

Artículo original

Estudio del crecimiento bacteriano. Enfoque de análisis de datos con medidas repetidas.

Study of bacterial growth. Analysis approach repeated measures data.

Peña G Jesús A^{1*}, Rosales Yolima², Giampaolo Orlandoni³.

¹Departamento de Bioanálisis Clínico, Cátedra de Bioestadística. ²Departamento de Microbiología y Parasitología, Cátedra de Microbiología Aplicada, ^{1,2}Facultad de Farmacia y Bioanálisis. ^{1,3}Instituto de Estadística Aplicada y Computación, Facultad de Ciencias Económicas y Sociales, ^{1,2,3}Universidad de Los Andes, Mérida, C.P. 5101, República Bolivariana de Venezuela. ³Departamento de Matemática y Estadística, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad de Santander, Bucaramanga, Santander, 4-72, Colombia.

Recibido febrero 2015 - Aceptado septiembre 2015

RESUMEN

Los análisis de datos para medidas repetidas se aplican a un conjunto de mediciones de la misma variable tomadas sobre la misma unidad experimental a lo largo de tiempo, donde los períodos de observación son establecidos por el investigador. Esta estructura de datos, por lo general causa que las varianzas y correlaciones sean diferentes en el tiempo y, por tal razón, no es apropiado analizar los datos mediante un análisis estadístico clásico. Este artículo explica la metodología adecuada haciendo uso de datos de crecimiento bacteriano en el jamón de pavo, mediante un modelo factorial con errores aleatorios dependientes del tiempo usando la estructura de covarianza ANTE(1) elegido por medio de los criterios AIC, BIC y AICC. El procesamiento de los datos se realizó con el PROC MIXED del SAS, que usa como método de estimación de los parámetros la Máxima Verosimilitud Residual (REML) por defecto, empleado para modelos de efectos mixtos, debido a que admite el modelamiento de la estructura de covarianza de los datos. Aplicar esta metodología estadística a la variable crecimiento bacteriano, nos llevó a la conclusión de que el uso de aceites esenciales con película favorece en la conservación del jamón de pavo en el tiempo.

PALABRAS CLAVE

Medidas repetidas, estructura de covarianza, crecimiento bacteriano.

ABSTRACT

Data analysis for repeated measures apply to a set of measurements of the same variable taken on the same experimental unit over time, where the observation periods are set by the researcher. This data structure usually causes the variances and correlations are different in time, and for this reason it is not appropriate to analyze the data using classical statistical analysis. This article explains the proper methodology using data of bacterial growth in turkey ham, using a factorial model with time-dependent random errors through a covariance structure model ANTE (1) chosen by the AIC criteria, BIC and AICC. The data processing was performed using PROC MIXED of SAS, using as a method of estimating the parameters Maximum Residual Likelihood (REML) by default, employed for mixed effects models because modeling supports the covariance structure of the data. Apply this to variable growth methodology, took us to the conclusion that the use of essential oils without film promotes the conservation of turkey ham in time.

KEY WORDS

Repeated measures, covariance structure, bacterial growth.

INTRODUCCIÓN

Los análisis de medidas repetidas se aplican a conjuntos de datos con múltiples medidas tomadas en secuencia de una variable dependiente sobre la misma unidad experimental, tal como, un paciente, una muestra de laboratorio, un animal, entre otros. La metodología consiste en registrar los valores de la variable dependiente como función del tiempo o espacio, por ejemplo, el conteo del crecimiento bacteriano en productos de charcutería durante una serie de días consecutivos. Este tipo de medición y otras similares pueden generar un conjunto de curvas de crecimiento que requieren ser analizadas mediante una estructura estadística distinta a los métodos clásicos más usados en el análisis de datos del área microbiológica. Los experimentos de medidas repetidas son un tipo especial de experimento factorial, donde tratamiento y tiempo son los factores principales, pero con la diferencia de que los patrones de correlación en la respuesta sobre la misma unidad experimental en el tiempo no son los mismos, lo que sí ocurre en un modelo factorial clásico, conocido como independencia de los errores aleatorios. Como en cualquier análisis factorial, lo que se pretende con las medidas repetidas es comparar tratamientos en tiempos específicos o establecer comparaciones de tiempos dentro de un mismo tratamiento.

Modelo Mixto Lineal General. Los modelos mixtos son “modelos matemáticos” o “modelos estadísticos” usados para representar observaciones que no son necesariamente independientes, esto es, modelos que describen una estructura de covarianza de los datos. Por consiguiente, incluyen tanto efectos fijos como aleatorios, de allí su nombre de modelos mixtos. Las categorías de modelos mixtos son: modelos de efectos aleatorios, modelos de coeficientes aleatorios y modelos que describen un patrón de covarianzas. Su elección dependerá de la estructura original de los datos y de los objetivos del estudio. La estimación de los parámetros de covarianza de este tipo de modelos puede llevarse a cabo mediante los métodos clásicos de estimación, pero en algunos casos existe la posibilidad de que resulten varianzas negativas. Este problema fue resuelto mediante la aplicación del Método de

Máxima Verosimilitud Residual (REML) propuesto por Patterson & Thompson (1971) [1].

