

# Hormonas en cerdas durante la lactancia

Hormonas en cerdas durante la lactancia

Elección del mejor modelo para el análisis de experimentos con medidas repetidas en tiempo

The objective of this study was to determine the effect of different covariance structures on the statistical significance and on the fixed effect estimates, using as an example the data of four hormones in a repeated measures experiment. (source: LRRD)

## Resumen

El objetivo de este estudio fue determinar el efecto de diferentes estructuras de covarianzas en la significancia estadística y en los estimadores de los efectos fijos, tomando como ejemplo los datos de cuatro hormonas en un experimento con mediciones repetidas. El estudio se realizó en una granja comercial de tres sitios con 2400 vientres en la Piedad Michoacán, México. Se utilizaron 32 cerdas de tercer parto pertenecientes a tres grupos genéticos (11 cerdas AZ, 11 cerdas B y 10 cerdas del grupo genético N) a las cuales se les tomó 10 ml de sangre para determinación de las concentraciones de las hormonas: progesterona (P4), estradiol (E2), prolactina (PRL) e insulina (INS). Las muestras de sangre se tomaron a los 3, 6, 9, 12, 15 días de lactancia. Las determinaciones de las hormonas se hicieron por medio de radioinmonoanálisis. Los datos fueron analizados utilizando anovas para medidas repetidas y procedimientos de modelos mixtos. Los anovas fueron realizados utilizando las opciones RANDOM o TEST del programa SAS. Los modelos mixtos consideraron nueve estructuras de covarianzas: componentes de varianza (CV), simetría compuesta (CS), autogresiva (AR(1)), antedependencia (ANTE(1)), Toeplitz, no estructurada (UN), CS heterogénea, AR(1) heterogénea y Toeplitz heterogénea.

Las mejores estructuras de covarianzas que describieron los datos fueron: ANTE(1), UN, CS y CS para P4, E2, PRL e INS, respectivamente. Se observaron diferencias en los niveles de significancia de los efectos fijos, principalmente de grupo genético, dependiendo de la estructura de covarianzas utilizada. Por ejemplo, para P4 el intervalo de las significancias para grupo genético fue de  $P = 0.0063$  para la estructura CV a  $P = 0.1318$  para Toeplitz. El efecto del tiempo de lactancia sobre las concentraciones de hormonas fue altamente significativo para todas las estructuras ( $P < 0.001$ ). Las medias de cuadrados mínimos generalizados y errores estándares (EE), fueron asimismo afectados por el tipo de matriz de covarianzas. Para P4 los EE más pequeños correspondieron a la estructura más simple (CV) y los mayores a las estructuras AR(1) y Toeplitz.

## Introducción

Los experimentos con medidas repetidas son frecuentes en las investigaciones pecuarias (Segura y Osorio 2002, ZooBell et al 2003, Wang y Goonewardene 2004) y se refieren a aquellos casos en que se hacen varias mediciones en una misma unidad experimental. En la mayoría de los casos, las múltiples observaciones por unidad experimental son tomadas a través del tiempo, aunque podrían ser en espacio, como es el caso de los experimentos en parcelas divididas. En el análisis de datos de experimentos con mediciones repetidas en tiempo, las suposiciones usuales acerca de independencia y homogeneidad de varianzas normalmente no

son válidas, ya que, a menudo, las medidas hechas en un mismo animal están correlacionadas entre sí y las varianzas entre mediciones pueden ser diferentes. La metodología de modelos mixtos permite analizar correctamente y eficientemente los datos de experimentos con medidas repetidas, a través del modelaje de la estructura de covarianzas que consideren las correlaciones entre medidas repetidas y la presencia de varianzas heterogéneas. El ignorar la importancia de la correlación dentro de sujetos utilizando modelos de efectos fijos (procedimientos ANOVA o GLM) o modelos mixtos con estructuras de covarianzas muy simple, podrían aumentar la tasa de error tipo I (rechazo de la hipótesis nula cuando debería ser aceptada) para la prueba de los efectos fijos del modelo, mientras que un modelo muy complicado conduciría a un sacrificio en el poder y eficiencia de la prueba para los efectos fijos (Littell et al 2000; Wang y Goonewardene 2004). La elección de la estructura de covarianzas apropiada resulta en estimadores más eficientes de los efectos fijos del modelo y consecuentemente pruebas más robustas de los efectos de las medidas repetidas. Sin embargo, la selección de la estructura de covarianzas más apropiada y parsimoniosa es complicada debido a la existencia de un gran número de estructuras posibles (SAS, 2000). Los criterios más utilizados para la elección de la mejor estructura son el criterio de información de Akaike y el criterio de información Bayesiano de Schwarz (Littell et al 1998; Wang y Goonewardene 2004).

