I. DISEÑOS FACTORIALES CON TRES FACTORES DE EFECTOS FIJOS EN UN DCA

El estudio factorial de tres factores, A, B y C, con efectos fijos, permite estudiar los efectos A, B, C (efectos principales), AB, AC, BC (efectos de las interacciones de a dos factores) y ABC (la triple interacción), donde el grado de detalle con el que se puede estudiar depende del número de niveles usados para cada factor. Por ejemplo, para un factor con dos niveles no es posible descomponer su efecto individual; pero si se toman tres niveles el efecto individual de un factor puede descomponerse en una parte lineal y una cuadrática (ver contrastes de polinomios ortogonales).

Sean **a, b,** y **c** el número de niveles usados de los factores A, B y C respectivamente y **n** el número de réplicas en cada tratamiento (combinaciones de niveles de los tres factores). Se tendrá entonces la siguiente tabla ANOVA:

ANOVA					
Fuente	**SC**	**GL**	**CM**	**F_0**	**Valor P**
A	SSA	a-1	MSA	F_A=MSA/MSE	$P(f_{a-1,abc(n-1)}>F_A)$
B	SSB	b-1	MSB	F_B=MSB/MSE	$P(f_{b-1,abc(n-1)}>F_B)$
C	SSC	c-1	MSC	F_C=MSC/MSE	$P(f_{c-1,abc(n-1)}>F_C)$
AB	SS(AB)	(a-1)(b-1)	MS(AB)	F_{AB}=MS(AB)/MSE	$P(f_{(a-1)(b-1),abc(n-1)}>F_{AB})$
AC	SS(AC)	(a-1)(c-1)	MS(AC)	F_{AC}=MS(AC)/MSE	$P(f_{(a-1)(c-1),abc(n-1)}>F_{AC})$
BC	SS(BC)	(b-1)(c-1)	MS(BC)	F_{BC}=MS(BC)/MSE	$P(f_{(b-1)(c-1),abc(n-1)}>F_{BC})$
ABC	SS(ABC)	(a-1)(b-1)(a-1)	MS(ABC)	F_{ABC}=MS(ABC)/MSE	$P(f_{(a-1)(b-1)(c-1),abc(n-1)}>F_{ABC})$
ERROR	SSE	abc(n-1)			
TOTAL	SST	abcn-1			

El modelo asociado a este ANOVA es dado por:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \lambda_k + (\alpha\beta)_{ij} + (\alpha\lambda)_{ik} + (\beta\lambda)_{jk} + (\alpha\beta\lambda)_{ijk} + \varepsilon_{ijk}$$

en el cual α_i, β_j y λ_k denotan los efectos principales sobre la respuesta promedio global de los factores A, B y C en sus niveles i, j k respectivamente, $(\alpha\beta)_{ij}$, $(\alpha\lambda)_{ik}$ y $(\beta\lambda)_{jk}$ los efectos de las interacciones de a dos factores y $(\alpha\beta\lambda)_{ijk}$ los efectos de la interacción triple, con $\varepsilon_{ijkl} \overset{iid}{\sim} N(0,\sigma^2)$, sujetos a las restricciones $\sum_{i=1}^{a}\alpha_i = \sum_{j=1}^{b}\beta_j = \sum_{k=1}^{c}\lambda_k = 0$ y además las sumas de las interacciones sobre cualesquiera de los índices i, j y k, son iguales a cero.

Las sumas de cuadrados se calculan de la siguiente manera:

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n}\left(Y_{ijkl} - \bar{Y}_{\bullet\bullet\bullet\bullet}\right)^2,\ SSA = bcn\sum_{i=1}^{a}\left(\bar{Y}_{i\bullet\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet\bullet}\right)^2,\ SSB = acn\sum_{j=1}^{b}\left(\bar{Y}_{\bullet j\bullet\bullet} - \bar{Y}_{\bullet\bullet\bullet\bullet}\right)^2,$$

$$SSC = abn\sum_{k=1}^{c}\left(\bar{Y}_{\bullet\bullet k\bullet} - \bar{Y}_{\bullet\bullet\bullet\bullet}\right)^2,\ SS(AB) = cn\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{Y}_{ij\bullet\bullet} - \bar{Y}_{i\bullet\bullet\bullet} - \bar{Y}_{\bullet j\bullet\bullet} + \bar{Y}_{\bullet\bullet\bullet\bullet}\right)^2$$

$$SS(AC)=bn\sum_{i=1}^{a}\sum_{k=1}^{c}\left(\bar{Y}_{i\bullet k\bullet}-\bar{Y}_{i\bullet\bullet\bullet}-\bar{Y}_{\bullet\bullet k\bullet}+\bar{Y}_{\bullet\bullet\bullet\bullet}\right)^2,\ SS(BC)=an\sum_{j=1}^{b}\sum_{k=1}^{c}\left(\bar{Y}_{\bullet jk\bullet}-\bar{Y}_{\bullet j\bullet\bullet}-\bar{Y}_{\bullet\bullet k\bullet}+\bar{Y}_{\bullet\bullet\bullet\bullet}\right)^2$$

$$SS(ABC)=n\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\left(\bar{Y}_{ijk\bullet}-\bar{Y}_{ij\bullet\bullet}-\bar{Y}_{i\bullet k\bullet}-\bar{Y}_{\bullet jk\bullet}+\bar{Y}_{i\bullet\bullet\bullet}+\bar{Y}_{\bullet j\bullet\bullet}+\bar{Y}_{\bullet\bullet k\bullet}-\bar{Y}_{\bullet\bullet\bullet\bullet}\right)^2$$

$$SSE=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n}\left(Y_{ijkl}-\bar{Y}_{ijk\bullet}\right)^2$$

donde:

$\bar{Y}_{\bullet\bullet\bullet\bullet}$ es el promedio de todas las **abcn** observaciones

$\bar{Y}_{i\bullet\bullet\bullet}$ es el promedio de las observaciones en el nivel i-ésimo del factor A

$\bar{Y}_{\bullet j\bullet\bullet}$ es el promedio de las observaciones en el nivel j-ésimo del factor B

$\bar{Y}_{\bullet\bullet k\bullet}$ es el promedio de las observaciones en el nivel k-ésimo del factor C

$\bar{Y}_{ij\bullet\bullet}$ es el promedio de las observaciones para el i-ésimo nivel de A y el j-ésimo nivel de B

$\bar{Y}_{i\bullet k\bullet}$ es el promedio de las observaciones para el i-ésimo nivel de A y el k-ésimo nivel de C

$\bar{Y}_{\bullet jk\bullet}$ es el promedio de las observaciones para el j-ésimo nivel de B y el k-ésimo nivel de C

$\bar{Y}_{ijk\bullet}$ es el promedio de las observaciones para el i-ésimo nivel de A , el j-ésimo nivel de B y el k-ésimo nivel de C, es decir, la media muestral de tratamiento ijk.

