



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN
EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

**PROPUESTA DE ANÁLISIS PARA EL CASO DE
RESPUESTAS MÚLTIPLES EN DISEÑOS
EXPERIMENTALES**

GLORIA SELENE HERRERA REYES

T E S I S

PRESENTADA COMO REQUISITO PARCIAL PARA
OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS

MONTECILLO, TEXCOCO, ESTADO DE MÉXICO
2019



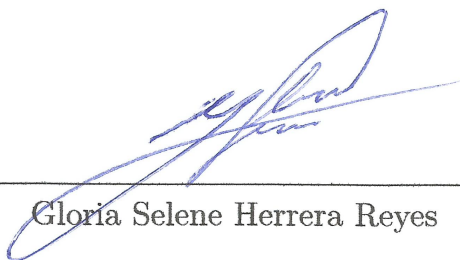
COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

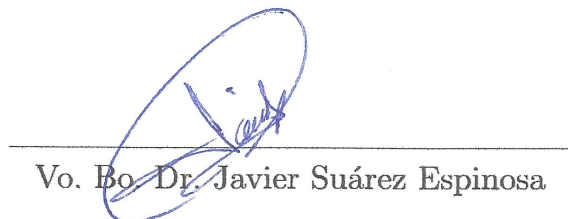
CARTA DE CONSENTIMIENTO DE USO DE LOS DERECHOS DE AUTOR Y DE LAS REGALÍAS COMERCIALES DE PRODUCTOS DE INVESTIGACIÓN

En adición al beneficio ético, moral y académico que he obtenido durante mis estudios en el Colegio de Postgraduados, la que suscribe, **“Gloria Selene Herrera Reyes”**, alumna de esta Institución, estoy de acuerdo en ser partícipe de las regalías económicas y/o académicas, de procedencia nacional e internacional, que se deriven del trabajo de investigación que realicé en esta institución, bajo la dirección del Profesor **“Dr. Javier Suárez Espinosa”**, por lo que otorgo los derechos de autor de mi tesis **“Propuesta de análisis para el caso de respuestas múltiples en diseños experimentales”**, y de los productos de dicha investigación al Colegio de Postgraduados. Las patentes y secretos industriales que se puedan derivar serán registrados a nombre del colegio de Postgraduados y las regalías económicas que se deriven serán distribuidas entre la Institución, el Consejero o Director de Tesis y el que suscribe, de acuerdo a las negociaciones entre las tres partes, por ello me comprometo a no realizar ninguna acción que dañe el proceso de explotación comercial de dichos productos a favor de esta Institución.

Colegio de Postgraduados, Montecillo, a 24 de junio de 2019.



Gloria Selene Herrera Reyes



Vo. Bo. Dr. Javier Suárez Espinosa

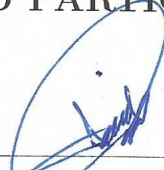
La presente tesis titulada: **Propuesta de análisis para el caso de respuestas múltiples en diseños experimentales**, realizada por la alumna: **Gloria Selene Herrera Reyes**, bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRO EN CIENCIAS

SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA ESTADÍSTICA

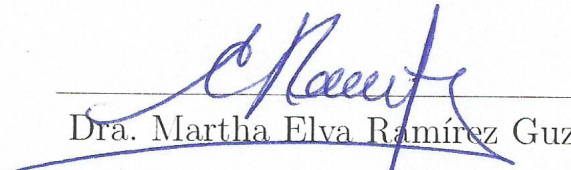
CONSEJO PARTICULAR

CONSEJERO



Dr. Javier Suárez Espinosa

ASESOR



Dra. Martha Elva Ramírez Guzmán

ASESOR



Dr. Gabriel A. Rodríguez Yam

Montecillo, Texcoco, Estado de México, 24 de junio de 2019

PROPUESTA DE ANÁLISIS PARA EL CASO DE RESPUESTAS MÚLTIPLES EN DISEÑOS EXPERIMENTALES

Gloria Selene Herrera Reyes, M. en C.

Colegio de Postgraduados, 2019

RESUMEN

En la investigación agropecuaria y forestal con frecuencia se generan respuestas múltiples, por ejemplo a la semilla de algodón se le mide el contenido de humedad, ácidos grasos, carbohidratos, tamaño, diámetro, longitud, forma, dureza, entre otras. Naturalmente, las respuestas múltiples se presentan en muchas otras áreas de investigación como la social, económica, ingenieril, salud, entre otras. El análisis de varianza multivariado (MANOVA) puede ser útil para el análisis de respuestas múltiples cuando se quiere determinar diferencias de efectos de tratamientos; sin embargo, esta metodología tiene ciertas limitaciones, tales como el cumplimiento de los supuestos de normalidad multivariada y homogeneidad de matrices de varianzas y covarianzas, supuestos que son difíciles de verificar si el tamaño de muestra es pequeño. Además, las pruebas post hoc actuales en este contexto, no son satisfactorias dada su dificultad de empleo por investigadores no expertos en la materia, o bien por limitaciones propias del método. Esta investigación propone una alternativa de análisis para contrastar la hipótesis de igualdad de efectos entre tratamientos en el caso de respuestas múltiples. Se demuestra un resultado asintótico para la variable aleatoria generada en la propuesta para el caso de variables normales no correlacionadas y queda abierto el caso para variables aleatorias normales correlacionadas. Un estudio de simulación muestra que el desempeño de la propuesta con muestras pequeñas es satisfactorio en términos de potencia y que tiene ventajas en comparación con el MANOVA. La propuesta metodológica es una buena alternativa para probar la hipótesis de igualdad de efectos de tratamientos con respuestas múltiples provenientes de un diseño experimental completamente al azar (DECA), incluso en el caso de muestras pequeñas y variables normales correlacionadas.

Palabras clave: ANOVA, MANOVA, Supuestos, Transformación de datos, Norma euclídeana

ANALYSIS PROPOSAL FOR THE CASE OF MULTIPLE RESPONSES IN EXPERIMENTAL DESIGNS

Gloria Selene Herrera Reyes, M. en C.

Colegio de Postgraduados, 2019

ABSTRACT

Multiple responses are often obtained in agricultural and forest research, for example, the cotton seed is measured for moisture content, fatty acids, carbohydrates, size, diameter, length, shape, hardness, etc. Naturally, multiple responses are presented in other areas of research, such as social, economic, engineering, health, among others. The multivariate analysis of variance (MANOVA) can be useful to analyze multiple responses when we would like to determine the differences of treatment effects; but this methodology has some limitations such as the accomplishment of assumptions like multivariate normality and homogeneity of variances and covariances matrices; these assumptions are difficult to verify if the sample size is small. In addition in this context, the current post hoc tests are not satisfactory given the difficulty in employment by researchers who are not experts in the field or due to limitations inherent to the method. An alternative analysis to test the hypothesis of equality of the effects in the multiple responses case is propose in this research. An asymptotic result for the random variable generated in the proposal has been demonstrated for the normal uncorrelated case, and the result for normal random correlated variables is leave open. A simulation study shows that the performance of the proposal with small samples is satisfactory in terms of power and that it has advantages compared with the MANOVA. The proposed methodology is a good alternative to test the hypothesis of equality of treatment effects for multiple responses from a completely random experimental design, even in the case of small samples and normal correlated variables.

Key words: ANOVA, MANOVA, Assumptions, Data transformation, Euclidean norm

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico brindado para la realización de mis estudios de maestría.

Al Colegio de Postgraduados, especialmente al postgrado en Estadística por permitirme formar parte de su comunidad y así continuar con mi formación académica.

Al Dr. Javier Suárez Espinosa le agradezco por su tiempo, apoyo, indicaciones y consejos para la realización de esta tesis.

A la Dra. Martha Elva Ramírez Guzmán por la disponibilidad para dedicar parte de su tiempo en la revisión de este trabajo y sus atinadas indicaciones.

Al Dr. Gabriel A. Rodríguez Yam por el tiempo brindado para la revisión, su calidez como persona y los consejos brindados.

A mis profesores del Colegio de Postgraduados.

Al personal administrativo del PSEI, por la amabilidad y apoyo brindado en este periodo, especialmente a Mary e Isa.

A mis compañeros de Maestría, en especial a Ricardo y Waldenia, por su amistad, apoyo y ayuda.

A mi familia, por su amor, apoyo y comprensión.

DEDICATORIA

A mi querida familia:

María Esther, Luna Mariana, Iván, Anahí, Ricardo y Damián,
gracias por ser un gran ejemplo, por su amor, amistad y apoyo incondicional.

CONTENIDO

RESUMEN	iv
ABSTRACT	iv
LISTA DE TABLAS	xiv
LISTA DE FIGURAS	xvi
1. Introducción	1
2. Objetivos	3
2.1. General	3
2.2. Específicos	3
3. Marco teórico	4
3.1. Análisis de varianza multivariado	4
3.1.1. Supuestos básicos	6
3.1.2. Comparaciones múltiples de vectores de medias	9
3.2. Resultados teóricos	11
3.2.1. Distancia euclideana	11

CONTENIDO

3.2.2.	Distribución normal multivariada	12
3.2.3.	Distribución de $\ \mathbf{Y}\ ^2 = \mathbf{Y}'\mathbf{Y}$: Distribución ji-cuadrada no central	12
3.2.4.	Teorema Central de Límites (TCL) para la suma de variables aleatorias	13
3.3.	Análisis de Varianza (ANOVA) de un factor	14
3.3.1.	Principales supuestos del ANOVA	16
3.3.2.	Procedimientos de comparaciones múltiples	19
3.3.3.	Potencia de la prueba	20
4.	Metodología propuesta	24
4.1.	Propuesta metodológica	24
4.2.	Evaluación del desempeño de la propuesta	29
4.2.1.	Preliminares	29
4.2.2.	Potencia de la prueba de igualdad de efectos de tratamientos usando la metodología propuesta	32
4.2.3.	Potencias del MANOVA vs ANOVA usando la metodología propuesta	37
4.3.	Comparaciones de resultados del MANOVA con el ANOVA usando la metodología propuesta	39
5.	Ejemplo de aplicación	45
6.	Discusión de resultados y conclusiones	49
6.1.	Potencia de la prueba	49
6.2.	Comparaciones de resultados del MANOVA con el ANOVA	50
6.3.	Conclusiones	50
	Referencias	50

CONTENIDO

Anexos	54
A.4. Códigos en el programa R	54
Códigos en el programa R	54
A.4.1. Demostración de convergencia para el caso de variables no correlacionadas	54
A.4.2. Caso 1. Variables no correlacionadas	54
A.4.3. Potencias del MANOVA vs ANOVA usando la metodología propuesta	56
A.4.4. Comparaciones de resultados del MANOVA con el ANOVA	57
A.4.5. Ejemplo de aplicación	59
A.5. Tablas de las comparaciones del MANOVA con el ANOVA	60

LISTA DE TABLAS

3.1. Estimación de de los estadísticos de prueba del MANOVA mediante SCP y <i>eigenvalores</i> , y transformación a estadístico $F_{(v_H, v_E)}$	6
3.2. Datos típicos de un experimento univariado de un solo factor	15
3.3. ANOVA correspondiente a un DECA	16
4.1. Uso de la norma euclidiana en el conjunto de datos	24
4.2. Comparaciones para el caso $t = 3, p = 3, r = 4$, pequeñas diferencias entre μ_i^s	42
4.3. Comparaciones para el caso $t = 3, p = 7, r = 4$, pequeñas diferencias entre μ_i^s	43
4.4. Comparaciones para el caso $t = 3, p = 3, r = 4$, grandes diferencias entre μ_i^s Variables correlacionadas	44
5.1. Funciones en el programa R para verificar supuestos del MANOVA	46
5.2. Resultados del análisis de normalidad multivariada	46
5.3. Resultado del análisis posterior al uso de la metodología propuesta	47
5.4. Contrastes de los rizomas	48
A.1. Comparaciones para el caso $t = 3, p = 3, r = 16$, pequeñas diferencias entre μ_i^s	61
A.2. Comparaciones para el caso $t = 3, p = 3, r = 4$, grandes diferencias entre μ_i^s	62
A.3. Comparaciones para el caso $t = 3, p = 3, r = 16$, grandes diferencias entre μ_i^s	63
A.4. Comparaciones para el caso $t = 3, p = 7, r = 16$, pequeñas diferencias entre μ_i^s	64

LISTA DE TABLAS

A.5. Comparaciones para el caso $t = 3, p = 7, r = 4$, grandes diferencias entre $\mu_i^{\prime s}$	65
A.6. Comparaciones para el caso $t = 3, p = 7, r = 16$, grandes diferencias entre $\mu_i^{\prime s}$	66
A.7. Comparaciones para el caso $t = 7, p = 3, r = 4$, pequeñas diferencias entre $\mu_i^{\prime s}$	67
A.8. Continuación: Comparaciones para el caso $t = 7, p = 3, r = 4$, pequeñas diferencias entre $\mu_i^{\prime s}$	68
A.9. Comparaciones para el caso $t = 7, p = 3, r = 16$, pequeñas diferencias entre $\mu_i^{\prime s}$	69
A.10. Continuación: Comparaciones para el caso $t = 7, p = 3, r = 16$, pequeñas diferencias entre $\mu_i^{\prime s}$	70
A.11. Comparaciones para el caso $t = 7, p = 3, r = 4$, grandes diferencias entre $\mu_i^{\prime s}$	71
A.12. Continuación: Comparaciones para el caso $t = 7, p = 3, r = 4$, grandes diferencias entre $\mu_i^{\prime s}$	72
A.13. Comparaciones para el caso $t = 7, p = 3, r = 16$, grandes diferencias entre $\mu_i^{\prime s}$	73
A.14. Continuación: Comparaciones para el caso $t = 7, p = 3, r = 16$, grandes diferencias entre $\mu_i^{\prime s}$	74
A.15. Comparaciones para el caso $t = 7, p = 7, r = 4$, pequeñas diferencias entre $\mu_i^{\prime s}$	75
A.16. Continuación: Comparaciones para el caso $t = 7, p = 7, r = 4$, pequeñas diferencias entre $\mu_i^{\prime s}$	76
A.17. Comparaciones para el caso $t = 7, p = 7, r = 16$, pequeñas diferencias entre $\mu_i^{\prime s}$	77
A.18. Continuación: Comparaciones para el caso $t = 7, p = 7, r = 16$, pequeñas diferencias entre $\mu_i^{\prime s}$	78
A.19. Comparaciones para el caso $t = 7, p = 7, r = 4$, grandes diferencias entre $\mu_i^{\prime s}$	79
A.20. Continuación: Comparaciones para el caso $t = 7, p = 7, r = 4$, grandes diferencias entre $\mu_i^{\prime s}$	80
A.21. Comparaciones para el caso $t = 7, p = 7, r = 16$, grandes diferencias entre $\mu_i^{\prime s}$	81
A.22. Continuación: Comparaciones para el caso $t = 7, p = 7, r = 16$, grandes diferencias entre $\mu_i^{\prime s}$	82

LISTA DE TABLAS

A.23.Comparaciones para el caso $t = 3, p = 3, r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas	83
A.24.Comparaciones para el caso $t = 3, p = 3, r = 16$, pequeñas diferencias entre μ_i^s Variables correlacionadas	84
A.25.Comparaciones para el caso $t = 3, p = 3, r = 16$, grandes diferencias entre μ_i^s Variables correlacionadas	85
A.26.Comparaciones para el caso $t = 3, p = 7, r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas	86
A.27.Comparaciones para el caso $t = 3, p = 7, r = 16$, pequeñas diferencias entre μ_i^s Variables correlacionadas	87
A.28.Comparaciones para el caso $t = 3, p = 7, r = 4$, grandes diferencias entre μ_i^s Variables correlacionadas	88
A.29.Comparaciones para el caso $t = 3, p = 7, r = 16$, grandes diferencias entre μ_i^s Variables correlacionadas	89
A.30.Comparaciones para el caso $t = 7, p = 3, r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas	90
A.31.Continuación: Comparaciones para el caso $t = 7, p = 3, r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas	91
A.32.Comparaciones para el caso $t = 7, p = 3, r = 16$, pequeñas diferencias entre μ_i^s Variables correlacionadas	92
A.33.Continuación: Comparaciones para el caso $t = 7, p = 3, r = 16$, pequeñas diferencias entre μ_i^s Variables correlacionadas	93
A.34.Comparaciones para el caso $t = 7, p = 3, r = 4$, grandes diferencias entre μ_i^s Variables correlacionadas	94
A.35.Continuación: Comparaciones para el caso $t = 7, p = 3, r = 4$, grandes diferencias entre μ_i^s Variables correlacionadas	95
A.36.Comparaciones para el caso $t = 7, p = 3, r = 16$, grandes diferencias entre μ_i^s Variables correlacionadas	96
A.37.Continuación: Comparaciones para el caso $t = 7, p = 3, r = 16$, grandes diferencias entre μ_i^s Variables correlacionadas	97

LISTA DE TABLAS

A.38.Comparaciones para el caso $t = 7, p = 7, r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas	98
A.39.Continuación: Comparaciones para el caso $t = 7, p = 7, r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas	99
A.40.Comparaciones para el caso $t = 7, p = 7, r = 16$, pequeñas diferencias entre μ_i^s Variables correlacionadas	100
A.41.Continuación: Comparaciones para el caso $t = 7, p = 7, r = 16$, pequeñas diferencias entre μ_i^s Variables correlacionadas	101
A.43.Continuación: Comparaciones para el caso $t = 7, p = 7, r = 4$, grandes diferencias entre μ_i^s Variables correlacionadas	102
A.44.Comparaciones para el caso $t = 7, p = 7, r = 16$, grandes diferencias entre μ_i^s Variables correlacionadas	103
A.45.Continuación: Comparaciones para el caso $t = 7, p = 7, r = 16$, grandes diferencias entre μ_i^s Variables correlacionadas	104

LISTA DE FIGURAS

3.1. Distribución del estadístico F para distintos valores del λ	23
4.1. Distribución $\chi_p^2(\lambda)$ vs aproximación asintótica a la normal. $t = 3, p = 2, 3$. .	27
4.2. Distribución $\chi_p^2(\lambda)$ vs aproximación asintótica a la normal. $t = 3, p = 5, 7$. .	27
4.3. Distribución $\chi_p^2(\lambda)$ vs aproximación asintótica a la normal. $t = 7, p = 2, 3$. .	28
4.4. Distribución $\chi_p^2(\lambda)$ vs aproximación asintótica a la normal. $t = 7, p = 5, 7$. .	28
4.5. Potencias estimadas cambiando la cantidad de vectores de medias en $t = 3$, $p = 3, r = 4$	31
4.6. Potencias estimadas cambiando la cantidad de vectores de medias en $t = 5$, $p = 3, r = 4$	31
4.7. Potencias estimadas cambian la cantidad de vectores de medias en $t = 7$, $p = 3, r = 4$	32
4.8. Potencias estimadas para $p = 3$ y $t = 3$	33
4.9. Potencias estimadas para $p = 5$ y $t = 3$	33
4.10. Potencias estimadas para $p = 7$ y $t = 3$	34
4.11. Potencias estimadas para $p = 3$ y $t = 5$	35
4.12. Potencias estimadas para $p = 5$ y $t = 5$	35
4.13. Potencias estimadas para $p = 7$ y $t = 5$	36
4.14. Potencias estimadas para $p = 3$ y $t = 7$	36

LISTA DE FIGURAS

4.15. Potencias estimadas para $p = 5$ y $t = 7$	37
4.16. Potencias estimadas para $p = 7$ y $t = 7$	37
4.17. Potencias estimadas del MANOVA y del ANOVA empleando la metodología propuesta. $p = 3$ y $t = 3$	38
4.18. Potencias estimadas del MANOVA y del ANOVA empleando la metodología propuesta. $p = 5$ y $t = 3$	38
4.19. Potencias estimadas del MANOVA y del ANOVA empleando la metodología propuesta. $p = 7$ y $t = 3$	39

Capítulo 1

Introducción

En diversas áreas de investigación con frecuencia se generan respuestas múltiples como producto de un diseño experimental, por ejemplo, [Pérez *et al.* \(2014\)](#) presentan un estudio de cultivares de haba donde se generaron las siguientes respuestas: altura de plantas, número de ramas, número de nudos florales, número de vainas por planta, peso de vaina por planta, entre otras. En otras investigaciones se han obtenido mediciones de contenido de proteína, grasas, vitaminas, peso, color y textura a una porción de carne; la cantidad de bacterias, pH, contenido de fibra y vitaminas a un nopalito; el contenido de humedad, ácidos grasos, carbohidratos, tamaño, diámetro, longitud, forma y dureza a la semilla de algodón, entre otros ejemplos. Esto debido a que muchas preguntas de investigación no se pueden responder adecuadamente con solo una variable dependiente y por lo tanto, es importante analizar el efecto de los tratamientos a través de las respuestas múltiples. Sin embargo, el análisis de respuestas múltiples frecuentemente presenta dificultades.

Una de las metodologías empleadas para analizar el efecto de tratamientos con respuestas múltiples, es el análisis de varianza multivariado (MANOVA por sus siglas en inglés). El MANOVA presenta algunas limitantes, tales como el incumplimiento de los supuestos de normalidad multivariada y homogeneidad de varianzas y covarianzas, supuestos que además son difíciles de verificar si el tamaño de la muestra es pequeño ([Hair *et al.*, 1999](#)), así como el número de repeticiones por tratamiento debe ser mayor al número de variables a estudiar, y si el conjunto de datos es obtenido de un gran número de variables puede presentar alta multicolinealidad, lo que puede generar una matriz de varianzas y covarianzas singular. Un ejemplo del problema de aplicación del MANOVA en un experimento con mediciones repetidas se enuncia por [Torres *et al.* \(2003\)](#), pues en su estudio no fue posible la aplicación de este método por no haber suficientes observaciones para la estimación de sus parámetros. Además existen otras limitaciones, por ejemplo, para poder verificar si los datos tienen distribución normal multivariada existen numerosas funciones en el programa **R**, sin embargo [Joensuu y Vogel \(2014\)](#) concluyeron en su artículo que muchas de esas funciones no son confiables dada su baja potencia, y por ello no deberían emplearse. Aunado al problema anterior, en áreas diferentes a la estadística aún sigue la práctica de verificar el supuesto de

normalidad únicamente en las distribuciones marginales, asumiendo que si éstas son normales se cumple el supuesto de normalidad multivariada, lo cual es incorrecto.

Otra limitante del MANOVA es que cuando se rechaza la hipótesis de igualdad de efectos de tratamientos, se requiere determinar cuales tratamientos presentan efectos diferentes entre sí, para ello las pruebas post hoc actuales en este caso no son satisfactorias dada su dificultad de empleo por investigadores no expertos en la materia ([Warne, 2014](#)), o bien por limitaciones propias del método; por ejemplo, utilizar contrastes con corrección de Bonferroni, trae como consecuencia una metodología muy conservadora; otra alternativa es emplear la conjetura generalizada de Tukey, la cual es una metodología desarrollada por [Seo et al. \(1994\)](#) y [Nishiyama y Seo \(2008\)](#), la cual consiste en realizar múltiples comparaciones por pares de vectores de medias, asumiendo que éstos tienen distribución normal multivariada; sin embargo, esta propuesta tiene la limitante de que sólo se probó para un limitado número de tratamientos y variables.

Otra alternativa por la que optan muchos investigadores es realizar el análisis de varianza (ANOVA) a cada una de las respuestas múltiples, lo cual en principio trae como consecuencia que el error tipo I aumente a niveles no aceptados en la investigación científica. Otro problema que presenta este enfoque es que se puede caer en contradicciones al interpretar los resultados del ANOVA de cada variable respuesta. Por lo anterior es deseable tener otra alternativa para el análisis de experimentos con respuestas múltiples.

Capítulo 2

Objetivos

2.1. General

Proponer una metodología alternativa para el análisis de experimentos con respuestas múltiples.

2.2. Específicos

1. Estudiar la potencia de la prueba del ANOVA usando la propuesta metodológica para distintos escenarios de interés.
2. Comparar los resultados del desempeño del MANOVA contra la metodología propuesta para el análisis de experimentos (DECA) con respuestas múltiples.

Capítulo 3

Marco teórico

3.1. Análisis de varianza multivariado

El análisis de varianza multivariado de un solo factor (One-Way MANOVA), es una metodología empleada para evaluar si existen diferencias en los efectos entre dos o más tratamientos (t) cuando tienen dos o más variables dependientes (p). (Johnson y Wichern, 2007) (Rencher y Christensen, 2012). El modelo empleado para comparar los efectos de los tratamientos en un diseño experimental completamente aleatorio (DECA) es:

$$\mathbf{Y}_{ij} = \boldsymbol{\mu} + \boldsymbol{\tau}_i + \boldsymbol{\varepsilon}_{ij},$$

donde cada componente del modelo es un vector p -dimensional: $\mathbf{Y}_{ij} = (Y_{ij1}, Y_{ij2}, \dots, Y_{ijp})'$ es el vector de k variables aleatorias correspondientes a la j -ésima repetición del i -ésimo tratamiento, $i=1,2,\dots,t$, $j=1,2,\dots,r$, Y_{ijk} pertenecen a la k -ésima variable en cada vector \mathbf{Y}_{ij} ; $\boldsymbol{\mu}$ es el vector de medias globales; $\boldsymbol{\tau}_i$ representa al vector de efectos del i -ésimo tratamiento, con $\sum_{i=1}^t \boldsymbol{\tau}_i = \mathbf{0}$; $\boldsymbol{\varepsilon}_{ij}$ es el vector de errores aleatorios, $\boldsymbol{\varepsilon}_{ij} \stackrel{iid}{\sim} N_p(\mathbf{0}, \boldsymbol{\Sigma})$, con $\boldsymbol{\Sigma}$ como la matriz de varianzas y covarianzas.

El juego de hipótesis empleado en el MANOVA es (Timm, 2002):

$$H_0 : \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \dots = \boldsymbol{\tau}_t \quad \text{vs} \quad H_a : \boldsymbol{\tau}_i \neq \boldsymbol{\tau}_j, \text{ para al menos un } i \neq j.$$

Para probar este juego de hipótesis, se obtienen las sumas de cuadrados y productos (SCP) (Peña, 2002), (Dattalo, 2013):

3.1. Análisis de varianza multivariado

$$\underbrace{SCP_{Total}}_{\mathbf{T}} = \underbrace{SCP_{Tratamiento}}_{\mathbf{H}} + \underbrace{SCP_{Error}}_{\mathbf{E}}$$

En la cual \mathbf{T} corresponde a la matriz de la suma total de cuadrados, \mathbf{H} es la matriz de la suma de cuadrados entre grupos (matriz de hipótesis o efecto) y \mathbf{E} es la matriz de la suma de cuadrados dentro de los grupos (matriz residual o de error). Las matrices \mathbf{H} y \mathbf{E} pueden expresarse como:

$$\mathbf{H} = r \sum_{i=1}^t (\bar{\mathbf{Y}}_i - \bar{\mathbf{Y}}_{..}) (\bar{\mathbf{Y}}_i - \bar{\mathbf{Y}}_{..})'$$

$$\mathbf{E} = \sum_{i=1}^t \sum_{j=1}^r (\mathbf{Y}_{ij} - \bar{\mathbf{Y}}_i) (\mathbf{Y}_{ij} - \bar{\mathbf{Y}}_i)'$$

donde $\bar{\mathbf{Y}}_i = \frac{1}{r} \sum_{j=1}^r \mathbf{Y}_{ij} = \frac{1}{r} (\mathbf{Y}_{i1} + \mathbf{Y}_{i2} + \dots + \mathbf{Y}_{ir})$ es el vector de medias muestrales del tratamiento i , y $\bar{\mathbf{Y}}_{..} = \frac{1}{tr} \sum_{i=1}^t \sum_{j=1}^r \mathbf{Y}_{ij} = \frac{1}{tr} (\mathbf{Y}_{11} + \dots + \mathbf{Y}_{tr})$ corresponde al vector de medias global.

