

JORNADAS NACIONALES

8, 9 y 10 / NOVIEMBRE / 2017

La ESTADÍSTICA como herramienta CIENTÍFICA



UNIVERSIDAD DE JAÉN



DETECCIÓN DE LA MULTICOLINEALIDAD NO ESENCIAL EN EL MODELO DE REGRESIÓN LINEAL GENERAL

Ainara Rodríguez Sánchez⁽¹⁾, Román Salmerón Gómez⁽²⁾, Catalina García García⁽³⁾

(1) Programa de Doctorado en Ciencias Económicas y Empresariales- Técnicas Cuantitativas Avanzadas en el Ámbito Económico y Empresarial, Universidad de Granada, Campus Universitario de La Cartuja, 18071 (Granada), e-mail: ainararodriguez@correo.ugr.es

(2) Departamento de Métodos Cuantitativos para la Economía y la Empresa, Universidad de Granada, Campus Universitario de La Cartuja, 18071 (Granada) , e-mail: romansg@ugr.es

(3) Departamento de Métodos Cuantitativos para la Economía y la Empresa, Universidad de Granada, Campus Universitario de La Cartuja, 18071 (Granada) , e-mail: cbgarcia@ugr.es

RESUMEN

Cuando existe una relación lineal aproximada entre la variable independiente que representa a la constante del modelo de regresión lineal y el resto de variables independientes, se produce un problema denominado multicolinealidad no esencial o micro. Este tipo de problema afecta a las estimaciones tanto como el ya conocido problema de multicolinealidad esencial (que se produce si existe relación entre el resto de variables independientes del modelo). Para detectar este problema de multicolinealidad entre el intercepto y las variables independientes, el Factor Inflación de la Varianza (FIV), no puede ser usado, debido a que no tiene en cuenta el término independiente del modelo. Por este motivo, se analiza como detectar este problema partiendo del Número de Condición (NC).

Palabras clave: Multicolinealidad, variable independiente, regresión lineal múltiple, constante.

ABSTRACT

When exist a relationship between the independent variable that represents to the constant of the lineal regression model and the rest of independent variables, it is said that there is a problem called nonessential or micro multicollinearity. This type of problem affects to the estimations as well as the already know essential multicollinearity problem (that exist when the rest of independent variables of the model are relationship). For diagnose this multicollinearity problem between intercept and independent variables, the Variance Inflator Factor (VIF), can not be used, due to not account with the independent term of the model. For this reason, it is analysed like to detect this problem starting from the Condition Number (CN).

Keywords: Multicollinearity, independet variable, multiple linear regression, constant.

1. INTRODUCCIÓN

El modelo de regresión lineal múltiple estudia la relación lineal entre una variable respuesta o explicada (y) y un conjunto de variables regresoras o explicativas ($\mathbf{X}_1, \dots, \mathbf{X}_p$, $p \geq 1$). Cuando existe una combinación lineal exacta entre dos o más variables regresoras nos encontramos ante un problema denominado multicolinealidad exacta, en este caso la estimación por Mínimos Cuadrados Ordinarios (MCO) no será posible. Por otro lado, por ejemplo en (Uriel, E. et al¹, 1994), se define el problema de la multicolinealidad aproximada como la existencia de relaciones lineales aproximadas entre los regresores del modelo lineal básico, cuando los estimadores obtenidos por MCO y la precisión de estos se ven sencillamente afectados.

Si se considera el siguiente modelo lineal general (MLG):

$$\mathbf{y}_t = \beta_1 + \beta_2 \mathbf{X}_{2t} + \beta_3 \mathbf{X}_{3t} + \dots + \beta_p \mathbf{X}_{pt} + \mathbf{u}_t \quad t = 1, \dots, n. \quad (1)$$

donde \mathbf{u} representa la perturbación aleatoria del modelo, $\mathbf{X}_{n \times p}$ recoge las observaciones de las variables independientes (encontrándose la constante representada por una variable independiente \mathbf{X}_1 , que toma el valor 1) e $\mathbf{Y}_{n \times 1}$ las observaciones de la variable dependiente.