El modelo mixto lineal general está dado por: $Y=X\beta+ZU+e$, donde Y es un vector columna de mediciones $n \times 1$, indicando la variable respuesta, β es un vector columna $p \times 1$ de coeficientes de regresión de efectos fijos, son los parámetros desconocidos a ser estimados a partir de los datos, X es una matriz del diseño $n \times p$ para las p variables predictoras de efectos fijos, U es un vector $q \times 1$ de efectos aleatorios no observables, no son parámetros, Z es una matriz del diseño $n \times q$ para los q efectos aleatorios, y e es un vector columna $n \times 1$ de errores aleatorios; y por último n es el número total de unidades estadísticas bajo observación. Es importante considerar algunos supuestos estadísticos que son necesarios para la validez del modelo: (1) el vector aleatorio U se asume normalmente distribuido con media $E(U)=0$ y varianza $V(U)=G$; (2) e representa la variación entre unidades experimentales, esto es, la parte no explicada por el modelo $X\beta+ZU$, y se asume normalmente distribuida con media $E(e)=0$ y varianza $V(e)=R$. Además, los vectores U y e no están correlacionados, esto es, $Cov[U,e]=0$. Como consecuencia, el vector de mediciones Y está normalmente distribuido con media $E(Y)=X\beta$ y varianza $V=V(Y)=V(X\beta+ZU+e)=ZGZ'+R$. Cabe mencionar que la matriz G es de dimensión $q \times q$ donde q es el número de parámetros de efectos aleatorios del modelo, cuya cantidad de elementos a partir de la diagonal superior o inferior de esta matriz está dada por $q(q+1)/2$ y R es una matriz de varianzas y covarianzas de los residuos cuyas dimensiones dependen del número de tiempos especificados por el investigador [2]. Para más detalles de la descripción matemática de este modelo pueden consultarse los libros de Littell (2000) [3], Brown (2006) [2], Littell (2006) [4], McCulloch (2001) [5].

Análisis de Medidas Repetidas. El análisis de medidas repetidas se aplica a situaciones donde la variable dependiente o de respuesta, sobre la misma unidad experimental, es observada en una secuencia de tiempos definidos por el investigador. En consecuencia, las mediciones sobre la misma unidad experimental tienden a estar relacionadas. Aquellas mediciones adyacentes en tiempo tienden a estar más altamente correlacionadas que las mediciones consideradas en tiempos distanciados. Además, las varianzas también cambian en el tiempo. Para un conjunto de datos con estas características, un modelo estadístico clásico no es recomendable

debido a que se estarían violando los supuestos de independencia e igualdad de varianzas de los errores aleatorios y por consiguiente las estimaciones de los parámetros obtenidas a partir de éste no son válidas. Ahora bien, al combinar los patrones de correlación y variación se da origen a un modelo de estructura de covarianza necesario para el análisis de datos con medidas repetidas, dentro de los Modelos Mixtos Lineal General, garantizando estimaciones válidas de los parámetros.

Hasta hoy día se han desarrollado cuatro métodos para analizar datos con medidas repetidas, como son: 1) análisis separado para cada tiempo de observación, 2) análisis univariante de la varianza, 3) análisis univariante y multivariante del tiempo respecto a las variables, y 4) la metodología de los modelos mixtos [6]. Littell y col. (1998) [6], describe la aplicación, ventajas y desventajas de cada método con respecto a la estructura de covarianza que surge a partir de la naturaleza del diseño experimental de donde provienen los datos. Este mismo autor, hace referencia a que se deben tomar en consideración dos pasos básicos en la ejecución de un análisis de medidas repetidas, primero, modelar la estructura de covarianza y segundo, analizar las tendencias del tiempo para los tratamientos por medio de la estimación, predicción o comparación de medias [6]. Así, cuando se aplica un análisis de medidas repetidas, el objetivo es hacer inferencias acerca de las medias de las poblaciones de donde provienen las unidades experimentales muestreadas, en términos de los efectos del tiempo y los tratamientos bajo estudio. Si el modelo es de efectos fijos, los resultados se expresan en la estimación de las diferencias entre medias de tratamiento en tiempos particulares o diferencias entre medias en tiempos diferentes para el mismo tratamiento [2].

La Estructura de Covarianza. En el análisis factorial clásico, las mediciones sobre diferentes unidades experimentales son independientes, aunque la validez de este supuesto depende del diseño del estudio. Para el análisis de medidas repetidas, también se considera la independencia entre unidades experimentales, pero el término covarianza se refiere sólo a las mediciones sobre la misma unidad experimental [3].

La estructura de covarianza representa las varianzas a partir de tiempos individuales y la correlación entre medidas en diferentes tiempos sobre la misma unidad experimental. En cualquier análisis estadístico factorial se controla la variabilidad

entre tratamientos y dentro de tratamientos (error aleatorio), sin tomar en consideración la covarianza entre mediciones dentro de la misma unidad experimental. En cambio, para un análisis de datos a partir de una estructura de covarianza, es importante considerar la covarianza de mediciones en diferentes tiempos pero dentro de la misma unidad experimental [3].

Ahora bien, para tratar éste particular ordenamiento de los datos, existen diferentes estructuras que se pueden modelar a fin de buscar un mejor ajuste del modelo a los datos y, de acuerdo a sus características, el investigador debe elegir bajo un criterio de selección la estructura de covarianza más adecuada. A continuación, se describen brevemente algunas de las estructuras disponibles:

La estructura componente simétrico (CS) asume que las mediciones en todos los tiempos de observación tienen la misma varianza, y que todos los pares de medidas sobre la misma unidad experimental tiene la misma correlación. Su importancia radica en que las mediciones presentan un elemento común que contribuye a la estructura de la covarianza: la unidad experimental. Además, mejora la estimación de los errores estándar y diferencias de medias, que pudieran resultar inválidos en un diseño factorial de errores independientes. CS es recomendable cuando se desea controlar la variabilidad debida a la unidad experimental o al número de réplicas generadas a partir de ésta. Este modelo requiere que sean estimados dos parámetros, el coeficiente de correlación y la varianza (ρ y σ^2) respectivamente [6].