[Rencher y Christensen \(2012\)](#) y [Dattalo \(2013\)](#) muestran como a partir de \mathbf{H} y \mathbf{E} se derivan los siguientes estadísticos de prueba del MANOVA:

- Lambda de Wilks (Λ)
- Traza de Hotelling-Lawley (T_0^2)
- Traza de Pillai ($V^{(s)}$)
- Mayor raíz característica de Roy (θ)

Cada uno de estos estadísticos se puede estimar con las matrices \mathbf{H} y \mathbf{E} , pero computacionalmente se calculan en términos de *eigenvalores* ($\lambda_i, i = 1, \dots, s$). En la práctica es común evaluar la significancia del MANOVA mediante una aproximación a un estadístico F , en lugar de emplear alguno de los estadísticos de prueba multivariados, y para dicha aproximación se requiere transformar la distribución de los estadísticos multivariados acorde a la Tabla 3.1.

[Rencher y Christensen \(2012\)](#) establecen que en términos de potencia, las pruebas tienen el siguiente orden para el caso colineal $\theta \geq T_0^2 \geq \Lambda \geq V^{(s)}$. En el caso difuso, en el que al menos un grupo difiere de los otros grupos en al menos una variable ([Olson, 1974](#)), y para estructuras intermedias entre colineal y difuso, la potencia tiene un orden inverso al caso colineal $V^{(s)} \geq \Lambda \geq T_0^2 \geq \theta$. Este último orden también se aplica a la precisión de la tasa de error Tipo I cuando las matrices de covarianza poblacional $\Sigma_1, \Sigma_2, \dots, \Sigma_t$ no son iguales.

3.1. Análisis de varianza multivariado

Tabla 3.1: Estimación de de los estadísticos de prueba del MANOVA mediante SCP y *eigenvalores*, y transformación a estadístico $F_{(v_H, v_E)}$

Estadístico de prueba	Λ	T_0^2	$V^{(s)}$	θ
SCP	$\ \mathbf{E}(\mathbf{H} + \mathbf{E})^{-1}\ $	traza $(\mathbf{E}^{-1}\mathbf{H})$	traza $((\mathbf{H} + \mathbf{E})^{-1}\mathbf{H})$	λ_{max} de $(\mathbf{E}^{-1}\mathbf{H})$
Eigenvalores	$\prod_{i=1}^s \frac{1}{1 + \lambda_i}$	$\sum_{i=1}^s \lambda_i$	$\sum_{i=1}^s \frac{\lambda_i}{1 + \lambda_i}$	$\frac{\lambda_{max}}{1 + \lambda_{max}}$
$F_{(v_H, v_E)}$	$\frac{1 - \Lambda^{\frac{1}{d}} v_E}{\Lambda^{\frac{1}{d}} v_H}$	$\frac{T_0^2 v_E}{s v_H}$	$\frac{V^{(s)} v_E}{s - V^{(s)} v_H}$	$\frac{\theta v_E}{1 - \theta v_H}$
v_H	$p(v_1)$	$p(v_1)$	$p(v_1)$	v_1
v_E	$1 + wd - \frac{pv_H}{2}$	$s(N - m - 2 - p) + 2$	$s(N - m + s - p)$	$N - m - 1$

Donde: v_H y v_E son los grados de libertad multivariados para la aproximación al estadístico F ,

$$w = N - 1 - \frac{p + v_1 + 1}{2}, \quad d = \sqrt{\frac{p^2 v_1^2 - 4}{p^2 + v_1^2 - 5}}, \quad v_1 = t - 1, \quad v_2 = t(r - 1), \quad N = tr, \quad m = \sum v_1, \quad y$$

$$s = \min(v_1, p).$$

En general, el uso de θ de Roy no es recomendada más que en el caso colineal bajo supuestos estándar. Si los datos provienen de una distribución no normal exhibiendo falta de simetría o curtosis positiva, bajo ciertas condiciones el desempeño de cualquiera de las otras tres pruebas es aceptable, aunque el estadístico de Pillai es el más empleado dada su robustez ante violaciones de los supuestos del MANOVA (Olson, 1974) (Olson, 1979). Pruebas F multivariadas estadísticamente significativas indican que hay diferencias entre los grupos respecto a sus vectores de medias, es decir, se rechaza H_0 (Dattalo, 2013).

Para que el resultado obtenido del MANOVA sea confiable, Johnson y Wichern (2007) establecen que se debe asumir que se tienen datos multivariados que cumplen con los siguientes supuestos básicos.

3.1.1. Supuestos básicos

1. Independencia

Las observaciones son muestras aleatorias independientes, de tamaño r del tratamiento i .

2. Normalidad

El supuesto de normalidad considera que los vectores de variables aleatorias tienen distribución normal multivariada con media $\boldsymbol{\mu}_i$ y matriz de varianzas y covarianzas $\boldsymbol{\Sigma}$ dentro de cada grupo (tratamiento) (Johnson y Wichern, 2007).

Si los tamaños de muestra son grandes, la violación de este supuesto no tiene mucha influencia; si son medianos, los incumplimientos pequeños pueden ser suavizados siempre y cuando las diferencias sean debidas a asimetría y no a datos anómalos.

3.1. Análisis de varianza multivariado

El supuesto de normalidad multivariada puede verificarse mediante algunos de los siguientes métodos: Mardia, Henze-Zirkler, Royston, Doornik-Hansen, entre otros. [Romeu y Ozturk \(1993\)](#) presentan un análisis de varias pruebas de bondad de ajuste para normalidad multivariada, y de sus estudios de simulación concluyen que la prueba de asimetría y curtosis propuesta por [Mardia \(1970\)](#) es de las más estables para valorar normalidad multivariada.

3. Igualdad de las matrices de varianzas y covarianzas

El supuesto de homogeneidad entre las matrices de varianzas y covarianzas, considera que tanto las varianzas como las covarianzas son constantes en los diferentes niveles de un factor, es decir, $\Sigma_1 = \Sigma_2 = \dots = \Sigma_t$.

En lugar de probar la igualdad de varianzas para una sola variable como en el ANOVA, la prueba en el caso del MANOVA examina todos los elementos de la matriz de varianzas y covarianzas.

El MANOVA es robusto a violaciones moderadas de este supuesto si los grupos son aproximadamente del mismo tamaño, es decir, el tamaño del grupo más grande dividido entre el tamaño del grupo más pequeño es menor de 1.5. Si los tamaños difieren más que esta medida, si es posible, el investigador debe corregir la posible existencia de varianzas diferentes mediante el uso de transformaciones que estabilicen la varianza ([Hair et al., 1999](#)).

Para realizar la prueba de igualdad de las matrices de varianzas y covarianzas, normalmente se emplea la metodología del contraste M de Box, sin embargo, [Hair et al. \(1999\)](#) establecen que esta metodología es muy sensible a la falta de normalidad, y por ello siempre se debe comprobar la normalidad multivariante antes de llevar a cabo esta prueba.

4. Multicolinealidad

Las variables dependientes deben de tener baja a moderada correlación entre ellas ([Dattalo, 2013](#)).

5. Sensibilidad a los datos atípicos

Los datos atípicos (outliers) influyen sobre el *error tipo I*. Si se encuentran datos aberrantes, se deben eliminar si es posible, ya que su influencia será muy importante en el conjunto de los datos.

Los valores atípicos en muestras multivariadas afectan no solo la media y la varianza de una variable, sino también su orientación en la muestra medida por la covarianza o la correlación con otras variables. Los residuos formados por el ajuste de un modelo multivariable a un conjunto de datos en presencia de extremos atípicos pueden conducir a la identificación de valores falsos ([Timm, 2002](#)).

Para verificar el cumplimiento de los supuestos del MANOVA se cuenta con los siguientes métodos analíticos:

3.1. Análisis de varianza multivariado

Pruebas para verificar los supuestos del MANOVA

Prueba de normalidad multivariada: Método de Mardia

[Mardia \(1970\)](#), propuso una extensión multivariada de los estadísticos de asimetría y curtosis, los cuales se calculan mediante el tercer y cuarto momento respectivamente.

Sea $\mathbf{X}_1, \dots, \mathbf{X}_n$ una muestra aleatoria de vectores p -variados, independientes e idénticamente distribuidos con vector de medias $\boldsymbol{\mu}$ y matriz de varianzas y covarianzas $\boldsymbol{\Sigma}$ desconocidos, se desea probar la hipótesis nula de que la muestra aleatoria proviene de una distribución normal p -variada.

Se establece que $\bar{\mathbf{X}}$ y \mathbf{S} son la media y la matriz de varianzas y covarianzas muestrales, esto es: $\bar{\mathbf{X}} = \frac{1}{r} \sum_{i=1}^r \bar{\mathbf{X}}_i$ y $\mathbf{S} = \frac{1}{r} \sum_{i=1}^r (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'$.

Los estadísticos de asimetría y curtosis p -variada se denotan respectivamente por $b_{1,p}$ y $b_{2,p}$, y se definen como:

$$b_{1,p} = \frac{1}{r^2} \sum_{i,j=1}^r [(\mathbf{X}_i - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X}_j - \bar{\mathbf{X}})]^3.$$
$$b_{2,p} = \frac{1}{r} \sum_{i=1}^r [(\mathbf{X}_i - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X}_i - \bar{\mathbf{X}})]^2.$$

Bajo H_0 , $\frac{1}{6} r b_{1,p}$ converge en distribución a una ji-cuadrada asintótica con $\frac{p(p+1)(p+2)}{6}$ grados de libertad.

Bajo H_0 , $\sqrt{r}(b_{2,p} - p(p+2)(r-1)/(r+1)) \xrightarrow{d} N(0, \frac{8p(p+2)}{n})$.

Para una distribución normal multivariada las medidas multivariadas de asimetría y curtosis tienen valores de $b_{1,p} = 0$ y $b_{2,p} = p(p+2)$.

Los valores críticos de asimetría y curtosis para muestras pequeñas se encuentran en [Mardia \(1974\)](#).

Prueba de homogeneidad de matrices de varianzas y covarianzas: M de Box

[Box \(1949\)](#) propuso el siguiente estadístico de prueba de razón de verosimilitud para probar la hipótesis de igualdad de matrices de varianzas y covarianzas:

$$M = (N - t) \ln |\mathbf{S}_p| - \sum_{i=1}^t (r_i - 1) \ln |\mathbf{S}_i|,$$

3.1. Análisis de varianza multivariado

en el cual $N = \sum_{i=1}^t r_i$, \mathbf{S}_i corresponde a la matriz de varianzas y covarianzas de cada tratamiento y $\mathbf{S}_p = (N-t)^{-1} \sum_{i=1}^t (r_i-1)\mathbf{S}_i$ es la matriz de varianzas y covarianzas agrupada.

Este método también es conocido como prueba de Bartlett multivariada.

Estableciendo:

$$u = \left[t \frac{1}{r-1} - \frac{1}{t(r-1)} \right] \left[\frac{2p^2 + 3p - 1}{6(p+1)(t-1)} \right],$$

entonces $C = (1-u)M$ tiene una distribución aproximada $\chi_{\frac{1}{2}p(p+1)(t-1)}^2$.

A un nivel de significancia dado, se rechaza la hipótesis de igualdad de matrices de varianzas y covarianzas si $C > \chi_{\frac{p(p+1)(t-1)}{2}}(\alpha)$.

Esta aproximación tiene las siguientes limitantes: requiere que $p < r$, es altamente sensible a violaciones en la distribución normal multivariada y a la presencia de outliers .

Como se puede observar, el cumplimiento y la medición de los supuestos del MANOVA presentan algunos problemas.

3.1.2. Comparaciones múltiples de vectores de medias

Si con base en el MANOVA se rechaza H_0 , se requiere identificar cuáles son los tratamientos que presentan efectos diferentes entre sí, para ello se realizan pruebas de seguimiento, a veces denominadas contrastes, las cuales son: **a priori** (planificadas) o **post hoc** (no planificadas). Una comparación **a priori** es aquella en la que se ha decidido evaluar antes del análisis de los datos; estas comparaciones son impulsadas por la teoría y forman parte de una estrategia de análisis confirmatorio de datos. La comparación **post hoc** es aquella en la que se decidió analizar después de obtener los datos y son parte de una estrategia de análisis exploratorio de datos (Anderson, 2003), (Rencher y Christensen, 2012).

Un método muy popular para realizar contrastes por pares de tratamientos es emplear la T^2 de Hotelling, y para mantener constante la tasa de error tipo I se emplea la corrección de Bonferroni, la cual consiste en someter a prueba cada uno de los contrastes a un nivel de significancia $\frac{\alpha}{c}$, donde c es el número de comparaciones a realizar y se obtiene mediante la fórmula $c = t(t-1)/2$. Por ejemplo, si se tiene un MANOVA de una vía con cuatro grupos ($t = 4$) y una variable, entonces hay $t(t-1)/2 = 6$ posibles comparaciones por pares de los tratamientos. Entonces, para limitar la tasa de error del estudio se emplea un valor corregido de la significancia, el cual será $P = \frac{0.05}{6} = 0.0083$, y por lo tanto ninguna de las 6 comparaciones se considerará significativa a menos que alcance un valor-p ≤ 0.0083 . Sin embargo este procedimiento es muy conservador (Dattalo, 2013).

3.1. Análisis de varianza multivariado

Otro candidato para evaluar los vectores de medias que son diferentes es la metodología de Conjetura generalizada de Tukey, la cual es la versión multivariada de la metodología Tukey-Kramer.

Conjetura generalizada de Tukey

Éste es un método desarrollado por Seo *et al.* (1994) y Nishiyama y Seo (2008), el cual es una generalización de la metodología Tukey-Kramer univariada, y en el cual, se obtienen intervalos de confianza simultáneos como sigue.

Partiendo de un modelo lineal multivariado donde $\mathbf{M} = [\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_t]$ es la matriz de t vectores de medias p -dimensionales, y $\hat{\mathbf{M}} = [\hat{\boldsymbol{\mu}}_1, \dots, \hat{\boldsymbol{\mu}}_t]$ es el estimador de \mathbf{M} , tal que:

$$\text{vec}(\hat{\mathbf{M}}) \sim N_{kp}(\text{vec}(\mathbf{M}), \mathbf{V} \otimes \boldsymbol{\Sigma}).$$

Donde:

$\text{vec}(\mathbf{A})$: Vector columna formada tras apilar las columnas de \mathbf{A} .

$\mathbf{V} = [v_{ij}]$: Matriz conocida definida positiva $k \times k$.

$\boldsymbol{\Sigma}$: Matriz desconocida definida positiva $p \times p$.

Si se considera a \mathbf{S} como un estimador insesgado de $\boldsymbol{\Sigma}$, se podrán realizar cálculos para la comparación entre vectores de medias, donde \mathbf{S} corresponde a la \mathbf{S}_{pl} de la T^2 de Hottelling.

En general el procedimiento para generar intervalos de confianza para las diferencias por pares de vectores de medias, y para todas las combinaciones lineales $\mathbf{a}'(\boldsymbol{\mu}_i - \boldsymbol{\mu}_{i'})$ está dado por:

$$\mathbf{a}'(\boldsymbol{\mu}_i - \boldsymbol{\mu}_{i'}) \in \left[\hat{\mathbf{a}}'(\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_{i'}) \pm q_{p,t,v}(\alpha) \sqrt{\frac{d_{ij}}{2} \mathbf{a}' \mathbf{S} \mathbf{a}} \right], \quad 1 \leq i < i' < t, \quad \mathbf{a} \in \mathbb{R}^p. \quad (3.1)$$

Donde:

\mathbf{a} : Vector p -dimensional de las combinaciones.

$(\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_{i'})$: Diferencia entre los vectores de medias estimados.

$q_{p,t,v}(\alpha)$: Punto α superior del rango p -variado Studentizado.

3.2. Resultados teóricos

d_{ij} : $v_{ii} - 2v_{ij} + v_{jj}$.

Se puede apreciar que la Ecuación (3.1) es semejante a:

$$\mathbf{a}'\mathbf{M}\mathbf{b} \in \left[\mathbf{a}'\hat{\mathbf{M}}\mathbf{b} \pm z(\mathbf{b}'\mathbf{V}\mathbf{b})^{\frac{1}{2}}(\mathbf{a}'\mathbf{S}\mathbf{a})^{\frac{1}{2}} \right], 1 \leq i < i' < t, \mathbf{a} \in \mathbb{R}^p, \mathbf{b} = \mathbf{e}_i - \mathbf{e}_{i'}.$$

Donde:

\mathbf{b} : Vector de contrastes, dado que \mathbf{e}_i es un vector k -dimensional que tiene un 1 en el i -ésimo componente, y cero en los otros.

$z = \frac{q_{p,t,v}(\alpha;\mathbf{I})}{\sqrt{2}}$: Punto α superior del rango p -variado Studentizado.

En general, $\frac{q_{p,t,v}^2(\alpha;\mathbf{V})}{\sqrt{2}}$ es el punto α superior de T_{max}^2 cuando $\mathbf{V} \neq v\mathbf{I}_t$, y cuando $\mathbf{V} = v\mathbf{I}_t$ es el único valor de T_{max}^2 , el cual está tabulado parcialmente para los casos de $t = 3$ y $t = 4$ en las tablas de los artículos de Seo *et al.* (1994) y Nishiyama y Seo (2008).

La limitante de estas propuestas es que solo se probaron para un limitado número de tratamientos, variables y grados de libertad. Es decir, existen problemas para realizar las pruebas de comparación de vectores de medias existentes y no existen muchas opciones en el caso multivariado.

3.2. Resultados teóricos

La propuesta metodológica se basa en algunos resultados teóricos que se presentan en este apartado.

3.2.1. Distancia euclideana

La distancia o norma euclideana corresponde a la longitud del vector unitario con p componentes, y es definida como la raíz cuadrada del producto punto como sigue (Strang, 2006):

$$\|\mathbf{y}\| = \sqrt{\mathbf{y}'\mathbf{y}} = \sqrt{y_1^2 + \dots + y_p^2}.$$

Todos los puntos con la misma distancia desde el origen $\|\mathbf{y}\| = c$ satisfacen $y_1^2 + \dots + y_p^2 = c^2$ que es la ecuación de un esferoide; lo que significa que todos los componentes del vector \mathbf{y}

3.2. Resultados teóricos

contribuyen igualmente a la distancia euclídeana.

3.2.2. Distribución normal multivariada

Asumiendo que $\mathbf{Y} = (Y_1, \dots, Y_p)'$ tiene distribución normal multivariada: $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, se establece que su densidad está dada por:

$$f(\mathbf{y}) = \frac{1}{(\sqrt{2\pi})^p |\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}, \mathbf{y} = (y_1, \dots, y_p)' \in \mathbb{R}^p,$$

donde p corresponde al número de variables, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)' \in \mathbb{R}^p$ es el vector de medias, y $\boldsymbol{\Sigma}$ es la matriz de varianzas y covarianzas definida positiva del vector \mathbf{Y} .

3.2.3. Distribución de $\|\mathbf{Y}\|^2 = \mathbf{Y}'\mathbf{Y}$: Distribución ji-cuadrada no central

Sea $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \mathbf{I}_p)$, entonces la variable aleatoria $X = \mathbf{Y}'\mathbf{Y} = \sum_{i=1}^p Y_i^2$, tiene distribución ji-cuadrada no central (Clarke, 2008):

$$\mathbf{Y}'\mathbf{Y} \sim \chi_p^2(\lambda).$$

Con parámetros: $p > 0$ que especifica los grados de libertad y $\lambda \geq 0$ que es el parámetro de no centralidad, dado por:

$$\lambda = \frac{1}{2} \boldsymbol{\mu}'\boldsymbol{\mu}.$$

Observaciones:

- La variable X tiene función de densidad (Clarke, 2008):

$$f(x) = \sum_{k=0}^{\infty} \frac{e^{-\frac{\lambda}{2}} \left(\frac{\lambda}{2}\right)^k}{k!} \frac{e^{-\frac{x}{2}} x^{\frac{p}{2}+k-1}}{\Gamma\left(\frac{p}{2} + k\right) 2^{\frac{p}{2}+k}}, \text{ si } x > 0.$$

3.2. Resultados teóricos

- La función generatriz de momentos de una distribución $\chi_p^2(\lambda)$ es:

$$M(t; p, \lambda) = \frac{\exp\left(-\lambda\left(1 - \frac{1}{1-2t}\right)\right)}{(1-2t)^{\frac{p}{2}}}, \quad 2t < 1.$$

- Los primeros cuatro momentos de la distribución $\chi_p^2(\lambda)$ son:
 - Media: $p + \lambda$
 - Varianza: $2(p + 2\lambda)$.
 - Coeficiente de Asimetría: $8(p + 3\lambda)$,
 - Coeficiente de Curtosis: $48(p + 4\lambda) + 12(p + 2\lambda)^2$.
- Cuando $\lambda = 0$ la distribución ji-cuadrada no central es igual a una distribución ji-cuadrada central, es decir: $\chi_p^2(0) \equiv \chi_p^2$.

COROLARIO 1: Si $\mathbf{Y} \sim N_p(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_p)$, entonces $\frac{\mathbf{Y}'\mathbf{Y}}{\sigma^2} \sim \chi_p^2(\lambda)$, siendo $\lambda = \frac{1}{2\sigma^2} \boldsymbol{\mu}'\boldsymbol{\mu}$ (Casella, 2008).

COROLARIO 2: Sean $\mathbf{X}_i \sim \chi_{p_i}^2(\lambda_i)$, $i = 1, \dots, t$ variables aleatorias independientes. Entonces:

$$\sum_{i=1}^t X_i \sim \chi_{\sum_{i=1}^t p_i}^2 \left(\sum_{i=1}^t \lambda_i \right).$$

3.2.4. Teorema Central de Límites (TCL) para la suma de variables aleatorias

DeGroot y Schervish (2012), mencionan que el TCL indica que la distribución de la suma de variables aleatorias puede aproximarse a la normal bajo ciertas condiciones.

Teorema 1. Sea W_1, W_2, \dots, W_n una sucesión de variables aleatorias independientes e idénticamente distribuidas de una distribución dada con media μ y varianza σ^2 finitas, entonces la suma $W_1 + \dots + W_n$ converge en distribución a una normal con media $n\mu$ y varianza $n\sigma^2$.

Dado lo anterior, podemos concluir que la suma de variables aleatorias independientes e idénticamente distribuidas ji-cuadradas no centrales ($\chi_p^2(\lambda)$) converge en distribución a una distribución normal, como sigue:

3.3. Análisis de Varianza (ANOVA) de un factor

$$\sum_{i=1}^n W_i \sim N(n(p + \lambda), 2n(p + 2\lambda)).$$

Teorema 2: Sea W_1, W_2, \dots, W_n una sucesión de variables aleatorias independientes con funciones de distribución F_1, F_2, \dots, F_n respectivamente, tales que $E(W_i) = \mu_i$ y $Var(W_i) = \sigma_i^2$ para $i = 1, \dots, n$, y $s_n^2 = \sum_{i=1}^n \sigma_i^2$ entonces:

$$S_n^* = \frac{\sum_{i=1}^n (W_i - \mu_i)}{s_n} \xrightarrow{d} Z, \quad (3.2)$$

donde $Z \sim N(0, 1)$, siempre que las F_i 's sean absolutamente continuas con función de densidad f_i , tales que se cumpla la siguiente condición:

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{i=1}^n \int_{|w - \mu_i| > \varepsilon s_n} (w - \mu_i)^2 f_i(w) dw = 0. \quad (3.3)$$

La ecuación 3.3 es conocida como la condición de Lindeberg ([Rohatgi y Saleh, 2001](#)).

Simultáneamente 3.2 permite identificar que la distribución de la suma de variables aleatorias independientes no idénticamente distribuidas, converge en distribución a una distribución normal de la siguiente manera:

$$\sum_{i=1}^n W_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right).$$

3.3. Análisis de Varianza (ANOVA) de un factor

El material de este apartado se presenta para comparar las bondades y limitaciones de esta metodología en comparación con las características del MANOVA.

[Montgomery \(2004\)](#) e [Infante y Zarate \(2012\)](#) establecen que el análisis de varianza de un solo factor (one-way ANOVA), se usa cuando se desea probar la hipótesis nula de igualdad de efectos de dos o más tratamientos. Para comparar los t tratamientos o niveles diferentes de un solo factor, el modelo estadístico más empleado es el modelo del DECA:

$$Y_{ij} = \mu + \tau_i + e_{ij},$$

3.3. Análisis de Varianza (ANOVA) de un factor

donde:

- Y_{ij} : Variable aleatoria correspondiente a la j -ésima repetición del i -ésimo tratamiento o nivel del factor, $i = 1, \dots, t$, $j = 1, \dots, r$.
- μ : Parámetro de escala común a todos los tratamientos.
- τ_i : Efecto del i -ésimo tratamiento $i = 1, \dots, t$.
- e_{ij} : Error aleatorio.

Se asume que $e_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$, con varianza σ^2 .

Montgomery (2004) y Gutiérrez y de la Vara (2012) indican que el modelo anterior implica que sobre la variable respuesta Y_{ij} actúa una media general y dos fuentes de variabilidad: los tratamientos (factor de estudio) y el error aleatorio (fuentes de variabilidad del experimento, como: mediciones, diferencias entre las unidades experimentales a las que se les aplican los tratamientos, factores ambientales, genéticos, etc.).

El juego de hipótesis a probar en un DECA es:

$$H_0 : \tau_1 = \tau_2 = \dots = \tau_t \quad \text{vs} \quad H_a : \tau_i \neq \tau_j, \text{ para al menos un } i \neq j.$$

Las observaciones obtenidas del experimento son consideradas realizaciones de la variable aleatoria Y_{ij} (a las que se les denotará y_{ij}), y se presentan en la Tabla 3.2.

Tabla 3.2: Datos típicos de un experimento univariado de un solo factor

Tratamiento	Observación				Total	Media
1	y_{11}	y_{12}	\dots	y_{1r}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	\dots	y_{2r}	$y_{2.}$	$\bar{y}_{2.}$
\vdots	\vdots	\vdots	\dots	\vdots	\vdots	\vdots
t	y_{t1}	y_{t2}	\dots	y_{tr}	$y_{t.}$	$\bar{y}_{t.}$
					$y_{..}$	$\bar{y}_{..}$

Donde $y_{i.} = \sum_{j=1}^r y_{ij}$ es el total del tratamiento i , $\bar{y}_{i.} = \frac{1}{r} \sum_{j=1}^r y_{ij}$ es la media de las observaciones del tratamiento i , $y_{..} = \sum_{i=1}^t y_{i.}$ es el gran total de las observaciones y $\bar{y}_{..} = \frac{1}{r} \sum_{i=1}^t \bar{y}_{i.}$ es la media global o promedio de todas las observaciones.

