**

  Realizando un análisis gráfico no se observa un comportamiento estrictamente normal, ya que los puntos tienden a alejarse de la recta. Es necesario realizar un análisis con un test formal para chequear la normalidad de los datos.

  Se plantea la siguiente prueba de hipótesis:

$$H_0 : \varepsilon_{ij} \sim \text{ Normal vs..}$$

$$H_1 : \varepsilon_{ij} \text{no son Normales}$$

  Usando el test de Shapiro Wilk y haciendo uso del software R se obtiene lo siguiente:

```
Shapiro-Wilk normality test
data:  rstandard(modeloanova)
W = 0.95652, p-value = 0.5993
```

  Por lo tanto, contrariamente a lo observado gráficamente, se concluye que no hay evidencia suficiente para rechazar H_0 con un valor-p de 0.5993, por lo cual no hay evidencia suficiente para rechazar normalidad con un nivel de confianza $0.01 < \alpha < 0.1$.

- **Igualdad de varianzas**

  Se plantea la siguiente prueba de hipótesis:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2$$

$$H_1 : \text{ algún par} \sigma_i^2 \neq \sigma_j^2$$

  Haciendo uso del test de Barlette y del software R se obtiene:

```
Bartlett test of homogeneity of variances
data:  Vida by Tipo
Bartlett’s K-squared = 2.7241, df = 3, p-value = 0.4361
```

  Acorde al valor p = 0.4361 del estadístico de Barlett, teniendo en cuenta que la hipótesis nula iba a favor de la igualdad de las varianzas, NO se rechaza la hipótesis nula y se cumple (de manera analítica) el supuesto de varianza constante.

Nota
En conclusión, después de analizar la normalidad y la igualdad de las varianzas sin tener en cuenta el costo de la baterías, se llegan a las mismas conclusiones sobre cuál es la mejor batería cuando el análisis se hace en la escala original de los tiempos de vida.

2.9. Respuesta la vida de las baterías raíz cuadrada

- **Normalidad**

  Se plantea la siguiente prueba de hipótesis:

$$H_0 : \varepsilon_{ij} \sim \text{ Normal vs..}$$

$$H_1 : \varepsilon_{ij} \text{no son Normales}$$

  Haciendo uso del test de Shapiro Wilk se obtiene:

```
Shapiro-Wilk normality test
data:  rstandard(modelo_sqrt)
W = 0.91237, p-value = 0.1271
```

Se concluye que no hay evidencia suficiente para rechazar H_0, por lo cual no se rechaza normalidad con un nivel de confianza 0.01.

- **Igualdad de varianzas raíz vida**
  Se plantea la siguiente prueba de hipótesis:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2$$

$$H_1 : \text{ algún par} \sigma_i^2 \neq \sigma_j^2$$

  Haciendo uso del test de Barlette y del software R se obtiene:

```
Bartlett test of homogeneity of variances
data:  raizVida by Tipo
Bartlett's K-squared = 0.26483, df = 3, p-value = 0.9665
```

  Según el valor p, NO se rechaza la hipótesis nula y por lo tanto se evidencia igualdad de varianza.

Nota
En conclusión, después de analizar la normalidad y la igualdad de las varianzas bajo la transformacion raiz cuadrada se llega a las mismas conclusiones sobre cual es la mejor batería cuando el análisis se hace en la escala original de los tiempos de vida.

3. Punto 13

13. Un fabricante sospecha que los lotes de materia prima despachados por su proveedor difieren significativamente en el contenido de calcio. Actualmente en la bodega hay en inventario un gran numero de lotes almacenados de este material. El fabricante decide escoger al azar cinco de estos lotes para un estudio. Un quimico realiza cinco determinaciones del contenido de calcio sobre cada uno de los lotes y obtiene los siguientes datos:

Lote	ContCalcio
1	23.46
2	23.59
3	23.51
4	23.28
5	23.29
1	23.48
2	23.46
3	23.64
4	23.40
5	23.46
1	23.56
2	23.42
3	23.46
4	23.37
5	23.37
1	23.39
2	23.49
3	23.52
4	23.46
5	23.32
1	23.40
2	23.50
3	23.49
4	23.39
5	23.38

(a) Proponga el plan experimental para este problema

(b) Escriba el modelo de analisis de varianza apropiado para el experimento, identifique sus terminos y los supuestos necesarios.

(c) Con base en el grafico de cajas haga un analisis descriptivo y formule conclusiones preliminares sobre la posible significancia de la variacion en el contenido de calcio, de lote a lote.