La estructura autorregresivo de primer orden [AR(1)], se recomienda cuando mediciones adyacentes tienden a estar más altamente correlacionadas que las mediciones que se encuentran más distanciadas en el tiempo; esto es, la correlación entre mediciones es una función de su retardo en el tiempo. Este tipo de estructura ayuda a explicar por qué los modelos de errores independientes tienden a subestimar la varianza dentro de la unidad experimental cuando la correlación entre los errores debe tomarse en consideración. Similar al modelo CS, se requieren estimar dos parámetros ρ y σ^2 [6].

La estructura toeplitz (TOEP) es similar a la estructura AR(1) en el sentido de que pares de errores dentro de la unidad experimental, separados por un tiempo común, tienen la misma correlación. Este modelo requiere estimar q parámetros, dados por σ^2 y $q - 1$ coeficientes de correlación, donde q se corresponde con el número de tiempos a ser observados [6].

La estructura ante-dependencia o de primer orden [ANTE(1)] consiste en la combinación de las estructuras AR(1) y TOEP, pero admite espacios de tiempo desiguales y cambios de la correlación en el tiempo. Esta estructura asume que las varianzas entre mediciones cambian en el tiempo y la correlación entre pares de mediciones es el producto de correlaciones entre tiempos adyacentes, por lo que la correlación no es la misma en el tiempo. Este modelo requiere estimar $2q - 1$ parámetros [6].

La estructura autorregresiva de primer orden heterogéneo [ARH(1)] también es una combinación y modificación de las estructuras AR(1) y TOEP, permitiendo ajustar el problema de las varianzas heterogéneas en el tiempo. [4].

Ahora bien, para seleccionar el modelo con mejor ajuste de los datos, nos basamos en el criterio de información de Akaike (*Akaike information Criterion*, AIC) [7], el criterio de Información Bayesiano Schwarz (BIC) [8], y el criterio de Burnham y Anderson (AICC o AIC corregido) [9], que miden el ajuste relativo de la competición de los modelos de estructuras de covarianza. Por consiguiente, el modelo con los valores más bajos de AIC, BIC y AICC es considerado el mejor para explicar el ajuste de los datos con el mínimo número de parámetros [10]. Por lo tanto, el modelo que minimiza AIC, AICC o BIC es la mejor estructura de covarianza a ser elegida. Sin embargo, la penalidad impuesta por el criterio BIC es más exigente que la impuesta por los criterios AIC y AICC.

MATERIAL Y MÉTODOS

Con el propósito de ejemplificar esta metodología estadística, se usaron los datos provenientes de un experimento donde se investigó el efecto de tres tratamientos para el control bacteriano en el jamón de pechuga de pavo, con el fin de garantizar la mejor conservación del producto en el tiempo. El problema bajo estudio es que las películas antimicrobianas representan una novedosa alternativa en la conservación de alimentos, incluyendo los derivados cárnicos, ya que estas películas pueden ser elaboradas de sustancias naturales, biodegradables, de bajo costo y, además, incluyen en sus formulaciones compuestos antimicrobianos que prolongan la vida útil del producto y garantizan la inocuidad de los mismos para el consumidor [11].

Es así, que desde una perspectiva de seguridad alimentaria, las películas o recubrimientos

comestibles representan una importante alternativa como material de empaque, ya que previenen la contaminación del alimento y aumentan la efectividad de los métodos de conservación. Las películas y recubrimientos comestibles se utilizan en una variedad de productos alimenticios, tales como frutas y hortalizas, productos lácteos, carnes, pescados, productos de panadería y repostería, con el fin de protegerlos de los efectos del entorno o ambiente. No obstante, estos materiales de empaque pueden cumplir también otras funciones, como, por ejemplo, servir de medio de transporte de agentes antimicrobianos, agentes antioxidantes, colorantes, saborizantes y nutrientes [11].

Por otra parte, las películas comestibles se adhieren a la superficie del alimento, creando una microatmósfera con bajos niveles de oxígeno. Por lo que, algunas propiedades de barrera, la frescura y vida útil del alimento se ven favorecidas. Estas propiedades de barrera creadas por las películas y recubrimientos comestibles dependen de los compuestos empleados en su fabricación, como los polisacáridos, lípidos y proteínas o combinaciones de ellos, logrando extender la vida útil del producto y previniendo los riesgos para el consumidor.

El objetivo general del estudio microbiológico consistió en elaborar películas antimicrobianas a base de alginato, quitosano e hidroxipropilmetil-celulosa (HPMC) incorporando nisina y aceites esenciales de orégano y ajo como agentes antimicrobianos para recubrir jamón de pechuga de pavo, realizando observaciones durante un periodo de almacenamiento de 9 días a 5 ± 2 °C [11].

Para determinar la estabilidad microbiana, se prepararon muestras de lonjas de jamón de pechuga de pavo sin inocular y fueron recubiertas con películas a base de alginato de sodio (2 %), glicerol (1 %), aceite de maíz (0,2 %) e incorporadas con nisina (1000 UI/ml), aceites esenciales de ajo (0,25 %) y orégano (0,25 %); otras, se recubrieron con películas sin antimicrobianos y otras últimas se dejaron sin películas. Todas las muestras se almacenaron durante 9 días a 5 ± 2 °C, para observar el recuento de bacterias aerobias mesófilas, psicrófilas, mohos y levaduras a los 0, 1, 3, 5, 7 y 9 días. Sin embargo, para este caso particular, sólo se consideró el recuento de bacterias aerobias mesófilas (BAM).