Algunos autores como Hinkelman y Kempthorne (2005), Gutiérrez y de la Vara (2012) y Dean *et al.* (2017) establecen que la metodología del ANOVA se basa en medir la variabilidad para determinar si los efectos de los tratamientos son diferentes; para ello, se comparan dos cálculos independientes: uno que refleja la variabilidad de las observaciones dentro de los grupos (SC_E) y otro que representa la variabilidad entre los grupos a los que se atribuyen los efectos de los tratamientos (SC_T). Al dividir las sumas de cuadrados entre sus respectivos

3.3. Análisis de Varianza (ANOVA) de un factor

grados de libertad se obtienen los cuadrados medios (CM), los cuales estiman la magnitud de su correspondiente fuente de variabilidad.

Para probar la hipótesis nula del DECA balanceado, se emplea la siguiente tabla:

Tabla 3.3: ANOVA correspondiente a un DECA

Fuente de variación FV	Grados de libertad GL	Suma de cuadrados SC	Cuadrado medio CM	$F_{calculada}$
Tratamientos(T)	$t - 1$	$\sum_{i=1}^t \frac{y_{i.}^2}{r} - \frac{y_{..}^2}{tr}$	$\frac{SC_T}{GL_T}$	$\frac{CM_T}{CM_E}$
Error(E)	$t(r - 1)$	$\sum_{i=1}^t \left(\sum_{j=1}^r y_{ij}^2 - \frac{y_{i.}^2}{r} \right)$	$\frac{SC_E}{GL_E}$	
Total	$tr - 1$	$\sum_{i=1}^t \sum_{j=1}^r y_{ij}^2 - \frac{y_{..}^2}{tr}$		

Usando la Tabla 3.3 se tiene la siguiente regla de decisión: rechazar H_0 al nivel α si $F_{calculada} > F_{crítica}$, donde la $F_{crítica}$ corresponde al cuantil $1 - \alpha$ de la distribución F_{v_1, v_2} , donde v_1 corresponde a los grados de libertad del tratamiento (GL_T) y v_2 a los grados de libertad asociados al error (GL_E).

Para que el resultado del ANOVA sea confiable, es necesario que los datos cumplan con algunos supuestos básicos, tales como independencia, normalidad y homogeneidad de varianzas.

3.3.1. Principales supuestos del ANOVA

Dean *et al.* (2017) y Gutiérrez y de la Vara (2012) establecen como principales supuestos:

1. Se debe tener una variable respuesta continua y un factor categórico con dos o más niveles.
2. Las observaciones deben ser estocásticamente independientes. Se puede mantener razonablemente si las unidades experimentales se eligen aleatoriamente y la medición se hace separadamente para cada unidad.
3. Los Y_{ij} deben tener una distribución aproximadamente normal con homogeneidad de varianzas entre los niveles de factores (es decir, cada tratamiento tendrá una distribución $N(\mu_i, \sigma^2)$).

3.3. Análisis de Varianza (ANOVA) de un factor

Pruebas para verificar normalidad y homocedasticidad

Para verificar el cumplimiento de los supuestos de normalidad y homocedasticidad se cuentan con numerosas pruebas. Para el caso de normalidad se encuentran las pruebas de Kolmogorov-Smirnov, Anderson-Darling, Lilliefors, de rango estudentizado, Shapiro-Wilk, etc. [Romo et al. \(2010\)](#) realizaron un estudio comparativo de diferentes pruebas para normalidad, y una conclusión fue que la prueba de Shapiro-Wilk fue la más potente y más sensible a alternativas asimétricas, por lo que será la prueba que se empleará en este estudio.

Prueba de normalidad: Shapiro-Wilk

Sea X_1, X_2, \dots, X_n una muestra aleatoria de una distribución desconocida $F(x)$, se desea probar la hipótesis nula de que $F(x)$ es una función de distribución normal con parámetros μ y σ^2 desconocidos. El estadístico de prueba es:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

donde $x_{(i)}$ es la i -ésima estadística de orden, obtenida de ordenar la muestra aleatoria de menor a mayor: $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$.

$$a' = (a_1, \dots, a_n) = \frac{m'V^{-1}}{(m'V^{-1}V^{-1}m)^{1/2}},$$

del cual $m' = (m_1, m_2, \dots, m_n)$ es el vector de esperanzas de las estadísticas de orden de una muestra aleatoria de variables normal estándar, y $V = (v_{ij})$ es la matriz de covarianzas de los estadísticos de orden.

Shapiro y Wilk construyeron tablas de valores de a' para distintos tamaños de muestra n así como tablas de valores críticos de la prueba $W_{n,\alpha}$, $\alpha \in [0, 1]$.

Se rechaza la hipótesis nula si $W < W_{n,\alpha}$.

Para probar homogeneidad de varianzas existen varias pruebas, algunas de las más empleadas son las de Bartlett, F-Max de Hartley, Levene, prueba de Levene modificada por Brown y Forsythe; ésta última es menos sensible a la falta de normalidad que la prueba de Bartlett y los resultados obtenidos por [Brown y Forsythe \(1974\)](#) muestran que al emplear la media o la mediana se tienen potencias muy semejantes cuando los datos tienen distribución normal, pero cuando los datos siguen una distribución asimétrica (χ_4^2) la mediana consigue mejorar la potencia de la prueba de Levene.

Prueba de homogeneidad de varianzas: Levene modificada por Brown y Forsythe

Consiste en contrastar la hipótesis nula de que los t tratamientos tienen la misma varianza,

3.3. Análisis de Varianza (ANOVA) de un factor

para lo cual se realiza un análisis de varianza de un factor sobre el valor absoluto de las desviaciones de los valores muestrales respecto a la mediana (\tilde{y}_i).

El estadístico de prueba es:

$$L_{med} = \frac{\frac{r \sum_{i=1}^t (\bar{w}_{1i.} - \bar{w}_{1..})^2}{t-1}}{\frac{\sum_{i=1}^t \sum_{j=1}^r (w_{1ij} - \bar{w}_{1i.})^2}{t(r-1)}},$$

donde $w_{1ij} = |y_{ij} - \tilde{y}_i|$, $\bar{w}_{1i.} = \frac{1}{r} \sum_{j=1}^r w_{1ij}$ y $\bar{w}_{1..} = \frac{1}{rt} \sum_{i=1}^t \sum_{j=1}^r w_{1ij}$.

La regla de decisión es rechazar H_0 si $L_{med} > F_{t(r-1)}^{t-1}(\alpha)$.

Aunque ligeras desviaciones de los supuestos de normalidad y homogeneidad de varianzas no afectan el ANOVA, se recomienda realizar la transformación de los datos con el fin de obtener una distribución aproximadamente normal y/o estabilizar la varianza.

Dean *et al.* (2017), Osborne (2010), Hinkelman y Kempthorne (2005) y Gutiérrez y de la Vara (2012) establecen algunas técnicas de transformación de los datos, por ejemplo: transformación mediante la raíz cuadrada, transformación logarítmica, transformación angular, método de Box-Cox, entre otros.

Método de Box-Cox

La metodología propuesta por Box y Cox (1964) consiste en una familia de funciones de potencia para la variable de respuesta, en la cual se estima un valor λ para elevar la variable hasta encontrar desviaciones estándar más cercanas a cero respecto a la media de los residuales, lo cual sirve para corregir sesgos en la distribución de errores, heterogeneidad de varianzas y mejorar correlación entre las variables. Sin embargo, no corrige todos los fallos en los supuestos, por ejemplo, datos que contienen outliers no pueden ser ajustados propiamente para cumplir el supuesto de normalidad.

Las transformaciones originales de Box y Cox están dadas por:

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \log y & \text{si } \lambda = 0 \end{cases}$$

3.3.2. Procedimientos de comparaciones múltiples

Si como resultado de un ANOVA se rechaza la hipótesis nula, se procederá a identificar cuales tratamientos presentan efectos diferentes entre sí, así como determinar cuál es el más conveniente de ellos. Para ello se cuenta con metodologías que permiten la comparación por pares de medias de tratamientos (también conocidos como métodos de comparaciones múltiples o *post-hoc*).

De manera general, el nuevo juego de hipótesis es:

$$H_0 : \tau_i = \tau_{i'} (i \neq i') \quad \text{vs} \quad H_a : \tau_i \neq \tau_{i'}.$$

Para t tratamientos, el total de comparaciones por pares de medias corresponde a la solución de la fórmula $\frac{t(t-1)}{2}$ (Gutiérrez y de la Vara, 2012); y H_0 se rechazará si cualquiera de los estadísticos de prueba es menor al valor absoluto de la correspondiente diferencia entre las medias muestrales $|\bar{Y}_i - \bar{Y}_{i'}|$.

Montgomery (2004) y Hinkelmann y Kempthorne (2005) mencionan que los principales métodos para las comparaciones de medias por pares son:

- Diferencia Mínima Significativa de Fisher (DMS o LSD por las siglas del inglés *Least Significant Difference*)
- Tukey
- Rango múltiple de Duncan
- Dunnett (DMS de Dunnett)
- SNK (Student, Neyman y Köel)

Para analizar todos los contrastes de medias de interés sin que se infle el error tipo I (detección de diferencias que no existen), el método de Sheffé es un método adecuado (Gutiérrez y de la Vara, 2012).

De manera general, cuando sólo se hacen comparaciones por parejas, el método de Tukey conduce a límites de confianza más estrechos que el método de Scheffé, por lo cual el método de Tukey encontrará más diferencias significativas, siendo en este caso el método preferido por la mayoría. En cambio cuando los contrastes son más complicados que la diferencia de medias, es el método de Scheffé el que tiene límites de confianza más estrechos (Dean *et al.*, 2017). Dado lo anterior, la presente investigación empleará el método de Tukey.

3.3. Análisis de Varianza (ANOVA) de un factor

Método de Tukey

Método conservador (es decir, indica que es probable que el nivel de confianza real sea mayor que el nivel de confianza que se muestra, o dicho de otra manera, el método podría aceptar menos diferencias significativas de las que realmente existen), también conocido como método DHS (Diferencia Honestamente Significativa) de Tukey, o procedimiento de rango Studentizado ([Hinkelman y Kempthorne, 2005](#)). Cuando el diseño es balanceado, el nivel de significación global será exactamente α , y si es desbalanceado, es al menos α . Cuando los tamaños de muestras son diferentes, este método es conocido como *Tukey-Kramer*.

Para realizar las comparaciones por pares de medias, se recomienda primero encontrar los promedios de cada tratamiento, para posteriormente ordenarlos en forma no creciente y proceder con la comparación por pares de medias.

Si el diseño es balanceado, el estadístico de prueba es:

$$DHS = q_{\alpha}(t, v_2) \sqrt{\frac{CM_E}{r}},$$

donde v_2 son los grados de libertad asociados al error, $q_{\alpha}(t, v_2)$ son puntos porcentuales α de la distribución del rango Studentizado que son superiores de t , y se obtienen de las tablas de Tukey, CM_E es el cuadrado medio del error (obtenido de la tabla de ANOVA) y r es el número de repeticiones.

Se rechaza H_0 al nivel α si: $|\bar{Y}_i - \bar{Y}_{i'}| > DHS$.

Si no se rechaza H_0 , al par de medias comparadas se les pone una misma letra (comenzando desde la a), y en caso contrario, diferentes letras. Para finalizar las letras enteramente contenidas se eliminan. Éste método alfabético permite comparar de una manera sencilla cuales pares de medias son diferentes.

3.3.3. Potencia de la prueba

Debido a que una de las formas de medir la eficiencia de una metodología estadística propuesta es a través de un estudio de potencia, a continuación se presenta material al respecto.

La decisión de rechazar o no la hipótesis nula de igualdad en los efectos de los tratamientos no está exenta de errores, y es por ello que es importante conocer la probabilidad de rechazar H_0 cuando ésta es falsa. A dicha probabilidad se le conoce como potencia de la prueba de hipótesis, y de alguna manera caracteriza la sensibilidad de la prueba para detectar diferencias significativas ([Menéndez y Jamed, 2006](#)).

3.3. Análisis de Varianza (ANOVA) de un factor

En el ANOVA cuando se llega a una conclusión a partir del estadístico de prueba F , existe la probabilidad de cometer dos tipos de error: Error tipo I, que consiste en rechazar H_0 cuando H_0 es cierta, y el Error tipo II, que es no rechazar H_0 cuando H_0 es falsa. Las probabilidades de dichos errores se denotan como α y β .

- $\alpha = P[\text{Error Tipo I}] = P[\text{rechazar } H_0 | H_0]$
- $\beta = P[\text{Error Tipo II}] = P[\text{no rechazar } H_0 | H_a]$

Lo ideal sería que los dos tipos de error tuvieran probabilidades pequeñas, sin embargo esto generalmente no es posible, ya que cuando se minimiza la probabilidad de alguno de los errores, aumenta la probabilidad de cometer el otro error. Es por ello, que en la práctica se fija α y se trata de minimizar β .

La potencia de la prueba se denota como $1 - \beta$. [Hinkelman y Kempthorne \(2005\)](#) mencionan que β depende de:

1. Nivel de significancia (α);
2. Grados de libertad ($t - 1$ y $t(r - 1)$);
3. Magnitud del efecto de tratamiento (δ o diferencia real entre las medias poblacionales);
4. Varianza (σ^2);
5. Tamaño de muestra (r).

Para evaluar la potencia se deben generar datos en situaciones donde se asume que H_0 es falsa, y posteriormente calcular su estadístico de prueba para proceder a concluir respecto al juego de hipótesis.

Para calcular la potencia del ANOVA de una vía, se emplea el estadístico F , el cual asumiendo que H_0 no es cierta, tiene una distribución ji-cuadrada no central, con $\lambda = \frac{1}{2}\mu'\mu$ como parámetro de no centralidad.

Distribución del estadístico de prueba

En la Tabla 3.3 se establece que el estadístico $F_{\text{calculada}}$ se estima como el cociente $\frac{CM_T}{CM_E}$, y que éstos a su vez se calculan a partir de la suma de cuadrados del tratamiento y del error, respectivamente.

3.3. Análisis de Varianza (ANOVA) de un factor

Sean $SC_T \sim \chi_{v_1}^2(\lambda)$ y $SC_E \sim \chi_{v_2}^2(0)$ dos variables aleatorias independientes, entonces la variable aleatoria $F_{\text{calculada}} = \frac{\frac{SC_T}{v_1}}{\frac{SC_E}{v_2}}$ se distribuye como una F -Snedecor no central, con parámetros v_1 , v_2 y λ como el parámetro de no centralidad o pnc; y se escribe que $F_{\text{calculada}} \sim F(v_1, v_2, \lambda)$.

En la práctica μ es un parámetro desconocido, por lo que considerando como base la ecuación (1) de [Liu y Raudenbush \(2004\)](#) y la teoría de la función de potencia de la prueba del ANOVA de [Patnaik \(1949\)](#), λ se calcula a partir de las funciones de observaciones de la siguiente manera:

$$\hat{\lambda} = \frac{r \sum_{i=1}^t (\hat{\mu}_i - \hat{\mu})^2}{\hat{\sigma}_e^2}.$$

En la cual:

r : Tamaño de muestra, es decir, número de repeticiones por tratamiento.

$\hat{\mu}_i$: Media estimada de cada tratamiento.

$\hat{\mu}$: Media general estimada.

$\hat{\sigma}_e^2$: Valor de CM_E .

Para calcular la potencia analítica [Clarke \(2008\)](#) y [Seltman \(2018\)](#) establecen probar el juego de hipótesis de H_0 contra H_a , mediante una prueba F a un nivel de significancia α , la cual, está sujeta por el criterio de aceptación: Rechazar H_0 si $F_{\text{calculada}} > F_{v_2}^{v_1}(\alpha)$. Donde $P_{H_0}[F_{\text{calculada}} > F_{v_2}^{v_1}(\alpha)|H_0] \leq \alpha$, y representa la probabilidad de rechazar H_0 asumiendo que H_0 es cierta, y se rechazará la hipótesis nula si el estadístico F se encuentra a la derecha del punto crítico $F_{v_2}^{v_1}(\alpha)$.

Entonces la potencia, que es la $Pr_{H_a}[F_{\text{calculada}} > F_{v_2}^{v_1}(\alpha)]$, corresponderá al área bajo la curva que está a la derecha del punto crítico bajo la distribución $F_{v_2}^{v_1}(\lambda)$, lo cual se muestra en la Figura 3.1.

3.3. Análisis de Varianza (ANOVA) de un factor

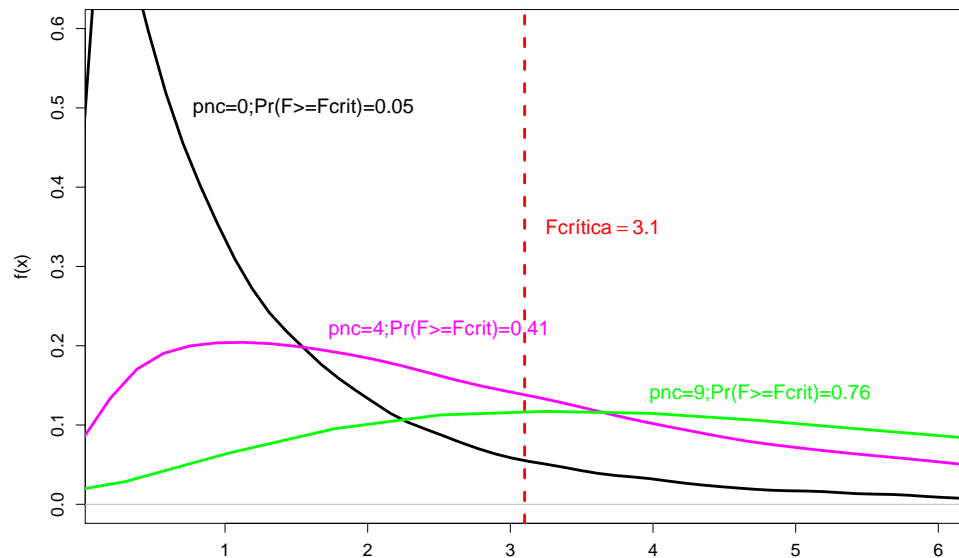


Figura 3.1: Distribución del estadístico F para distintos valores del λ .

Capítulo 4

Metodología propuesta

4.1. Propuesta metodológica

La propuesta plantea reducir a un escalar cada vector de variables respuesta con el fin de obtener datos que permitan ser analizados mediante un ANOVA y posteriormente mediante algún método post hoc hacer la comparación de medias.

Se propone que cada vector p -variado \mathbf{Y}_{ij} sea transformado con la función cuadrática de la norma euclídeana de la siguiente manera: $X_{ij} = \mathbf{Y}'_{ij}\mathbf{Y}_{ij} = \sum_{k=1}^p Y_{ijk}^2$, donde Y_{ijk} es una observación de la k -ésima variable ($k = 1, \dots, p$) del vector \mathbf{Y}_{ij} , con $i = 1, \dots, t$ y $j = 1, \dots, r$, donde t es el número de tratamientos y r es el número de repeticiones. La idea anterior se plasma en la Tabla 4.1.

Tabla 4.1: Uso de la norma euclídeana en el conjunto de datos

Tratamiento	Variable 1	...	Variable p	Cuadrado de la norma
1	y_{111}	...	y_{11p}	$y_{111}^2 + \dots + y_{11p}^2 = \ \mathbf{y}_{11}\ ^2 = x_{11}$
	y_{121}	...	y_{12p}	$y_{121}^2 + \dots + y_{12p}^2 = \ \mathbf{y}_{12}\ ^2 = x_{12}$
	\vdots	\vdots	\vdots	\vdots
\vdots	y_{1r1}	...	y_{1rp}	$y_{1r1}^2 + \dots + y_{1rp}^2 = \ \mathbf{y}_{1r}\ ^2 = x_{1r}$
	\vdots	\vdots	\vdots	\vdots
t	y_{t11}	...	y_{t1p}	$y_{t11}^2 + \dots + y_{t1p}^2 = \ \mathbf{y}_{t1}\ ^2 = x_{t1}$
	y_{t21}	...	y_{t2p}	$y_{t21}^2 + \dots + y_{t2p}^2 = \ \mathbf{y}_{t2}\ ^2 = x_{t2}$
	\vdots	\vdots	\vdots	\vdots
	y_{tr1}	...	y_{trp}	$y_{tr1}^2 + \dots + y_{trp}^2 = \ \mathbf{y}_{tr}\ ^2 = x_{tr}$

Al realizar la transformación de los datos con respuestas múltiples a un valor escalar, se

4.1. Propuesta metodológica

generan variables aleatorias X_{ij} , las cuales podrán ser analizadas mediante un ANOVA si se cumple con sus supuestos.

Ahora bien, note que X_{ij} es una suma de variables aleatorias, por lo que se puede aplicar el Teorema Central de Límites para la suma de variables aleatorias (Teorema 2).

Por lo anterior siempre que se cumpla la condición de Lindeberg y las varianzas de las X_{ij} sean homogéneas para todo i , estas variables pueden ser analizadas mediante un ANOVA.

Aunque potencialmente existen muchas distribuciones multivariadas, tales que el vector p -variado \mathbf{Y}_{ij} bajo la transformación pueda cumplir con las condiciones anteriores, se explorará el caso más típico.

Caso 1. Variables normales no correlacionadas (resultado teórico)

Puesto que el MANOVA funciona bajo los supuestos de normalidad multivariada y homogeneidad de varianzas y covarianzas, se partirá de estos supuestos para aplicar la propuesta metodológica.

Asumiendo que \mathbf{Y}_{ij} sigue una distribución $N_p(\boldsymbol{\mu}_i, \sigma^2 \mathbf{I})$, entonces la variable aleatoria generada de la función cuadrada de la norma X_{ij} tiene una distribución ji-cuadrada no central, con media $p + \lambda$ y varianza $2(p + 2\lambda)$, para $p > 0$ que especifica los grados de libertad y con parámetro de no centralidad ($\lambda = \frac{\boldsymbol{\mu}'_i \boldsymbol{\mu}_i}{2\sigma^2}$):

$$\frac{\mathbf{Y}'_{ij} \mathbf{Y}_{ij}}{\sigma^2} \sim \chi_p^2(\lambda).$$

Note que aun cuando se asume la misma varianza de \mathbf{Y}_{ij} para todo i , las X_{ij} tienen varianzas diferentes, debido a que éstas dependen de la media de cada tratamiento. Bajo este escenario, las variables aleatorias X_{ij} no sigue una distribución normal y tampoco tienen varianzas homogéneas. Por lo anterior, si las X_{ij} son usadas en el ANOVA los resultados no serán válidos.

Sin embargo, a continuación se demuestra que X_{ij} converge a la normal, por lo que es factible usar la propuesta metodológica en este caso.

Demostración de convergencia para el caso de variables no correlacionadas:

Sea $X_{ij} = \sum_{k=1}^p Y_{ijk}^2$, $E(Y_{ijk}^2) = \mu_k$ y $Var(Y_{ijk}^2) = a_k^2$, note que para este caso se puede asumir que existe una constante a , tal que: $|a_k| \leq a$, ya que a_k depende de μ_k y σ^2 , por lo que es razonable este supuesto para el Caso 1. Por otra parte, $\sum_{k=1}^p a_k^2 \rightarrow \infty, p \rightarrow \infty$ dado que las varianzas son siempre positivas.

4.1. Propuesta metodológica

Bajo estas consideraciones se puede ver que:

$$\begin{aligned}
 \frac{1}{s_p^2} \sum_{k=1}^p \int_{|Y_{ijk}^2| > \varepsilon s_p} (Y_{ijk}^2 - \mu_k)^2 f_Y(y_{ijk}^2) dy &\leq \frac{a^2}{s_p^2} \sum_{k=1}^p P(|Y_{ijk}^2 - \mu_k| > \varepsilon s_p) \\
 &\leq \frac{a^2}{s_p^2} \sum_{k=1}^p \frac{\text{Var}(Y_{ijk}^2)}{\varepsilon^2 s_p^2} \\
 &\leq \frac{a^2}{\varepsilon^2 s_p^2} \rightarrow 0, p \rightarrow \infty.
 \end{aligned}$$

Por lo tanto se cumple la condición de Lindeberg y entonces:

$$X_{ij} \sim N \left(\sum_{i=1}^t \left(p + \frac{\boldsymbol{\mu}'_i \boldsymbol{\mu}_i}{2\sigma^2} \right), 2 \sum_{i=1}^t \left(p + \frac{\boldsymbol{\mu}'_i \boldsymbol{\mu}_i}{\sigma^2} \right) \right),$$

si p es suficientemente grande.

Además, note que bajo la hipótesis nula las varianzas son homogéneas, y por lo tanto, si p es suficientemente grande las X_{ij} cumplen con los supuestos del ANOVA y puede usarse la propuesta metodológica.

Con el objeto de conocer que tan grande debe ser p para lograr una buena aproximación se hizo una comparación de la distribución real de los datos transformados ($\chi_p^2(\lambda)$) con una aproximación asintótica a la distribución normal empleando el código presentado en el Anexo A.4.1. En las Figuras 4.1, 4.2, 4.3 y 4.4 se puede apreciar que la distribución de los datos transformados presenta una muy buena aproximación asintótica a la normal desde que $p = 2$, por lo que el método propuesto teóricamente puede generar variables que pueden ser empleadas directamente en un ANOVA.