El análisis de la significancia de los efectos de este modelo debe realizarse en forma jerárquica partiendo de los efectos de interacción de mayor orden. Por ello, ningún efecto principal o de interacción de dos factores puede eliminarse de un modelo mientras la triple interacción sea significativa. ***En general, ningún factor deberá descartarse del modelo mientras participe en efectos de interacción de cualquier orden que sean significativos*** y caso que ni sus efectos de interacción ni sus efectos principales resulten significativos, se podrá concluir que tal factor no incide sobre el promedio de la respuesta:

Paso 1: $H_0:(\alpha\beta\lambda)_{111}=\cdots=(\alpha\beta\lambda)_{abc}=0$. Si no se rechaza H_0, eliminarlos del modelo y realizar las pruebas en el paso 2.
Paso 2: $H_0:(\alpha\beta)_{11}=\cdots=(\alpha\beta)_{ab}=0$, $H_0:(\alpha\lambda)_{11}=\cdots=(\alpha\lambda)_{ac}=0$, $H_0:(\beta\lambda)_{11}=\cdots=(\beta\lambda)_{bc}=0$. En los casos en donde no se rechace H_0, eliminar las interacciones dobles no significativas y realizar las pruebas del paso 3 sólo para los factores no involucrados en interacciones significativas de ningún orden (dobles o triple).
Paso 3: $H_0:\alpha_1=\alpha_2=\cdots=\alpha_a=0$, $H_0:\beta_1=\beta_2=\cdots=\beta_b=0$, $H_0:\lambda_1=\lambda_2=\cdots=\lambda_c=0$

Gráficamente se pueden evaluar las interacciones dobles a través de los gráficos de perfiles de medias de sólo dos subíndices. Por ejemplo, para evaluar la interacción de los factores A y B podemos graficar las medias $\bar{Y}_{ij\bullet\bullet}$ vs. los niveles del factor A (ó B) y unimos mediante líneas las medias con el mismo subíndice j (ó i si los niveles de B aparecen en el eje horizontal del gráfico). Para evaluar la triple interacción en forma gráfica, podemos graficar por ejemplo, los perfiles de las medias $\bar{Y}_{ijk\bullet}$ vs. A (ó B) para cada nivel de B (ó A si B se toma en el eje horizontal) en cada nivel del factor C.

La validación de los supuestos para los errores del modelo de nuevo se realiza gráficamente:

Para chequear:	**Graficar residuales contra:**
Independencia	Orden de las observaciones (según espacio o tiempo)
Varianza igual y outliers	Valores ajustados, niveles de cada factor
Normalidad	Scores normales y test Shapiro Wilk

Ante la presencia de interacciones significativas, las comparaciones de medias por lo general, se centran en las medias de tratamiento más que en las medias de nivel de los factores, así que los métodos de comparación múltiple (Tukey, LSD, etc.) se pueden aplicar para construir intervalos de confianza simultáneos del tipo:

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} d_{ijk}\mu_{ijk} \in \left(\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} d_{ijk}\bar{Y}_{ijk\bullet} \pm \varpi\sqrt{MSE\,\frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} d_{ijk}^{2}}{n}} \right),$$

$$\text{con } \sum_{i=1}^{a} d_{ijk} = 0\ \forall_{jk},\quad \sum_{j=1}^{b} d_{ijk} = 0\ \forall_{ik},\quad \sum_{k=1}^{c} d_{ijk} = 0,\ \forall_{ij}$$

Donde ϖ es el coeficiente crítico apropiado para cada método, con los grados de libertad del **MSE** del modelo ANOVA de tres factores y teniendo en cuenta que el número de medias de tratamientos que se comparan es **abc**.

II. DISEÑO FACTORIAL GENERAL

Considere **m** factores A, B, C, ..., K con niveles a, b, c, ..., k respectivamente, donde K es el último factor del conjunto de factores de estudio, no necesariamente el undécimo, que es el lugar que esta letra ocupa en el alfabeto. Con estos niveles y factores se puede construir el diseño factorial general que consiste de a×b×c×...×k tratamientos. Si **m** es el número de factores, se pueden estudiar:

$\binom{m}{1} = m$ efectos principales (A, B, ..., K)

$\binom{m}{2}$ interacciones dobles (AB, AC, ...,(K-1)K)

$\vdots$

$\binom{m}{m-1}$ interacciones de **m-1** factores (AB...K-1,....)

$\binom{m}{m} = 1$ interacción de todos los **m** factores (ABC...K)

Por tanto en este modelo general se pueden plantear $2^m - 1$ pruebas de hipótesis mediante el ANOVA. Sin embargo para medir los efectos de interacciones se requieren al menos 2 réplicas por tratamiento, por ende a mayor número de tratamientos más costosa resultará la experimentación. En la práctica en un factorial completo casi nunca interesan todos los posibles efectos, puesto que por lo general sólo algunos de ellos son significativos. *Se aplica un principio pareto, el cual nos dice que la mayoría de la variabilidad observada se debe a unos pocos de los efectos posibles; por lo común se debe a algunos de los efectos principales e interacciones dobles. En general no influyen las interacciones de tres o más factores, por lo que puede resultar inútil en muchos casos gastar recursos experimentales para estudiar estos efectos.*

En resumen, sólo en el caso irreal de que todos los posibles efectos en el diseño factorial general estén activos, es necesario realizar al menos dos réplicas por tratamiento. En las situaciones reales se replican al menos dos veces (y no siempre) sólo los diseños factoriales consistentes de 16 o menos puntos de prueba o tratamientos [1].

REGLAS PARA SUMAS DE CUADRADOS DEL ANOVA CON TAMAÑOS DE MUESTRAS IGUALES (DISEÑOS BALANCEADOS)

1. Escriba el nombre del efecto principal (A, B, ...) o de interacción de interés y los correspondientes números de niveles y subíndices. Por ejemplo con los factores A, B, C, D, para la interacción ABD, los números de niveles son a, b, y d; los subíndices i, j, l.
2. El número de grados de libertad ν del efecto considerado es el producto del número de niveles disminuido cada uno en 1, para cada uno de los factores incluidos en el efecto. Por ejemplo, para ABD, ν=(a-1)(b-1)(d-1).
3. Realizar el producto indicado en los grados de libertad ν y reemplazar cada letra con el correspondiente subíndice. Por ejemplo para ABD ν=(a-1)(b-1)(d-1)=abd-ab-ad-bd+a+b+d-1, lo cual reemplazando por los subíndices correspondientes da $ijl - ij - il - jl + i + j + l - 1$
4. La suma de cuadrados asociada a un efecto específico se obtiene usando cada grupo de subíndices del paso anterior como los subíndices de un término $\bar{Y}$ promediado sobre todos los subíndices no presentes y manteniendo los mismos signos. Colocar lo que resulte entre paréntesis, elevar al cuadrado y sumar sobre todos los posibles subíndices. Por ejemplo para la interacción ABD,