Para realizar el recuento de bacterias, se utilizaron placas petrifilm (3M) específicas para cada uno de los grupos microbianos, siguiendo la metodología establecida en la Norma Covenin

3338:1997 902:1987 y la Norma ISO 7954:1987. Las placas de microorganismos mesófilos se incubaron a 35-37 °C durante 48 horas.

Es importante evaluar la estabilidad microbiana del jamón de pechuga de pavo, debido a que este producto una vez que es rebanado en lonjas se deteriora más rápidamente por existir una mayor exposición al oxígeno, lo cual sumado a los factores como el pH ligeramente neutro, la alta actividad de agua y el alto contenido de nutrientes, así como su manipulación, favorece el deterioro del producto disminuyendo su calidad. Por esta razón, se evaluó la estabilidad microbiana de las lonjas de jamón de pechuga de pavo recubiertas con películas a base de alginato de sodio con antimicrobianos (nisina, aceite esencial de ajo y aceite esencial de orégano) (C); con películas sin adición de antimicrobianos (B) y sin recubrir (A, control) (Sin película) de forma tal de poder comparar y evaluar la efectividad de las películas con antimicrobianos sobre la extensión de la estabilidad microbiana de este producto.

Modelo estadístico propuesto. En este estudio, la variable dependiente es el crecimiento bacteriano en el jamón de pechuga de pavo, que fue analizado a partir de lonjas de jamón recubiertas con película y otras sin película. Se consideraron dos factores que influyen en el crecimiento bacteriano: un factor entre unidades experimentales, que se refiere a las distintas formas en que fue tratado el producto, a saber: sin película (A), con película sin aceites esenciales (B) y con película con aceites esenciales (C). Un segundo factor, dentro de las unidades experimentales, es el tiempo. Por lo tanto, el factor tratamiento tiene tres niveles o categorías (A, B, C) y cada lonja de jamón del mismo producto recibió sólo uno de estos tratamientos. El factor tiempo consta de seis niveles (día 0, día 1, día 3, día 5, día 7 y día 9); además, representa las medidas repetidas (tiempos no equidistantes) para cada lonja de jamón, debido a que para cada una de ellas se tienen tres mediciones, que corresponden al crecimiento bacteriano en los diferentes intervalos de tiempo. Para cada lonja de jamón se realizaron dos réplicas.

Con base en lo anterior, el modelo estadístico propuesto es el siguiente:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}; i = 1,2,3; j = 1,2, \dots, 6; k = 1,2$$

Donde: Y_{ijk} representa el crecimiento bacteriano de la k – ésima réplica de la lonja de jamón durante el j – ésimo tiempo con el i –

ésimo tratamiento; μ es el crecimiento bacteriano medio general; α_i , es el efecto del i – ésimo tratamiento; β_j constituye el efecto del j – ésimo tiempo de observación; $(\alpha\beta)_{ij}$ es el efecto de la ij – ésima interacción entre el tratamiento y tiempo de observación; y, por último, ϵ_{ijk} representa el término de error aleatorio.

Para este caso de medidas repetidas, los errores aleatorios no están igualmente correlacionados y, por lo tanto, es necesario encontrar una estructura de covarianza de tal manera que se tengan inferencias válidas acerca de los parámetros a ser estimados. Para el ajuste del modelo estadístico y su estructura de covarianza se usó el procedimiento PROC MIXED del paquete estadístico SAS [4].

Posterior al proceso de elegir la mejor estructura de covarianza, se analizaron e interpretaron los resultados del Análisis de Varianza, donde se expone el efecto de cada factor considerado. Adicionalmente se evaluó el supuesto de normalidad de los residuos a través de los estadísticos Shapiro-Wilk, Kolmogorov-Sminrnov, Cramer-von Mises y Anderson-Draling.

RESULTADOS Y DISCUSIÓN

Los alimentos de charcutería poseen un alto riesgo de ser contaminados con bacterias a medida que transcurre el tiempo y como consecuencia, la adquisición de enfermedades por parte de quienes los consumen. Los análisis de este tipo de datos consisten en desarrollar métodos que permitan disminuir al máximo el riesgo del crecimiento bacteriano y así garantizar alimentos en buen estado y de mejor calidad.

Desde el punto de vista clínico, el crecimiento bacteriano es una característica que no solo forma parte de los alimentos, también es de interés su estudio en plantas o humanos, como por ejemplo, el crecimiento bacteriano *in vitro* del Plasma Rico en Plaquetas (PRP).

Desde el punto de vista estadístico, el crecimiento bacteriano es una variable aleatoria que requiere se describa la distribución de probabilidad a la cual se ajusta y así proponer el mejor método de análisis para este tipo de datos. De acuerdo a la naturaleza de esta variable, lo más recomendable no es el uso de los métodos estadísticos paramétricos clásicos, debido a que en algunos casos no se cumple el supuesto de homogeneidad de las varianzas.

Dentro de los métodos estadísticos clásicos de mayor uso están, el análisis de varianza (ANOVA),

pruebas de esfericidad de Mauchly, prueba de Scheffe y la Corrección de Bonferroni, como fue el caso del análisis de datos para observar el efecto inhibitor del Crecimiento Bacteriano *in vitro* del Plasma Rico en Plaquetas (PRP) [12].

Sin embargo, la observación de la variable crecimiento bacteriano como función del tiempo ha cobrado mayor importancia en investigaciones más recientes. Los chorizos son un producto de charcutería que ha sido objeto de análisis desde el punto de vista microbiológico y estadístico. Para su análisis estadístico se realizó un diseño de Bloques Completos al Azar (BCA) con arreglo factorial 2x3 con tres repeticiones para un total de 18 unidades experimentales y con medidas repetidas en el tiempo (0, 7 y 14 días) [13].