(d) Realice el test anova pertinente a este experimento y concluya sobre significancia de la variacion en el contenido de calcio, de lote a lote

(e) Halle un intervalo de confianza del 95 % para proporcion de varianza debida a los lotes e interprete los resultados Valide supuestos

3.1. (a) Plan Experimental

- Estructura de tratamiento: Un solo factor de efecto aleatorio
- Estructura de diseño: Completamente aleatorizado (DCA)
- Variable respuesta: Contenido de calcio en cada lote
- Unidad experimental: Lote
- Tipo de modelo: ANOVA de un factor de efectos aleatorios

3.2. (b) Modelo y test ANOVA

El modelo corresponde a un modelo ANOVA de un factor de efectos aleatorios en una estructura de diseño DCA:

$$Y_{ij} = \mu + A_i + E_{ij}, \text{con } E_{ij} \sim IIDN\left(0, \sigma^2\right), A_i \sim IIDN\left(0, \sigma_\alpha^2\right) \tag{3}$$

Con los A_i y los E_{ij} mutuamente independientes $\forall i, j, i = 1, 2, \ldots, a = 5$

Donde: Y_{ij} Es el j-ésimo contenido de calcio en el i-ésimo lote. A_i es el efecto aleatorio del lote i sobre la media de la variable respuesta(cantidad de calcio). μ corresponde a la media global de la cantidad de calcio y ε_{ij} es el error de la j-esima repetición i-ésimo tratamiento. Con fines prácticos se reorganiza la tabla de la siguiente manera:

Cuadro 8: Observaciones experimentales

NIVEL	ContCalcio					$\bar{Y}_{i\bullet}^2$	$\bar{S}_i^2$
1	23.46	23.48	23.56	23.39	23.4	23.458	0.00472
2	23.59	23.46	23.42	23.49	23.5	23.492	0.00397
3	23.51	23.64	23.46	23.52	23.49	23.524	0.00473
4	23.28	23.4	23.37	23.46	23.39	23.38	0.00425
5	23.29	23.46	23.37	23.32	23.38	23.364	0.00423
						$\bar{Y}_{\bullet\bullet}$=23.4436	$\bar{S}^2$=0.0000001114

3.3. (c) Análisis descriptivo

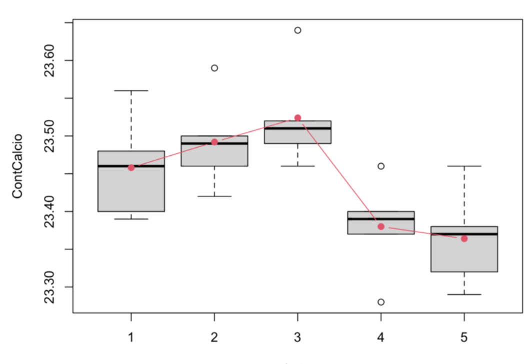


Figura 4: Diagrama de cajas Lote por Contenido de Calcio

En la figura 4 se observa que los lotes 1, 2 y 3 tienen una cantidad de calcio similar. También se evidencia que en el lote 2 y 3 existe un dato atípico en cada uno. Por su parte, el lote 4 y 5 tienen una cantidad de calcio menor a diferencia de los tres primeros. De la misma forma, se encuentra un dato atípico en el lote 4 tanto inferior como superior.

3.4. (d) Estimaciones e inferencias sobre parámetros del modelo ANOVA

$$H_0 : \sigma_\alpha^2 = 0 \text{ vs. } H_1 : \sigma_\alpha^2 > 0$$

De esta manera entonces, se realiza un anális ANOVA:

Cuadro 9: Tabla ANOVA

Fuente	gl.	SC	CM	F	P
Tipo	a-1 = 4	0.096976	0.024244	5.5352	0.003626
Error	N-a = 20	0.087600	0.004380	-	-

Donde,

$$SSA = \sum_i n_i \bar{Y}_{i\bullet}^2 - N\bar{Y}_{\bullet\bullet}^2 = 5 \times \left(\sum_{i=1}^{5} \bar{Y}_{i\bullet}^2 \right) - 25 \times (23.4436)^2 = 0.096976$$

$$SST = \sum_i \sum_j Y_{ij}^2 - N\bar{Y}_{\bullet\bullet}^2 = \sum_i \sum_j Y_{ij}^2 - 25 \times (23.4436)^2 = 0.184576$$

$$SSE = SST - SSA = 0.0876$$

Por lo cual,

$$MSA = \frac{SSA}{a-1} = \frac{0.096976}{5-1} = 0.024244$$

$$MSE = \frac{SSE}{N-a} = \frac{0.0876}{25-5} = 0.004380$$

Así entonces, el estadístico de prueba resulta ser $F_0 = \frac{MSA}{MSE} = 5.535160 \sim f_{a-1=4, n-a=20}$. Por lo tanto, $P\left(f_{4,20} > 5.535160\right) = 0.003626022$, es decir que, se rechaza la hipótesis nula a favor de la alterna dado que el valor p es menor que el nivel de significancia $0.01 < \alpha < 0.1$. Esto implica que la cantidad de calcio si contribuye significativamente a la varianza total de la respuesta, es decir, si hay componente de varianza debida la cantidad de calcio:

. En el software R se obtienen los siguientes resultados:

```
Analysis of Variance Table

Response: ContCalcio
          Df   Sum Sq  Mean Sq F value   Pr(>F)
Lote       4 0.096976 0.024244  5.5352 0.003626 **
Residuals 20 0.087600 0.004380
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.5. (e) Intervalos de confianza

El intervalo de confianza de $1 - \gamma 100\,\%$para σ^2 viene dada por:

$$\frac{SSE}{\chi^2_{\gamma/2, N-a}} < \sigma^2 < \frac{SSE}{\chi^2_{1-\gamma/2, N-a}}$$

$$\frac{0.0876}{\chi^2_{0.025, 25-5}} < \sigma^2 < \frac{0.0876}{\chi^2_{0.95/2, 25-5}}$$

$$\frac{0.0876}{34.16961} < \sigma^2 < \frac{0.0876}{10.85081}$$

$$0.0026 < \sigma^2 < 0.00807$$

4. Punto 22

Una empresa dedicada a la prueba de productos de consumo, realizó un experimento para comparar el consumo anual de energía de cinco marcas de deshumidificadores. Como el consumo de energía depende del nivel de humedad presente, se decidió monitorear cada marca en cuatro niveles de humedad (desde humedad moderada hasta intensa, codificadas por 1, 2, 3, y 4). Dentro de cada nivel de humedad se asignaron al azar las marcas entre cinco lugares seleccionados. Las observaciones resultantes (KWh anuales) fueron los siguientes:

Marcas	Nivelhumedad	KWh
1	1	685.00
1	2	792.00
1	3	838.00
1	4	875.00
2	1	722.00
2	2	806.00
2	3	893.00
2	4	953.00
3	1	733.00
3	2	802.00
3	3	880.00
3	4	941.00
4	1	811.00
4	2	888.00
4	3	952.00
4	4	1005.00
5	1	828.00
5	2	920.00
5	3	978.00
5	4	1023.00

(a) Describa claramente la estructura de tratamiento, la estructura de diseño aplicado, las unidades experimentales y la variable respuesta medida. ¿Se debe considerar o no interacción entre las marcas de los deshumidificadores y el nivel de humedad?

(b) Haga un análisis descriptivo de los datos (según marcas y segun nivel de humedad).

(c) Escriba un modelo estadístico para este experimento, explique los terminos. ¿Cuáles son las suposiciones necesarias para un análisis de varianza válido?

(d) Calcule el análisis de varianza y pruebe con una significancia de 0.05 la hipotesis de interes en este estudio. Escriba claramente las hipotesis, el estadistico de prueba y su valor, y el valor P.

(e) Valide los supuestos de varianza constante con graficos de residuales estandarizados (vs. valores ajustados, marcas y nivel de humedad) y normalidad con test de Shapiro Wilk y grafico de probabilidad con residuales estandarizados. Escriba las hipotesis y pruebas correspondientes. Concluya.

(f) Defina los efectos de tratamientos como un contraste de medias de tratamientos y estimelos junto con un I.C.

(g) Asumiendo validos supuestos sobre el error, agrupe las medias del factor de interes mediante el metodo de Tukey. Interprete en terminos del problema. ¿cual marca recomendaria?

(h) Evalue la eficiencia del diseno empleado mediante la seudo prueba F sobre el factor de bloqueo al 5 % de significancia y mediante la formula de la EER. ¿Valio la pena bloquear? Explique que se gano o que se perdio, segun la eficiencia del diseno. Formule un contraste apropiado y desarrolle test para probar que los deshumidificadores de marcas 4 y 5 exceden el consumo de los deshumidificadores de las marcas 2, y 3 en mas de 60KWh anuales.

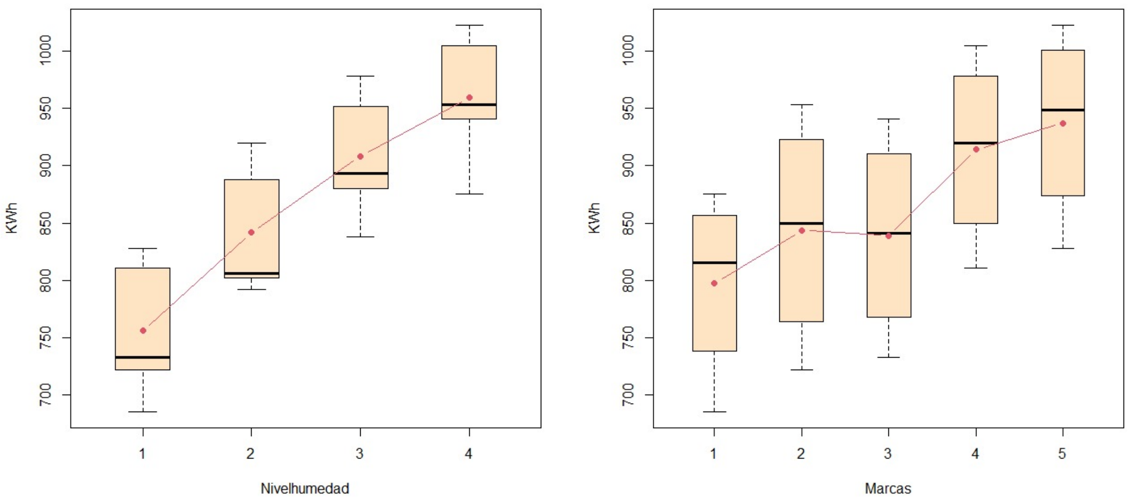
4.1. (a) Plan Experimental

- Estructura de tratamiento: Un solo factor de efectos fijos
- Estructura de diseño: Un Factor de Bloques Completos Aleatorizados (DBCA)

- Bloqueo:Nivel de Humedad
- Variable respuesta: Consumo anual de energía en el nivel de humedad j de la marca i
- Tipo de modelo: Modelo de diseño de bloques completos aleatorizados DBCA
- Unidad Experimental: Cada uno de los deshumificadores a los cuales se le realiza la toma de medida (corrida de experimento)

Tenga presente que la interacción entre las marcas de deshumificadores y el nivel de humedad hace referencia a si los segundos mencionados representan elementos a bloquear.