4.1. Propuesta metodológica

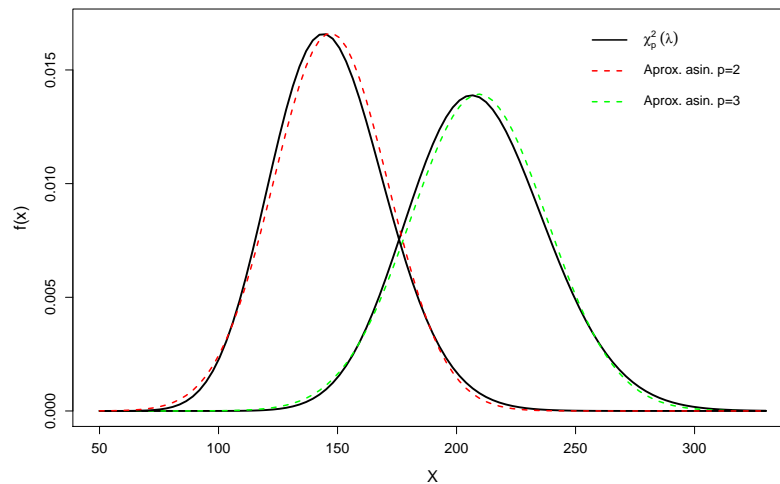


Figura 4.1: Distribución $\chi_p^2(\lambda)$ vs aproximación asintótica a la normal. $t = 3, p = 2, 3$

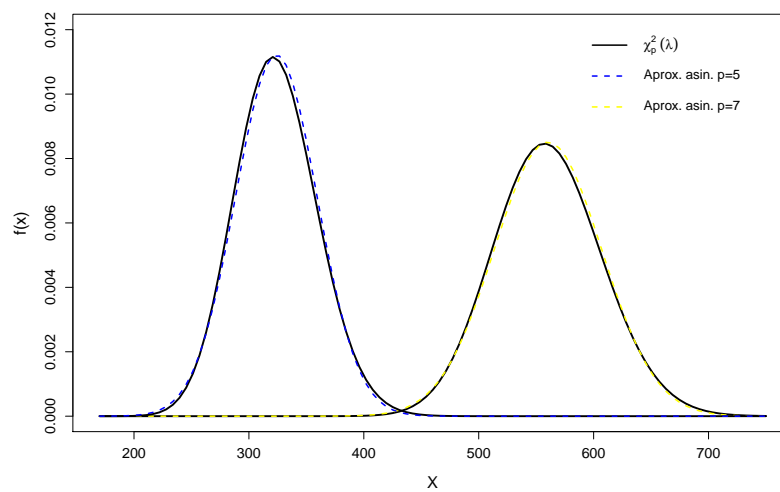


Figura 4.2: Distribución $\chi_p^2(\lambda)$ vs aproximación asintótica a la normal. $t = 3, p = 5, 7$

4.1. Propuesta metodológica

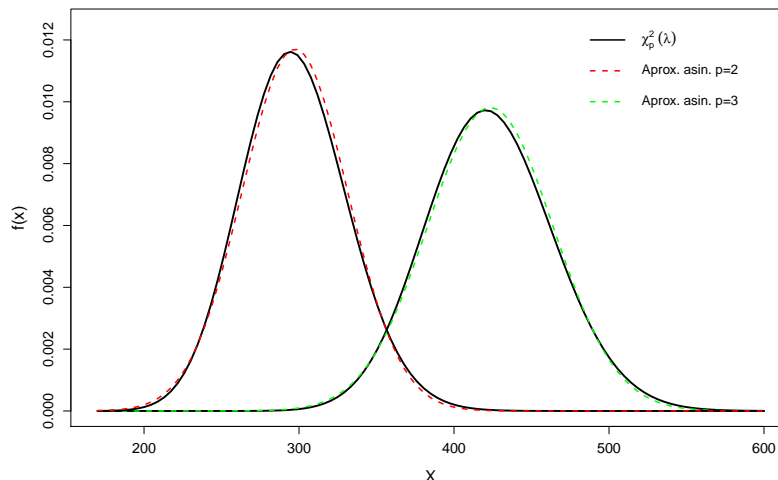


Figura 4.3: Distribución $\chi_p^2(\lambda)$ vs aproximación asintótica a la normal. $t = 7$, $p = 2, 3$

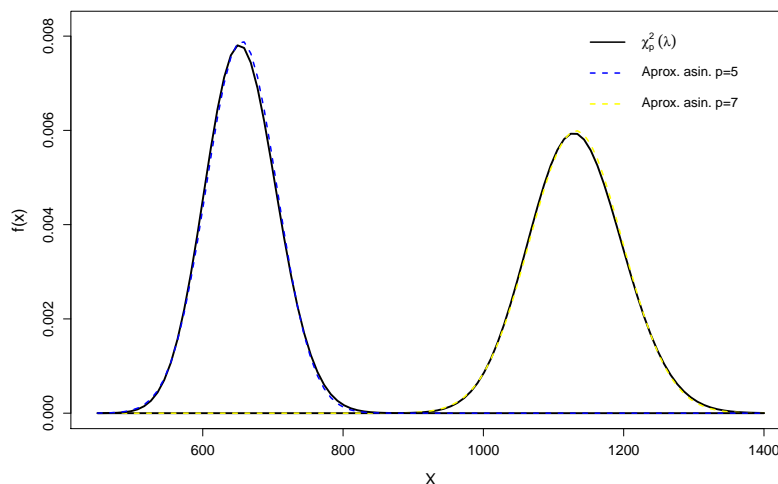


Figura 4.4: Distribución $\chi_p^2(\lambda)$ vs aproximación asintótica a la normal. $t = 7$, $p = 5, 7$

Caso 2. Variables normales correlacionadas

Si el vector \mathbf{Y}_{ij} tiene distribución $N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$, entonces no se conoce la distribución de la variable aleatoria $X_{ij} = \mathbf{Y}'_{ij} \mathbf{Y}_{ij} = \sum_{k=1}^p Y_{ijk}^2$ y queda abierta la incógnita de cuál será su distribución. Sin embargo, se hizo un estudio de simulación para investigar el desempeño de la propuesta contra el MANOVA, usando variables normales correlacionadas.

4.2. Evaluación del desempeño de la propuesta

Como muestran las Figuras 4.1, 4.2, 4.3 y 4.4, el número de variables no requiere ser tan grande para obtener resultados de convergencia satisfactorios, por lo que a continuación se presenta un estudio de simulación para evaluar el desempeño de la propuesta con un número de variables pequeño y con tamaños de muestra que usualmente aparecen en la práctica para el caso en el que un DECA genera respuestas múltiples. .

4.2.1. Preliminares

Para evaluar la potencia del ANOVA usando los datos transformados con el método propuesto, se realizó un estudio de simulación Monte Carlo, empleando 2 000 réplicas (B) y un nivel de significancia de 0.05, bajo el supuesto de que las respuestas múltiples provienen de una distribución normal multivariada.

Los parámetros que pueden modificarse en la simulación de datos con distribución normal multivariada fueron:

- Vectores de medias de cada tratamiento: El vector principal será $\boldsymbol{\mu}_1$ y a partir de él se generarán las diferencias entre vectores de medias.
- Matriz de varianzas: Se consideraron matrices de variables no correlacionadas ($\boldsymbol{\Sigma} = \mathbf{I}_p$).
- Número de tratamientos: $t = 3, 5, 7$
- Número de variables: $p = 3, 5, 7$
- Número de repeticiones por tratamiento: $r = 4, 8, 12, 16$.

Para evaluar el desempeño de la metodología propuesta a través de la estimación de su potencia, se empleó el siguiente algoritmo:

1. Seleccionar p, t, r .
2. Generar una muestra aleatoria con distribución normal multivariada para cada tratamiento, con vector de medias $\boldsymbol{\mu}_i \in [1, 10]$ y una matriz de varianzas $\boldsymbol{\Sigma} = \mathbf{I}_p$ común en todos los tratamientos.
3. Obtener las variables transformadas X_{ij} a partir de la muestra.
4. Realizar un ANOVA sobre las variables X_{ij} y obtener los valores de los grados de libertad (GL_T y GL_E) y del cuadrado medio del error (CM_E).

4.2. Evaluación del desempeño de la propuesta

5. Calcular las medias de cada tratamiento como: $p + \hat{\mu}'_i \hat{\mu}_i$
6. Calcular $\hat{\mu}$ como el promedio de las medias de los tratamientos.
7. Calcular el parámetro de no centralidad: $\hat{\lambda} = \frac{r \sum_{i=1}^t (\hat{\mu}_i - \hat{\mu})^2}{\hat{\sigma}_e^2}$ con los valores obtenidos en los pasos 4-6.
8. Obtener pf como la función de distribución de la F no central ($F_{v_2}^{v_1}(\hat{\lambda})$).
9. Obtener la $F_{crítica}$ como el cuantil $1 - \alpha$ de la distribución F central.
10. Estimar la potencia del ANOVA como $1 - pf$ evaluado en la $F_{crítica}$.
11. Repetir B veces los pasos 2 - 10.
12. Estimar la potencia como la proporción de veces que se rechaza H_0 en la simulación.

El estudio de simulación se realizó con el software R-Studio, versión 1.1.383, en un equipo HP con procesador AMD A8-4555 M APU con 8.00 GB de memoria RAM, sistema operativo de 64 bits y windows 8.1.

El programa donde se implementa el algoritmo anterior para estimar la potencia se encuentra en el Anexo A.4.

Para la determinación de las potencias se previeron los ‘peores escenarios’, es decir, se contemplaron los números mínimos de: repeticiones, variables y tratamientos para la generación de datos; así como: que cada variable presentara valores muy cercanos entre sí y que variaran todos los vectores de medias por tratamiento.

Potencias estimadas para diferente número de vectores de medias

Acorde a lo establecido previamente, el vector μ_1 es el vector que se fija, y es a partir del cual los otros vectores de medias cambian. Las Figuras 4.5, 4.6 y 4.7 fueron simuladas con 1000 repeticiones Monte Carlo, y en ellos se puede apreciar que las potencias estimadas respecto a las diferencias de medias, son menores conforme aumenta la cantidad de vectores de medias que cambian. Por ejemplo, la línea amarilla corresponde a la diferencia que hay entre un vector de medias y todos los demás vectores quedan fijos como vectores nulos, la línea azul es cuando dos vectores de medias cambian y los demás quedan fijos, y así sucesivamente hasta llegar a la línea color magenta, la cual indica que todos los vectores de medias cambian.

4.2. Evaluación del desempeño de la propuesta

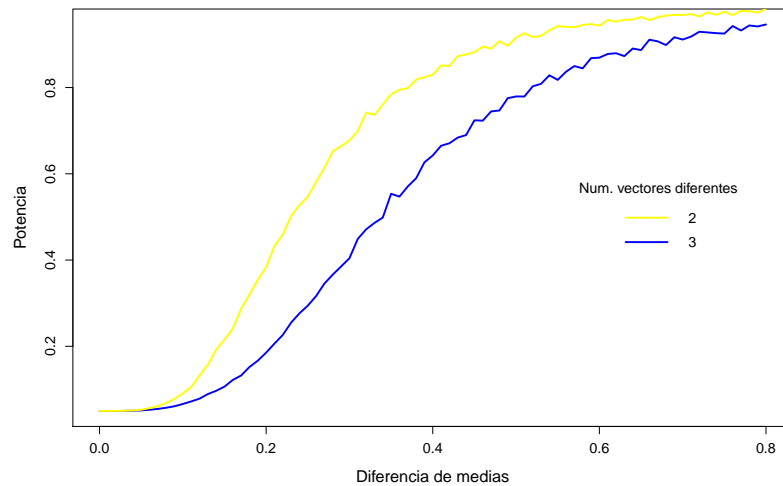


Figura 4.5: Potencias estimadas cambiando la cantidad de vectores de medias en $t = 3$, $p = 3$, $r = 4$.

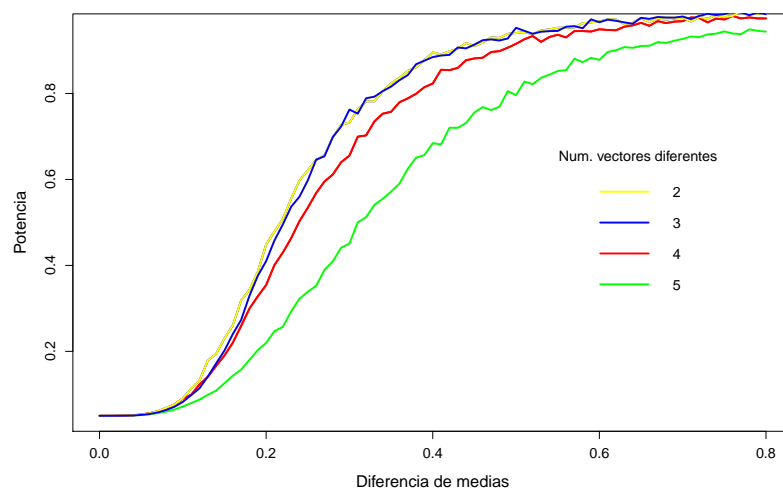


Figura 4.6: Potencias estimadas cambiando la cantidad de vectores de medias en $t = 5$, $p = 3$, $r = 4$.

4.2. Evaluación del desempeño de la propuesta

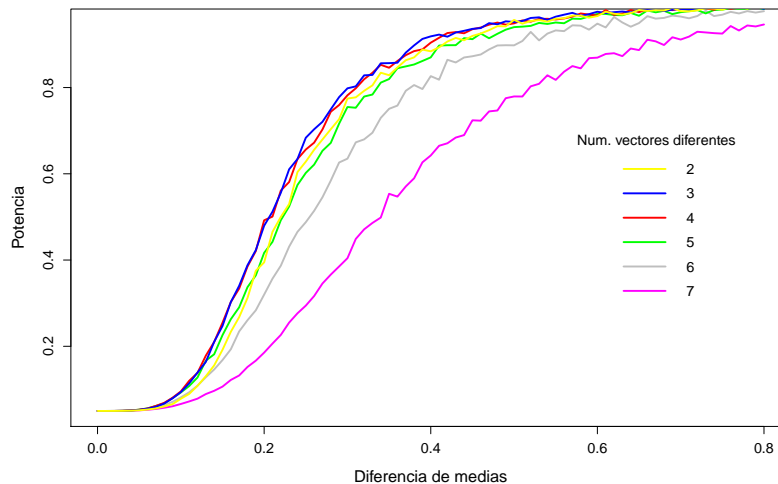


Figura 4.7: Potencias estimadas cambian la cantidad de vectores de medias en $t = 7$, $p = 3$, $r = 4$.

Dado lo anterior, para la simulación se emplearán variaciones entre todos los vectores de medias, con el objetivo de identificar condiciones en las cuales las potencias son bajas, y con esto asegurar potencias aceptables en la práctica.

En las figuras subsecuentes, se observará que se grafican la potencia contra la diferencia máxima de medias, éste último corresponde al porcentaje que tendrán de diferencia entre sí los vectores de medias, es decir, el valor de 0.2 que se observa en el eje de las abscisas corresponde a un 20 %.

4.2.2. Potencia de la prueba de igualdad de efectos de tratamientos usando la metodología propuesta

En todas las figuras presentadas en esta sección se emplea la metodología propuesta con datos no correlacionados, se fija el número de tratamientos, y se observa el efecto que hay en la potencia cuando hay un incremento de variables y de repeticiones por tratamiento. El código empleado para realizar esta simulación se encuentra en el Anexo [A.4.2](#).

■ 3 Tratamientos

4.2. Evaluación del desempeño de la propuesta

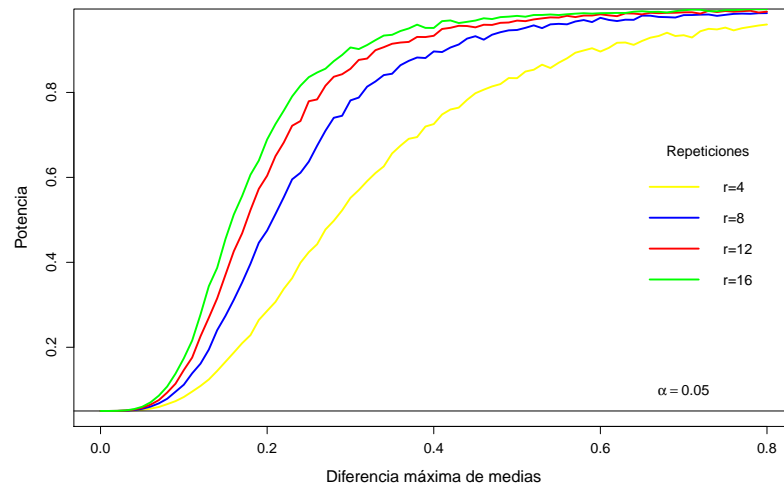


Figura 4.8: Potencias estimadas para $p = 3$ y $t = 3$

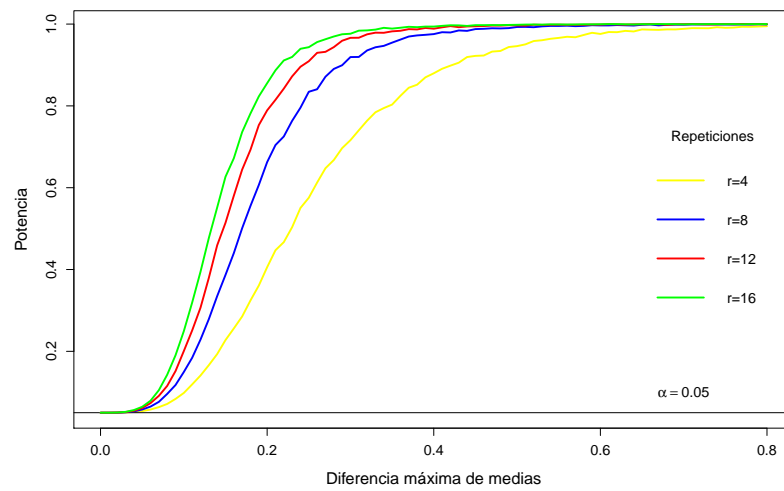


Figura 4.9: Potencias estimadas para $p = 5$ y $t = 3$

4.2. Evaluación del desempeño de la propuesta

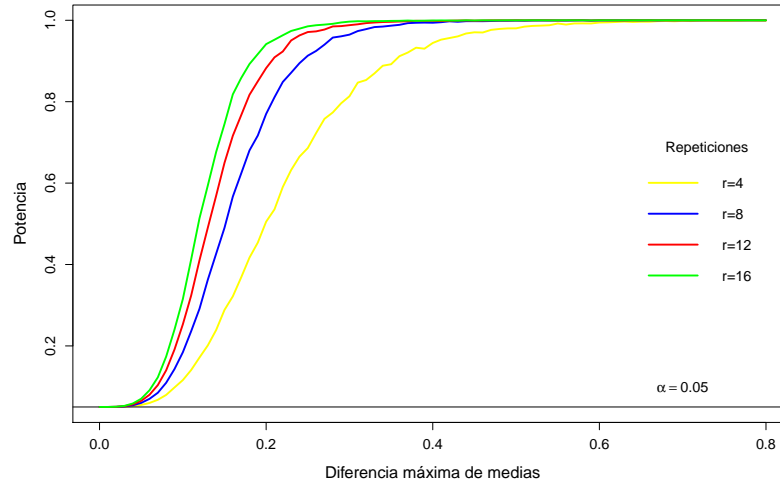


Figura 4.10: Potencias estimadas para $p = 7$ y $t = 3$

Para los casos estudiados en las Figuras 4.8, 4.9 y 4.10 se observa un buen comportamiento de la potencia de la prueba al aumentar el tamaño de muestra, resultado que es el esperado de acuerdo a la teoría estadística ((Casella, 2008), (Hinkelman y Kempthorne, 2005)). Cabe hacer notar que al aumentar el número de variables (p), la potencia de la prueba también aumenta, lo que indiza que el resultado asintótico de convergencia de X_{ij} a la distribución normal funciona con tamaños de muestras pequeños. Los resultados muestran que la potencia de la prueba aumenta al aumentar las diferencias máximas entre los vectores, lo cual también es el comportamiento esperado (Casella, 2008).

También se observa que cuando se tienen cuatro repeticiones para lograr una potencia mayor a 0.8 se requiere que la diferencia entre los vectores de medias sea al menos del 40%. Para los casos de 8, 12 y 16 repeticiones, dicha potencia se obtiene cuando existe entre un 20% y 30% de diferencia entre los vectores de medias.

■ 5 Tratamientos

4.2. Evaluación del desempeño de la propuesta

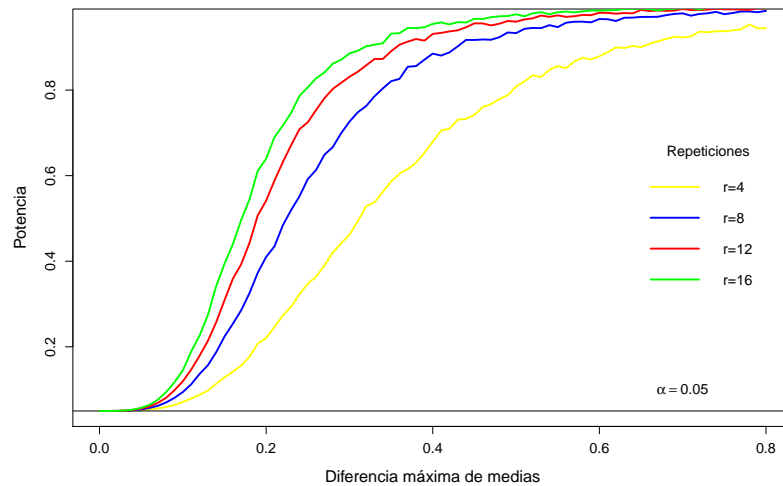


Figura 4.11: Potencias estimadas para $p = 3$ y $t = 5$

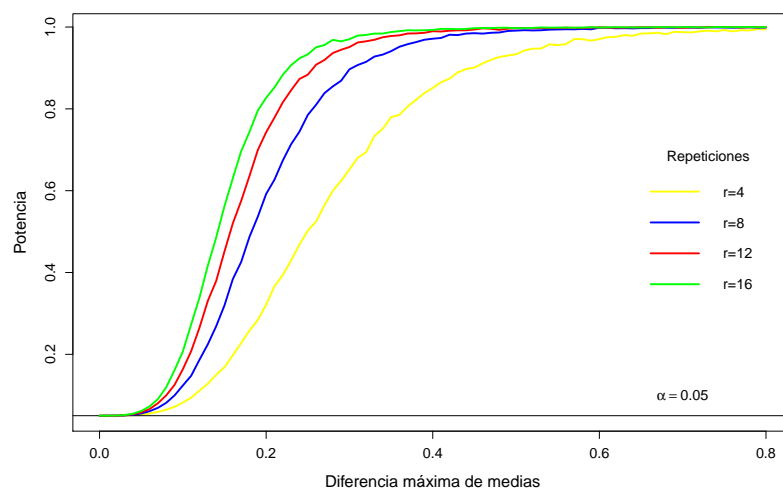


Figura 4.12: Potencias estimadas para $p = 5$ y $t = 5$

4.2. Evaluación del desempeño de la propuesta

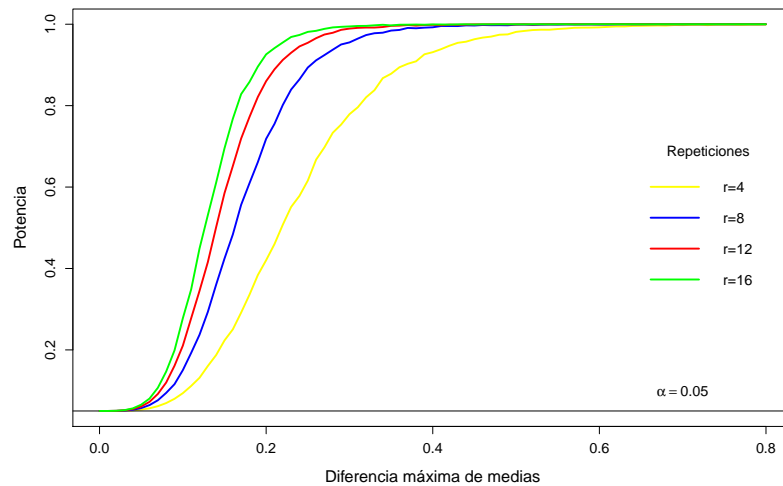


Figura 4.13: Potencias estimadas para $p = 7$ y $t = 5$

■ 7 Tratamientos

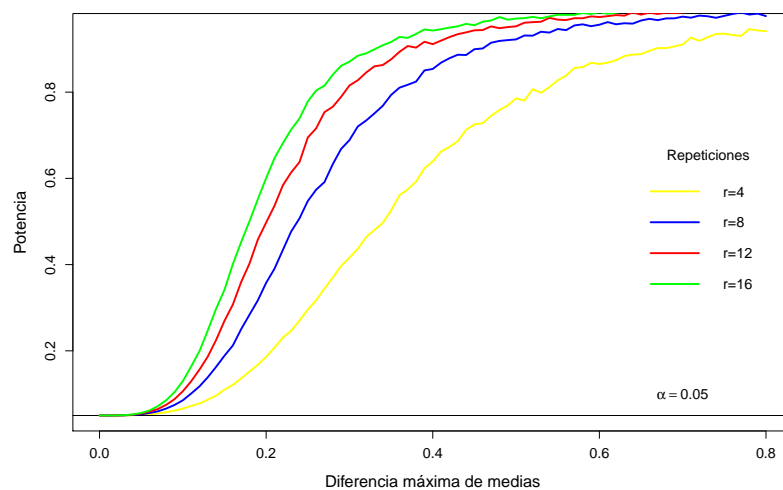


Figura 4.14: Potencias estimadas para $p = 3$ y $t = 7$

4.2. Evaluación del desempeño de la propuesta

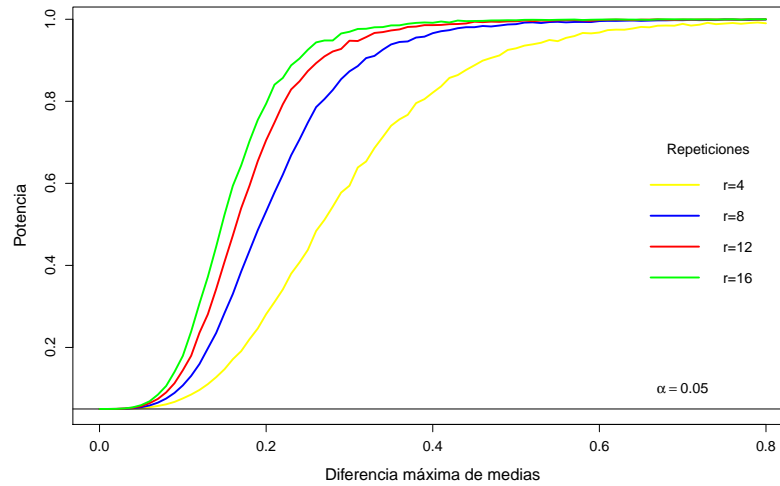


Figura 4.15: Potencias estimadas para $p = 5$ y $t = 7$

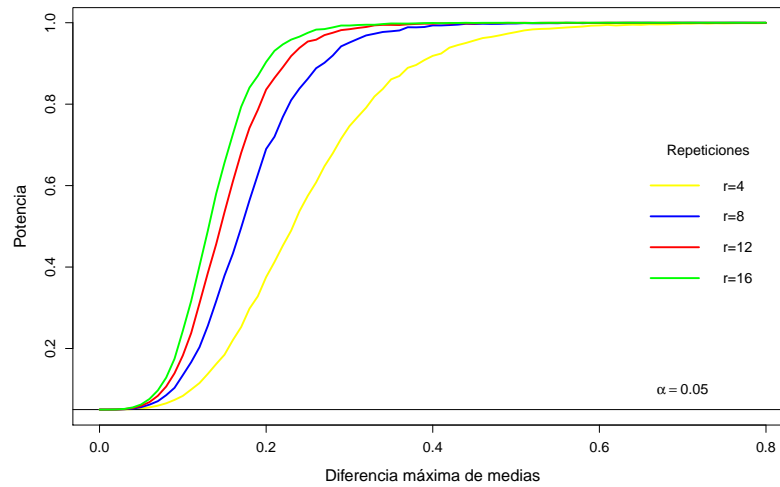


Figura 4.16: Potencias estimadas para $p = 7$ y $t = 7$

Se puede ver que prácticamente no hay diferencia entre las figuras observadas en 3, 5 o 7 tratamientos, por lo que no hay efecto evidente del número de tratamientos en la potencia estimada.

4.2.3. Potencias del MANOVA vs ANOVA usando la metodología propuesta

Para evaluar si la potencia del ANOVA empleando la metodología propuesta es semejante a la del MANOVA, se presentan las siguientes figuras, las cuales corresponden a las potencias estimadas para el caso de tres tratamientos. El código usado se encuentra en el Anexo A.4.3.