$$SS(ABD) = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}\sum_{m=1}^{n}\left(\bar{Y}_{ij\bullet l\bullet} - \bar{Y}_{ij\bullet\bullet\bullet} - \bar{Y}_{i\bullet\bullet l\bullet} - \bar{Y}_{\bullet j\bullet l\bullet} + \bar{Y}_{i\bullet\bullet\bullet\bullet} + \bar{Y}_{\bullet j\bullet\bullet\bullet} + \bar{Y}_{\bullet\bullet\bullet l\bullet} - \bar{Y}_{\bullet\bullet\bullet\bullet\bullet}\right)^2$$

$$SS(ABD) = cn\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{l=1}^{d}\left(\bar{Y}_{ij\bullet l\bullet} - \bar{Y}_{ij\bullet\bullet\bullet} - \bar{Y}_{i\bullet\bullet l\bullet} - \bar{Y}_{\bullet j\bullet l\bullet} + \bar{Y}_{i\bullet\bullet\bullet\bullet} + \bar{Y}_{\bullet j\bullet\bullet\bullet} + \bar{Y}_{\bullet\bullet\bullet l\bullet} - \bar{Y}_{\bullet\bullet\bullet\bullet\bullet}\right)^2$$

5. La suma de cuadrados totales es la suma de los cuadrados de las desviaciones de los valores de los datos con respecto a la media global estimada, y sus grados de libertad es el producto de niveles de todos los factores por el número de réplicas menos 1. Por ejemplo, con cuatro factores y **n** replicaciones de cada tratamiento,

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}\sum_{m=1}^{n}\left(Y_{ijklm} - \bar{Y}_{\bullet\bullet\bullet\bullet\bullet}\right)^2 \text{ con g.l=abcdn-1}$$

6. El SSE es el SST menos las sumas de cuadrados de los otros efectos en el modelo. Los grados de libertad del SSE es igual a los grados de libertad del SST menos los grados de libertad de los otros efectos en el modelo.

7. La suma de cuadrados medios de un efecto es la correspondiente suma de cuadrados dividida por los respectivos grados de libertad.

8. La regla de decisión para probar la hipótesis nula de que los efectos debidos a un término específico del modelo son cero contra la hipótesis alternativa de que alguno de tales efectos no es cero, a un nivel de significancia γ, es:

$$\textbf{rechazar } \mathbf{H_0} \textbf{ si } \frac{SS/\nu}{SSE/df} = \frac{MS}{MSE} > f_{\gamma,\nu,df}$$

Por ejemplo, en un diseño con los factores A, B, C, D, para probar:

H_0 **: los efectos de la interacción ABD no son significativos**

H_1 **: Al menos uno de los efectos de la interacción ABD es significativos**

la regla de decisión es:

$\textbf{rechazar } \mathbf{H_0} \textbf{ si } \dfrac{\mathbf{MS(ABD)}}{\mathbf{MSE}} > f_{\gamma,(a-1)(b-1)(d-1),df}$ donde **df** son los respectivos grados de libertad del MSE.

9. Se pueden construir intervalos de confianza simultáneos para contrastes de medias de tratamientos, con los métodos antes vistos de Bonferroni, Scheffé, Tukey, Dunnett, utilizando los coeficientes críticos apropiados y el número de observaciones sobre las cuales se obtiene cada promedio en el contraste. También pueden construirse intervalos de confianza simultáneos para comparar medias de nivel de un factor teniendo en cuenta el número de observaciones que resultan para cada nivel en el diseño específico.

EJEMPLO

Se realizó un experimento sobre la duración de tela recubierta sujeta a pruebas con abrasivos normales. El diseño factorial de 3 factores incluyó dos sustancias distintas (F1, F2) en tres proporciones diferentes (25%, 50%, 75%) con y sin tratamiento de superficie (S1, S2); se probaron dos especímenes (réplicas) de cada una de las 12 combinaciones en un diseño totalmente aleatorizado. Los datos corresponden a la pérdida de peso (mg.) de los especímenes de tela por la prueba de abrasión.

Tratamiento de superficie

	S1		S2	
Proporción sustancias	**sustancia**		**sustancia**	
	F1	**F2**	**F1**	**F2**
	194	239	155	137
25%	208	187	173	160
	233	224	198	129
50%	241	243	177	98
	265	243	235	155
75%	269	226	229	132

1. Plantee el modelo estadístico
2. Realice un análisis descriptivo, tenga en cuenta que debe priorizar el análisis de las interacciones triples sobre cualquier otro efecto y que lo que se busca en este problema es el tratamiento (combinación de los tres factores) que brinde la mayor resistencia de la tela a la abrasión, es decir, la menor pérdida de peso.

3. Realice el análisis de varianza. Depure el modelo hasta hallar los términos significativos. Con base en el modelo final, recomiende el mejor tratamiento posible y tenga en cuenta también los análisis descriptivos previos.
4. Para el modelo final, valide supuestos sobre el término de error (normalidad y varianza constante)

1. MODELO ESTADÍSTICO

Considere como factores (todos de efectos fijos)
A: Proporción en que se aplica la sustancia de recubrimiento, niveles i=1, para 25%, i=2 para 50%, i=3 para 75%
B: Tratamiento de la superficie, niveles j=1 con tratamiento (S1) y j=2, sin tratamiento (S2)
C: tipo de sustancia de recubrimiento, niveles k=1 para tipo F1 y k=2 para tipo F2

MODELO 1: El modelo estadístico es (con **a=3, b=2, c=2, n=2**)

$$\boldsymbol{Y}_{ijkl} = \mu + \alpha_i + \beta_j + \lambda_k + (\alpha\beta)_{ij} + (\alpha\lambda)_{ik} + (\beta\lambda)_{jk} + (\alpha\beta\lambda)_{ijk} + \varepsilon_{ijk}$$

Donde α_i, β_j y λ_k denotan respectivamente los efectos fijos de los factores A, B y C en sus correspondientes niveles i, j, k, sobre el promedio global de la pérdida de peso de la tela por abrasión, en las condiciones simuladas en el experimento.

$(\alpha\beta)_{ij}$ es el efecto sobre el promedio global de la pérdida de peso, de la interacción de los factores proporción de la sustancia y tratamiento de la superficie, en sus respectivos niveles i, j.

$(\alpha\lambda)_{ik}$ es el efecto sobre el promedio global de la pérdida de peso, de la interacción de los factores proporción de la sustancia y tipo de sustancia de recubrimiento, en sus respectivos niveles i, k.

$(\beta\lambda)_{jk}$ es el efecto sobre el promedio global de la pérdida de peso, de la interacción de los factores tratamiento de la superficie y tipo de sustancia de recubrimiento, en sus respectivos niveles j, k.

$(\alpha\beta\lambda)_{ijk}$ es el efecto sobre el promedio global de la pérdida de peso, debido a la triple interacción de los tres factores, en sus respectivos niveles i, j, y k.