Claramente, el factor humedad no es la variable de interés pero si puede afectar de manera significativa los resultados de las corridas experimentales y finalmente, la variable respuesta. Por tanto, es recomendable bloquear el factor humedad con el fin de que, la interacción con los deshumidificadores se vea reducida.

4.2. (b) Análisis descriptivo



(a) Gráfico de caja humedad

(b) Gráfico de caja marca

Figura 5: Gráficos de caja de los factores

En un principio, el supuesto de varianza constante dado el análisis de boxplot para el consumo de energía anual por marca, parece no incumplirse. La longitud de los bigotes de los boxplot en apariencia, parecen tener la misma magnitud. Esto, al menos en un primer análisis descriptivo, apunta a favor del supuesto de varianza constante.

Acerca de las medias por marcas: Las medias son diferentes para cada marca indicando una tendencia a influir en los resultados. Entonces, se puede concluir que el consumo anual de energía varia en función de la marca usada. Se concluye que hay efectos significativos según los tratamientos. Teniendo esto presente, parece que hay similitud entre las medias de las marcas 2 y 3. Y una similitud mas ligera entre las medias de las marcas 4 y 5. Además, estas dos superan significativamente en el consumo de energía a las otras 3 marcas.

La media de la marca 1 parece ser diferente respecto a todas las demás (diferente en un sentido estadístico); puesto que tiene el valor mas bajo de todas las medias. Vale la pena verificar la igualdad estadística de las medias entre si, dado el anterior análisis. Se esperarían 3 grupos; dados por las marcas 2 y 3 (grupo 1), las marcas 4 y 5 (grupo 2) y la marca 1 (grupo 3). Esto será verificado de manera analítica. La variabilidad dentro de los tratamientos es similar, siendo la marca 2 el tratamiento con mayor dispersión.

Sobre el boxplot de bloqueo Conforme el nivel de humedad se intensifica, también se intensifica el consumo de energía anual por lo tanto, el análisis de bloqueo resulta favorable para el control de la varianza. Esto, dado los resultados de las medias en cada bloque.

Paralelo a esto, se tiene que al igual que en el inciso anterior, la varianza en cada bloque aparenta ser similar, por tanto, el supuesto de varianza constante dentro de los bloques se espera, sea cumplido.

4.3. (c,d) Modelo y test ANOVA

Modelo
En este caso se implementa un modelo de diseño de bloques completos aleatorizados DBCA:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \quad \varepsilon_{ij} \overset{iid}{\sim} N\left(0, \sigma^2\right) \tag{4}$$

Donde μ es la media global de consumo de energía, $Y_{i,j}$ es el consumo anual de energía en el nivel de humedad j (bloque) de la marca i, con i = 1,2,3,4,5 y j = 1,2,3,4. α_i representa el efecto sobre el consumo anual de energía debido a la i-ésima marca. β_j es el efecto del j-ésimo nivel de humedad sobre el consumo anual de energía y $\epsilon_i j$ es el error aleatorio en la observación del tratamiento i en el bloque j. Restricciones: $\sum_{i=1}^{5} \alpha_i = \sum_{j=1}^{4} \beta_j = 0$

test ANOVA Se plantea la siguiente prueba de hipótesis:

$$H_0 : \mu_{1,\bullet} = \mu_{2,\bullet} = \mu_{3,\bullet} = \mu_{4,\bullet} = \mu_{5,\bullet} \quad \text{vs} \quad H_1 : \text{algún par} \quad \mu i_{\bullet} \neq \mu_{k,\bullet}$$

Obteniendo:

Cuadro 10: ANOVA

Df	Sum Sq	Mean Sq	F value	Pr(>F)
4	53231.00	13307.75	95.57	0.0000
3	116217.75	38739.25	278.20	0.0000
12	1671.00	139.25		

Donde:

$$F = \frac{MSA}{MSE} \sim f_{a-1,(a-1)*(b-1)}$$

$$F = \frac{13308}{38739} \simeq 95.6$$

Se evalúa la eficiencia con los bloques

$$ER \Leftrightarrow H = \frac{MSB}{MSE}$$

$$H : 278.199 > 1 \quad \therefore$$

Al parecer el DCBA usando bloqueo según el nivel de humedad fue eficiente. En R se obtienen los siguientes resultados:

```
Analysis of Variance Table

Response: KWh
             Df Sum Sq Mean Sq F value    Pr(>F)
Marcas        4  53231   13308  95.567 5.419e-09 ***
Nivelhumedad  3 116218   38739 278.199 2.364e-11 ***
Residuals    12   1671     139
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4.4. (e) Análisis de residuos y validación de supuestos

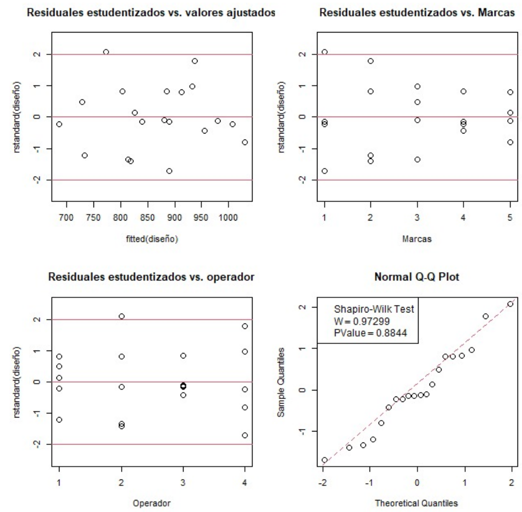


Figura 6: Análisis de residuos

De los residuales estandarizados vs los valores ajustados se observa que no hay evidencia suficiente en contra de la varianza constante ni contra la media igual cero. Tampoco hay motivos para sospechar carencia de ajuste pues no se observa ningun patron en la distribucion de los datos en este grafico.

De los residuales estudentizados vs el factor de tratamiento (marca) se observa a datos en forma similar a una campana donde a la izquierda del gráfico se ven datos bastante dispersos y a medida que se realiza un cambio de marca (valores se 1 al 4) la dispersión se va disminuyendo paulatinamente, por lo tanto solo se cumple que la media es igual a cero.

De los residuales estandarizados vs el factor de bloqueo es notable que el supuesto de varianza, almenos en un primer analisis descriptivo; no se cumple. Ya que los datos asociados al nivel de humedad 3 tiene la menor dispersión seguido por el nivel de humedad 1, mientras que entre los datos asociados al nivel de humedad 2 y 4 se observa una varianza mayor.

Finalmente, respecto a la normalidad de los errores, se nota que el ajuste de normalidad en los errores puede no cumplirse. En un principio, los datos deberían ajustarse a la recta de normalidad para asegurar esto. Sin embargo, si bien los datos no se ajustan perfectamente, tampoco se ve un patrón exagerado (una parábola o curva).
Al verificar el valor p de la prueba de Shaphiro-Wilk se constata el hecho de que los errores no distribuyen normal y dicho supuesto se ve refutado. Ahora, tenga presente que si bien, tanto gráfica como analíticamente la normalidad se ha visto afectada, no ha sido una grave afectación puesto que no se observan patrones o curvas muy marcadas. Tenga presente que los datos son relativamente pocos para tomar conclusiones con mayor confianza. Se recomienda aumentar el tamaño de la muestra y correr nuevamente las pruebas de supuestos.

4.5. (f) Contraste de medias e intervalos

Contrastes de medias Sea $W = \sum_{i=1}^{5} c_i \mu_i$, tal que: $\sum_{i=1}^{5} c_i = 0.n_i = 4, \quad \forall_i$

$$\hat{W} = \sum_{i=1}^{5} c_i \bar{Y} i \bullet \sim N\left(\sum_{i=1}^{5} c_i \mu_{i\bullet}, \sum_{i=1}^{5} c_i^2 \frac{\sigma^2}{4}\right)$$

Tenemos que:

$$W_1 = 797.5, W_2 = 843.5, W_3 = 839, W_4 = 914, W_5 = 937.25$$

Primer I.C de $(1-0.05)100\,\%$ $W_1 \pm t_{0,025,15}\sqrt{\text{MSE } \sum_{i=1}^{2}\frac{ci^2}{4}}$

$$797.5 \pm 2.13\sqrt{139 \cdot \frac{1}{4}} \Rightarrow (784.935, 810.065)$$

Segundo I.C de $(1-0.05)100\,\%$

$$843.5 \pm 2.13\sqrt{139 \cdot \frac{1}{4}} \Rightarrow (830.935, 856.065)$$

Tercer I.C de $(1-0.05)100\,\%$

$$8.39 \pm 2.13\sqrt{139 \cdot \frac{1}{4}} \Rightarrow (826.435, 851.565)$$

Cuarto I.C de $(1-0.05)100\,\%$

$$914 \pm 2.13\sqrt{139 \cdot \frac{1}{4}} \Rightarrow (901.435, 926.565)$$

Quinto I.C de $(1-0.05)100\,\%$

$$937 \pm 2.13\sqrt{139 \cdot \frac{1}{4}} \Rightarrow (924.685, 949.815)$$

4.6. Agrupación de medias

Para la agrupación estadística de medias, se procede a realizar la prueba Tukey, que permite verificar cuales de todas las medias que se tienen, son estadísticamente iguales. Para cada

$$\mu_{i\bullet} - \mu_{k\bullet}.$$

se construye un intervalo:

$$\bar{y}_{i,\bullet} - \bar{Y}_{k,\bullet} \pm HSD_{ik}$$

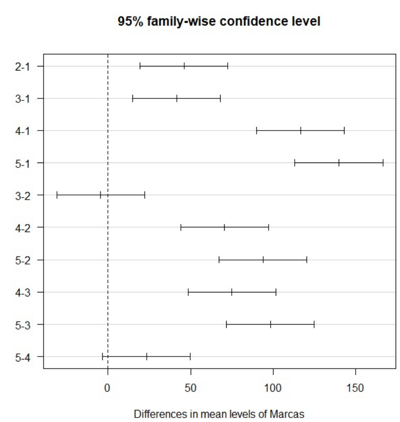


Figura 7: Gráficos de caja de los factores

Cuadro 11: Comparaciones por pares e IC simultáneos $1(1-\gamma)100\,\%$

	Diferencia	Inferior	superior	P-value
2-1	46.00	19.403556	72.59644	0.0010268
3-1	41.50	14.903556	68.09644	0.0024312
4-1	116.50	89.903556	143.09644	0.0000001
5-1	139.75	113.153556	166.34644	0.0000000
3-2	-4.50	-31.096444	22.09644	0.9812528
4-2	70.50	43.903556	97.09644	0.0000175
5-2	93.75	67.153556	120.34644	0.0000008
4-3	75.00	48.403556	101.59644	0.0000092
5-3	98.25	71.653556	124.84644	0.0000005
5-4	23.25	-3.346444	49.84644	0.0978028

Al analizar el p-value calculado con el software R, este es consistente respecto a lo observado en el grafico. Entonces se concluye igualdad entre igualdad entre las medias 5 y 4, 3 y 2.
Los cálculos para estas conclusiones son:
Donde HSD es:

$$HSD_{i,k} = q_{0.05}(5,12)\sqrt{\frac{MSE}{4}}$$

Se tiene entonces que

$$\mu_{3,\bullet} = \mu_{2\bullet} \Rightarrow \text{ Grupo 1}$$
$$\mu_{5\bullet} = \mu_{4\bullet} \Rightarrow \text{ Grupo 2}$$

Formando únicamente 2 grupos, prefiriendo el grupo 1 (Marcas 3 y 2) ya que en promedio consumen una menor cantidad de energía que las Marcas 4 y 5.

4.7. Eficiencia del diseño

Para el factor de bloqueo se evalúa la eficiencia:

$$ER = k + (1-k)H = \frac{s_2^2}{s_1^2} = \frac{(b-1)MSB + b(a-1)s_1^2}{(ab-1)s_1^2}$$

Se puede mostrar que:

$$\text{ER} = 1 \Leftrightarrow H = 1$$
$$\text{ER} < 1 \Leftrightarrow H < 1$$
$$\text{ER} > 1 \Leftrightarrow H > 1$$

$$\text{ER} = \frac{(b-1)MSB + b(a-1)s_1^2}{(ab-1)s_1^2} = \frac{3\times 38739 + 4\times 4\times 139}{(5\times 4-1)\times 139} = 44.847$$

$$EER \times 100\,\% = \frac{(12+1)(15+3)}{(15+1)(12+3)} * 44.847 * 100\,\% = 4372.5825\,\%$$

Dado el análisis anterior del DBCA, se concluye fue altamente eficiente ya que para un DCA se requeriría n = b*EER= 4*43.7259=174.90, redondeando a 175 replicas por tratamiento para obtener la misma información que este DBCA.
Esto implica que el hecho de no bloquear afectaría de manera significativa (drástica, al tener en cuenta el valor tan alto del EER) y se perdería información con el tamaño de muestra dado en este experimento si la metodología fuese DCA. Sumado a que, bloquear, ayuda a controlar factores externos y ahorrar en cuando a la recolección de datos para un tamaño de muestra grande; se concluye que el análisis DBCA ha sido satisfactorio y la metodología usada es adecúa para concluir acerca del experimento.

5. Códigos en el software R

5.1. Punto 5

```
rm(list=ls(all=TRUE))
datos5=data.frame(scan(what=list(Tipo="",Vida=0)))
```

```
1 602
2 863
1 529
4 235
1 534
1 585
2 743
3 232
4 282
2 773
2 840
3 255
4 238
3 200
4 228
3 215

datos5
attach(datos5)
costo=ifelse(Tipo=="1",0.985,ifelse(Tipo=="2",0.935,ifelse(Tipo=="3",0.520,0.495)))
Vida.costo=Vida/costo
raizVida=sqrt(Vida)

# MODELO ANOVA
modeloanova=aov(Vida.costo~Tipo)
summary(modeloanova)

# TEST DE INDEPENDENCIA
Box.test(Vida.costo, lag = 1, type = c("Ljung-Box"), fitdf = 0)

# TEST DE NORMALIDAD CON RESIDUOS INTERNAMENTE ESTUDENTIZADOS
shapiro.test(rstandard(modeloanova))


# TESTS PARA HOMOGENEIDAD DE VARIANZA
bartlett.test(Vida.costo~Tipo)
leveneTest(Vida.costo~Tipo)
cochran.test(Vida.costo~Tipo)

#TAMANO DE MUESTRA
#funcon
diftukey2=function(n,a,sigma2,difsig,alpha=0.05){
  gl=a*(n-1)
  dif=qtukey(alpha,nmeans=a,df=gl,lower.tail=FALSE)^2-(n*difsig^2)/sigma2
  dif
}
#Buscando secuencialmente a n, desde n=2
n=2
#sigma2=MSE
res=diftukey2(n,a = 4,sigma2 = 2374.497,difsig=125,alpha = 0.05)
```

```
while(res>=0){
  n=n+1
  res=diftukey2(n,4,2374.497,difsig=125,alpha = 0.05)
}
n.optim=n; n.optim


tabla=data.frame(n=2:6,difer=diftukey2(2:6,4,2374.497,difsig=125),
                 Acción=ifelse(diftukey2(2:6,4,2374.497,difsig=125)>=0,
                               "Incrementar n","No incrementar n"))
tabla
```