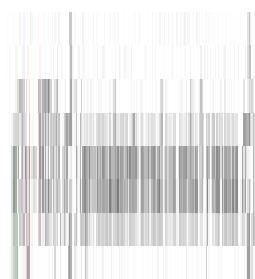
El objetivo de este estudio fue determinar el efecto de la estructura de covarianzas en los niveles de significancia de los efectos fijos y en los estimadores de los efectos fijos, tomando como ejemplo los datos de un experimento con mediciones repetidas de las hormonas progesterona, estrógenos, prolactina e insulina.

## Material y métodos

### Análisis estadístico

El modelo estadístico que describió los resultados de las hormonas fue:

donde:



Los datos de las hormonas se analizaron utilizando procedimientos de análisis de varianza para mediciones repetidas, que utilizan el cuadrado medio del error de unidad experimental dentro de tratamiento para probar la hipótesis de igualdad de tratamientos y modelos mixtos, considerando diferentes estructuras de covarianza (correlaciones) de las medidas repetidas. Las estructuras de covarianzas evaluadas fueron: componentes de varianza (CV) la cual es la estructura más simple ya que asume que todas las observaciones son independientes una de otra y que no hay correlación (covarianza) entre pares de observaciones, aun cuando las mediciones se hicieron en el mismo individuo.

Esta estructura tiene varianzas iguales en la diagonal principal

$$\begin{pmatrix} \sigma^2 & & & \\ & \sigma^2 & & \\ & & \sigma^2 & \\ & & & \sigma^2 \end{pmatrix}$$

y ceros en los demás elementos de la matriz. La estructura de simetría compuesta (CS) considera varianzas iguales en la diagonal principal e iguales covarianzas en los demás elementos de la matriz, es decir asume una misma correlación entre observaciones independientemente de la distancia entre tiempos de medición. La estructura autoregresiva de primer orden (AR(1)) considera varianzas homogéneas

$$\begin{pmatrix} \sigma^2 & & & \\ & \sigma^2 & & \\ & & \sigma^2 & \\ & & & \sigma^2 \end{pmatrix}$$

. Asimismo considera que la correlaciones entre dos medidas adyacentes son iguales

$$\begin{pmatrix} \sigma^2 & & & \\ & \sigma^2 & & \\ & & \sigma^2 & \\ & & & \sigma^2 \end{pmatrix} + \begin{pmatrix} \sigma^2 & & & \\ & \sigma^2 & & \\ & & \sigma^2 & \\ & & & \sigma^2 \end{pmatrix}$$

y que éstas declinan exponencialmente con la distancia entre mediciones

$$\begin{pmatrix} \sigma^2 & & & \\ & \sigma^2 & & \\ & & \sigma^2 & \\ & & & \sigma^2 \end{pmatrix} + \begin{pmatrix} \sigma^2 & & & \\ & \sigma^2 & & \\ & & \sigma^2 & \\ & & & \sigma^2 \end{pmatrix}$$

L  
a  
e  
s  
t  
r  
u  
c  
t

ura de TOEPLITZ (TOEP) es similar a la AR(1) en que todas las mediciones próximas a la siguiente tienen la misma correlación, medidas separadas por una medición tienen la misma correlación, pero diferente de la primera, medidas separadas por dos mediciones tienen la misma correlación pero diferente de la primera y segunda y así sucesivamente. La estructura de covarianzas de primer orden y covarianzas dependientes (ANTE(1)) permite varianzas diferentes para cada medición en tiempo y correlaciones y covarianzas diferentes entre diferentes pares de medidas. La estructura de covarianzas no estructurada (UN) permite que cada término de la matriz de covarianzas sea diferente. Asimismo, se corrieron las versiones: CSH, ARH(1), TOEPH que son las versiones de las estructuras CS, AR(1) y TOEP pero considerando varianzas heterogéneas. También se probaron las estructuras CS + AR(1), CS+ARH(1) y CS+TOEP, las cuales se desecharon por tener peores valores de AICC (criterio de información de Akaike) y BIC (criterio de información bayesiano) que algunas de las estructuras anteriormente citadas.

Los anovas con medidas repetidas se corrieron utilizando las opciones RANDOM y TEST del procedimiento GLM y las nueve estructuras de covarianzas evaluadas, se corrieron utilizando la sentencia REPEATED del procedimiento MIXED (SAS 2000) la cual controla la estructura de covarianzas de los errores residuales. Las sentencias SAS utilizadas para correr los diferentes modelos se proporcionan en el Apéndice. Los datos de concentraciones de hormona se transformaron usando la función raíz cuadrada para aproximarlos a la normalidad, pero al no afectar la jerarquía de las estructuras comparadas se prefirió presentar los resultados de los datos sin transformar. Los datos transformados redujeron en 2 a 3% los niveles de significancia. La hipótesis de homogeneidad de varianzas entre mediciones se probó utilizando la prueba de Bartlett (Snedecor y Cochran 1978).