Se asume errores aleatorios $\varepsilon_{ijkl} \overset{iid}{\sim} N\left(0,\sigma^2\right)$ y para efectos de estimación por mínimos cuadrados se impone como restricciones que $\sum_{i=1}^{3}\alpha_i = \sum_{j=1}^{2}\beta_j = \sum_{k=1}^{2}\lambda_k = 0$ y además las sumas de las tres interacciones dobles y de la triple sobre cualquiera de los índices i, j y k, son iguales a cero.

2. ANALISIS DESCRIPTIVO

Con base en las gráficas que se proporcionan, se realizan los análisis gráficos de los efectos de los factores, teniendo presente que primero deben considerarse las interacciones triples y dobles antes de interpretar efectos individuales.

Análisis de la triple interacción (fijando sustancia)

Figura 1

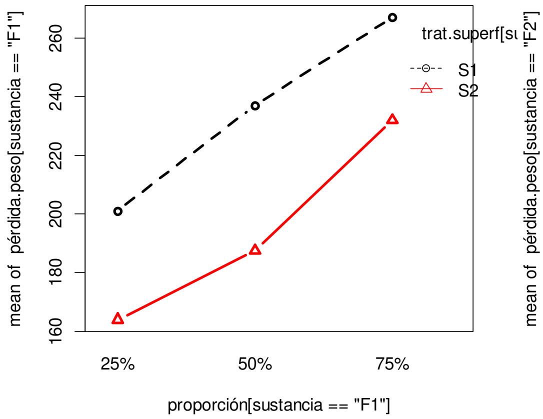
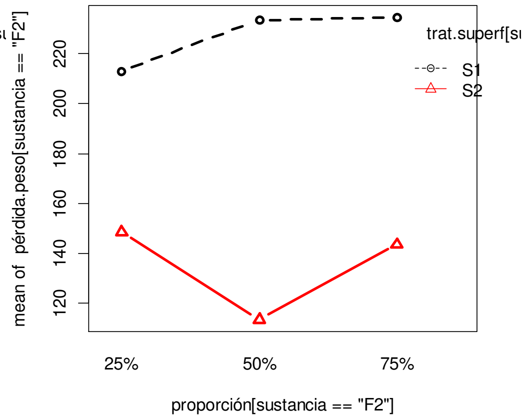


Figura 2



Análisis de interacciones dobles (si resultara apropiado)

Figura 3

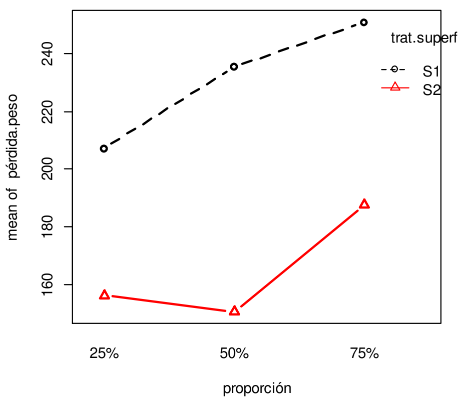


Figura 4

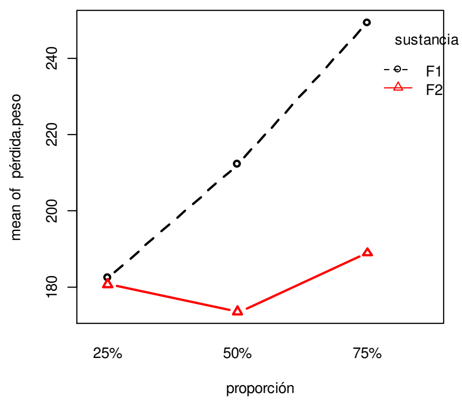
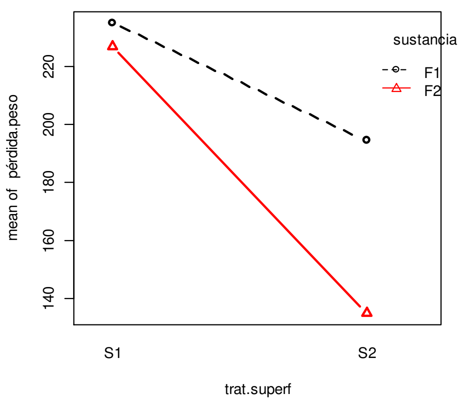


Figura 5



Análisis de efectos marginales o principales (si resultara apropiado)

Figura 6

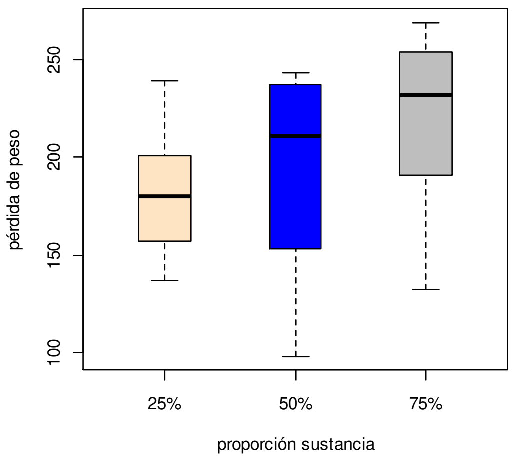


Figura 7

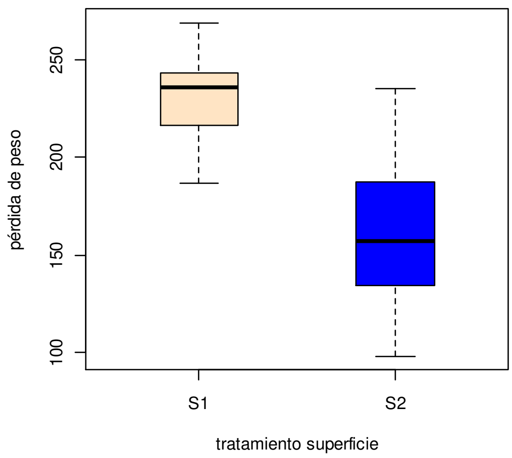
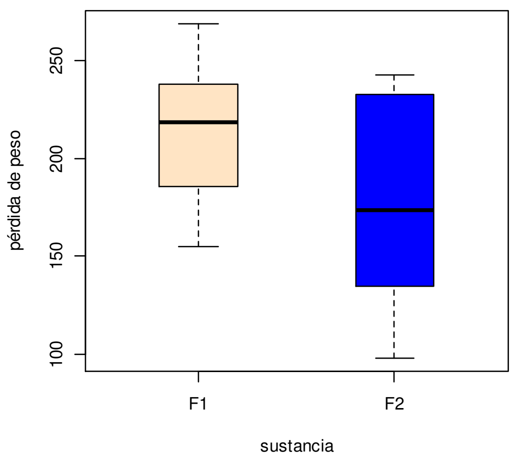