5.2. Punto 13

```
rm(list=ls(all=TRUE))
datos13=data.frame(Lote=factor(rep(1:5,times=5)),ContCalcio=scan())
23.46 23.59 23.51 23.28 23.29
23.48 23.46 23.64 23.40 23.46
23.56 23.42 23.46 23.37 23.37
23.39 23.49 23.52 23.46 23.32
23.40 23.50 23.49 23.39 23.38


datos = datos13[order(datos13$Lote),]

datos13
attach(datos13)
mediasLot=sapply(split(ContCalcio,Lote),mean);mediasLot
boxplot(ContCalcio~Lote)
lines(1:5,mediasLot,col=2,type="b",pch=19)


#AJUSTANDO EFECTOS ALEATORIOS PARA OBTENER COMPONENTES DE VARIANZA
library(lme4)
diseño=lmer(ContCalcio~1|Lote)


#LO SIGUIENTE NOS DA ALGUNAS MEDIDAS DE AJUSTE Y LAS ESTIMACIONES DE LAS COMPONENTES
#DE VARIANZA OBSERVE QUE LO QUE APARECE EN LA SALIDA COMO INTERCEPTO
#ES SIMPLEMENTE EL VALOR DE LA MEDIA GLOBAL MUESTRAL
summary(diseño)

#INTERVALOS DE CONFIANZA ASINTÓTICOS PARA LAS RAÍCES CUADRADAS DE LAS VARIANZAS Y MEDIA GLOBAL
#sólo son buenas aproximaciones con grados de libertad mayores a 45,
#en cada componente de varianza!!
confint(diseño)

#OBTENIENDO LA ANOVA
summary(aov(ContCalcio~Error(Lote)))
dis2=aov(ContCalcio~Error(Lote))
summary(dis2)
```

```
anova(aov(ContCalcio~Lote))
## se rechaza la hipotesis nula por tanto

#------Probabilidad F test
pf(q = 5.5352,df1 =4 ,df2 = 20, lower.tail = F)

#OBTENEMOS RESIDUALES INTERNAMENTE ESTUDENTIZADOS Y VALORES AJUSTADOS COMO SI EL MODELO
#FUESE DE EFECTOS FIJOS Y CON ELLOS HACEMOS GRÁFICOS DE RESIDUALES
res.estudent=stdres(aov(ContCalcio~Lote))
Yhat=fitted(aov(ContCalcio~Lote))

#GRÁFICO DE RESIDUOS INTERNAMENTE ESTUDENTIZADOS VS. VALORES AJUSTADOS
layout(rbind(c(1,1,2,2),c(3,3,4,4)))
plot(Yhat,res.estudent,xlab="Respuesta ajustada",ylab="Residuales internamente estudentizados",
pch=1:5,col=c(1:5),cex=1.5)
abline(h=c(-2,0,2),lty=2)
legend("topright",legend=c(1:5),col=c(1:5),pch=c(1:5))

#GRÁFICO DE RESIDUOS VS. NIVELES FACTOR FARDO
stripchart(res.estudent~fardos,xlab="fardos",ylab="Residuales internamente
estudentizados",vertical=T, pch=1:5,col=c(1:5),cex=1.5)
abline(h=c(-2,0,2),lty=2)
legend("topright",legend=c(1:5),col=c(1:5),pch=c(1:5))

#NORMALIDAD SOBRE LOS ERRORES
shapiro.test(res.estudent)
qqnorm(res.estudent,
main="Gráfico probabilidad normal\nsobre residuos internamente estudentizados")
qqline(res.estudent,lty=2)

#NORMALIDAD SOBRE LOS EFECTOS ALEATORIOS
#Calcular medias de tratamientos estandarizadas
medias.fardos.estand=scale(mediasLot)
medias.fardos.estand
qqnorm(medias.fardos.estand,main="Gráfico de probabilidad normal\npara los efectos")
qqline(medias.fardos.estand,lty=2,col=2)
abline(h=0,lty=2)
abline(v=0,lty=2)




detach(datos13)
```