4.2. Evaluación del desempeño de la propuesta

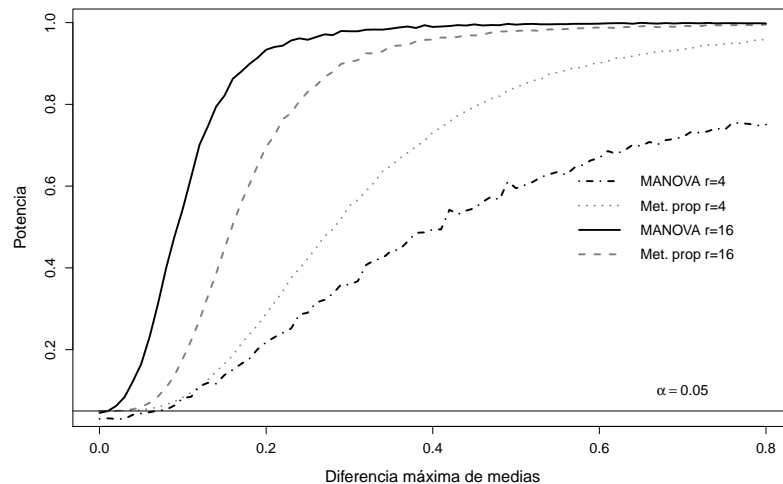


Figura 4.17: Potencias estimadas del MANOVA y del ANOVA empleando la metodología propuesta. $p = 3$ y $t = 3$

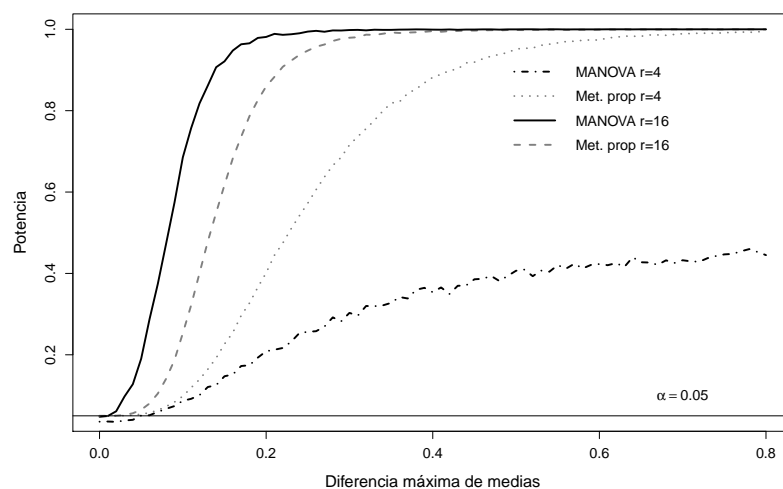


Figura 4.18: Potencias estimadas del MANOVA y del ANOVA empleando la metodología propuesta. $p = 5$ y $t = 3$

4.3. Comparaciones de resultados del MANOVA con el ANOVA usando la metodología propuesta

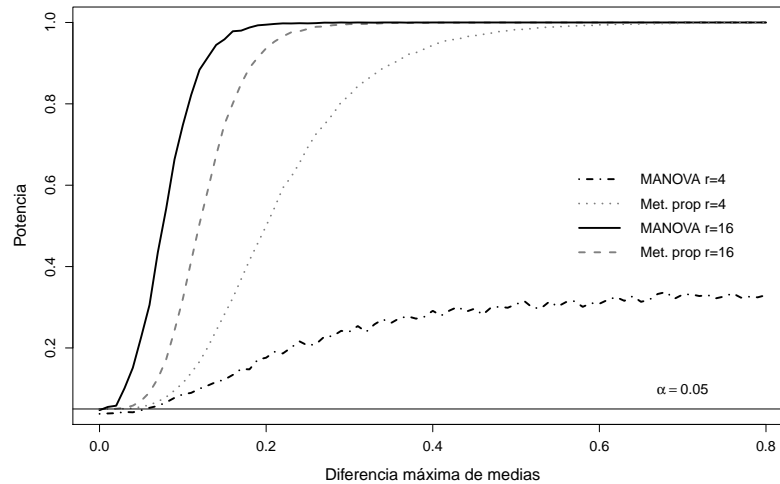


Figura 4.19: Potencias estimadas del MANOVA y del ANOVA empleando la metodología propuesta. $p = 7$ y $t = 3$

Con cuatro repeticiones se puede observar en todos los casos que la potencia del MANOVA es menor que la del ANOVA con el uso de los datos transformados, sin embargo, con 16 repeticiones por tratamiento, la potencia del MANOVA es mayor.

4.3. Comparaciones de resultados del MANOVA con el ANOVA usando la metodología propuesta

Como se estableció previamente, tanto el MANOVA como el ANOVA son metodologías que sólo permiten aceptar o rechazar la hipótesis nula de igualdad de los efectos de los tratamientos. Parte de lo que se pretende en esta sección es demostrar que si con el MANOVA el efecto de algún tratamiento es significativo, cuando se transforman los datos mediante la metodología propuesta, el resultado con el ANOVA también será el mismo.

Si el resultado con el MANOVA es significativo, un procedimiento de seguimiento para identificar las diferencias entre los efectos de los tratamientos es la *Conjetura generalizada de Tukey* (la cual sólo se ha utilizado con tres y cuatro tratamientos), mientras que para un ANOVA un procedimiento de seguimiento es el *HSD de Tukey*. Otro objetivo de esta investigación es mostrar que los resultados al aplicar estas metodologías son los mismos en términos de separación de tratamientos.

Para cumplir con dichos objetivos, se simularon datos con distribución normal multivariada (empleando el código del Anexo [A.4.4](#)) cambiando los siguientes parámetros de interés:

- μ

4.3. Comparaciones de resultados del MANOVA con el ANOVA usando la metodología propuesta

- Distancia entre las medias de los elementos en μ_1 . Para ello se consideran números aleatorios obtenidos de la distribución uniforme mediante la función *runif* en \mathbf{R} , de tal manera que los componentes de μ_1 estén contenidos en alguno de los siguientes intervalos:
 - μ_1 con cada uno de sus componentes en el intervalo $[1,10]$. Para el caso de $p = 3$ $\mu_1 = [2.9, 8.3, 5.7]$, y para $p = 7$ se consideraron los valores para $p = 3$ junto con otros valores aleatorios, con lo cual $\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$.
 - μ_1 con cada uno de sus componentes en el intervalo $[5,25]$. Cuando $p = 3$ $\mu_1 = [21.6, 5.9, 14.1]$, y para $p = 7$ $\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$.
 - μ_1 con cada uno de sus componentes en el intervalo $[50,100]$. Cuando $p = 3$ $\mu_1 = [90.3, 57.6, 82.3]$, y para $p = 7$ $\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$.
- Diferencia entre los vectores de medias. Se consideran dos diferencias entre las μ_i 's.
 - Pequeñas diferencias. Consisten en diferencias contenidas entre el 1% y el 20% respecto a μ_1 , con lo cual para $t = 3$ $\mu_2 = 1.1\mu_1$ (diferencia del 10% respecto a μ_1) y $\mu_3 = 1.2\mu_1$ (diferencia del 20% respecto a μ_1); cuando $t = 7$, $\mu_2 = 0.99\mu_1$, $\mu_3 = 1.10\mu_1$, $\mu_4 = 0.80\mu_1$, $\mu_5 = 1.01\mu_1$, $\mu_6 = 0.90\mu_1$, $\mu_7 = 1.20\mu_1$. Dado lo anterior, se pronostica que con el HSD de Tukey el acomodo de los tratamientos sea: para $t = 3$: T3(μ_3), T2(μ_2), T1(μ_1), y para $t = 7$: T7, T3, T5, T1, T2, T6, T4, y que dependiendo de la naturaleza de los datos, sean las diferencias que se encuentren entre los efectos de los tratamientos.
 - Grandes diferencias. Las cuales contemplan porcentajes muy altos entre las diferencias entre los vectores de medias, que van desde el 50% al 300% respecto a μ_1 . Entonces para $t = 3$ $\mu_2 = 3\mu_1$ y $\mu_3 = 2\mu_1$; y para $t = 7$, $\mu_2 = 2.5\mu_1$, $\mu_3 = 0.5\mu_1$, $\mu_4 = 2.0\mu_1$, $\mu_5 = 0.7\mu_1$, $\mu_6 = 3.0\mu_1$, $\mu_7 = 1.5\mu_1$, y por lo tanto se espera que el acomodo de los tratamientos mediante HSD de Tukey sea: $t = 3$: T2, T3, T1, y para $t = 7$: T6, T2, T4, T7, T1, T5, T3.
- Σ .
 - Variables no correlacionadas (**Caso 1**): En este caso $\Sigma = \sigma\mathbf{I}_p$ considerando $\sigma=1$ ó 10.
 - Variables correlacionadas (**Caso 2**): Para generar matrices de varianzas y covarianzas que sean aleatorias y definidas positivas, se empleó la función *genPositiveDefMat* de la librería *clusterGeneration*. Con esta función se generaron dos matrices: una con varianzas en un intervalo de $[1, 2]$ (junto con sus respectivas covarianzas) y otra con varianzas en el intervalo $[8, 12]$ para emular las varianzas en las matrices estudiadas en el caso de variables no correlacionadas. Para comprobar si son definidas positivas se usó la función *is.positive.definite* de la librería *matrixcalc*.
- Número de tratamientos. $t = 3, 7$.
- Número de variables. $p = 3, 7$.

4.3. Comparaciones de resultados del MANOVA con el ANOVA usando la metodología propuesta

- Número de repeticiones: $r = 4, 16$.

Primeramente los datos fueron analizados con un MANOVA, se obtuvo el valor de la aproximación al estadístico $F_{(v_H, v_E)}$ proveniente de la Traza de Pillai y el valor-p. En los casos de tres tratamientos con tres variables y cuatro repeticiones se obtuvieron los intervalos de confianza mediante la conjetura generalizada de Tukey, en los demás casos no se pudo realizar esta determinación debido a la limitante de grados de libertad.

Posteriormente se transformaron los datos mediante la metodología propuesta y éstos fueron analizados con el ANOVA para el modelo del DECA. En este caso se verificó el cumplimiento de las supuestos de normalidad y homocedasticidad mediante las pruebas de Shapiro-Wilk (SW) y Levene (L). Si los datos transformados no cumplieron alguno de los supuestos, se hizo la transformación de Box-Cox.

Si los datos transformados con la metodología de Box-Cox no cumplen con el supuesto de normalidad, el resultado correspondiente a la $F_{v_1}^{v_2}$ del ANOVA se verá de color rojo; si el ANOVA rechaza la hipótesis nula de igualdad de efectos de tratamientos, se procedió a la comparación de medias mediante la prueba de Tukey, pues como establecen [Driscoll \(1996\)](#) y [Salkind \(2010\)](#), éstas pruebas son robustas ante la falta de normalidad.

Para ejemplificar las tablas generadas para realizar las comparaciones entre metodologías, se muestran las Tablas [4.2](#), [4.3](#) y [4.4](#), las cuales sólo corresponden a ejemplos de tres tratamientos, y con ellas se pretende observar el efecto de aumentar el número de variables en los tratamientos o el efecto de tener matrices de varianzas y covarianzas con datos correlacionados.

Las demás tablas que muestran los resultados obtenidos al cambiar los parámetros previos, se encuentran en el Anexo [A.5](#).

4.3. Comparaciones de resultados del MANOVA con el ANOVA usando la metodología propuesta

Tabla 4.2: Comparaciones para el caso $t = 3$, $p = 3$, $r = 4$, pequeñas diferencias entre μ_i^s

μ	Σ	MANOVA:		CG Tukey		Datos trans		ANOVA:		HSD Tukey
		$\sim F_{v_H, v_E}$	$Pr(> F)$	CG Tukey	valor-p	F_{v_1, v_2}	$Pr(> F)$	HSD Tukey		
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		1.6376	0.2009	T1vsT2: [-10.8, 7.4] T1vsT3: [-12.5, 5.7] T2vsT3: [-10.8, 7.4]	0.0201	0.9601	5	0.0347	T3 662.0251 a T2 519.3456 ab T1 398.0759 b	
		2.608	0.05871	T1vsT2: [-13.3, 4.9] T1vsT3: [-17.4, 0.8] T2vsT3: [-13.3, 4.9]	0.1417	0.9437	30.78	9.46e-05	T3 953.9957 a T2 797.7182 b T1 655.4444 c	
		2.6633	0.05491	T1vsT2: [-32.1, -13.8] T1vsT3: [-55.1, -36.8] T2vsT3: [-32.1, -13.8]	0.0318	0.9542	597.2	2.7e-10	T3 338111233 a T2 238289208 b T1 162395826 c	
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		0.3661	0.8898	T1vsT2: [-30.6, 27.2] T1vsT3: [-32.3, 25.5] T2vsT3: [-30.6, 27.2]	0.0654	0.9090	0.705	0.519	No hay diferencia significativa	
		2.1768	0.1002	T1vsT2: [-33.1, 24.7] T1vsT3: [-37.2, 20.6] T2vsT3: [-33.1, 24.7]	0.2397	0.9249	3.237	0.0873	No hay diferencia significativa	
		2.6334	0.05693	T1vsT2: [-51.9, 5.9] T1vsT3: [-74.9, -17.0] T2vsT3: [-51.9, 5.9]	0.0351	0.9517	61.29	5.73e-06	T3 338111233 a T2 238289208 b T1 162395826 c	

1. Cuando las diferencias entre los vectores de medias son de 10 % y 20 % y existen pocas repeticiones, el MANOVA no rechaza igualdad de los efectos de los tratamientos, en cambio usando la metodología propuesta sí se identifican dichas diferencias, igualmente el método HSD de Tukey ordena e identifica correctamente los efectos de los tratamientos.
2. El valor-p tanto del MANOVA como del ANOVA, aumenta conforme aumenta la distancia entre las medias de las variables.
3. La metodología de la conjetura generalizada de Tukey no encontró diferencias entre los efectos de los tratamientos en los casos de pequeñas y medianas distancias entre las medias.
4. En todos los casos estudiados en esta tabla, los datos obtenidos empleando la metodología propuesta cumplen con el supuesto de homocedasticidad, y sólo en algunos casos no cumple con el supuesto de normalidad.

4.3. Comparaciones de resultados del MANOVA con el ANOVA usando la metodología propuesta

Tabla 4.3: Comparaciones para el caso $t = 3$, $p = 7$, $r = 4$, pequeñas diferencias entre μ_i 's

μ	Σ	MANOVA:	Datos trans		ANOVA:	HSD Tukey
		$\sim F_{v_H, v_E}$ $Pr(> F)$	valor-p SW	L	F_{v_1, v_2} $Pr(> F)$	
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		residuals have rank 3<7	0.0558	0.9713	28.14 0.000134	T3 411.3858 a T2 345.1291 b T1 284.7821 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	\mathbf{I}_7	residuals have rank 3<7	0.2872	0.9494	413.5 1.39e-09	T3 2332.021 a T2 1954.931 b T1 1611.166 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		residuals have rank 3<7	0.0013	0.9841	2909 2.24e-13	T3 0.3448276 a T2 0.3448276 b T1 0.3448276 c
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		residuals have rank 3<7	0.1120	0.9728	2.111 0.177	No hay diferencia significativa
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	$10 * \mathbf{I}_7$	residuals have rank 3<7	0.0500	0.9054	32.34 7.78e-05	T3 2229.853 a T2 1866.041 b T1 1535.553 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		residuals have rank 3<7	0.0021	0.9827	282.1 7.61e-09	T3 0.3846154 a T2 0.3846154 b T1 0.3846154 c

1. El MANOVA tiene la particularidad de que sólo puede realizarse si $r - 1 \geq p$, y como puede observarse, en caso contrario no puede calcularse el estadístico de prueba.
2. Para este caso, cuando se tienen grades distancias entre variables, no se obtiene el cumplimiento de la normalidad, pero si el de homocedasticidad. Para las otras distancias entre variables se cumplen ambos supuestos.

4.3. Comparaciones de resultados del MANOVA con el ANOVA usando la metodología propuesta

Tabla 4.4: Comparaciones para el caso $t = 3$, $p = 3$, $r = 4$, grandes diferencias entre μ_i^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	CG Tukey	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		2.6656 0.05476	T1vsT2: [-38.9, -28.6] T1vsT3: [-22.0, -11.7] T2vsT3: [11.1, 22.6]	0.4984 0.2948	183.4 5.09e-08	T2 1010.8554 a T3 456.4263 b T1 121.5772 c
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$\begin{pmatrix} 1.5 & -0.6 & -1.1 \\ -0.6 & 1.7 & 0.3 \\ -1.1 & 0.3 & 1.2 \end{pmatrix}$	2.6666 0.05469	T1vsT2: [-88.3, -78.0] T1vsT3: [-46.7, -36.4] T2vsT3: [35.8, 47.3]	0.1310 0.4363	12514 3.17e-16	T2 6235.4150 a T3 2758.6860 b T1 682.3171 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		2.6667 0.05469	T1vsT2: [-465.5, -455.2] T1vsT3: [-235.3, -225.0] T2vsT3: [224.4, 235.9]	0.4394 0.5424	654865 <2e-16	T2 164033.52 a T3 72867.50 b T1 18191.76 c
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		2.659 0.0552	T1vsT2: [-47.9, -19.6] T1vsT3: [-31.0, -2.7] T2vsT3: [2.7, 31.0]	0.7305 0.2447	166.9 7.71e-08	T2 1010.4536 a T3 467.5679 b T1 144.2623 c
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$\begin{pmatrix} 8.6 & 3.0 & -8.3 \\ 3.0 & 10.7 & 0.1 \\ -8.3 & 0.1 & 10.1 \end{pmatrix}$	2.6664 0.0547	T1vsT2: [-97.3, -69.0] T1vsT3: [-55.7, -27.4] T2vsT3: [25.8, 57.3]	0.5144 0.2234	742.1 1.02e-10	T2 6350.252 a T3 2846.653 b T1 743.415 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		2.6667 0.05469	T1vsT2: [-474.5, -446.2] T1vsT3: [-244.3, -216.0] T2vsT3: [214.4, 24.9]	0.3327 0.4508	32511 <2e-16	T2 164116.56 a T3 72934.27 b T1 18242.26 c

1. Se vuelve a observar que cuando se tienen cuatro repeticiones, aún cuando son grandes diferencias entre los vectores de medias, el MANOVA no identifica diferencia entre los efectos de los tratamientos.
2. La CG de Tukey identificó las diferencias entre los vectores de medias acorde a los establecido en lo parámetros de simulación.
3. En todos estos casos los resultados del ANOVA y HSD de Tukey son acordes a lo pronosticado.

Capítulo 5

Ejemplo de aplicación

Para ilustrar la aplicación del método propuesto en un modelo balanceado se emplea la base de datos *RootStock*, la cual puede ser descargada de la librería *heplots* de **R**.

Esta base de datos proviene de un experimento del crecimiento de 8 árboles de manzana provenientes de seis diferentes rizomas ($r = 8$ por cada rizoma-R), los cuales se compararon mediante cuatro medidas (variables dependientes):

- *girth4*=Circunferencia del tronco a los 4 años (mm x 100).
- *ext4*=Crecimiento de la extensión a los 4 años (m).
- *girth15*=Circunferencia del tronco a los 15 años (mm x 100).
- *weight15*=Peso del árbol sobre el suelo a los 15 años (lb x 1000).

El código para realizar los cálculos necesarios para aplicar la metodología propuesta se encuentran en el Anexo [A.4.5](#).

Lo primero a realizar es el MANOVA, y posteriormente realizar la verificación de cada supuesto.

No hay métodos confiables para verificar independencia en estos casos, pero se asegura mediante la aleatorización en el diseño experimental.

Para las demás suposiciones, en la Tabla [5.1](#) se enlistan algunas funciones en el programa R.

Cabe mencionar que la función *mvn* con el argumento *mardia*, sólo acepta el cumplimiento del supuesto de normalidad multivariada si se cumplen tanto la asimetría como la curtosis, otra ventaja de esta función es que identifica los datos atípicos.

5. Ejemplo de aplicación

Tabla 5.1: Funciones en el programa R para verificar supuestos del MANOVA

Supuesto	Librería	Función	Argumentos
Normalidad multivariada	MVN	mvn	mvnTest= 'mardia'
Homogeneidad de matrices de varianzas y covarianzas	heplots	boxM	No aplica
Datos atípicos	MVN	mvn	multivariateOutlierMethod= 'adj' showOutliers=TRUE (Opcional) showNewData=TRUE (Opcional)

El criterio que se empleará para aceptar o rechazar H_0 en todas las pruebas es el valor-p, recordando que la hipótesis nula varía dependiendo de cada prueba, y que se rechaza H_0 si $\text{valor-p} < \alpha$, $\alpha = 0.05$.

Valor-p del MANOVA: 1.983e-07

Lo anterior indica que el MANOVA rechaza H_0 a un nivel $\alpha = 0.05$, y por lo tanto los efectos de los tratamientos son diferentes.

Para verificar si este resultado es confiables, en la Tabla 5.2 se muestran los resultados del análisis de normalidad multivariada y posteriormente la M de Box.

Tabla 5.2: Resultados del análisis de normalidad multivariada

	valor-p					
	R1	R2	R3	R4	R5	R6
Asimetría de Mardia	0.2647	0.7344	0.1538	0.5559	0.6616	0.1144
Cutosis de Mardia	0.4005	0.2276	0.3733	0.2311	0.1994	0.4306
Cumple NMV	Si	Si	Si	Si	Si	Si

Valor-p (M de Box): 0.711

Dado que los $\text{valor-p} > 0.05$ se cumplen ambos supuestos, por lo que se puede establecer que el resultado del MANOVA es confiable.

Para identificar qué pares de medias son distintas, se procede a realizar contrastes empleando la corrección de Bonferroni. Empleando la fórmula para calcular en numero de contrastes $t(t-1)/2$, se tienen 15 contrastes por realizar, y el valor ajustado es $0.05/15 = 0.0033$. Por lo tanto, se considerará que hay diferencia en los efectos de los tratamientos si el valor-

5. Ejemplo de aplicación

$p \leq 0.0033$. La Tabla 5.4 muestra los resultados de los valores-p obtenidos tras realizar los contrastes de cada par de tratamientos y si es que con la corrección de Bonferroni se tienen diferencias entre los tratamientos o no, dicha tabla se mostrará posteriormente para poder comparar la identificación de diferencias en los efectos de los tratamientos mediante el método de contrastes y el HSD de Tukey.

En la Tabla 5.3 se muestran los resultados del análisis realizado con la metodología propuesta.

Tabla 5.3: Resultado del análisis posterior al uso de la metodología propuesta

valor-p			
ANOVA	SW	L	HSD Tukey
			2 33.45173 a
			3 31.20606 ab
			5 28.23406 ab
0.000167	0.4625	0.7134	4 26.17140 abc
			1 25.12841 bc
			6 19.96439 c

Como puede observarse, tanto el MANOVA como el ANOVA rechazan la hipótesis nula de igualdad de los efectos de los tratamientos, y al comparar los resultados obtenidos con los contrastes y del HSD de Tukey, nos podemos percatar de primera instancia que la segunda metodología tiene la ventaja de poder ordenar los efectos para poder identificar al tratamiento más conveniente. En la Tabla 5.4 se pueden comparar las diferencias entre tratamientos que encuentra cada metodología.

En la Tabla 5.4 se aprecia que la comparación de los efectos de los tratamientos con la metodología de contrastes aplicados en datos con respuestas múltiples son muy semejantes a los obtenidos con la metodología HSD de Tukey aplicado posterior a la transformación de los datos con la metodología propuesta.

5. Ejemplo de aplicación

Tabla 5.4: Contrastes de los rizomas

Comparacion Rizomas	valor-p	Diferencias por contrastes	Diferencias por HSD Tukey
1-2	9.507e-05	Si	Si
1-3	0.0014	Si	No
1-4	0.3063	No	No
1-5	0.0028	Si	Si
1-6	0.1252	No	No
2-3	0.1471	No	No
2-4	0.0059	No	No
2-5	0.5560	No	No
2-6	0.0038	No	Si
3-4	0.0313	No	No
3-5	0.0149	No	No
3-6	0.0006	Si	Si
4-5	0.0020	Si	No
4-6	0.0022	Si	Si
5-6	0.0197	No	Si

Capítulo 6

Discusión de resultados y conclusiones

6.1. Potencia de la prueba

La propuesta de transformar los vectores generados en un DECA a escalares y probar la hipótesis de igualdad de efectos de tratamientos con el ANOVA presenta un desempeño satisfactorio en términos de potencia. De la Figura 4.8 a la Figura 4.16 se observa un buen comportamiento de la potencia de la prueba al aumentar el tamaño de muestra, resultado que es el esperado de acuerdo a la teoría estadística (Casella, 2008) (Hinkelman y Kempthorne, 2005). Cabe hacer notar que al aumentar el número de variables (p), la potencia de la prueba también aumenta, lo que indica que el resultado asintótico de convergencia de X_{ij} a la distribución normal funciona con tamaños de muestras pequeños. Finalmente, los resultados muestran que al aumentar las diferencias máximas entre los vectores, la potencia de la prueba aumenta, lo cual también es el comportamiento esperado (Casella, 2008).

En las Figuras 4.1, 4.2, 4.3 y 4.4 se puede ver que cuando se tienen cuatro repeticiones el ANOVA posterior al uso de la metodología propuesta presenta mayor potencia que el MANOVA, por lo que dicha metodología es una buena opción para el análisis de respuestas múltiples en este caso. Para el caso de 16 repeticiones el MANOVA presenta una mayor potencia, sin embargo, el método propuesto tiene potencias muy semejantes al MANOVA y considerando las limitantes que presenta el MANOVA, la propuesta de este trabajo es una buena alternativa.

6.2. Comparaciones de resultados del MANOVA con el ANOVA

Los resultados de la comparación del desempeño del MANOVA y la metodología propuesta en la prueba de la hipótesis de igualdad de efectos de tratamientos son también satisfactorios en el caso de un DECA con respuestas múltiples.

La propuesta metodológica en la mayoría de los casos detecta diferencias más pequeñas que el MANOVA (Tabla 4.2). La Tabla 4.3 evidencia que en algunos casos el MANOVA no puede analizar los datos debido a limitaciones propias de la metodología (cuando $r - 1 > p$), pues en dicho caso la matriz no es de rango completo (Strang, 2006), mientras que la propuesta metodológica no presenta dicho problema y el análisis en este caso puede ser llevado a cabo. La Tabla 4.4 muestra un desempeño satisfactorio de la propuesta metodológica cuando se tienen observaciones correlacionadas. Cuando las diferencias máximas entre los vectores de medias es mayor al 100% el MANOVA y la propuesta metodológica rechazan la hipótesis de igualdad de efectos de tratamientos, aunque el nivel de significancia es mayor en el MANOVA (4.4). Se observó que en algunos casos los datos transformados no cumplieron el supuesto de normalidad y generalmente cumplieron con el supuesto de homocedasticidad. Frecuentemente el problema del no cumplimiento de los supuestos se solucionó con la transformación de Box-Cox. No se observan cambios en la significancia del MANOVA y del ANOVA al cambiar la matriz de varianzas y covarianzas. En todos los casos estudiados, la metodología de HSD de Tukey presentó resultados acordes a los parámetros de simulación. Este estudio consistió en estudiar los parámetros establecidos en la metodología, generando resultados con simulación en 46 cuadros.

Por lo anterior, la metodología propuesta es una buena alternativa para probar la hipótesis de igualdad de tratamientos con respuestas múltiples generadas por un DECA con variables normales correlacionadas. Potencialmente la propuesta metodológica puede funcionar para modelos más complejos y otro tipo de distribuciones multivariadas.