Figura 8



2.1 Diagnóstico de posible interacción triple

Para realizar este análisis es necesario considerar la interacción doble de dos de los factores, en cada nivel del tercero. Por ejemplo, en las figuras 1 y 2 se presentan los perfiles de medias para la interacción de los factores proporción y tratamiento de la superficie (interacción **AB**) en los niveles 1 y 2 del tipo de sustancia (factor **C**), respectivamente. *Para diagnosticar la posible presencia de una triple interacción, se determina si la interacción AB es o no similar en cada nivel de C*, es decir, si las figuras 1 y 2 son similares o no. Si son similares entonces decimos que la triple interacción no es significativa, pero si difieren, afirmamos que es posible que haya una triple interacción significativa. Para este problema notamos que la forma en que interactúan los factores proporción y tratamiento de superficie no parece igual cuando se usa la sustancia 1 que cuando se usa la sustancia 2; es más, a nivel de la sustancia 1 parece que no hay interacción entre la proporción y el tratamiento de la superficie, porque los dos perfiles de medias son casi paralelos, en tanto que a nivel de la sustancia 2 parece que hay cierta interacción entre estos dos factores puesto que los perfiles de medias no son paralelos. Sin embargo, es necesario determinar si estadísticamente se encuentra evidencia de interacción significativa.

Por otra parte, al parecer, comparando las medias de las figuras 1 y 2, la menor pérdida de peso promedio se logra cuando se aplica la sustancia 2 en una proporción de 50% y no se realiza tratamiento a la superficie de la tela. En esta combinación de los factores la media está alrededor de 115 (ver figura 2).

2.2 Diagnóstico de posible interacción entre factores proporción de sustancia y tratamiento de la superficie de la tela

La figura 3 permite diagnosticar en términos globales la interacción entre los factores proporción de la sustancia de recubrimiento y tratamiento de la superficie. Note la similitud de esta gráfica con la presentada en la figura 2, aunque los valores máximo y mínimo de las escalas del eje vertical no coinciden, siendo mayores en la figura 3. Dado que los perfiles de medias no son paralelos, puede decirse en principio que la interacción proporción de la sustancia y tratamiento de la superficie (interacción **AB**) puede ser significativa. Pero tenga en cuenta que si la triple interacción resulta significativa, entonces de hecho cualquier efecto de interacción doble es significativo a través de esa triple interacción y tal interacción doble variará según el nivel del tercer factor. Si este gráfico hubiese presentado los perfiles paralelos, de entrada no se puede concluir que la interacción doble que estamos examinando no va a ser significativa, porque puede ser que esté enmascarada por la triple interacción, si de hecho ésta última es significativa.

Además, en esta figura parece que la mejor combinación de los factores para reducir la pérdida de peso es no aplicar tratamiento a la superficie de la tela y usar una proporción del 50% de cualquiera de las dos sustancias (porque en este gráfico se hace una generalización respecto a la sustancias). Esta combinación para estos dos factores es coherente con la que arrojó el análisis sobre la triple interacción. Si no hubiese sido así, se tomaría la conclusión a la cual se llega con las gráficas 1 y 2 (las de la triple interacción) si la triple interacción resulta significativa.

2.3 Diagnóstico de posible interacción entre factores proporción de sustancia y tipo de sustancia

La figura 4 permite diagnosticar esta interacción. Es evidente que estos dos factores interactúan (la interacción **AC** parece significativa) dado que los perfiles de medias no son en absoluto paralelos. Por tanto, es necesario evaluar el efecto conjunto de estos factores sobre la respuesta.

También en esta figura parece que la mejor combinación para reducir el promedio de la pérdida de peso de la tela es usar la sustancia 2 en una proporción de 50%. Aquí de nuevo encontramos con esta gráfica una conclusión coherente a la obtenida con las figuras 1 y 2. Si no hubiese sido así, se tomaría la conclusión a la cual se llega con las figuras 1 y 2 si la triple interacción fuese significativa.

2.4 Diagnóstico de posible interacción entre factores tratamiento de la superficie de la tela y tipo de sustancia

En la figura 5 se aprecian los perfiles de medias de sustancia en los niveles de tratamiento de la superficie. Aquí también es evidente que los dos perfiles no son paralelos. Por tanto, es necesario evaluar el efecto conjunto de estos factores sobre la respuesta.

Respecto a la mejor combinación de estos dos factores, esta gráfica indica que para reducir el promedio de pérdida de peso en la tela, es recomendable usar la sustancia 2 y no realizar tratamiento a la superficie de la tela. De nuevo, encontramos con esta gráfica una conclusión coherente a la obtenida con las figuras 1 y 2. Si no hubiese sido así, se tomaría la conclusión a la cual se llega con las figuras 1 y 2 si la triple interacción fuese significativa.

2.5 Diagnóstico de efectos principales

En este punto es claro que los tres factores sí tienen efecto sobre los promedios de la variable respuesta, por tanto no es necesario evaluar lo que se presenta en las figuras 6, 7 y 8, porque de acuerdo a los análisis previos, se ha concluido que los efectos de un factor dependerán de los otros factores. Sin embargo veamos qué conclusiones se obtienen de cada gráfico. De la figura 6, aparentemente a medida que aumenta la proporción de la sustancia aplicada a la tela (sea cual sea el tipo de sustancia), tiende a aumentar el promedio de pérdida de peso; de la figura 7, se pierde en promedio más peso si se aplica tratamiento superficial (S1) que cuando no se aplica (S2); y de la figura 8 es mayor la pérdida de peso con la sustancia F1 que con F2.

Ahora bien, mirando estas tres gráficas se observa que para un menor promedio de pérdida de peso en la tela el factor proporción debe fijarse en un 25%, el factor tratamiento de la superficie debe fijarse en "no tratamiento de la superficie" y el factor sustancia debe fijarse en sustancia 2. De estas tres recomendaciones sólo la primera no coincide con el resultado arrojado por el análisis de las figuras 1 y 2, es decir las usadas para evaluar la triple interacción y de las figuras 3 a 5, es decir, de las interacciones dobles. ¿Cuál es la recomendación final a dar? Pues la que se obtuvo cuando se mira la combinación de los tres factores y no ésta última en la cual se considera individualmente a cada factor!!!.