Actualmente esta metodología estadística, se aplica tanto en investigaciones agropecuarias como en ciencias de la salud, por ser las mediciones repetidas un tipo particular de datos que aportan información estadística y del fenómeno bajo análisis, debido a que la variable dependiente medida sobre un mismo individuo en diferentes instantes de tiempo, hace que no se cumplen los supuestos de independencia, igualdad de varianzas y linealidad; por lo que surge la necesidad de un método alternativo cuyo enfoque son los modelos mixtos con estructuras de dependencia, desbalance y falta de normalidad [14].

A continuación se presentan los resultados más relevantes en esta investigación. En la Figura 1, se muestra el logaritmo del crecimiento de las bacterias aerobias mesófilas [BAM(ufc/g)] en lonjas de jamón de pechuga de pavo. La representación de la Figura 1 a través del logaritmo [$\log(x+10)$] fue debido a que el crecimiento bacteriano vario entre 0 y 1250 (ufc/g). Por consiguiente, se logró un mejor perfil del grafico pero aún manteniendo diferentes grados de variabilidad y crecimiento medio para cada instante de tiempo observado (días).

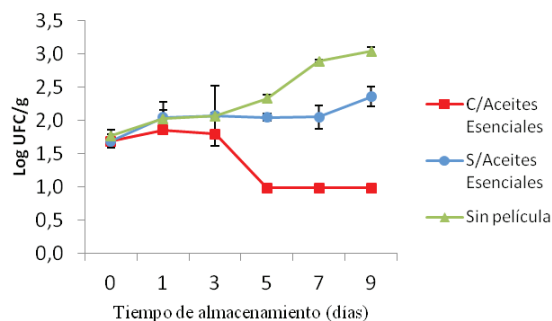


Fig. 1. Descripción del Logaritmo del crecimiento bacteriano. C/Aceites Esenciales (C): Con Aceites Esenciales, S/Aceites Esenciales (B): Sin Aceites Esenciales. Sin película (A).

El análisis se inició con la matriz de correlación mediante el Coeficiente de Correlación de Pearson, así como las varianzas en los tiempos de observación para los datos originales. Los resultados obtenidos, se muestran en la siguiente matriz de correlación:

	t_0	t_1	t_3	t_5	t_7	t_9
t_0	1,0000	0,1908	0,0595	0,5092	0,4293	0,3588
t_1	0,1908	1,0000	0,0143	-0,1408	-0,0289	0,0567
t_3	0,0595	0,0143	1,0000	0,7598	0,8064	0,8643
t_5	0,5092	-0,1408	0,7598	1,0000	0,9769	0,9428
t_7	0,4293	-0,0289	0,8068	0,9769	1,0000	0,9893
t_9	0,3588	0,0567	0,8643	0,9428	0,9893	1,0000

Los siguientes valores son las varianzas en cada tiempo de observación:

t_0	t_1	t_3	t_5	t_7	t_9
146,67	1626,67	2136,67	10546,67	142120,00	278586,67

Las correlaciones entre las mediciones cambian en el tiempo. Para los tiempos de observación más cercanos, la correlación es más alta que para los tiempos de observación más alejados. Con respecto a las varianzas, se tiene que a mayor tiempo mayor es la varianza, esto es, no hay homogeneidad de varianzas a medida que se incrementa el tiempo.

Como consecuencia, es necesario generar un modelo en función de la estructura de covarianza. Aquí se plantean las estructuras de covarianza CS, AR(1), TOEP, ANTE(1) y ARH(1). Posteriormente, se usaron los criterios de ajuste AIC, BIC y AICC para elegir la estructura de covarianza que produce el mejor ajuste a los datos, tal como se indica en la Tabla 1.

TABLA 1

Estadísticos de Ajuste para las Estructuras de Covarianza Modeladas.

Estructura de Covarianza	AIC	BIC	AICC
CS	215,5	215,0	216,3
AR(1)	214,9	214,5	215,7
TOEP	218,2	216,9	225,8
ANTE(1)	197,9	195,6	241,9
ARH(1)	205,5	204,1	216,7

De acuerdo a los criterios antes mencionados se selecciona como estructura de covarianza más adecuada para los datos originales, el modelo ante-dependencia o de primer orden [ANTE(1)].

Adicionalmente, se usó la prueba de razón de verosimilitud para contrastar si el modelo con la estructura de covarianza específica ajusta mejor que un modelo con errores independientes (modelo

nulo), esto es, $\Sigma = \sigma^2 I$. Obteniendo la información mostrada en la Tabla 2:

TABLA 2

Pruebas de Razón de Verosimilitud del Modelo Nulo vs el Modelo de Estructura de Covarianza.

Estructura de Covarianza	G1	Chi-cuadrado	p_valor > Chisq
CS	1	1,15	0,2839
AR(1)	1	1,66	0,1972
TOEP	5	6,43	0,2664
ANTE(1)	10	36,71	< 0,0001
ARH(1)	6	21,10	0,0018

El $p_valor < 0,0001$ muestra que el modelo de errores independientes, $iid N(0, \sigma^2 I)$ es claramente inadecuado y, por tanto, el modelo que hace un mejor ajuste a los datos es ANTE(1). Similarmente, ocurre con el modelo ARH(1), cuyo $p_valor = 0,0018$ indica que este modelo también pudiera ser adecuado para el ajuste de los datos. Respecto a los modelos CS, AR(1) y TOEP podemos decir que es preferible el modelo de errores independientes a uno de estos, pues presentan $p_valores$ no significativos, lo que implica que, las estimaciones obtenidas a partir de ambos modelos (modelo nulo vs modelo de estructura de covarianza) son equivalentes.