6.3. Conclusiones

La metodología propuesta cumple teóricamente con los supuestos del ANOVA para el caso de respuestas múltiples generadas con un DECA suponiendo variables normales no correlacionadas. Dicha metodología es una buena alternativa para probar la hipótesis de igualdad de efectos de tratamientos para respuestas múltiples provenientes de un DECA balanceado, incluso en el caso de muestras pequeñas y variables normales correlacionadas. Así mismo, es satisfactorio el desempeño de la transformación propuesta en la metodología HSD de Tukey. Los datos transformados tienen problemas para ajustarse al supuesto de normalidad cuando el número de variables (p) es relativamente pequeño, lo que generalmente se ve solucionado con la transformación de Box-Cox.

Referencias

- Anderson, T. (2003). An introduction to multivariate statistical analysis. 7–21.
- Box, G. y Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B*, 26, 2, 211–252.
- Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika*, 36, 3–4, 317–346.
- Brown, M. B. y Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69, 346, 364–367.
- Casella, G. (2008). Statistical design. 5, 57–59, 80–85.
- Clarke, B. (2008). Linear models: the theory and application of analysis of variance. 55–61.
- Dattalo, P. (2013). Analysis of multiple dependent variables. 22–65.
- Dean, A., Voss, D., Draguljić, D. *et al.* (2017). Design and analysis of experiments. 41–49, 81–92.
- DeGroot, M. y Schervish, M. (2012). Probability and statistics. 360–369.
- Driscoll, W. C. (1996). Robustness of the ANOVA and Tukey-Kramer statistical tests. *Computers & Industrial Engineering*, 31, 1-2, 265–268.
- Gutiérrez, H. y de la Vara, R. (2012). Análisis y diseño de experimentos. *Mc Graw Hill*, 52–75.
- Hair, J., Anderson, R., Tatham, R. y Black, W. (1999). Análisis multivarnte de datos. 345–397.
- Hinkelmann, K. y Kempthorne, O. (2005). Design and analysis of experiments. 182–196, 224–228.
- Infante, S. y Zarate, G. (2012). Métodos Estadísticos: Un Enfoque Interdisciplinario. *La Gaya Ciencia-COLPOS-SAGARPA*, 392–426.
- Joenssen, D. W. y Vogel, J. (2014). A power study of goodness-of-fit tests for multivariate normality implemented in R. *Journal of Statistical Computation and Simulation*, 84, 5, 1055–1078.
- Johnson, R. A. y Wichern, D. (2007). Applied multivariate statistical analysis. *Pearson Prentice Hall*, 149–195, 273–307.
- Liu, X. y Raudenbush, S. (2004). A note on the noncentrality parameter and effect size estimates for the F test in ANOVA. *Journal of Educational and Behavioral Statistics*, 29, 2, 251–255.

Referencias

- Mardia, K. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519–530.
- Mardia, K. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhyā: The Indian Journal of Statistics, Series B*, 36, 2, 115–128.
- Menéndez, E. y Jamed, L. (2006). Una propuesta para el cálculo de la potencia en el ANOVA. *Revista investigación operacional*, 27, 2, 194–205.
- Montgomery, D. (2004). Diseño y análisis de experimentos. *Limusa Wiley*, 60–110.
- Nishiyama, T. y Seo, T. (2008). The Multivariate Tukey-Kramer Multiple Comparison Procedure Among Four Correlated Mean Vectors. *American Journal of Mathematical and Management Sciences*, 28, 1-2, 115–130.
- Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. *Journal of the American Statistical Association*, 69, 348, 894–908.
- Olson, C. L. (1979). Practical considerations in choosing a MANOVA test statistic: A rejoinder to Stevens.
- Osborne, J. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical assessment, research and evaluation*, 15, 12.
- Patnaik, P. (1949). The non-central χ^2 - and F-distribution and their applications. *Biometrika*, 36, 1/2, 202–232.
- Peña, D. (2002). Análisis de Datos Multivariantes. 295–320.
- Pérez, D., González, A., Franco, O., Rubí, M., Ramírez, J., neda A., C. y J.G., A. (2014). Aplicación de métodos multivariados para identificar cultivares sobresalientes de haba para el Estado de México, México. *Revista Mexicana de Ciencias Agrícolas*, 5, 2, 265–278.
- Rencher, A. y Christensen, W. (2012). Methods of Multivariate Analysis. 91–195.
- Rohatgi, V. K. y Saleh, A. M. E. (2001). An introduction to probability and statistics. 298–300.
- Romeu, J. y Ozturk, A. (1993). A comparative study of goodness-of-fit tests for multivariate normality. *Journal of Multivariate Analysis*, 46, 309–334.
- Romo, X., Delgado, R. y Costa, A. (2010). An empirical power comparison of univariate goodness-of-fit tests for normality. *Journal of Statistical Computation and Simulation*, 80, 5, 545–591.
- Salkind, N. J. (2010). Encyclopedia of research design. 3, 1565–1570.
- Seltman, H. (2018). Experimental design and analysis.
- Seo, T., Mano, S. y Fujikoshi, Y. (1994). A generalized Tukey conjecture for multiple comparisons among mean vectors. *Journal of the American Statistical Association*, 89, 426, 676–679. ISSN 01621459.
- Strang, G. (2006). Linear algebra and its applications. 141–143.

Referencias

Timm, N. (2002). Applied multivariate analysis. 118–159, 218–225.

Torres, V., Navarro, J. y Pérez, T. (2003). Modelos estadísticos para el procesamiento de experimentos con mediciones repetidas en la misma unidad experimental. *Revista Cubana de Ciencia Agrícola*, 37, 3, 227–232.

Warne, R. T. (2014). A primer on multivariate analysis of variance (MANOVA) for behavioral scientists. *Practical Assessment, Research & Evaluation*, 19.

Anexos

A.4. Códigos en el programa R

A.4.1. Demostración de convergencia para el caso de variables no correlacionadas

```
S=diag(7) #Cambiar respecto al numero de variables
p=7 #Num. de variables
mu1= c(2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5) #runif(p,1,10)
mu2= mu1*1.1
mu3= mu1*1.2
m1=p+((t(mu1)\ %*\ %mu1)/2)
m2=p+((t(mu2)\ %*\ %mu2)/2)
m3=p+((t(mu3)\ %*\ %mu3)/2) #Agregar o quitar respecto al numero de tratamientos
v1=p+(t(mu1)\ %*\ %mu1)
v2=p+(t(mu2)\ %*\ %mu2)
v3=p+(t(mu3)\ %*\ %mu3)
med=as.numeric(m1+m2+m3)
var=as.numeric(2*(v1+v2+v3))

curve(dchisq(x,p,(med-p)),main="t=3",xlim = c(170,750),ylim = c(0,0.012),
type="l", lty=1, lwd=2.5,xlab="X",ylab="f(x)", cex.lab=1.3, cex.axis=1.2)
curve(dnorm(x,med,sqrt(var)),add = TRUE,col="yellow", type="l", pch=2,
lty=2, lwd=2)
```

A.4.2. Caso 1. Variables no correlacionadas

El siguiente código permite obtener los datos transformados mediante la metodología propuesta a partir de datos aleatorios multivariados. Este código debe ejecutarse previo a calcular las potencias de los casos 1 y 2. Para emplearlo es requerido instalar el paquete *MASS*.

Códigos en el programa R

```
dat_t=function(r,mu,S)
{
  library(MASS)
  T1= mvrnorm(n= r, mu= mu[,1], Sigma= S)
  T2= mvrnorm(n= r, mu= mu[,2], Sigma= S)
  T3= mvrnorm(n= r, mu= mu[,3], Sigma= S)
  #Agregar tantos tratamientos como se requiera
  gpo= gl(t,r)
  variables=rbind(T1,T2,T3) #Si se agregaron mas tratamientos, enlistarlos
  return(as.data.frame(cbind(gpo,rowSums(variables[,]^2))))
}
```

Para poder generar las figuras de potencia se emplea el siguiente código, en el cual t , p y r son los parámetros a variar para la simulación.

```
t=3
p=7
r=16

escal=seq(from=0,to=0.8,by=0.01)
lescal=length(escal)
potencias=c()
pot=c()
for(j in 1:lescal)
{
  for(i in 1:2000)
  {
    mu1=runif(p,0,10)
    mu=cbind(mu1*0,mu1*escal[j],mu1*escal[j])
    S=diag(p)
    datos=dat_t(r,mu,S)
    gpo=as.factor(datos[,1])
    anava=summary(aov(datos[,2]~gpo))
    CME=anava[[1]][, 'Mean Sq'][2]
    m1_est=(t(mu[,1])%*%mu[,1])+p
    m2_est=(t(mu[,2])%*%mu[,2])+p
    m3_est=(t(mu[,3])%*%mu[,3])+p
    #Adicional las estimaciones de mu respecto al numero de tratamientos
    m_pob=(m1_est+m2_est+m3_est)/t #Agregar las estimaciones de mu
    scc=r*((m1_est-m_pob)^2)+((m2_est-m_pob)^2)+((m3_est-m_pob)^2)
    nc=scc/CME
    pot[i]=1-pf(qf(1-0.05,anava[[1]][, 'Df'][1],anava[[1]][, 'Df'][2]),
               anava[[1]][, 'Df'][1],anava[[1]][, 'Df'][2],nc)
  }
}
```

```
potencias[j]=mean(pot)
}

plot(escal, potencias, type="l", col="yellow", lty=1,
      xlab="Diferencia_maxima_de_medias", ylab="Potencia")
abline(h=0.05)
text(0.6, 0.075, expression(alpha==0.05))
```

Para agregar las líneas a la figura se tiene que cambiar el número de repeticiones (r) y volver a correr el código sólo para el ciclo 'for' para posteriormente correr la línea correspondiente al número de repeticiones.

```
lines(escal, potencias, col="blue", lty=1) #r=8
lines(escal, potencias, col="red", lty=1) #r=12
lines(escal, potencias, col="green", lty=1) #r=16
legend("right", legend = c("r=4", "r=8", "r=12", "r=16"), lty = c(1, 1, 1, 1),
      col=c("yellow", "blue", "red", "green"), bty="n", title="Repeticiones_por_tratamiento")
```

A.4.3. Potencias del MANOVA vs ANOVA usando la metodología propuesta

El siguiente código corresponde únicamente a la generación de las potencias estimadas del MANOVA, por lo que para realizar las comparaciones de las potencias de este método contra las del ANOVA posterior al uso de la metodología propuesta, hay que emplear el siguiente código, junto con el código del anexo [A.4.2](#).

```
dat_tM=function(r, mu, S)
{
  library(MASS)
  T1= mvrnorm(n= r, mu= mu[,1], Sigma= S)
  T2= mvrnorm(n= r, mu= mu[,2], Sigma= S)
  T3= mvrnorm(n= r, mu= mu[,3], Sigma= S)
  gpo= gl(t, r)
  variables=rbind(T1, T2, T3)
  return(as.data.frame(cbind(gpo, variables)))
}

phi_an=function(datos, alpha=0.05)
{
  gpo=as.factor(datos[,1])
  variables=as.matrix(datos[, -1])
  library(car)
```



```
man=summary(manova(variables~gpo)) #str(man$stats)
pval=man$stats[11]
return(list(pval=pval,indicadora=ifelse(pval<alpha,1,0)))
}

escal=seq(from=0,to=0.8,by=0.01)
B=5000
power_16=rep(0,length(escal))

for (j in 1:length(escal))
{
  for (k in 1:B)
  {
    mu1=runif(p,0,10)
    mu=cbind(mu1*0,mu1*escal[j],mu1*escal[j])
    S=diag(p)
    datos=dat_tM(r,mu,S)
    power_16[j]=power_16[j]+phi_an(datos)$indicadora
  }
  power_16[j]=power_16[j]/B
}
lines(escal,power_16,lty=1,lwd=2.5)
```

A.4.4. Comparaciones de resultados del MANOVA con el ANOVA

En el siguiente código se requiere que se realicen los cambios necesarios en cada parámetro, dependiendo del caso de interés (ver apartado 4.3). También se adjunta el código empleado para aplicar la conjetura generalizada de Tukey.

```
rm(list=ls())

p=3
r=4
S= genPositiveDefMat("onion",dim=p,rangeVar=c(1,10))$Sigma
mu1= c(2.9, 8.3, 5.7) #round(runif(3,1,10),1)
mu2= mu1*3
mu3= mu1*2
library(MASS)
set.seed(1234)
T1= mvrnorm(n= r, mu= mu1, Sigma= S)
set.seed(1234)
T2= mvrnorm(n= r, mu= mu2, Sigma= S)
set.seed(1234)
T3= mvrnorm(n= r, mu= mu3, Sigma= S)
```

Códigos en el programa R

```
gpo= gl(p,r,labels= c("T1", "T2", "T3"))
variables=rbind(T1,T2,T3)
datos= as.data.frame(cbind(gpo, variables))

#MANOVA
library(car)
an_man=manova( variables ~ gpo, data=datos)
summary(an_man)

#          CONJETURA GENERALIZADA DE TUKEY
#Realizando las comparaciones anteriores con el modelo:
#aÂ´M(est)b +/- z(bÂ´Vb)^1/2 (aÂ´Sa)^1/2
#z=q_(p,k,v)(0.05), v=N-k-1
#z=q_(3,3,32)(0.05)=3.5628

a=c(1,1,1) #vector de 1xp
V=diag(3) #txt

MediasV1= tapply(datos$V1, datos$gpo, mean)
MediasV2= tapply(datos$V2, datos$gpo, mean)
MediasV3= tapply(datos$V3, datos$gpo, mean)
M_est=t(matrix(c(MediasV1,MediasV2,MediasV3),3,3,
                dimnames = list(c("G1", "G2", "G3"),c("V1", "V2", "V3"))))

#Para realizar las comparaciones
nT1=length(T1)
nT2=length(T2)
nT3=length(T3)

ST1=cov(T1)
ST2=cov(T2)
ST3=cov(T3)

####Comparacion T1vsT2
S_12=(1/(nT1+nT2-2))*(((nT1-1)*ST1)+((nT2-1)*ST2))
b12=c(1,-1,0)
aMb_12=t(a)%*%M_est%*%b12
cal12=sqrt((t(b12)%*%V%*%b12))%*%sqrt((t(a)%*%S_12%*%a))
est12=(3.5628)*as.numeric(cal12)
LS_12=as.numeric(aMb_12)+est12
LI_12=as.numeric(aMb_12)-est12

####1-3
S_13=(1/(nT1+nT3-2))*(((nT1-1)*ST1)+((nT3-1)*ST3))
b13=c(1,0,-1)
aMb_13=t(a)%*%M_est%*%b13
cal13=sqrt((t(b13)%*%V%*%b13))%*%sqrt((t(a)%*%S_13%*%a))
```

Códigos en el programa R

```
est13=(3.5628)*as.numeric(cal13)
LS_13=as.numeric(aMb_13)+est13
LI_13=as.numeric(aMb_13)-est13

####2-3
S_23=(1/(nT2+nT3-2))*(((nT2-1)*ST2)+((nT3-1)*ST3))
b23=c(0,1,-1)
aMb_23=t(a)%*%M_est%*%b23
cal23=sqrt((t(b23)%*%V%*%b23))%*%sqrt((t(a)%*%S_23%*%a))
est23=(3.5628)*as.numeric(cal23)
LS_23=as.numeric(aMb_23)+est23
LI_23=as.numeric(aMb_23)-est23
```

```
##METODOLOGIA PROPUESTA
dat_t2=rowSums(variables[,]^2)
ajustado2=aov(dat_t2~gpo)
summary(ajustado2)
shapiro.test(ajustado2$residuals)
leveneTest(dat_t2~gpo)
library(agricolae)
HSD.test(ajustado2, "gpo", console = TRUE)
```

A.4.5. Ejemplo de aplicación

```
library(heplots)
datos=RootStock
variab=as.matrix(RootStock[, -1])
riz=RootStock$rootstock

##MANOVA
library(car)
man_r=manova(variab~riz, data=RootStock)
summary(man_r)

#Normalidad multivariada
library(MVN)
mvn(data=RootStock, subset = "rootstock", mvnTest = "mardia")

#Homogeneidad matrices
boxM(variab~riz, data=RootStock)

#CONTRASTES
summary(manova(variab~riz, subset=RootStock$rootstock%in%("1", "2")))
```

A.5 Tablas de las comparaciones del MANOVA con el ANOVA

```
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("1", "3")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("1", "4")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("1", "5")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("1", "6")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("2", "3")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("2", "4")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("2", "5")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("2", "6")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("3", "4")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("3", "5")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("3", "6")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("4", "5")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("4", "6")))
summary(manova(variab~riz, subset=RootStock$rootstock %in% ("5", "6")))
```

```
#METODOLOGIA PROPUESTA
```

```
dat_t=rowSums(variab[,]^2)
ajustado=aov(dat_t~riz)
summary(ajustado)
shapiro.test(ajustado$residuals)
leveneTest(dat_t~riz)
library(agricolae)
HSD.test(ajustado, "riz", console = TRUE)
```

A.5. Tablas de las comparaciones del MANOVA con el ANOVA

Tabla A.1: Comparaciones para el caso $t = 3$, $p = 3$, $r = 16$, pequeñas diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $P_T(> F)$	Datos trans valor-p SW	ANOVA: F_{v_1, v_2} $P_T(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		7.2273 2.727e-06	0.0039 0.8741	28.78 8.93e-09	T3 9.995981 a T2 9.374889 b T1 8.724961 c
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	\mathbf{I}_3	12.881 2.145e-10	0.0004 0.8755	295.3 <2e-16	T3 0.6060532 a T2 0.6060507 b T1 0.6060469 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		14.592 1.711e-11	0.0045 0.8487	8056 <2e-16	T3 0.8333291 a T2 0.8333282 b T1 0.8333268 c
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		1.2987 0.2662	0.0003 0.9197	1.902 0.161	No hay diferencias significativas
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	$10 * \mathbf{I}_3$	6.1467 2.056e-05	3.453e-05 0.8789	36.45 3.87e-10	T3 0.9988930 a T2 0.9986464 b T1 0.9983146 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		13.955 4.318e-11	0.0037 0.8496	831.8 <2e-16	T3 1.110991 a T2 1.110970 b T1 1.110943 c

1. Aumentar las repeticiones tuvo el efecto de que el MANOVA pudiera identificar la diferencia entre los efectos, y que el ANOVA aumentara su significancia.
2. En la columna SW, se puede observar que los datos transformados mediante la metodología propuesta no cumplieron con el supuesto de normalidad, inclusive algunos datos tampoco cumplieron con dicho supuesto aunque se hayan transformado mediante Box-Cox. Sin embargo, tanto el ANOVA como el HSD de Tukey pudieron identificar las diferencias entre los efectos y con el HSD el orden de los tratamientos fue adecuado.
3. En todos estos casos, se cumple con el supuesto de homocedasticidad.

Tabla A.2: Comparaciones para el caso $t = 3$, $p = 3$, $r = 4$, grandes diferencias entre μ_i^s

μ	Σ	MANOVA:		CG Tukey		Datos trans		ANOVA:		HSD Tukey
		$\sim F_{v_H, v_E}$	$Pr(> F)$			valor-p	L	F_{v_1, v_2}	$Pr(> F)$	
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	\mathbf{I}_3	2.65		T1vsT2: [-42.9, -24.6]	0.1752	0.4381	392.3	T2 957.2182 a		
		0.0558		T1vsT3: [-26.0, -7.7]			1.76e-09	T3 419.5336 b		
				T2vsT3: [7.7, 26.0]					T1 101.4289 c	
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	\mathbf{I}_3	2.6661		T1vsT2: [-92.3, -74.0]	0.8342	0.2918	2663	T2 6161.6048 a		
		0.05473		T1vsT3: [-50.7, -32.4]			3.33e-13	T3 2708.3446 b		
				T2vsT3: [32.4, 50.7]					T1 655.4444 c	
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	\mathbf{I}_3	2.6666		T1vsT2: [-469.5, -451.2]	0.4166	0.3415	54772	T2 163525.73 a		
		0.05469		T1vsT3: [-239.3, -221.0]			<2e-16	T3 72527.84 b		
				T2vsT3: [221.0, 239.3]					T1 18020.23 c	
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$10 * \mathbf{I}_3$	2.509		T1vsT2: [-62.7, -4.8]	0.2429	0.3321	44.6	T2 910.2799 a		
		0.06624		T1vsT3: [-45.8, 12.0]			2.14e-05	T3 396.9540 b		
				T2vsT3: [-12.0, 45.8]					T1 103.2082 c	
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$10 * \mathbf{I}_3$	2.6607		T1vsT2: [-112.1, -54.2]	0.8977	0.2502	269.8	T2 5878.711 a		
		0.05508		T1vsT3: [-70.5, -12.6]			9.27e-09	T3 2528.462 b		
				T2vsT3: [12.6, 70.5]					T1 578.572 c	
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$10 * \mathbf{I}_3$	2.663		T1vsT2: [-489.3, -431.4]	0.4302	0.3333	5490	T2 162074.11 a		
		0.05471		T1vsT3: [-259.1, -201.2]			1.29e-14	T3 71568.80 b		
				T2vsT3: [201.2, 259.1]					T1 17553.78 c	

1. A pesar de que las diferencias entre los vectores de medias son muy grandes, cuando se tiene $r = 4$, el MANOVA no identifica diferencia entre los efectos de los tratamientos.
2. La CG de Tukey identificó las diferencias entre los vectores de medias acorde a los establecido en lo parámetros de simulación de los datos, por ejemplo, no identificó las diferencias entre los vectores cuando los valores de μ_1 son pequeños con una Σ grande.
3. En todos estos casos los resultados del ANOVA y HSD de Tukey son acordes a lo pronosticado.

Tabla A.3: Comparaciones para el caso $t = 3$, $p = 3$, $r = 16$, grandes diferencias entre μ_i^s

μ	Σ	MANOVA:		Datos trans		ANOVA:		HSD Tukey
		$\sim F_{v_H, v_E}$ $P_T(> F)$	valor-p	SW	L	F_{v_1, v_2} $P_T(> F)$		
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		14.517 1.906e-11	0.0316	0.0316	0.0316	3073 <2e-16	T2 62.28887 a T3 40.26543 b T1 18.50204 c	
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	\mathbf{I}_3	14.646 1.583e-11	0.0005	0.0410	0.0410	28347 <2e-16	T2 42.35327 a T3 32.36821 b T1 19.99037 c	
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		14.666 1.539e-11	0.0051	0.0184	0.0184	805555 <2e-16	T2 118.83477 a T3 92.39375 b T1 59.70650 c	
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		13.298 1.144e-10	0.0152	0.0517	0.0517	227.5 <2e-16	T2 55.84143 a T3 36.21607 b T1 17.04625 c	
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$10 * \mathbf{I}_3$	14.466 2.052e-11	0.0001	0.0395	0.0395	2921 <2e-16	T2 41.70839 a T3 31.62904 b T1 19.10090 c	
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		14.659 1.554e-11	0.0041	0.0190	0.0190	80117 <2e-16	T2 118.50988 a T3 92.01359 b T1 59.21266 c	

1. Cuando aumenta el número de repeticiones, el MANOVA es capaz de detectar las diferencias entre los vectores de medias.
2. Cuando se tienen las condiciones del título de la tabla, los datos generados con el método propuesto no cumplen con los supuestos de normalidad ni de homocedasticidad, por lo que es requerido transformarlos previo a realizar el ANOVA y el HSD de Tukey, posterior a ello, los resultados obtenidos de dichas pruebas son acordes a lo planeado.
3. En los resultados del valor-p tanto de MANOVA como de ANOVA, no se observa algún efecto significativo del uso de diferentes Σ .

A.5 Tablas de las comparaciones del MANOVA con el ANOVA

Tabla A.4: Comparaciones para el caso $t = 3$, $p = 7$, $r = 16$, pequeñas diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW	L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		3.86 5.486e-05	0.0402	0.7999	53.3 1.35e-12	T3 89812.02 a T2 63612.05 b T1 43648.37 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	I_7	5.3463 4.741e-07	0.0002	0.8692	300.9 <2e-16	T3 11.45148 a T2 11.09005 b T1 10.70077 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		5.7029 1.606e-07	0.0191	0.6388	12331 <2e-16	T3 682959080 a T2 490358532 b T1 341100428 c
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		0.98458 0.4766	0.0279	0.7915	4.232 0.0207	T3 113973.79 a T2 84956.38 ab T1 62264.05 b
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	$10 * I_7$	3.3846 0.0002688	4.488e-05	0.9008	29.9 5.48e-09	T3 1.247443 a T2 1.247055 b T1 1.246562 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		5.6021 2.175e-07	0.0177	0.6875	1229 <2e-16	T3 10314495 a T2 7933429 b T1 5950491 c

1. En todos estos casos, posterior a la transformación de los datos, los resultados de ANOVA y el HSD de Tukey son consistentes a lo esperado.

Tabla A.5: Comparaciones para el caso $t = 3$, $p = 7$, $r = 4$, grandes diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		residuals have rank 3<7	0.4410 0.4313	2485 4.54e-13	T2 2614.5834 a T3 1154.1928 b T1 284.7821 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	\mathbf{I}_7	residuals have rank 3<7	0.9549 0.3679	37519 <2e-16	T2 14818.114 a T3 6548.420 b T1 1611.166 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		residuals have rank 3<7	0.0159 0.6747	270626 <2e-16	T2 4286682.1 a T3 1617239.1 b T1 304886.1 c
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		residuals have rank 3<7	0.8444 0.6408	233.2 1.77e-08	T2 2561.0854 a T3 1137.5817 b T1 305.0581 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$10 * \mathbf{I}_7$	residuals have rank 3<7	0.9776 0.2568	3434 1.06e-13	T2 14476.949 a T3 6340.031 b T1 1535.553 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		residuals have rank 3<7	0.0194 0.6649	24413 <2e-16	T2 13.73690 a T3 12.81341 b T1 11.24617 c

1. Aplica comentario 1 de la tabla 4.3.
2. Se observa que en el caso de que las distancias entre las medias de las variables sean grandes, los datos obtenidos con la metodología propuesta no cumplen con el supuesto de normalidad, pero en los demás casos si cumplen. Se observa un dato en rojo, que corresponde al caso de los datos que no cumplieron con el supuesto de normalidad aun con varias transformaciones, sin embargo, nótese, que tanto el ANOVA como el HSD de Tukey dieron los resultados acordes a los parámetros de simulación.
3. En todos estos casos, posterior a la transformación de los datos para que al menos cumplieran con el supuesto de homocedasticidad, los resultados de ANOVA y el HSD de Tukey son acordes a los parámetros para la simulación de los datos.

Tabla A.6: Comparaciones para el caso $t = 3$, $p = 7$, $r = 16$, grandes diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		5.687 1.684e-07	0.4606 0.0077	5083 <2e-16	T2 32.09715 a T3 24.42884 b T1 14.98228 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	I_7	5.7104 1.57e-07	0.0009 0.0272	25672 <2e-16	T2 29.16433 a T3 24.03098 b T1 16.96405 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		5.7142 1.552e-07	0.9332 0.0002	1186977 <2e-16	T2 155.47117 a T3 121.14834 b T1 78.73674 c
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		5.4524 3.428e-07	0.8443 0.0047	442.8 <2e-16	T2 42.15306 a T3 31.16056 b T1 18.72594 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$10 * I_7$	5.6752 1.745e-07	0.0004 0.0458	2580 <2e-16	T2 55.98265 a T3 43.07584 b T1 27.18802 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		5.7131 1.557e-07	0.9236 0.0003	115748 <2e-16	T2 155.3113 a T3 120.9638 b T1 78.5074 c

1. Se observa que en el caso de que las distancias entre las medias de las variables sean grandes, los datos obtenidos con la metodología propuesta no cumplen con el supuesto de normalidad, pero en los demás casos sí cumplen. Se observa un dato en rojo, que corresponde al caso de los datos que no cumplieron con el supuesto de normalidad aun con varias transformaciones, sin embargo, nótese, que tanto el ANOVA como el HSD de Tukey dieron los resultados acordes a los parámetros de simulación.
2. En todos estos casos, posterior a la transformación de los datos para que al menos cumplieran con el supuesto de homocedasticidad, los resultados de ANOVA y el HSD de Tukey son acordes a los parámetros para la simulación de los datos.