3. ANALISIS DE VARIANZA

3.1 Modelo ANOVA con triple y dobles interacciones

A continuación se presenta la tabla ANOVA para el modelo inicial planteado en el numeral 1. Se procede primero con la prueba sobre la significancia de los efectos de interacción triple:

$$\boldsymbol{H}_0 : (\alpha\beta\lambda)_{ijk} = 0 \textit{ para todo i,j,k, } i = 1,\ 2,\ 3 \quad j = 1,\ 2 \quad k = 1,\ 2$$

$$\boldsymbol{H}_1 : \textit{ algún } (\alpha\beta\lambda)_{ijk} \neq 0$$

```
Analysis of Variance Table
Response: pérdida.peso
                                 Df  Sum Sq Mean Sq F value    Pr(>F)
proporción                        2  5967.6  2983.8 11.1025  0.001864 **
trat.superf                       1 26268.2 26268.2 97.7420 4.051e-07 ***
sustancia                         1  6800.7  6800.7 25.3048  0.000294 ***
proporción:trat.superf            2  1186.1   593.0  2.2067  0.152724
proporción:sustancia              2  3529.1  1764.5  6.5657  0.011852 *
trat.superf:sustancia             1  3952.7  3952.7 14.7076  0.002374 **
proporción:trat.superf:sustancia  2   478.6   239.3  0.8904  0.435959
Residuals                        12  3225.0   268.8
---
```

En esta tabla tenemos que el estadístico de prueba $\boldsymbol{F}_0 = \boldsymbol{MS}(\boldsymbol{ABC})/\boldsymbol{MSE} = 0.89,\ \boldsymbol{F}_0 \overset{H_0}{\sim} f_{2,12}$ y con $\boldsymbol{VP} = \boldsymbol{P}(f_{2,12} > 0.89) = 0.4360$, razón por la cual se puede concluir que no hay evidencia de interacción triple entre los factores proporción, tratamiento de la superficie y sustancia aplicada, a pesar de lo que se vio en las figuras 1 y 2. Antes de proceder con la evaluación de las interacciones dobles, es necesario replantear el modelo sin triple interacción.

NOTA: Si la triple interacción hubiese resultado significativa entonces el análisis puede terminar aquí y nos centraríamos en estudiar y comparar las medias μ_{ijk} para hallar el mejor tratamiento, que en este caso es el que proporcione un menor promedio de pérdida de peso de la tela.

MODELO 2: El modelo sin triple interacción es:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \lambda_k + (\alpha\beta)_{ij} + (\alpha\lambda)_{ik} + (\beta\lambda)_{jk} + \varepsilon_{ijk},$$

donde los términos quedan definidos de la misma forma como se hizo en el modelo 1.

Se asume errores aleatorios $\varepsilon_{ijkl} \overset{iid}{\sim} N(0,\sigma^2)$ y para efectos de estimación por mínimos cuadrados se impone como restricciones que $\sum_{i=1}^{3}\alpha_i = \sum_{j=1}^{2}\beta_j = \sum_{k=1}^{2}\lambda_k = 0$ y además las sumas de las tres interacciones dobles sobre cualquiera de los índices i, j y k, son iguales a cero.

```
   Analysis of Variance Table
   Response: pérdida.peso
                              Df  Sum Sq Mean Sq F value    Pr(>F)
A  proporción                  2  5967.6  2983.8 11.2791 0.0012078 **
C  trat.superf                 1 26268.2 26268.2 99.2969 9.758e-08 ***
B  sustancia                   1  6800.7  6800.7 25.7074 0.0001708 ***
AC proporción:trat.superf      2  1186.1   593.0  2.2418 0.1430211
AB proporción:sustancia        2  3529.1  1764.5  6.6702 0.0092314 **
CB trat.superf:sustancia       1  3952.7  3952.7 14.9416 0.0017146 **
   Residuals                  14  3703.6   264.5
   ---
```

Con la tabla ANOVA previa, se realizan los tests correspondientes a las dobles interacciones. Note en esta tabla que el SSE suma los valores del SSE y del SS(ABC) del modelo 1, y lo mismo respecto a sus grados de libertad.

Interacción Proporción×Tratamiento de la superficie	Interacción Proporción×Sustancia	Interacción Tratamiento superfic×Sustancia
$H_0: (\alpha\beta)_{ij} = 0$ *para todo i,j* $H_1:$ *algún* $(\alpha\beta)_{ij} \neq 0$ $i = 1,\ 2,\ 3 \quad j = 1,\ 2$	$H_0: (\alpha\lambda)_{ik} = 0$ *para todo i,k* $H_1:$ *algún* $(\alpha\lambda)_{ik} \neq 0$ $i = 1,\ 2,\ 3 \quad k = 1,\ 2$	$H_0: (\beta\lambda)_{jk} = 0$ *para todo j,k* $H_1:$ *algún* $(\beta\lambda)_{jk} \neq 0$ $j = 1,\ 2, \quad k = 1,\ 2$
$F_0 = MS(AB)/MSE = 2.24$ $F_0 \overset{H_0}{\sim} f_{2,14},\ VP = P(f_{2,14} > 2.24) = 0.1430$	$F_0 = MS(AC)/MSE = 6.67$ $F_0 \overset{H_0}{\sim} f_{2,14},\ VP = P(f_{2,14} > 6.67) = 0.0092$	$F_0 = MS(BC)/MSE = 14.94$ $F_0 \overset{H_0}{\sim} f_{1,14},\ VP = P(f_{1,14} > 14.94) = 0.0017$
Decisión: No hay interacción significativa entre los factores proporción de la sustancia y tratamiento de la superficie	**Decisión:** Sí hay interacción significativa entre los factores proporción de la sustancia y tipo de sustancia	**Decisión:** Sí hay interacción significativa entre los factores tratamiento de la superficie y tipo de sustancia

De lo anterior es claro que puede eliminarse del modelo los términos $(\alpha\beta)_{ij}$, hacemos esto antes de proseguir con otros análisis.

MODELO 3: $Y_{ijkl} = \mu + \alpha_i + \beta_j + \lambda_k + (\alpha\lambda)_{ik} + (\beta\lambda)_{jk} + \varepsilon_{ijk}$

De nuevo, se asume errores aleatorios $\varepsilon_{ijkl} \overset{iid}{\sim} N(0,\sigma^2)$ y para efectos de estimación por mínimos cuadrados se impone como restricciones que $\sum_{i=1}^{3}\alpha_i = \sum_{j=1}^{2}\beta_j = \sum_{k=1}^{2}\lambda_k = 0$ y además las sumas de las dos interacciones dobles sobre cualquiera de los índices i, j y k, son iguales a cero.

El ANOVA correspondiente se da a continuación:

```
Analysis of Variance Table
Response: pérdida.peso
                        Df  Sum Sq Mean Sq F value    Pr(>F)
proporción               2  5967.6  2983.8  9.7636 0.0016923 **
sustancia                1  6800.7  6800.7 22.2532 0.0002325 ***
trat.superf              1 26268.2 26268.2 85.9549 7.789e-08 ***
proporción:sustancia     2  3529.1  1764.5  5.7739 0.0129495 *
sustancia:trat.superf    1  3952.7  3952.7 12.9339 0.0024176 **
Residuals               16  4889.7   305.6
---
```

Note que todo lo que se va eliminando del modelo, se va agregando al término de error, de ahí que finalmente el error tenga un SSE=4889.6667 y 16 grados de libertad.