Ahora bien, para que las estimaciones a partir del modelo ANTE(1) sean válidas y confiables, se evaluó el supuesto de normalidad de los residuos, obteniendo que no siguen una distribución normal, tal como se indica en la Tabla 3. Por esta razón fue necesario aplicar una transformación logarítmica a los valores de crecimiento bacteriano, mediante $\text{Log}(x+10)$ donde x representa el crecimiento bacteriano del producto bajo análisis.

TABLA 3

Prueba de Normalidad de los Residuos.

Prueba estadística	Datos originales (p_valor)	Datos $\text{Log}(x+10)$ (p_valor)
Shapiro-Wilk	< 0,0001	0,9744
Kolmogorov-Smirnov	< 0,0100	> 0,1500
Cramer-von Mises	< 0,0050	> 0,2500
Anderson-Darling	< 0,0050	> 0,2500

Nivel de significación: $\alpha=0,05$

Como se observa en la Tabla 3, una vez transformados los datos mediante el logaritmo, el ajuste distribucional mejora considerablemente

($p_valor = 0,9744$). A partir de esta transformación, se analizan las correlaciones y varianzas de las mediciones en el tiempo. Se observa que para las mediciones en tiempos adyacentes se tiene una mayor correlación que entre mediciones de tiempos distanciados. Como se indica en la siguiente matriz de correlación para los datos transformados:

	t_0	t_1	t_3	t_5	t_7	t_9
t_0	1,0000	0,2385	0,1149	0,3325	0,3385	0,3016
t_1	0,2385	1,0000	-0,074	-0,086	-0,030	-0,028
t_3	0,1149	-0,074	1,0000	0,9547	0,9567	0,9624
t_5	0,3325	-0,086	0,9547	1,0000	0,9984	0,9955
t_7	0,3385	-0,030	0,9567	0,9984	1,0000	0,9972
t_9	0,3016	-0,028	0,9624	0,9955	0,9972	1,0000

Similarmente ocurre en las varianzas, a mayor tiempo de observación se tiene una mayor varianza del logaritmo del crecimiento bacteriano, como se muestra a continuación:

t_0	t_1	t_3	t_5	t_7	t_9
0,0517	0,1204	0,1974	6,1200	9,4471	10,8354

De acuerdo a lo mostrado por estos resultados, es necesario nuevamente determinar un modelo de estructura de covarianza que nos permita mostrar estimaciones válidas y confiables en función del logaritmo del crecimiento bacteriano.

Similar al caso de los datos originales, se selecciona la estructura de covarianza más adecuada, en base a los criterios de decisión AIC, BIC y AICC, tal como se indica en la Tabla 4.

TABLA 4

Estadísticos de ajuste para las estructuras de covarianza modeladas en base al logaritmo del crecimiento bacteriano, $\text{log}(x+10)$.

Estructura de Covarianza	AIC	BIC	AICC
CS	17,6	17,2	18,4
AR(1)	19,9	19,5	20,7
TOEP	18,7	17,5	26,4
ANTE(1)	17,3	15,0	61,3
ARH(1)	24,6	23,2	35,8

De acuerdo a la información mostrada en la Tabla 4, se selecciona como estructura de covarianza para los datos transformados el modelo ANTE(1). Este modelo coincide con el seleccionado para los datos originales. Así mismo, se incluye la prueba de razón de verosimilitud para contrastar si el modelo

con la estructura de covarianza con transformación logarítmica ajusta mejor que un modelo con errores independientes (modelo nulo), esto es, $\Sigma = \sigma^2 I$. Los resultados se muestran en la Tabla 5:

TABLA 5

Pruebas de Razón de Verosimilitud del modelo nulo vs el modelo de estructura de covarianza con transformación $\log(x+10)$.

Estructura de Covarianza	gl	Chi-cuadrado	p_valor > Chisq
CS	1	3,71	0,0540
AR(1)	1	1,42	0,2341
TOEP	5	10,6	0,0600
ANTE(1)	10	22,06	0,0148
ARH(1)	6	6,72	0,3479

En este caso se observa que, a partir de la transformación logarítmica, el modelo de estructura de covarianza cuyos datos se ajustan mejor continua siendo ANTE(1) (p_valor = 0,0148), aunque marginalmente mejoró el ajuste del modelo CS (p_valor = 0,054) y desmejoró el modelo ARH(1) (p_valor = 0,3479). Un modelo a elegir para los datos transformados podría ser el modelo CS, aunque no es el correcto, debido a que asume que las mediciones en todos los tiempos de observación tienen la misma varianza, y que todos los pares de medidas sobre la misma unidad experimental tiene la misma correlación, lo cual es contradictorio con las estimaciones de la matriz de varianzas, covarianzas y correlaciones para los datos transformados. Por lo tanto, el mejor modelo para los datos transformados mediante el logaritmo es ANTE(1). Ahora bien, en base a este nuevo modelo se evalúa el supuesto de normalidad a partir de los residuos y se verifica que este supuesto se cumple para los datos logaritmo del crecimiento bacteriano (Tabla 3).

A continuación se muestran los resultados obtenidos a partir del modelo ANTE(1) para los datos transformados en base al logaritmo, $\text{Log}(x+10)$. Los resultados del Análisis de Varianza para un diseño factorial con errores dependientes se muestran en la Tabla 6.

TABLA 6

Análisis de varianza de datos transformados en base al modelo ANTE(1).