Tabla A.7: Comparaciones para el caso $t = 7$, $p = 3$, $r = 4$, pequeñas diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		1.435 0.147	0.0005 0.9972	8.43 9.87e-05	T7 1125811.56 a T3 664516.25 ab T5 396587.74 bc T1 373459.11 bc T2 351473.56 bc T6 197904.29 bc T4 97627.29 c
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	I_3	1.7364 0.05595	0.0139 0.9931	48.31 3.25e-11	T7 456348.37 a T3 319257.02 b T5 224711.60 c T1 215688.29 cd T2 206939.95 cd T6 139665.14 de T4 85798.82 e
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		1.7492 0.0536	0.0010 0.9962	963.2 <2e-16	T7 338111233 a T3 238289208 b T5 169030334 c T1 162395826 c T2 155958580 c T6 106261672 d T4 66116795 e

Continúa en la siguiente página

1. Comparando la tabla 4.2 con ésta (cuya única diferencia es el número de tratamientos), el MANOVA sigue sin rechazar H_0 , pero al transformar los datos, primero con la metodología propuesta y posteriormente con Box-Cox hasta cumplir con los supuestos del ANOVA, se obtiene que se encuentran diferencias entre los efectos de los tratamientos, lo cual se confirma con los resultados del HSD de Tukey.
2. Al igual que para $t = 3$, los datos transformados con la metodología propuesta no cumplen con el supuesto de normalidad, pero sí con el de homogeneidad de varianzas.

Tabla A.8: Continuación: Comparaciones para el caso $t = 7$, $p = 3$, $r = 4$, pequeñas diferencias entre μ'_i 's

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		0.54773 0.92223	0.0065	1.455 0.241	No hay diferencias significativas
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	$10 * \mathbf{I}_3$	1.6224 0.0814	0.0444	5.381 0.00168	T7 377659.66 a T3 261330.18 ab T5 182025.32 ab T1 174508.36 b T2 167230.54 b T6 111649.90 b T4 67798.83 b
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		1.7423 0.05486	0.0012	98.92 2.71e-14	T7 323974746 a T3 227494848 b T5 160752260 c T1 154370386 c T2 148180582 c T6 100485869 d T4 62125270 d

1. Aplica los mismos comentarios que los de la tabla A.7.
2. Con una Σ muy grande en comparación con los valores observados, es obvio que los datos presentarán resultados en la simulación muy semejantes, y por lo tanto no habrá diferencia entre los efectos de los tratamientos, lo que se refleja en los resultados obtenidos en el MANOVA, ANOVA y HSD de Tukey.

Tabla A.9: Comparaciones para el caso $t = 7$, $p = 3$, $r = 16$, pequeñas diferencias entre μ_i 's

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $P^r(> F)$	Datos trans valor-p SW	L	ANOVA: F_{v_1, v_2} $P^r(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		6.436 1.404e-13	0.0004	0.9514	46.17 <2e-16	T7 3.153771 a T3 3.085787 a T5 3.016463 b T1 3.008207 b T2 2.999830 b T6 2.918459 c T4 2.812877 d
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	I_3	8.3459 < 2.2e-16	2.163e-07	0.9510	524.1 <2e-16	T7 7.893902 a T3 7.659230 b T5 7.430095 c T1 7.403453 c T2 7.376557 c T6 7.122187 d T4 6.809505 e
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		8.7344 < 2.2e-16	2.213e-05	0.9174	14487 <2e-16	T7 10.700699 a T3 10.507192 b T5 10.317596 c T1 10.295510 d T2 10.273205 e T6 10.061836 f T4 9.801001 g

Continúa en la siguiente página

1. Nuevamente se ve el efecto de que al aumentar el número de repeticiones a 16, el MANOVA tendrá la capacidad de detectar las diferencias entre los efectos de los tratamientos.
2. Al igual que para $t = 3$, los datos transformados con la metodología propuesta no cumplen con el supuesto de normalidad, pero sí con el de homogeneidad de varianzas.

Tabla A.10: Continuación: Comparaciones para el caso $t = 7$, $p = 3$, $r = 16$, pequeñas diferencias entre μ_i 's

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		1.904 0.01514	5.828e-08 0.9876	3.085 0.00804	T7 6.441601 a T3 6.129590 ab T5 5.829858 ab T1 5.795339 ab T2 5.760565 ab T6 5.435514 ab T4 5.046998 b
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	$10 * \mathbf{I}_3$	5.8956 3.091e-12	4.599e-09 0.9597	55.24 <2e-16	T7 9.636435 a T3 9.270933 b T5 8.917276 c T1 8.876364 c T2 8.835108 c T6 8.447186 d T4 7.976115 e
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		8.5969 < 2.2e-16	1.497e-05 0.9217	1448 <2e-16	T7 17.57714 a T3 17.09355 b T5 16.62673 c T1 16.57279 cd T2 16.51841 d T6 16.00774 e T4 15.38893 f

1. Aplica mismos comentarios que los de la tabla A.9.

Tabla A.11: Comparaciones para el caso $t = 7$, $p = 3$, $r = 4$, grandes diferencias entre μ_i 's

μ	Σ	MANOVA: $\sim F^{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		1.7428 0.05476	0.0110 0.3701	396.1 <2e-16	T6 504.28545 a T2 362.33208 b T4 241.38264 c T7 142.54080 d T1 67.34899 e T5 34.54749 ef T3 18.30716 f
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	I_3	1.7497 0.05351	0.2455 0.1523	2493 <2e-16	T6 6161.6048 a T2 4259.9297 b T4 2708.3446 c T7 1506.8495 d T1 655.4444 e T5 312.6446 f T3 154.1293 f
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		1.75 0.05347	0.0779 0.2218	51278 <2e-16	T6 163525.734 a T2 113465.504 b T4 72527.844 c T7 40712.754 d T1 18020.234 e T5 8783.556 f T3 4450.284 g

Continúa en la siguiente página

1. Aún con aumentar los valores en la diferencias entre vectores de medias de los tratamientos, cuando se tienen 4 repeticiones por grupo, el MANOVA no identifica diferencia entre los efectos de los tratamientos.

Tabla A.12: Continuación: Comparaciones para el caso $t = 7$, $p = 3$, $r = 4$, grandes diferencias entre μ_i 's

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		1.6805 0.06732	0.0087 0.1863	36.78 4.41e-10	T6 3.268106 a T2 3.194630 ab T4 3.094074 ab T7 2.946051 bc T1 2.710882 cd T5 2.518810 de T3 2.384221 e
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	$10 * \mathbf{I}_3$	1.7474 0.05393	0.2418 0.1011	252.4 <2e-16	T6 5878.7114 a T2 4028.5416 b T4 2528.4617 c T7 1378.4719 d T1 578.5720 e T5 266.6753 e T3 128.7622 e
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		1.7498 0.05349	0.0808 0.2086	5140 <2e-16	T6 162074.105 a T2 112260.170 b T4 71568.804 c T7 40000.009 d T1 17553.783 e T5 8464.882 f T3 4230.128 g

1. No se observa un efecto significativo en los resultados del MANOVA y ANOVA al aumentar Σ , se observa que la significancia depende principalmente por la relación que tenga la matriz de varianzas y covarianzas con los vectores de medias.

Tabla A.13: Comparaciones para el caso $t = 7$, $p = 3$, $r = 16$, grandes diferencias entre μ_i 's

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		8.6911 < 2.2e-16	7.107e-06 0.0003	2098 < 2e-16	T6 28.690833 a T2 24.849466 b T4 20.766511 c T7 16.359248 d T1 11.477953 e T5 8.200230 f T3 5.793209 g
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	I_3	8.742 < 2.2e-16	1.91e-08 0.0007	17023 < 2e-16	T6 18.014331 a T2 16.684995 b T4 15.151106 c T7 13.313952 d T1 10.967635 e T5 9.112839 f T3 7.521743 g
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		8.7497 < 2.2e-16	9.459e-07 6.142e-05	623933 < 2e-16	T6 118.83477 a T2 106.14842 b T4 92.39375 c T7 77.16815 d T1 59.70650 e T5 47.48146 f T3 38.10543 g

Continúa en la siguiente página

1. Nuevamente se observa que con $r = 16$ el MANOVA es capaz de identificar las diferencias entre los efectos de los tratamientos.
2. Respecto a la metodología propuesta, se observa que al aumentar las repeticiones los datos transformados no cumplen con los supuestos de normalidad y homocedasticidad, e incluso cuando se aplica una segunda transformación, no hay cumplimiento del supuesto de normalidad. Sin embargo, los resultados del ANOVA y del HSD de Tukey dan los resultados acorde a lo planteado.

Tabla A.14: Continuación: Comparaciones para el caso $t = 7$, $p = 3$, $r = 16$, grandes diferencias entre μ_i^s

μ	Σ	MANOVA:		Datos trans		ANOVA:		HSD Tukey
		$\sim F_{v_H, v_E}$ $Pr(> F)$	valor-p SW	L	F_{v_1, v_2} $Pr(> F)$			
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		8.195 < 2.2e-16	4.842e-06	0.0022	150.6 < 2e-16	T6 28.030802 a T2 24.193652 b T4 20.130316 c T7 15.779959 d T1 11.067303 e T5 8.088992 f T3 6.283701 f		
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	$10 * \mathbf{I}_3$	8.6709 < 2.2e-16	2.657e-09	0.0009	2709 < 2e-16	T6 139.433366 a T2 115.60105 b T4 91.67180 c T7 67.63940 d T1 43.53132 e T5 29.12027 f T3 19.68962 g		
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		8.7471 < 2.2e-16	6.339e-07	6.826e-05	66286 < 2e-16	T6 187.40042 a T2 164.45311 b T4 140.06468 c T7 113.73463 d T1 84.53563 e T5 64.84502 f T3 50.24690 g		

1. Respecto a la tabla A.13 aumentar el valor de Σ no muestra algún efecto significativo en los resultados de las pruebas.

Tabla A.15: Comparaciones para el caso $t = 7$, $p = 7$, $r = 4$, pequeñas diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW	L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		residuals have rank 3<7	0.0030	0.9995	42.22 1.19e-10	T7 84871.90 a T3 59769.05 b T5 42388.43 c T1 40725.41 c T2 39112.21 c T6 26671.08 cd T4 16640.91 d
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	I_7	residuals have rank 3<7	0.0478	0.9946	682.3 <2e-16	T7 2332.021 a T3 1954.931 b T5 1644.043 c T1 1611.166 c T2 1578.622 c T6 1300.725 d T4 1023.608 e
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		residuals have rank 3<7	7.388e-06	0.9998	2645 <2e-16	T7 0.3448276 a T3 0.3448276 b T5 0.3448276 c T1 0.3448276 cd T2 0.3448276 d T6 0.3448276 e T4 0.3448276 f

Continúa en la siguiente página

1. Aplica comentario 1 de la tabla 4.3.
2. Se observa que los datos transformados no cumplen con el supuesto de normalidad pero sí con el de homocedasticidad. Al realizar una segunda transformación sólo los datos con grandes distancias entre las medias de las variables no cumplieron con el supuesto de normalidad.
3. Aún con pequeñas diferencias, el ANOVA fue capaz de rechazar H_0 , y el HSD de Tukey de acomodar correctamente los efectos de los tratamientos.

Tabla A.16: Continuación: Comparaciones para el caso $t = 7$, $p = 7$, $r = 4$, pequeñas diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		residuals have rank 3<7	0.0029 0.9992	3.451 0.0157	T7 10157.703 a T3 7893.742 ab T5 6204.202 ab T1 6035.280 ab T2 5869.952 ab T6 4536.251 b T4 3351.876 b
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	$10 * \mathbf{I}_7$	residuals have rank 3<7	0.0009 0.9714	49.85 2.4e-11	T7 592153.7 a T3 429917.1 b T5 314148.8 c T1 302878.0 cd T2 291906.9 cd T6 205834.6 de T4 133908.9 e
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		residuals have rank 3<7	1.604e-05 0.9998	531.2 <2e-16	T7 0.9999837 a T3 0.9999806 b T5 0.9999769 c T1 0.9999765 c T2 0.9999760 c T6 0.9999709 d T4 0.9999630 e

1. Comparando con la tabla A.15, el único valor en el que se ve una gran influencia de aumentar el valor Σ fue el obtenido en SW de medianas distancias entre las medias de las variables, sin embargo, esto se atribuye a la naturaleza de los datos.
2. En el primer renglón de resultados, se aprecia que a pesar de que Σ es grande en comparación con las pequeñas distancias entre medias de variables, tanto el ANOVA como el HSD de Tukey pudieron identificar diferencias entre los efectos de los tratamientos.

Tabla A.17: Comparaciones para el caso $t = 7$, $p = 7$, $r = 16$, pequeñas diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		2.6198 3.038e-07	0.0018 0.8406	88.61 <2e-16	T7 5800.039 a T3 4475.455 b T5 3472.600 c T1 3371.563 c T2 3272.530 c T6 2468.048 d T4 1744.597 e
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	I_7	2.9154 9.34e-09	1.169e-07 0.9467	497.6 <2e-16	T7 23.88506 a T3 22.68974 b T5 21.56646 c T1 21.43861 c T2 21.31012 c T6 20.12311 d T4 18.73221 e
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		2.9698 4.86e-09	0.0004 0.5959	20685 <2e-16	T7 3.212004 a T3 3.205488 b T5 3.198755 c T1 3.197947 d T2 3.197126 e T6 3.189090 f T4 3.178498 g

Continúa en la siguiente página

1. Al comparar con la tabla A.4 se ve que al aumentar el número de tratamientos, disminuye el ajuste de los datos al supuesto de normalidad, y que el supuesto de homogeneidad de varianzas sigue cumpliéndose.
2. Tanto el ANOVA como el HSD de Tukey muestran los resultados esperados.

Tabla A.18: Continuación: Comparaciones para el caso $t = 7$, $p = 7$, $r = 16$, pequeñas diferencias entre μ_i 's

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		1.2686 0.1238	0.0004 0.8391	6.617 5.89e-06	T7 2260.4396 a T3 1856.9831 ab T5 1539.5140 bc T1 1506.8049 bc T2 1474.5971 bc T6 1206.7547 bc T4 953.5569 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	$10 * \mathbf{I}_7$	2.4921 1.311e-06	9.906e-09 0.9743	46.09 <2e-16	T7 7.453257 a T3 7.290680 b T5 7.131112 c T1 7.112510 c T2 7.093721 c T6 6.915576 d T4 6.695676 e
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		2.9549 5.812e-09	0.0004 0.6071	2040 <2e-16	T7 1309096.2 a T3 1042747.2 b T5 834112.7 c T1 812681.5 cd T2 791593.0 d T6 616866.1 e T4 453183.9 f

1. No se ve un efecto significativo en los resultados obtenidos al aumentar el valor de Σ , excepto, obviamente en el caso de pequeños valores en las medias de las variables, en el cual el MANOVA no identifica diferencia entre los efectos de los tratamientos. Sin embargo al realizar la transformación de los datos con la metodología propuesta, y posteriormente con el método Box-Cox, el ANOVA rechaza H_0 y con el HSD se identifican y acomodan correctamente dichas diferencias, lo cual aplica para todos los casos.
2. Comparando con la tabla A.16 ($r = 4$) se observa que en ambas los datos no cumplen con el supuesto de normalidad y sí con el de homocedasticidad. El HSD identificó más diferencias entre efectos de los tratamientos al aumentar la repeticiones.

Tabla A.19: Comparaciones para el caso $t = 7$, $p = 7$, $r = 4$, grandes diferencias entre μ_i 's

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		residuals have rank 3<7	0.04301	1488 <2e-16	T6 15.035649 a T2 13.871435 b T4 12.529792 c T7 10.926841 d T1 8.891797 e T5 7.303645 f T3 5.972567 g
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	I_7	residuals have rank 3<7	0.2995	34953 <2e-16	T6 14818.1145 a T2 10266.7123 b T4 6548.4201 c T7 3663.2380 d T1 1611.1658 e T5 779.8153 f T3 392.2036 g
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		residuals have rank 3<7	0.0005	138653 <2e-16	T6 12.877786 a T2 12.512582 b T4 12.065456 c T7 11.488697 d T1 10.674994 e T5 9.958111 f T3 9.280509 g

Continúa en la siguiente página

1. Al tener grandes diferencias entre las μ_i , y con una Σ baja comparada con los valores de las medias de las variables, se obtienen resultados que cumplen con los supuestos de normalidad y homocedasticidad. Tanto ANOVA como HSD dan los resultados esperados.
2. Cuando se compara con la tabla A.11 se observa que el aumento en el número de variables influyó en los casos de pequeñas y medianas distancias entre medias de variables, pues el valor-p de SW aumentó, y para grandes distancias disminuyó.

Tabla A.20: Continuación: Comparaciones para el caso $t = 7$, $p = 7$, $r = 4$, grandes diferencias entre μ_i 's

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		residuals have rank $3 < 7$	0.3745	209.6 $< 2e-16$	T6 2561.0854 a T2 1775.4610 b T4 1137.5817 c T7 647.4474 d T1 305.0581 e T5 170.5421 e T3 110.4138 e
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	$10 * \mathbf{I}_7$	residuals have rank $3 < 7$	0.5917	3122 $< 2e-16$	T6 14476.9492 a T2 9991.9351 b T4 6340.0310 c T7 3521.2368 d T1 1535.5527 e T5 744.0350 f T3 382.9786 f
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		residuals have rank $3 < 7$	0.0002	7664 $< 2e-16$	T6 3.839850 a T2 3.824489 b T4 3.803645 c T7 3.773025 d T1 3.721428 e T5 3.666166 f T3 3.603673 g

1. Aumentar el valor de Σ tuvo el efecto de que para pequeñas y medianas distancias entre las variables, los datos transformados cumplieron directamente con los supuestos de normalidad y homocedasticidad.

Tabla A.21: Comparaciones para el caso $t = 7$, $p = 7$, $r = 16$, grandes diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		2.9639 5.213e-09	0.0173 3.973e-06	4012 <2e-16	T6 39.902754 a T2 34.895035 b T4 29.556103 c T7 23.772936 d T1 17.342925 e T5 13.013671 f T3 9.835032 g
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	I_7	2.9703 4.826e-09	9.786e-08 0.0002	19396 <2e-16	T6 40.18900 a T2 36.32592 b T4 32.05172 c T7 27.19841 d T1 21.43861 e T5 17.25141 f T3 13.93249 g
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		2.9714 4.766e-09	0.3779 1.348e-10	983459 <2e-16	T6 256.31008 a T2 225.23229 b T4 192.21553 c T7 156.59143 d T1 117.12664 e T5 90.55348 f T3 70.88645 g

Continúa en la siguiente página

1. El efecto de aumentar el número de repeticiones (comparación con la tabla A.19) fue el que los datos transformados no cumplieran con el supuesto de homocedasticidad, y que aunque se aplicó una segunda transformación mediante Box-Cox, los datos no cumplieran con el supuesto de normalidad, no obstante, el ANOVA rechaza H_0 y la HSD realiza el correcto acomodo e identificación de los efectos de los tratamientos.
2. El aumento en el número de variables muestra el efecto de aumentar la significancia del MANOVA.

Tabla A.22: Continuación: Comparaciones para el caso $t = 7$, $p = 7$, $r = 16$, grandes diferencias entre μ_i^s

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		2.8981 1.147e-08	0.3680 1.131e-06	327.1 <2e-16	T6 39.88211 a T2 34.92952 b T4 29.68275 c T7 24.07141 d T1 18.02517 e T5 14.22078 f T3 11.73612 f
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	$10 * \mathbf{I}_7$	2.9607 5.42e-09	3.482e-08 0.0012	2041 <2e-16	T6 79.29440 a T2 69.37575 b T4 58.84464 c T7 47.49934 d T1 34.98964 e T5 26.66764 f T3 20.65709 g
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		2.9711 4.782e-09	0.4246 1.57e-10	95844 <2e-16	T6 283.47758 a T2 248.17584 b T4 210.82016 c T7 170.71494 d T1 126.58482 e T5 97.09828 f T3 75.43424 g

1. Aumentar el valor de Σ no causó mucha diferencia en los valor-p del MANOVA en comparación con los resultados obtenidos en la tabla A.22.
2. Al comparar con la tabla A.16 se nota que aumentar las repeticiones tiene el efecto de hacer que los datos transformados no cumplan con el supuesto de homocedasticidad.

Tabla A.23: Comparaciones para el caso $t = 3$, $p = 3$, $r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	CG Tukey	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		2.5661 0.06178	T1vsT2: [-6.8, 3.5] T1vsT3: [-8.5, 1.8] T2vsT3: [-6.8, 3.5]	0.0165 0.9604	2.109 0.177	T3 1.919675 a No hay diferencias
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	$\begin{pmatrix} 1.5 & -0.6 & -1.1 \\ -0.6 & 1.7 & 0.3 \\ -1.1 & 0.3 & 1.2 \end{pmatrix}$	2.6582 0.05525	T1vsT2: [-9.3, 1.0] T1vsT3: [-51.2, -40.8] T2vsT3: [-28.2, -17.8]	0.0020 0.9702	127.7 2.48e-07	significativas T2 0.3225806 b T1 0.3225806 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		2.6663 0.05471	T1vsT2: [-28.2, -17.8] T1vsT3: [-51.2, -7.83] T2vsT3: [-28.2, -17.8]	0.0682 0.9768	6951 4.46e-15	T3 26207.69 a T2 22017.28 b T1 18191.76 c
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		2.0699 0.1148	T1vsT2: [-15.7, 12.4] T1vsT3: [-17.4, 10.7] T2vsT3: [-15.7, 12.4]	0.0111 0.9999	1.919 0.202	No hay diferencias significativas
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	$\begin{pmatrix} 8.6 & 3.0 & -8.3 \\ 3.0 & 10.7 & 0.1 \\ -8.3 & 0.1 & 10.1 \end{pmatrix}$	2.6446 0.05617	T1vsT2: [-18.2, 9.9] T1vsT3: [-22.4, 5.7] T2vsT3: [-18.2, 9.9]	0.0876 0.9219	6.519 0.0178	T3 1052.0338 a T2 890.7226 ab T1 743.4150 b
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		2.6656 0.05476	T1vsT2: [-37.1, -8.9] T1vsT3: [-60.1, -31.9] T2vsT3: [-37.1, -8.9]	0.0302 0.9712	352.9 2.82e-09	T3 344909228 a T2 243594474 b T1 166444078 c

1. Con pequeñas distancias entre las medias de las variables, pocas repeticiones y pequeñas diferencias entre los vectores de medias, para este caso, tanto el MANOVA como el ANOVA no rechazan H_0 . Comparándolo con los resultados en la tabla 4.2 se observa que si hay un efecto de tener variables correlacionadas, pues con variables no correlacionadas el ANOVA rechaza H_0 . Para los demás casos de distancias entre las medias de las variables, el ANOVA rechaza H_0 .
2. La significancia del valor-p tanto del MANOVA como del ANOVA aumenta conforme aumenta la distancia entre las medias de las variables.
3. La metodología de conjetura generalizada de Tukey no encontró diferencias entre los efectos de los tratamientos en los casos de pequeñas y medianas distancias entre las medias.
4. En todos los casos estudiados en esta tabla, los datos obtenidos tras emplear la metodología propuesta cumplen con el supuesto de homocedasticidad, y sólo en algunos no cumple con el supuesto de normalidad, incluso después de una segunda transformación.

Tabla A.24: Comparaciones para el caso $t = 3$, $p = 3$, $r = 16$, pequeñas diferencias entre μ_i 's
Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		8.5114 2.711e-07	0.0451 0.8233	19.07 1e-06	T3 382.2467 a T2 312.5028 b T1 250.9472 c
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	$\begin{pmatrix} 1.5 & -0.6 & -1.1 \\ -0.6 & 1.7 & 0.3 \\ -1.1 & 0.3 & 1.2 \end{pmatrix}$	14.014 3.962e-11	0.0127 0.8712	565.9 <2e-16	T3 1.417105 a T2 1.415604 b T1 1.413731 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		14.637 1.603e-11	0.1766 0.8121	9801 <2e-16	T3 26195.37 a T2 22005.78 b T1 18181.09 c
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		3.6247 0.002926	3.605e-09 0.9636	3.324 0.045	T3 1.390646 a T2 1.384145 ab T1 1.375971 b
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	$\begin{pmatrix} 8.6 & 3.0 & -8.3 \\ 3.0 & 10.7 & 0.1 \\ -8.3 & 0.1 & 10.1 \end{pmatrix}$	13.084 1.579e-10	6.561e-08 0.9378	31.8 2.46e-09	T3 0.6666452 a T2 0.6666387 b T1 0.6666294 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		14.585 1.73e-11	8.456e-06 0.8929	580.9 <2e-16	T3 0.6666665 a T2 0.6666665 b T1 0.6666664 c

1. Aumentar las repeticiones tuvo el efecto de que el MANOVA pudiera identificar la diferencia entre los efectos, y que el ANOVA aumentara su significancia.
2. En general se observa que los datos transformados no cumplieron con el supuesto de normalidad, incluso con una segunda transformación. Sin embargo, tanto el ANOVA como el HSD de Tukey pudieron identificar las diferencias entre los efectos y con el HSD el orden de los tratamientos fue adecuado.

Tabla A.25: Comparaciones para el caso $t = 3$, $p = 3$, $r = 16$, grandes diferencias entre μ_i^s
Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $P_r(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $P_r(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		14.561 1.789e-11	0.2913 0.0114	1717 <2e-16	T2 23.13454 a T3 17.46425 b T1 10.50658 c
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$\begin{pmatrix} 1.5 & -0.6 & -1.1 \\ -0.6 & 1.7 & 0.3 \\ -1.1 & 0.3 & 1.2 \end{pmatrix}$	14.66 1.553e-11	0.0002 0.0587	49642 <2e-16	T2 31.55075 a T3 24.99404 b T1 16.44042 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		14.666 1.538e-11	0.2282 0.0079	899192 <2e-16	T2 118.94139 a T3 92.51939 b T1 59.87319 c
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		14.233 2.877e-11	4.505e-07 0.2166	158.9 <2e-16	T2 9.847029 a T3 8.274053 b T1 5.880131 c
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	$\begin{pmatrix} 8.6 & 3.0 & -8.3 \\ 3.0 & 10.7 & 0.1 \\ -8.3 & 0.1 & 10.1 \end{pmatrix}$	14.649 1.577e-11	3.707e-06 0.0944	2334 <2e-16	T2 31.57045 a T3 25.03169 b T1 16.54647 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		14.666 1.539e-11	2.968e-05 0.0505	56445 <2e-16	T2 118.87371 a T3 92.44178 b T1 59.77858 c

1. Cuando aumenta el número de repeticiones, el MANOVA es capaz de detectar las diferencias entre los vectores de medias, y aumenta la significancia tanto en MANOVA como en ANOVA.
2. En la mayoría de los casos, los datos transformados no cumplen con los supuestos de normalidad ni de homocedasticidad, por lo que es requerido una segunda transformación previo a realizar el ANOVA y el HSD de Tukey, sin embargo, los datos siguen presentando problemas de ajuste al supuesto de normalidad.
3. En los resultados del valor-p tanto de MANOVA como de ANOVA, no se observa algún efecto significativo del uso de diferentes Σ .