La selección de la mejor estructura de covarianzas se basó en la comparación de los criterios AICC y BIC. Estos estadísticos son funciones del logaritmo de la verosimilitudes y cuando se comparan dos estructuras aquella con los valores de los criterios más bajos indican una mejor estructura.

## Resultados

Los resultados de los ANOVAs con medidas repetidas fueron similares a los obtenidos con la

estructura de covarianzas CS del procedimiento MIXED. Lo que indica que para experimentos balanceados o ligeramente desbalanceados (como en este estudio) con varianzas homogéneas y similares correlaciones entre pares de medidas repetidas, los procedimientos de cuadrados mínimos (utilizados por GLM) proporcionan los mismos resultados que los procedimientos de máxima verosimilitud (utilizados por MIXED). Los resultados de los criterios de información (AICC y BIC) y los niveles de significancia obtenidos por los efectos de grupo genético y número de medición, para las estructuras comparadas, se presentan en los Cuadros 2-5. Con base en los criterios AICC y BIC las mejores estructuras de covarianzas que describieron los datos fueron ANTE(1), UN, CS y CS para las hormonas P4, E2, PRL e INS, respectivamente. Asimismo se observa, que las hormonas PRL e INS pudieron ser analizadas utilizando ANOVAs con medidas repetidas.

Las medias de cuadrados mínimos generalizados y errores estándares (EE), fueron asimismo afectados por el tipo de estructura de matriz de covarianzas. Por brevedad sólo se presentan los datos para la hormona progesterona, sin embargo las medias y EE para otras hormonas también se vieron afectadas por la estructura de covarianzas seleccionada. En el Cuadro 6 se observa que los EE más pequeños correspondieron a la estructura de covarianzas más simple (CV) y los mayores a las estructuras AR(1) y TOEP.

## Discusión

### Selección de la mejor estructura de covarianzas

Como se aprecia en los Cuadros 2-5, con base en los criterios AICC y BIC, la estructura que peor describe los datos fue la estructura de covarianzas denominada componentes de varianzas (VC), la cual corresponde a la estructura más simple. Esta estructura da valores similares a los obtenidos por el procedimiento GLM, considerando las medidas repetidas como independientes, por lo que no se recomienda su uso, para experimentos con medidas repetidas en tiempo. Sin embargo, para diseños de parcelas divididas en espacio es el más apropiado (Gil 2001, Segura datos inéditos). Para P4 la mejor estructura de covarianzas fue ANTE(1). Esta estructura permite varianzas heterogéneas y correlaciones diferentes entre pares de mediciones que disminuyen con el tiempo, lo cual es compatible con la matriz de covarianzas del Cuadro 1. Algunos estudios con mediciones repetidas de progesterona han analizado la información usando modelos mixtos con estructuras simples, posiblemente CV que es la estructura predeterminada por el programa SAS (Mejía-Guadarrama et al 2002; Stegner et al 2004) lo cual podría haber resultado en el rechazo de la hipótesis nula cuando debería ser aceptada y con estimadores de efectos fijos con EE más pequeños. Wang and Goonewardene (2004) encontraron que la estructura ANTE(1) fue la mejor en un experimento balanceado de crecimiento en novillas en la que compararon cinco estructuras de covarianzas.

Para E2 la mejor estructura fue UN lo que indica diferentes varianzas para cada medición y diferentes correlaciones entre medidas repetidas. Stegner et al (2004) en un experimento en el que midieron E2 en vacas utilizaron el procedimiento MIXED, no mencionan el tipo de estructura de covarianzas por lo que muy probablemente utilizaron la estructura CV. La mejor estructura de covarianzas para las hormonas PRL e INS fue CS, la cual supone homogeneidad de varianzas y covarianzas. Mejía-Guadarrama et al (2002) utilizaron el procedimiento MIXED para analizar la concentración de insulina en cerdas después del destete, con mediciones cada 15 min de 0815 a 1615 h pero no mencionan que estructura de covarianzas utilizaron por lo que posiblemente

hayan utilizando la estructura predeterminada de SAS y sus resultados pudieran estar sesgados a menos que tuvieran varianzas homogéneas y no existieran correlaciones entre medidas repetidas.

La estructura de la matriz de covarianzas cambia para cada tipo de estudio, la frecuencia de mediciones, cuando existen datos perdidos o pérdida de combinaciones de tratamientos, varianzas heterogéneas etc. Esto indica que utilizar la estructura CV predeterminada por el programa SAS y otros programas pudiera resultar en sesgo en los niveles de significancia de la prueba de F y en los estimadores de efectos fijos.