5.3. Punto 22

```
    rm(list=ls(all=TRUE))
datos22=data.frame(Marcas=factor(rep(1:5,each=4)),
Nivelhumedad=factor(rep(1:4,times=5)),KWh=scan())
685 792 838 875
```

```
722 806 893 953
733 802 880 941
811 888 952 1005
828 920 978 1023

datos22

attach(datos22)

mediasMarcas=sapply(split(KWh,Marcas),mean)
mediasMarcas
mediasNivelhumedad=sapply(split(KWh,Nivelhumedad),mean)
mediasNivelhumedad

boxplot(KWh~Marcas,boxwex=0.5,xlab="Marcas",col="bisque")
lines(mediasMarcas,col=2,lty=1,type="b",pch=19)

boxplot(KWh~Nivelhumedad,boxwex=0.5,xlab="Nivelhumedad",col="bisque")
lines(mediasNivelhumedad,col=2,lty=1,type="b",pch=19)

#escriba aquí el código R que sea necesario


#AJUSTANDO EL MODELO ANOVA Y OBTENCIÓN DE LA TABLA ANOVA
diseño=aov(KWh~Marcas+Nivelhumedad)
anova(diseño)

#OBTENCIÓN DE MEDIAS DE TRATAMIENTO CON SUS I.C DEL 95%
lsmeans(diseño,"Marcas")

#OBTENIENDO EFECTOS DE TRATAMIENTOS, RESULTADOS PARA TEST DE SIGNIFICANCIA Y SUS I.C DEL 95%
efect.MarcasA=fit.contrast(diseño,"Marcas",
rbind(":efecto Marcas A"=c(4/5,-1/5,-1/5,-1/5,-1/5)),conf=0.95)
efect.MarcasB=fit.contrast(diseño,"Marcas",
rbind(":efecto Marcas B"=c(-1/5,4/5,-1/5,-1/5,-1/5)),conf=0.95)
efect.MarcasC=fit.contrast(diseño,"Marcas",
rbind(":efecto Marcas C"=c(-1/5,-1/5,4/5,-1/5,-1/5)),conf=0.95)
efect.MarcasD=fit.contrast(diseño,"Marcas",
rbind(":efecto Marcas D"=c(-1/5,-1/5,-1/5,4/5,-1/5)),conf=0.95)
efect.MarcasE=fit.contrast(diseño,"Marcas",
rbind(":efecto Marcas D"=c(-1/5,-1/5,-1/5,-1/5,4/5)),conf=0.95)
rbind(efect.MarcasA,efect.MarcasB,efect.MarcasC,efect.MarcasD,efect.MarcasE)

#INTERVALOS DE TUKEY PARA LAS DIFERENCIAS DE MEDIAS DEL FACTOR DE TRATAMIENTOS
TukeyHSD(diseño,"Marcas",conf.level=0.95)
HSD.test(diseño,"Marcas", group=TRUE,console=TRUE) #Comparaciones de Tukey



#GRÁFICOS DE INTERVALOS DE TUKEY
plot(TukeyHSD(diseño,"Marcas",conf.level = 0.95),cex.lab=0.8,las=1)
```

```
#OBTENIENDO TEST DE NORMALIDAD SOBRE RESIDUALES ESTUDENTIZADOS INTERNAMENTE
shapiro.test(rstandard(diseño))


#OBTENIENDO DE GRÁFICOS PARA VALIDACIÓN DE SUPUESTOS CON RESIDUOS ESTUDENTIZADOS INTERNAMENTE
nf=layout(rbind(c(1,1,2,2),c(3,3,4,4)))
plot(fitted(diseño),rstandard(diseño),ylim=c(-2.5,2.5),cex=1.5,
main="Residuales estudentizados vs. valores ajustados")
abline(h=c(-2,0,2),col=2)

stripchart(rstandard(diseño)~Marcas,vertical=TRUE,ylim=c(-2.5,2.5),
pch=1,cex=1.5,xlab="Marcas",main="Residuales estudentizados vs. Marcas")
abline(h=c(-2,0,2),col=2)

stripchart(rstandard(diseño)~Nivelhumedad,vertical=TRUE,
ylim=c(-2.5,2.5),pch=1,cex=1.5,xlab="Humedad",main="Residuales estudentizados vs. operador")
abline(h=c(-2,0,2),col=2)

qqnorm(rstandard(diseño),cex=1.5)
qqline(rstandard(diseño),col=2,lty=2)
legend("topleft",legend=c("Shapiro-Wilk Test",expression(W==0.97299),expression(PValue== 0.8844)

detach(ensamble)


MSE22<-139
w1=797.50
w2=843.50
w3=839.00
w4=914.00
w5=937.25
qt(0.025,15,lower.tail = F)
caro=qt(0.025,15,lower.tail = F)*sqrt(MSE22*(1/4))
ic1<-w5+caro
ic2<-w5-caro
ic1
ic2
MSB=38739
f1=12
f2=15
```

6. Referencias

[1] Gonzalez Alvarez, N. (2021). Notas de Clase - Diseño de Experimentos - Partes I, II y III del curso 3007340.