En este modelo también resultan significativas las interacciones Proporción×Sustancia y Tratamiento superficie×Sustancia, veamos los respectivos tests:

Interacción Proporción×Sustancia	Interacción Tratamiento superfic×Sustancia
$H_0: (\alpha\lambda)_{ik} = 0$ *para todo i,k* $H_1:$ *algún* $(\alpha\lambda)_{ik} \neq 0$ $i = 1, 2, 3 \quad k = 1, 2$	$H_0: (\beta\lambda)_{jk} = 0$ *para todo j,k* $H_1:$ *algún* $(\beta\lambda)_{jk} \neq 0$ $j = 1, 2, \quad k = 1, 2$
$F_0 = MS(AC)/MSE = 5.77$ $F_0 \overset{H_0}{\sim} f_{2,16}, \quad VP = P(f_{2,16} > 5.77) = 0.0129$	$F_0 = MS(BC)/MSE = 12.93$ $F_0 \overset{H_0}{\sim} f_{1,16}, \quad VP = P(f_{1,16} > 12.93) = 0.0024$
Decisión: Sí hay interacción significativa entre los factores proporción de la sustancia y tipo de sustancia	**Decisión:** Sí hay interacción significativa entre los factores tratamiento de la superficie y tipo de sustancia

Puesto que estas dos interacciones resultan significativas, ninguno de los efectos principales (los α_i, β_j, λ_k) pueden eliminarse del modelo. Podemos verificar si los efectos correspondientes resultan o no enmascarados:

Efectos de la proporción de aplicación de las sustancias	Efectos del tratamiento de la superficie de la tela	Efectos del tipo de la sustancia
$H_0: \alpha_1 = \alpha_2 = \alpha_3 = 0$ $H_1:$ *algún* $\alpha_i \neq 0 \quad i = 1, 2, 3$	$H_0: \beta_1 = \beta_2 = 0$ $H_1:$ *algún* $\beta_j \neq 0 \quad j = 1, 2$	$H_0: \lambda_1 = \lambda_2 = 0$ $H_1:$ *algún* $\lambda_k \neq 0 \quad k = 1, 2$
$F_0 = MSA/MSE = 9.76$ $F_0 \overset{H_0}{\sim} f_{2,16}, \quad VP = P(f_{2,16} > 9.76) = 0.0017$	$F_0 = MSB/MSE = 85.95$ $F_0 \overset{H_0}{\sim} f_{1,16}, \quad VP = P(f_{1,16} > 85.95) < 0.0001$	$F_0 = MSC/MSE = 22.25$ $F_0 \overset{H_0}{\sim} f_{1,16}, \quad VP = P(f_{1,16} > 22.25) = 0.0002$
Decisión: Los efectos del factor proporción de la sustancia no resultan enmascarados por la interacción Proporción×Sustancia	**Decisión:** Los efectos del factor tratamiento de la superficie no resultan enmascarados por la interacción Tratamiento de la superficie×Sustancia	**Decisión:** Los efectos del factor sustancia no resultan enmascarados por la interacción Tratamiento de la superficie×Sustancia

A pesar que no resultan enmascarados los efectos de los tres factores, para determinar el mejor tratamiento con base en este último modelo, se analizarán las medias $\bar{y}_{i\bullet k\bullet}$ y $\bar{y}_{\bullet jk\bullet}$, es decir, usando las figuras 4 y 5 en las cuales se representa respectivamente las interacciones Proporción×Sustancia y Tratamiento de la superficie×Sustancia.

Como se determinó en los numerales 2.3 y 2.4 la mejor combinación de niveles de los tres factores es: Aplicar en un 50% la sustancia 2 y no realizar tratamiento a la superficie de la tela.

4. VALIDACIÓN DE SUPUESTOS EN EL MODELO FINAL

4.1 Varianza constante

Las figuras 9 a 12 presentan los gráficos de residuales estandarizados obtenidos del modelo 3. Sin considerar los outliers (observaciones con residuales por fuera de las bandas ±2):

Figura 9: Residuales estandarizados vs. valores ajustados. No hay una evidencia gráfica contundente en contra del supuesto de varianza constante de los errores del modelo.

Figura 10: Residuales estandarizados vs. proporción de aplicación de las sustancias. Parece que a nivel de la proporción de 75% hay una menor dispersión en relación a los otros dos niveles de este factor

Figura 11: Residuales estandarizados vs. tratamiento de la superficie de la tela. Sin considerar los dos outliers, para las observaciones restantes, no se observa una evidencia fuerte contra el supuesto de varianza constante.

Figura 12: Residuales estandarizados vs. sustancia. A nivel de la sustancia 1 hay menor dispersión en relación con la sustancia 2.

En conclusión parece haber problemas con este supuesto.