## Significancia de los efectos fijos

Como se mencionó anteriormente, los niveles de significancia de las pruebas de hipótesis (principalmente de grupo genético) se afectaron por el tipo de estructura de covarianzas utilizada. Esto demuestra la importancia de elegir la correcta estructura de covarianzas para la prueba de hipótesis de los efectos fijos en los experimentos. Similarmente, Littell et al (1998) y Wang y Goonewardene (2004) encontraron diferencias en los niveles de significancia según la estructura de covarianzas utilizada. La estructura CV rechazó más fácilmente la hipótesis de igualdad de grupos genéticos en comparación con la estructura ANTE(1), para el caso de la hormona P4, como ejemplo. Sin embargo, el uso de estructuras de covarianza más complejas no es garantía de mayor exactitud y precisión en los experimentos con mediciones repetidas. Por ejemplo, según Littell et al (1998) una desventaja de la estructura UN es que no permite tendencias en las covarianzas con el tiempo y a menudo resulta en patrones erráticos de los errores estándares. La estructura CV, proporciona resultados similares a los obtenidos con los procedimientos GLM sin considerar la existencia de medidas repetidas.

## Medias y errores estándares (EE)

La medias de cuadrados mínimos fueron ligeramente diferentes para las distintas estructuras de covarianzas debido a que el experimento no estaba balanceado. Experimentos balanceados y sin datos perdidos dan similares medias de cuadrados mínimos (Cnaan et al 1997, Wang y Goonewardene 2004). Sin embargo, los EE de las medias de cuadrados mínimos para datos balanceados o no, son diferentes para las distintas estructuras de covarianzas, porque son ajustadas por los parámetros de covarianzas en el modelo mixto (Littell et al 1998). Los errores estándares para CV fueron los más bajos debido a que esta estructura supone varianzas constantes con el tiempo y observaciones independientes.

## Conclusiones

Los resultados de este estudio indican que la elección de la estructura de covarianzas para el análisis de datos de experimentos con medidas repetidas en tiempo afecta principalmente el nivel de significancia de la prueba de F y los errores estándares de las medias de cuadrados mínimos. El análisis de datos de las hormonas aquí estudiadas pero con frecuencias de medición diferentes o por periodos más cortos o prolongados a los aquí investigados, así como el estudio de diferentes hormonas u otras variables de respuesta, requería el examen de diferentes estructura de covarianzas para encontrar la más apropiada.

# Literatura citada

**Cnaan A, Laird N M and Slasor P 1997**

Using the general linear mixed model to analyze unbalanced repeated measures and longitudinal data. *Statistics in Medicine* 16:2349-2380.

**Gil J L 2001**

Comparación de los procedimientos GLM y MIXED del SAS para analizar diseños de parcelas divididas con bloques al azar *Zootecnia Tropical* 19(1):43-58.

**Littell R C, Henry P R and Ammerman C B 1998**

Statistical analysis of repeated measures data using SAS procedures. *Journal of Animal Science* 76:1216-1231.

**Mejia-Guadarrama C A, Pasquier A, Dourmad J Y, Prunier A and Quesnel H 2002**

Protein (lisine) restriction in primiparous lactating sows: Effects on metabolic state, somatotropic axis, and reproductive performance after weaning. *Journal of Animal Science* 80:3286-3300.

**SAS 2000**

SAS/STAT User's Guide. Version 8. SAS Institute Inc. Cary NC, USA.

**Pérez Sánchez E R, Gutiérrez Vazquez E, Montes-Pérez R C, Aké López R, Centurión Castro F y Segura Correa J C 2005**

(en prensa) Efecto del genotipo, peso al parto y concentración de prolactina en el intervalo destete-estro en cerdas lactantes. *Veterinaria México* (Enviado).

**Wang Z and Goonewardene L A 2004**

The use of Mixed models in the analysis of animal experiments with repeated measures data. *Canadian Journal of Animal Science* 84:1-11.

**ZoBell D R, Goonewardene L A, Olson K C, Stonecipher C A and Weidmeier R D 2003**

Effects of feeding wheat middlings on production, digestibility, ruminal fermentation and carcass characteristics in beef cattle. *Canadian Journal of Animal Science* 83:551-557.

## Source



Instituto de Investigaciones Agropecuarias y Forestales, Universidad Michoacana de San Nicolás de Hidalgo, km 9.5, carretera Morelia - Zinápecuaro, municipio de Tarímbaro, Michoacán

\*Instituto de Investigaciones Agropecuarias y Forestales, Universidad Michoacana de San Nicolás de Hidalgo, km 9.5, carretera Morelia - Zinápecuaro, municipio de Tarímbaro, Michoacán

\*\* Facultad de Medicina Veterinaria y Zootecnia, Universidad Autónoma de Yucatán. Km 15.5 carretera Mérida-Xmatkuil, Mérida, Yucatán

Sí