Efecto	Gl Num	Gl Den	F	p_valor > F
Tratamiento	2	3	350,37	0,0003
Tiempo	5	15	84,79	<0,0001
Trat*Tiempo	10	15	56,45	<0,0001

De la Tabla 6 tenemos que tanto los efectos principales como la interacción son estadísticamente significativos (p_valor < 0,05) esto es, la interacción del tratamiento con el tiempo influyen en el logaritmo del crecimiento bacteriano del jamón de pechuga de pavo. Por consiguiente, los efectos principales (Tratamiento y Tiempo) tienen poco significado práctico por ser su interacción altamente significativa (p_valor < 0,0001). El conteo del número de bacterias tipo BAM (ufc/g) por nivel de tratamiento varía considerablemente para cada uno de los tiempos de observación definidos por el investigador del área microbiológica.

La Tabla 7, correspondiente a las "Estimaciones de Parámetros de Covarianza" muestra la estimación de los once (11) parámetros del modelo ANTE(1). En ella se observa que los únicos resultados estadísticamente significativos son los coeficientes de correlación correspondientes a los periodos de observación 3 y 5 (p_valor < 0,0001) respectivamente. Esto implica que los cambios en el tiempo del logaritmo del crecimiento bacteriano, $[\log(x+10)]$, es una condición importante para el modelo propuesto en el análisis de estos datos en particular.

TABLA 7

Estimación de parámetros de covarianza.

Parm Cov	Est	E_E	Valor Z	Pr Z	Inf	Sup
Var(1)	0.07	0.06	1.22	0.11	0.02	0.99
Var(2)	0.20	0.16	1.22	0.11	0.06	2.76
Var(3)	0.02	0.02	1.22	0.11	0.01	0.31
Var(4)	0.02	0.01	1.22	0.11	0.01	0.24
Var(5)	0.08	0.07	1.22	0.11	0.03	1.11
Var(6)	0.07	0.06	1.22	0.11	0.02	1.00
Rho(1)	0.31	0.52	0.60	0.55	-0.71	1.33
Rho(2)	0.06	0.58	0.11	0.92	-1.07	1.19
Rho(3)	-0.96	0.04	-21.76	<.0001	-1.05	-0.87
Rho(4)	0.60	0.37	1.62	0.11	-0.13	1.33
Rho(5)	0.95	0.06	16.79	<.0001	0.84	1.06

Dentro del conjunto de resultados a posteriori al rechazo de la hipótesis nula para el efecto de la interacción, se muestran algunos resultados parciales como por ejemplo, el comportamiento de los tratamientos bajo análisis a lo largo del tiempo, cuyos resultados se muestran en la Tabla 8. Por lo

tanto, evaluando cuáles pares de medias referentes a la interacción de los factores difieren entre sí, se tiene que manteniendo fijo el tiempo de observación el investigador puede decidir en qué momento del tiempo el producto presenta el mayor o menor crecimiento bacteriano, así como encontrar el tratamiento que garantiza la mayor conservación del producto.

Ahora bien, de acuerdo a la Tabla 8, para cada uno de los dos primeros periodos de observación (día 0, día 1) no se encontraron diferencias estadísticamente significativas entre los tratamientos A, B y C respecto al logaritmo del crecimiento bacteriano del producto bajo evaluación ($p_valor > 0,05$). Sin embargo, a partir del tercer periodo de observación (día 3) se comienza a ver el efecto de los tratamientos para el control bacteriano. Específicamente, para el día 3, no se observaron diferencias estadísticamente significativas entre los tratamientos A y B ($p_valor = 0,1200$). Aunque para ese mismo día de observación, se encontraron diferencias estadísticamente significativas entre los tratamientos A y C ($p_valor < 0,0001$) y entre los tratamientos B y C ($p_valor = 0,0004$). Esto es, en promedio se observó un mayor crecimiento bacteriano en las muestras de jamón de pavo tratadas sin película que en las muestras tratadas sin película pero con aceites esenciales. La presencia de aceites esenciales en la muestra de jamón tiene un efecto favorable, en cuanto al crecimiento bacteriano durante el tercer periodo de observación, siendo las estimaciones del logaritmo del crecimiento bacteriano para las comparaciones por pares muy similares, entre 0,24 y 0,93.

TABLA 8

Análisis de Varianza con datos transformados para Estructura de Covarianza ANTE(1).

Tiempo	Trat	Est	EE	gl	t	p_valor
Día 0	A vs B	0,19	0,27	15	0,70	0,4976
	A vs C	0,19	0,27	15	0,70	0,4976
	B vs C	0,00	0,27	15	0,70	1,0000
Día 1	A vs B	-0,06	0,45	15	-0,13	0,8964
	A vs C	-0,08	0,45	15	-0,18	0,8580
	B vs C	-0,02	0,45	15	-0,05	0,9611
Día 3	A vs B	0,24	0,15	15	1,65	0,1200
	A vs C	0,93	0,15	15	6,24	< 0,0001
	B vs C	0,68	0,15	15	4,60	0,0004
Día 5	A vs B	15,46	0,13	15	11,87	< 0,0001
	A vs C	58,69	0,13	15	41,24	< 0,0001
	B vs C	38,23	0,13	15	29,36	< 0,0001
Día 7	A vs B	19,22	0,28	15	6,81	< 0,0001
	A vs C	66,60	0,28	15	23,60	< 0,0001
	B vs C	47,38	0,28	15	16,79	< 0,0001
Día 9	A vs B	15,54	0,27	15	5,78	< 0,0001
	A vs C	69,95	0,27	15	26,04	< 0,0001
	B vs C	54,41	0,27	15	20,25	< 0,0001

A: Muestras de jamón de pavo tratadas sin película.

B: Muestras de jamón de pavo tratadas con película sin aceites esenciales.

C: Muestras de jamón de pavo tratadas con película con aceites esenciales.