Figura 9

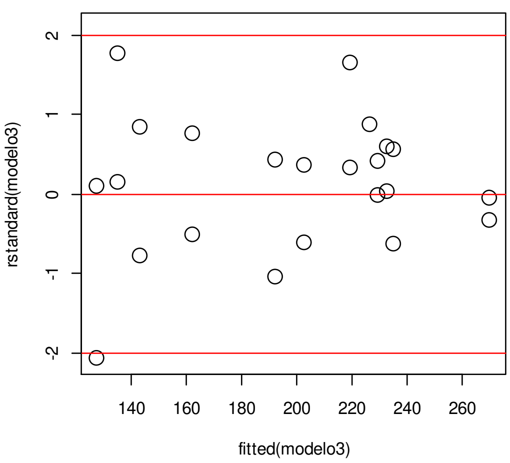


Figura 10

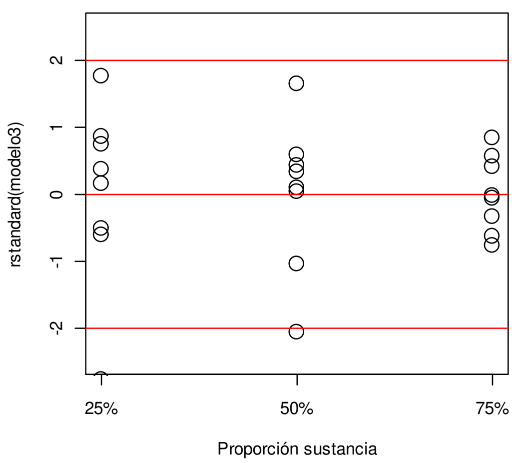


Figura 11

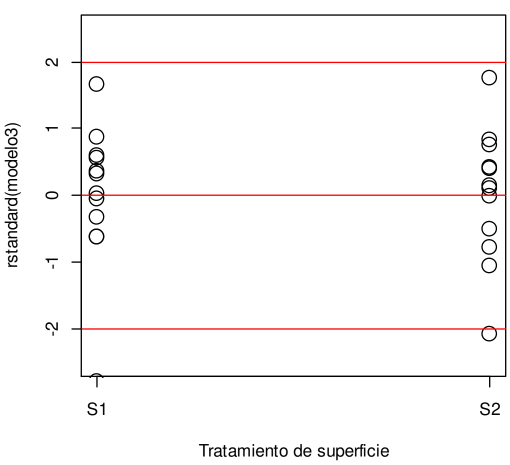
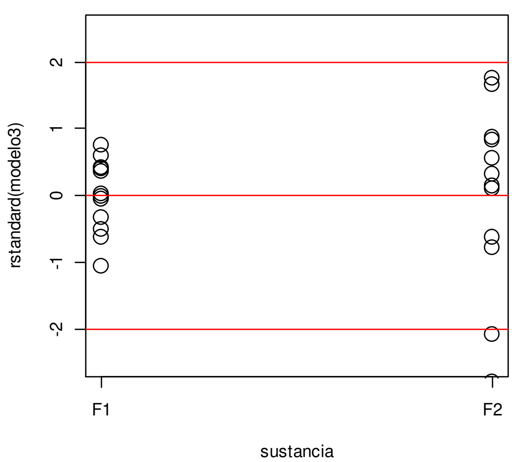
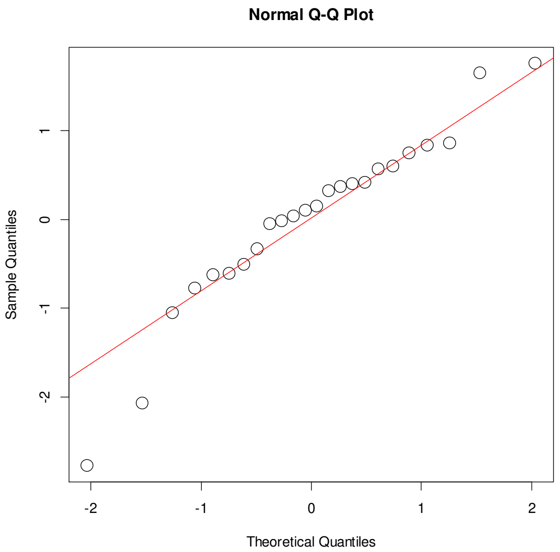


Figura 12



4.2 Normalidad

Respecto a la normalidad de los errores, a continuación se presentan los resultados del test de Shapiro Wilk sobre residuos estandarizados del modelo 3, que con un $W = 0.93534$ y un $VP = P(W < 0.93534) = 0.1283$ nos indica que es razonable asumir normalidad para el término de error del modelo.




```
Shapiro-Wilk normality test
data:  rstandard(modelo3)
W = 0.93534, p-value = 0.1283
```

PROGRAMA R USADO

```
#LECTURA DE LOS DATOS
datos.tela=data.frame(trat.superf=factor(rep(c("S1","S1","S2","S2"),times=6)),
sustancia=factor(rep(c("F1","F2","F1","F2"),times=6)),
proporción=factor(c(rep("25%",8),rep("50%",8),rep("75%",8))),pérdida.peso=scan())
194 239 155 137
208 187 173 160
233 224 198 129
241 243 177 98
265 243 235 155
269 226 229 132

datos.tela
attach(datos.tela)

#INTERACCIÓN TRIPLE
win.graph(width=10,height=5)
layout(cbind(c(1),c(2)))

interaction.plot(proporción[sustancia=="F1"],trat.superf[sustancia=="F1"],
pérdida.peso[sustancia=="F1"],type="b",pch=c(1,2),col=c("black","red"),lwd=3,
                 main="Figura 1")

interaction.plot(proporción[sustancia=="F2"],trat.superf[sustancia=="F2"],
pérdida.peso[sustancia=="F2"],type="b",pch=c(1,2),col=c("black","red"),lwd=3,
                 main="Figura 2")

#INTERACCIONES DOBLES
layout(rbind(c(1,1,2,2),c(0,3,3,0)))
interaction.plot(proporción,trat.superf,pérdida.peso,type="b",pch=c(1,2),
                  col=c("black","red"),lwd=2,main="Figura 3")
interaction.plot(proporción,sustancia,pérdida.peso,type="b",pch=c(1,2),col=c("black","red"),
                 lwd=2,main="Figura 4")
```

```
interaction.plot(trat.superf,sustancia,pérdida.peso,type="b",pch=c(1,2),
                 col=c("black","red"),lwd=2,main="Figura 5")

#BOXPLOTS EFECTOS PRINCIPALES
layout(rbind(c(1,1,2,2),c(0,3,3,0)))
boxplot(pérdida.peso~proporción,boxwex=0.4,ylab="pérdida de peso",
         xlab="proporción sustancia",col=c("bisque","blue","gray"),main="Figura 6")

boxplot(pérdida.peso~trat.superf,boxwex=0.4,ylab="pérdida de peso",
         xlab="tratamiento superficie",col=c("bisque","blue"),main="Figura 7")

boxplot(pérdida.peso~sustancia,boxwex=0.4,
         ylab="pérdida de peso",xlab="sustancia",col=c("bisque","blue"),main="Figura 8")

#MODELO ANOVA CON LA INTERACCIÓN TRIPLE
modelo1=aov(pérdida.peso~proporción*trat.superf*sustancia)
anova(modelo1)

#MODELO ANOVA SIN LA INTERACCIÓN TRIPLE
modelo2=aov(pérdida.peso~proporción*trat.superf+proporción*sustancia+trat.superf*sustancia)
anova(modelo2)

#CORRIENDO MODELO SIN INTERACCIÓN TRIPLE Y SIN EFECTO DE INTERACCIÓN PROPORCIÓN:TRATAMIENTO
#SUPERFICIAL
modelo3=aov(pérdida.peso~proporción*sustancia+trat.superf*sustancia)
anova(modelo3)

#ANÁLISIS RESIDUOS ESTANDARIZADOS MODELO 2
layout(rbind(c(1,1,2,2),c(3,3,4,4)))
plot(fitted(modelo3),rstandard(modelo3),cex=2,ylim=c(-2.1,2.1),main="Figura 9")
abline(h=c(-2,0,2),col=2)

stripchart(rstandard(modelo3)~proporción,cex=2,vertical=TRUE,
            ylim=c(-2.5,2.5),pch=1,xlab="Proporción sustancia",main="Figura 10")
abline(h=c(-2,0,2),col=2)

stripchart(rstandard(modelo3)~trat.superf,cex=2,vertical=TRUE,
            ylim=c(-2.5,2.5),pch=1,xlab="Tratamiento de superficie",main="Figura 11")
abline(h=c(-2,0,2),col=2)

stripchart(rstandard(modelo3)~sustancia,cex=2,vertical=TRUE,
           ylim=c(-2.5,2.5),pch=1,xlab="sustancia",main="Figura 12")
abline(h=c(-2,0,2),col=2)

win.graph()
qqnorm(rstandard(modelo3),cex=2)
qqline(rstandard(modelo3),col=2)

shapiro.test(rstandard(modelo3))

detach(datos.tela)
```

BIBLIOGRAFÍA

[1] Gutiérrez, Pulido H. y de la Vara Salazar, R. (2004). *Análisis y diseño de experimentos.* McGraw-Hill Interamericana.