Tabla A.26: Comparaciones para el caso $t = 3$, $p = 7$, $r = 4$, pequeñas diferencias entre μ_i^s
 Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		residuals have rank 3<7	0.0350 0.8856	12.6 0.00246	T3 90814.79 a T2 64554.88 b T1 44507.37 b
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	A	residuals have rank 3<7	0.0446 0.8834	29.29 0.000115	T3 76737.41 a T2 59050.01 b T1 44322.47 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		residuals have rank 3<7	0.1320 0.8793	348.9 2.97e-09	T3 62469.82 a T2 52466.42 b T1 43335.52 c
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		residuals have rank 3<7	0.0017 0.9788	3.782 0.0643	No hay diferencias significativas
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	B	residuals have rank 3<7	0.3311 0.8614	9.728 0.00563	T3 2429.582 a T2 2050.895 ab T1 1705.534 b
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		residuals have rank 3<7	0.0527 0.9687	640.1 1.98e-10	T3 62520.41 a T2 52518.73 b T1 43389.55 c

$$\mathbf{A} = \begin{pmatrix} 2.0 & -0.5 & 0.2 & 1.2 & 0.3 & -0.5 & 1.3 \\ -0.5 & 1.5 & -0.6 & -0.1 & -0.9 & -0.2 & -0.7 \\ 0.2 & -0.6 & 1.4 & 0.0 & 0.2 & -0.1 & 0.8 \\ 1.2 & -0.1 & 0.0 & 1.9 & 0.7 & 0.0 & 0.7 \\ 0.3 & -0.9 & 0.2 & 0.7 & 1.7 & 0.5 & 0.4 \\ -0.5 & -0.2 & -0.1 & 0.0 & 0.5 & 1.4 & -0.3 \\ 1.3 & -0.7 & 0.8 & 0.7 & 0.4 & -0.3 & 1.4 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 10.4 & 4.3 & -0.5 & -2.8 & 1.7 & -6.2 & -3.2 \\ 4.3 & 9.2 & -7.7 & -5.2 & 6.0 & 0.6 & -1.5 \\ -0.5 & -7.7 & 10.6 & 4.1 & -9.3 & -4.5 & 3.7 \\ -2.8 & -5.2 & 4.1 & 9.5 & -4.6 & -1.0 & -0.8 \\ 1.7 & 6.0 & -9.3 & -4.6 & 10.4 & 5.1 & -5.3 \\ -6.2 & 0.6 & -4.5 & -1.0 & 5.1 & 11.2 & -1.5 \\ -3.2 & -1.5 & 3.7 & -0.8 & -5.3 & -1.5 & 9.2 \end{pmatrix}$$

1. Aplican comentarios 1 y 3 de la tabla 4.3.

Tabla A.27: Comparaciones para el caso $t = 3$, $p = 7$, $r = 16$, pequeñas diferencias entre μ'_i
Variables correlacionadas

μ	Σ	MANOVA:		Datos trans		ANOVA:		HSD Tukey
		$\sim F_{v_H, v_E}$ $P_T(> F)$	valor-p SW	L	F_{v_1, v_2} $P_T(> F)$			
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		3.8099	0.0973	0.8484	22.14	T3 432.3655 a		
		6.478e-05			2.02e-07	T2 364.7405 b		
						T1 303.0253 c		
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	A	5.4231	0.1212	0.8030	110.3	T3 2413.390 a		
		3.748e-07			<2e-16	T2 2029.900 b		
						T1 1679.733 c		
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		5.6909	0.0519	0.8631	2546	T3 62676.00 a		
		1.664e-07			<2e-16	T2 52655.21 b		
						T1 43506.91 c		
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		3.9909	6.032e-06	0.9703	9.446	T3 0.9977721 a		
		3.56e-05			0.000376	T2 0.9973850 a		
						T1 0.9968924 b		
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$	B	5.107	0.0359	0.7563	99.04	T3 11.80706 a		
		9.932e-07			<2e-16	T2 11.43897 b		
						T1 11.04488 c		
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 1.1$ $\mu_3 = \mu_1 * 1.2$		5.7027	0.0619	0.7556	3877	T3 62622.92 a		
		1.606e-07			<2e-16	T2 52611.22 b		
						T1 43472.02 c		

1. En todos estos casos, posterior a la transformación de los datos, los resultados de ANOVA y el HSD de Tukey son consistentes a lo esperado.

Tabla A.28: Comparaciones para el caso $t = 3$, $p = 7$, $r = 4$, grandes diferencias entre μ_i^s
Variables correlacionadas

μ	Σ	MANOVA:		Datos trans		ANOVA:		HSD Tukey
		$\sim F_{v_H, v_E}$ $Pr(> F)$	residuals have rank $3 < 7$	valor-p SW	L	F_{v_1, v_2} $Pr(> F)$		
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$				0.5243	0.2215	813.2		T2 2637.3866 a T3 1171.7671 b T1 297.1276 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	A			0.9235	0.0777	2509		T2 14888.996 a T3 6598.047 b T1 1639.537 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$				0.9848	0.0832	30420		T2 391726.98 a T3 173906.69 b T1 43335.52 c
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$				0.1732	0.9298	633.6		T2 2694.6651 a T3 1233.7158 b T1 363.7465 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	B			0.9355	0.1480	639.8		T2 14944.406 a T3 6658.750 b T1 1705.534 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$				0.2805	0.5037	51074		T2 391746.51 a T3 173943.47 b T1 43389.55 c

1. Al comparar con la tabla A.26 se observa que cuando los datos tienen grandes diferencias entre los vectores de medias, los datos transformados cumplen con los supuestos de normalidad y homocedasticidad.
2. No se observa un efecto evidente de la dimensión que tome Σ respecto a los supuestos de SW y de Levene.
3. Aumentar el número de variables de 3 a 7, no muestra algún efecto significativo.

Tabla A.29: Comparaciones para el caso $t = 3$, $p = 7$, $r = 16$, grandes diferencias entre μ_i 's
Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $P_T(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $P_T(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		5.6859 1.69e-07	0.2856 0.0142	2136 <2e-16	T2 32.16607 a T3 24.51659 b T1 15.12477 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	A	5.7112 1.566e-07	0.6681 0.0062	10203 <2e-16	T2 56.34350 a T3 43.46916 b T1 27.58466 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		5.7141 1.552e-07	0.1596 0.0285	234612 <2e-16	T2 155.50130 a T3 121.18389 b T1 78.78413 c
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		5.6897 1.67e-07	0.0937 0.0827	1203 <2e-16	T2 2658.3461 a T3 1203.5937 b T1 339.8213 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$	B	5.7075 1.583e-07	0.7772 0.0012	8743 <2e-16	T2 56.31107 a T3 43.47024 b T1 27.74028 c
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 3$ $\mu_3 = \mu_1 * 2$		5.7142 1.552e-07	0.1184 0.0034	347911 <2e-16	T2 155.47507 a T3 121.15713 b T1 78.76507 c

1. Comparando con pequeñas diferencias entre vectores de medias (tabla A.27), cuando se tienen grandes diferencias, aumenta el valor del valor-p de la prueba de SW y disminuye el de Levene.

Tabla A.30: Comparaciones para el caso $t = 7$, $p = 3$, $r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		1.7264 0.05785	0.0003 0.9978	3.567 0.0136	T7 0.9939269 a T3 0.9927954 a T5 0.9914818 ab T1 0.9913140 ab T2 0.9911411 ab T6 0.9893212 ab T4 0.9865493 b
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	$\begin{pmatrix} 1.5 & -0.6 & -1.1 \\ -0.6 & 1.7 & 0.3 \\ -1.1 & 0.3 & 1.2 \end{pmatrix}$	1.748 0.05381	9.721e-06 0.9990	196.9 <2e-16	T7 616336853 a T3 357844068 b T5 209919847 c T1 197266286 c T2 185259879 c T6 102139934 d T4 48955493 d
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		1.7499 0.05348	0.0031 0.9995	10587 <2e-16	T7 343425543 a T3 242383653 b T5 172200669 c T1 165473030 d T2 158944496 e T6 108505854 f T4 67693769 g

Continúa en la siguiente página

1. Para este caso, al aumentar el número de tratamientos los datos transformados tuvieron problemas de cumplir el supuesto de normalidad.

Tabla A.31: Continuación: Comparaciones para el caso $t = 7$, $p = 3$, $r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F^{v_{H,vE}}$ $Pr(>F)$	Datos trans valor-p SW L	ANOVA: F_{v_1,v_2} $Pr(>F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		1.5899 0.09043	0.0015 1.0000	2.989 0.0285	T7 2551185.5 a T3 1717157.8 ab T5 1188869.1 ab T1 1140646.6 ab T2 1094277.0 ab T6 750778.3 ab T4 493198.1 b
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	$\begin{pmatrix} 8.6 & 3.0 & -8.3 \\ 3.0 & 10.7 & 0.1 \\ -8.3 & 0.1 & 10.1 \end{pmatrix}$	1.7449 0.05438	0.0050 0.9847	10.1 2.72e-05	T7 559616.9 a T3 402156.8 ab T5 291736.6 bc T1 281086.0 bc T2 270737.0 bc T6 190216.2 c T4 123947.0 c
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		1.7497 0.05351	0.0007 0.9991	524.2 <2e-16	T7 0.9999619 a T3 0.9999547 b T5 0.9999462 c T1 0.9999452 c T2 0.9999441 c T6 0.9999323 d T4 0.9999144 e

1. En este caso, aún con una Σ muy grande en comparación con los valores de μ_1 del primer renglón, el ANOVA tuvo un resultado en pro de H_a .

Tabla A.32: Comparaciones para el caso $t = 7$, $p = 3$, $r = 16$, pequeñas diferencias entre μ_i^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		6.9854 6.246e-15	0.0009 0.8856	31.32 <2e-16	T7 1782.071 a T3 1375.916 b T5 1069.369 c T1 1038.532 c T2 1008.316 c T6 763.168 cd T4 543.140 d
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	$\begin{pmatrix} 1.5 & -0.6 & -1.1 \\ -0.6 & 1.7 & 0.3 \\ -1.1 & 0.3 & 1.2 \end{pmatrix}$	8.6099 < 2.2e-16	7.808e-05 0.9430	984.8 <2e-16	T7 2.911835 a T3 2.889011 b T5 2.865406 c T1 2.862573 cd T2 2.859695 d T6 2.831495 e T4 2.794295 f
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		8.7439 < 2.2e-16	0.0177 0.8647	16139 <2e-16	T7 166989.14 a T3 135476.11 b T5 110344.52 c T1 107736.37 d T2 105164.55 e T6 83628.37 f T4 63001.78 g

Continúa en la siguiente página

1. Se observa que al aumentar el número de tratamientos los datos transformados no cumplen con el supuesto de normalidad, incluso después de una segunda transformación.

Tabla A.33: Continuación: Comparaciones para el caso $t = 7$, $p = 3$, $r = 16$, pequeñas diferencias entre μ_i^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW	L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		4.2392 4.478e-08	7.244e-15	0.9987	5.163 0.000109	T7 1.830421 a T3 1.814069 ab T5 1.796494 ab T1 1.794341 abc T2 1.792144 abc T6 1.770141 bc T4 1.739874 c
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	$\begin{pmatrix} 8.6 & 3.0 & -8.3 \\ 3.0 & 10.7 & 0.1 \\ -8.3 & 0.1 & 10.1 \end{pmatrix}$	8.3952 < 2.2e-16	4.424e-13	0.9936	46.61 < 2e-16	T7 1.936574 a T3 1.930802 ab T5 1.924635 bc T1 1.923882 c T2 1.923113 c T6 1.915436 d T4 1.904903 e
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		8.7328 < 2.2e-16	8.259e-10	0.9680	1008 < 2e-16	T7 10.703970 a T3 10.510813 b T5 10.321603 c T1 10.299565 cd T2 10.277309 d T6 10.066437 e T4 9.806313 f

1. Se observan casos en los cuales al aplicar una segunda transformación, los datos siguen sin cumplir el supuesto de normalidad, pero los resultados del ANOVA y del HSD de Tukey siguen siendo acorde a lo planteado.

Tabla A.34: Comparaciones para el caso $t = 7$, $p = 3$, $r = 4$, grandes diferencias entre μ_j^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F^{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F^{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		1.7495 0.05354	0.1967	168.2 <2e-16	T6 1010.85541 a T2 706.19337 b T4 456.42632 c T7 261.55427 d T1 121.57722 e T5 63.94059 e T3 36.49517 e
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	$\begin{pmatrix} 1.5 & -0.6 & -1.1 \\ -0.6 & 1.7 & 0.3 \\ -1.1 & 0.3 & 1.2 \end{pmatrix}$	1.75 0.05347	0.0045	0.4043 12727 <2e-16	T6 2890.4638 a T2 2077.9709 b T4 1386.8762 c T7 822.8405 d T1 393.6467 e T5 205.5770 f T3 111.4593 g
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		1.75 0.05346	0.0345	0.5735 392270 <2e-16	T6 50.20221 a T2 46.31761 b T4 41.93282 c T7 36.82726 d T1 30.55858 e T5 25.82422 f T3 21.93614 g

Continúa en la siguiente página

1. Aún con aumentar los valores en la diferencias entre vectores de medias de los tratamientos, cuando se tienen 4 repeticiones por grupo, el MANOVA no identifica diferencia entre los efectos de los tratamientos.

Tabla A.35: Continuación: Comparaciones para el caso $t = 7$, $p = 3$, $r = 4$, grandes diferencias entre μ_i 's
Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		1.7467 0.05406	0.4063 0.3373	146.6 5e-16	T6 1010.45357 a T2 711.56325 b T4 467.56794 c T7 278.46763 d T1 144.26231 de T5 90.08872 e T3 64.95200 e
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	$\begin{pmatrix} 8.6 & 3.0 & -8.3 \\ 3.0 & 10.7 & 0.1 \\ -8.3 & 0.1 & 10.1 \end{pmatrix}$	1.7499 0.05348	0.2261 0.0813	670.1 <2e-16	T6 6350.2515 a T2 4423.4074 b T4 2846.6532 c T7 1619.9891 d T1 743.4150 e T5 385.5137 ef T3 216.9308 f
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		1.75 0.05346	0.0192 0.4240	33113 <2e-16	T6 31874.516 a T2 23378.182 b T4 15997.009 c T7 9809.212 d T1 4924.762 e T5 2688.023 f T3 1520.383 g

1. En este caso, aumentar Σ muestra resultados muy semejantes en los valores-p obtenidos en el MANOVA y en el ANOVA, comparado con los obtenidos en la tabla A.34.

Tabla A.36: Comparaciones para el caso $t = 7$, $p = 3$, $r = 16$, grandes diferencias entre μ_i^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		8.7086 < 2.2e-16	0.0022 1.459e-05	1278 < 2e-16	T6 24.219973 a T2 21.307517 b T4 18.167700 c T7 14.720071 d T1 10.819229 e T5 8.144130 f T3 6.148328 g
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	$\begin{pmatrix} 1.5 & -0.6 & -1.1 \\ -0.6 & 1.7 & 0.3 \\ -1.1 & 0.3 & 1.2 \end{pmatrix}$	8.7473 < 2.2e-16	1.812e-09 0.0017	48536 < 2e-16	T6 79.91452 a T2 68.67664 b T4 56.97599 c T7 44.66821 d T1 31.49702 e T5 22.97413 f T3 16.90171 g
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		8.7499 < 2.2e-16	0.0004 4.622e-06	704626 < 2e-16	T6 118.94139 a T2 106.26319 b T4 92.51939 c T7 77.30936 d T1 59.87319 e T5 47.67462 f T3 38.32780 g

Continúa en la siguiente página

1. Al aumentar las repeticiones los datos transformados no cumplen con los supuestos de normalidad y homocedasticidad, e incluso cuando se aplica una segunda transformación no hay cumplimiento del supuesto de normalidad. Sin embargo, los resultados del ANOVA y del HSD de Tukey dan los resultados acorde a lo planteado.
2. Al comparar con A.32 se observa que al aumentar las diferencias entre los vectores de medias aumenta la significancia en las pruebas de MANOVA y ANOVA cuando las distancias entre medias son pequeñas.

Tabla A.37: Continuación: Comparaciones para el caso $t = 7$, $p = 3$, $r = 16$, grandes diferencias entre μ_i 's
Variables correlacionadas

μ	Σ	MANOVA: $\sim F^{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F^{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		8.5781 < 2.2e-16	2.694e-12 0.1019	221.54 < 2e-16	T6 14.702104 a T2 13.295897 a T4 11.710170 b T7 9.870531 c T1 7.645436 d T5 6.043572 e T3 4.854997 e
$\mu_1 = [21.6, 5.9, 14.1]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	$\begin{pmatrix} 8.6 & 3.0 & -8.3 \\ 3.0 & 10.7 & 0.1 \\ -8.3 & 0.1 & 10.1 \end{pmatrix}$	8.743 < 2.2e-16	4.752e-11 0.0112	2041.4 < 2e-16	T6 42.55700 a T2 37.79214 b T4 32.63090 c T7 26.92891 d T1 20.42241 e T5 15.91836 f T3 12.53410 g
$\mu_1 = [90.3, 57.6, 82.3]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		8.7497 < 2.2e-16	7.468e-10 0.0014	57058 < 2e-16	T6 6377.5210 a T2 4938.4932 b T4 3610.9156 c T7 2411.0470 d T1 1363.6956 e T5 825.4611 f T3 513.7135 g

1. Respecto a la tabla A.36 aumentar el valor de Σ no muestra algún efecto significativo en los resultados de los valores-p del MANOVA y del ANOVA.

Tabla A.38: Comparaciones para el caso $t = 7$, $p = 7$, $r = 4$, pequeñas diferencias entre μ_i 's
Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		residuals have rank 3<7	0.0022 0.9419	21.12 6.93e-08	T7 90814.79 a T3 64554.88 b T5 46264.14 bc T1 44507.37 bc T2 42801.84 bc T6 29593.05 cd T4 18837.54 d
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	A	residuals have rank 3<7	0.0019 0.9544	45.44 5.87e-11	T7 2803582.5 a T3 1977459.2 b T5 1404256.4 c T1 1349343.5 cd T2 1296062.2 cd T6 884655.6 de T4 552147.3 e
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		residuals have rank 3<7	0.0122 0.9483	535.9 <2e-16	T7 1951706166 a T3 1376754402 b T5 977562624 c T1 939306455 c T2 902184441 c T6 615462418 d T6 615462418 d

Continúa en la siguiente página

1. Se observa que las datos transformados no cumplen con el supuesto de normalidad pero sí con el de homocedasticidad.
2. Aún con pequeñas diferencias, el ANOVA fue capaz de rechazar H_0 , y el HSD de Tukey de acomodar correctamente los efectos de los tratamientos.

Tabla A.39: Continuación: Comparaciones para el caso $t = 7$, $p = 7$, $r = 4$, pequeñas diferencias entre μ_i^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		residuals have rank 3<7	2.189e-05	5.078 0.00232	T7 8.569439 a T3 8.298159 ab T5 8.043145 abc T1 8.014150 abc T2 7.985020 abc T6 7.716706 bc T4 7.405337 c
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	B	residuals have rank 3<7	0.1003	0.9155 17 4.46e-07	T7 2429.582 a T3 2050.895 ab T5 1738.570 bc T1 1705.534 bc T2 1672.830 bc T6 1393.496 cd T4 1114.783 d
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		residuals have rank 3<7	0.0025	0.9987 941.9 <2e-16	T7 0.8695626 a T3 0.8695620 b T5 0.8695613 c T1 0.8695612 c T2 0.8695611 c T6 0.8695601 d T4 0.8695585 e

1. Comparando con la tabla A.38, se ve que al aumentar el valor de las varianzas en Σ aumenta la significancia del ANOVA.
2. En el primer renglón de resultados, se aprecia que a pesar de que Σ es grande en comparación con las pequeñas distancias entre medias de variables, tanto el ANOVA como el HSD de Tukey pudieron identificar diferencias entre los efectos de los tratamientos.

Tabla A.40: Comparaciones para el caso $t = 7$, $p = 7$, $r = 16$, pequeñas diferencias entre μ_i^s
Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		2.6073 3.51e-07	0.0047 0.9129	38.59 <2e-16	T7 9.296527 a T3 8.918438 b T5 8.556928 bc T1 8.515391 c T2 8.473564 c T6 8.083288 d T4 7.616928 e
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	A	2.9275 8.075e-09	0.0052 0.8487	175.8 <2e-16	T7 79150.13 a T3 61070.63 b T5 47367.34 c T1 45985.88 c T2 44631.64 c T6 33623.83 d T4 23712.44 e
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		2.968 4.963e-09	0.0012 0.9377	4065 <2e-16	T7 1.992011 a T3 1.991284 b T5 1.990506 c T1 1.990411 d T2 1.990314 e T6 1.989344 f T4 1.988009 g

Continúa en la siguiente página

1. Aumentar las repeticiones tiene el efecto de aumentar ligeramente la significancia del ANOVA.
2. Al comparar con la tabla A.27 se ve que al aumentar el número de tratamientos, disminuye el ajuste de los datos al supuesto de normalidad, y que el supuesto de homogeneidad de varianzas sigue cumpliéndose.
3. Tanto el ANOVA como el HSD de Tukey muestran los resultados esperados.

Tabla A.41: Continuación: Comparaciones para el caso $t = 7$, $p = 7$, $r = 16$, pequeñas diferencias entre μ_i^s
Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		2.6515 2.104e-07	7.372e-11 0.9994	14.23 8.35e-12	T7 1.906002 a T3 1.898254 ab T5 1.890183 bc T1 1.889213 bc T2 1.888226 bc T6 1.878560 cd T4 1.865884 d
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$	B	2.8759 1.497e-08	0.0007 0.8168	161 <2e-16	T7 1.959420 a T3 1.955811 b T5 1.951976 c T1 1.951509 c T2 1.951033 c T6 1.946299 d T4 1.939870 e
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 0.99$ $\mu_3 = \mu_1 * 1.10$ $\mu_4 = \mu_1 * 0.80$ $\mu_5 = \mu_1 * 1.01$ $\mu_6 = \mu_1 * 0.90$ $\mu_7 = \mu_1 * 1.20$		2.9697 4.861e-09	0.0034 0.7486	6346 <2e-16	T7 3709468 a T3 2906679 b T5 2288209 c T1 2225278 d T2 2163471 e T6 1656326 f T4 1190708 g

1. Aumentar las repeticiones aumentó la significancia del ANOVA.
2. No se ve un efecto significativo en los resultados obtenidos al aumentar los valores de Σ .
3. Al aumentar las diferencias entre los elementos de μ_1 , aumenta la significancia del MANOVA y del ANOVA.

Tabla A.43: Continuación: Comparaciones para el caso $t = 7$, $p = 7$, $r = 4$, grandes diferencias entre μ_i 's
Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		residuals have rank 3<7	0.0181 0.9984	491.6 <2e-16	T6 15534.7536 a T2 9974.0026 b T4 5851.6337 c T7 3011.3168 d T1 1274.7465 e T5 676.6363 e T3 430.9701 e
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	B	residuals have rank 3<7	0.3452 0.0254	729 <2e-16	T6 21.51891 a T2 20.02008 b T4 18.29702 c T7 16.24784 d T1 13.67501 e T5 11.71457 f T3 10.13847 g
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		residuals have rank 3<7	0.0140 0.4913	51788 <2e-16	T6 120076.393 a T2 86452.068 b T4 57824.977 c T7 34426.139 d T1 16571.867 e T5 8711.209 f T3 4751.674 g

1. No se ve un efecto evidente de aumentar los valores valor de Σ .

Tabla A.44: Comparaciones para el caso $t = 7$, $p = 7$, $r = 16$, grandes diferencias entre μ_i^s Variables correlacionadas

μ	Σ	MANOVA: $\sim F_{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW L	ANOVA: F_{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		2.9636 5.233e-09	0.0025 3.699e-05	1775 <2e-16	T6 42.27739 a T2 36.87835 b T4 31.14913 c T7 24.98212 d T1 18.19126 e T5 13.68055 f T3 10.42406 g
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	A	2.9706 4.812e-09	0.0267 2.19e-06	10306 <2e-16	T6 1196.51285 a T2 926.89312 b T4 678.13256 c T7 453.27871 d T1 256.99377 e T5 156.14476 f T3 97.78565 g
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		2.9714 4.768e-09	0.0005 0.0002	239086 <2e-16	T6 11758.1547 a T2 9107.7999 b T4 6662.3991 c T7 4451.7765 d T1 2521.3668 e T5 1528.6101 f T3 952.9879 g

Continúa en la siguiente página

1. El efecto de aumentar el número de repeticiones (comparación con la tabla ??) fue el que los datos transformados no cumplieran con los supuestos de normalidad y homocedasticidad, y que aunque se aplicó una segunda la transformación mediante Box-Cox, cuando se tienen grandes distancias entre los elementos de μ_1 , los datos no cumplen con el supuesto de normalidad, no obstante, el ANOVA rechaza H_0 y la HSD realiza el correcto acomodo de los efectos de los tratamientos.
2. El aumento en el número de variables muestra el efecto de disminuir la significancia del MANOVA.
3. Comparando con la tabla A.40 se observa que se tiene el problema de no cumplir con el supuesto de homocedasticidad.

Tabla A.45: Continuación: Comparaciones para el caso $t = 7$, $p = 7$, $r = 16$, grandes diferencias entre μ_i 's
Variables correlacionadas

μ	Σ	MANOVA: $\sim F^{v_H, v_E}$ $Pr(> F)$	Datos trans valor-p SW	L	ANOVA: F^{v_1, v_2} $Pr(> F)$	HSD Tukey
$\mu_1 = [2.9, 8.3, 5.7, 7.6, 1.3, 6.0, 9.5]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		2.9647 5.166e-09	1.478e-05	0.0265	938.7 <2e-16	T6 354.98060 a T2 275.77956 b T4 203.09952 c T7 138.05745 d T1 82.52844 e T5 55.24253 f T3 40.51224 f
$\mu_1 = [21.6, 5.9, 14.1, 8.1, 17.5, 23.7, 5.7]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$	B	2.9696 4.872e-09	0.0260	4.619e-08	8482 <2e-16	T6 531.98812 a T2 427.35541 b T4 326.96718 c T7 231.73807 d T1 143.24269 e T5 94.66794 f T3 65.09402 g
$\mu_1 = [90.3, 57.6, 82.3, 55.0, 97.5, 60.9, 95.6]$ $\mu_2 = \mu_1 * 2.5$ $\mu_3 = \mu_1 * 0.5$ $\mu_4 = \mu_1 * 2.0$ $\mu_5 = \mu_1 * 0.7$ $\mu_6 = \mu_1 * 3.0$ $\mu_7 = \mu_1 * 1.5$		2.9714 4.766e-09	0.0002	2.991e-07	283846 <2e-16	T6 155.47507 a T2 139.00779 b T4 121.15713 c T7 101.40386 d T1 78.76507 e T5 62.93621 f T3 50.82224 g

1. Aumentar el valor de Σ no causó mucha diferencia en los valores-p del MANOVA.