Para el cuarto periodo de observación (día 5), se encontraron diferencias significativas para todos los pares de combinaciones de los tratamientos ($p_valor < 0,0001$). La estimación del logaritmo del crecimiento bacteriano para las comparaciones por pares es mayor que para el periodo de observación 3, este oscila entre 15,5 y 58,7. Al comparar los tratamientos A y B, hubo un mayor crecimiento bacteriano medio con el tratamiento A, pues su diferencia es positiva; al comparar A y C junto con B y C, se observó en promedio un mayor crecimiento bacteriano con los tratamientos A y C. Por lo tanto, para el cuarto periodo de observación el tratamiento que garantizó la mejor conservación del producto bajo análisis es el C, es decir, muestras de jamón de pavo tratadas con películas con aceites esenciales.

Similarmente ocurrió para los periodos de observación cinco (día 7) y seis (día 9), las comparaciones por pares de los tratamientos resultaron estadísticamente significativas ($p_valor < 0,0001$), siendo para ambos periodos el tratamiento C quien garantiza la mayor conservación del producto bajo análisis en el tiempo.

CONCLUSIONES

Si la estructura original de los datos es de mediciones repetidas sobre la misma unidad estadística en función del tiempo o el espacio para la variable dependiente, entonces el análisis de los datos debería ser a partir de una Estructura de Covarianza debidamente seleccionada, en la que previamente se debe evaluar la matriz de varianzas y covarianzas para determinar el cambio de la varianza en el tiempo y considerar si es razonable suponer que las varianzas en el tiempo son iguales. Luego, se evalúa la matriz de correlaciones para determinar si estas cambian en el tiempo. En el primer caso, es razonable proponer una estructura de covarianza para el ajuste de los datos al modelo. Si para algunas de las unidades experimentales no todas sus mediciones en el tiempo están registradas, esta metodología estadística tiene la ventaja de ejecutar el análisis sin que esta condición afecte los resultados obtenidos.

Mediante el análisis de datos de medidas repetidas, a partir del modelo de Estructura de Covarianza ANTE(1), se determinó que el tiempo afecta significativamente el logaritmo del crecimiento bacteriano del jamón de pechuga de pavo. Y a medida que se incrementa el tiempo de observación, el crecimiento bacteriano en muestras de jamón de

pavo tratado sin película y con película pero sin aceites esenciales fue cada vez mayor, mientras que al tratar las muestras con película con aceites esenciales no se observó crecimiento bacteriano a partir del tercer periodo de observación. El logaritmo del crecimiento bacteriano depende simultáneamente tanto del tratamiento como del tiempo de observación, cuyos periodos fueron establecidos por el investigador. Por lo que, el crecimiento bacteriano en productos de esta naturaleza es inevitable; sin embargo, tratar el producto con película con aceites esenciales garantiza una mayor conservación del mismo en el tiempo.

El análisis de medidas repetidas es una metodología estadística apropiada para el estudio de datos de procesamientos de alimentos donde el factor tiempo siempre está presente en la mayoría de los casos, como determinante del comportamiento microbiológico de las bacterias.

REFERENCIAS BIBLIOGRÁFICAS

- [1] Kendward M, Roger J. Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. *Biometrics*. 1997; 53: 983-997.
- [2] Brown H, Prescott R. *Applied Mixed Models in Medicine*. New York: Jhon Wiley & Sons, Inc; 2006.
- [3] Littell R, Pendergast J, Natarajan R. Tutorial in Biostatistics: Modelling covariance structure in the analysis of repeated measures data. *Statist Med*. 2000; 19: 1793-1819.
- [4] Littell R, Milliken G, Stroup W, Wolfinger R, Schabenberger O. *SAS for Mixed Models*. SAS Institute Inc; 2006.
- [5] McCulloch C, Searle S. *Generalized, Linear, and Mixed Models*. New York: Jhon Wiley & Sons, Inc; 2001.
- [6] Littell R, Henry P, Ammerman CB. Statistical Analysis of Repeated Measures Data Using SAS Procedures. *J Anim Sci*. 1998; 76(4): 1216-1231.
- [7] Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974; 19(6):716-723.
- [8] Schwarz G. Estimating the dimension of a model. *Ann Stat*. 1978; 6:461-464.
- [9] Burnham K, Anderson D. *Model Selection and Inference: A Practical Information-Theoretic Approach*. New York: Springer-Verlag; 1998.
- [10] Posada S, Rosero R. Comparación de Modelos Matemáticos: una aplicación en la evaluación de alimentos para animales. *Revista Colombiana de Ciencias Pecuarias, Universidad de Antioquia*. 2007; 20:141-148.
- [11] Rosales Y. Desarrollo de películas Biodegradables con propiedades antimicrobianas como una alternativa novedosa para la conservación de Jamón Cocido y Pechuga de Pavo. [Tesis Doctoral]. Universidad Central de Venezuela, Facultad de Ciencias, Instituto de Ciencia y Tecnología de Alimentos. Doctorado en Ciencias y Tecnología de Alimentos; 2013.
- [12] Huapaya C, Noriega J. Efecto inhibitor del Crecimiento Bacteriano in vitro del Plasma Rico en Plaquetas. *Kiru*, 2008; 5(1): 2-23.
- [13] Peñalba F. Evaluación del efecto de tratamientos de sal y soluciones antimicrobiales en las propiedades físicas, microbiológicas y sensoriales de la tripa de cerdo en el chorizo criollo Zamorano. [Tesis de Grado]. Honduras, 2011.
- [14] Gómez S, Torres V, García Y, Navarro JA. Procedimientos estadísticos más utilizados en el análisis de medidas repetidas en el tiempo en el sector agropecuario. *Rev Cub Cs Agr*. 2012; 46(1): 1-7.