Teniendo en cuenta el modelo (1), podría existir un posible problema de relación lineal entre la variable explicativa que representa el término constante (\mathbf{X}_1) y las demás variables independientes de la matriz \mathbf{X} . Según (Aiken, L. et al², 1991) si existe esta relación nos encontramos ante un problema denominado multicolinealidad no esencial. En (Marquardt, D. et al³, 1975), se defiende el uso de la estandarización cuando en el modelo existe un término constante, debido a que centrando los datos se tiende a solucionar el problema de la multicolinealidad no esencial. Por otro lado, los problemas de multicolinealidad producidos por relaciones entre el resto de variables se conocen como esenciales.

En (Iacobucci, D. et al⁴, 2016) a este tipo de multicolinealidad no esencial se le denomina también multicolinealidad micro. Además, también se tiene la multicolinealidad macro o esencial. Los investigadores encargados de estudiar la multicolinealidad a nivel macro, no hacen caso del posible problema de multicolinealidad a nivel micro, porque según estos no afecta a los resultados. Sin embargo, como se demuestra en el siguiente apartado los resultados de las estimaciones se ven afectados.

¹ Contreras, D.; Moltó, L.; Peiró, M.

² West, S.; Reno, R.

³ Snee, R.

⁴ Schneider, M.J.; Popovich, D.L.; Bakamistos, G.A.

El objetivo del presente trabajo consistirá en detectar cuando existe este tipo de multicolinealidad no esencial o micro. Para esto existen dos herramientas muy utilizadas: el FIV y el NC.

Según (Belsley, D.A. et al⁵, 2004) para la matriz $\mathbf{X}_{n \times p}$ se define el NC como:

$$\mathbf{k}(\mathbf{X}) = \frac{\mu_{\max}}{\mu_{\min}} = \sqrt{\frac{\epsilon_{\max}}{\epsilon_{\min}}} \quad (2)$$

donde μ_{\max} y μ_{\min} son, respectivamente, los valores singulares máximo y mínimo de la matriz \mathbf{X} . Por otra parte, ϵ_{\max} y ϵ_{\min} , representan, los autovalores máximo y mínimo de la matriz $\mathbf{X}^t\mathbf{X}$. Según esta expresión, cuando menor sea μ_{\min} , mayor dependencia lineal existirá entre las columnas de \mathbf{X} asociadas con este valor singular (por lo tanto, mayor NC) produciéndose un posible problema de multicolinealidad en el modelo (1).

Partiendo de (2), la varianza de los estimadores β obtenidos por MCO se pueden descomponer de la siguiente forma:

$$\text{var}(\beta_k) = \sigma^2 \sum_j \frac{\vartheta_{kj}^2}{\mu_j^2}$$

donde σ^2 representa la varianza del error de estimación del modelo (1), μ_j los valores singulares y ϑ_{kj}^2 son los componentes del autovector correspondiente al valor singular al cuadrado μ_j^2 . Cuando el valor singular sea menor, existirá una mayor varianza en los estimadores y por lo tanto, mayor inestabilidad en las estimaciones (uno de los muchos síntomas de posible existencia de multicolinealidad).

Como se ha observado en (2), el Número de Condición, puede servir para detectar el problema de multicolinealidad no esencial, debido a que se tiene en cuenta la variable (\mathbf{X}_1) que acompaña al intercepto.

Por otra parte, el FIV se define como sigue a continuación:

$$FIV(i) = \frac{\text{var}(\hat{\beta}_i)}{\text{var}(\hat{\beta}_i^0)} = \frac{1}{1 - R_i^2} \quad i = 2, \dots, p.$$

Siendo $\hat{\beta}$ el estimador por MCO del modelo (1), $\hat{\beta}^0$ el estimador por MCO del modelo (1) suponiendo que las variables explicativas son ortogonales, y R_i^2 el coeficiente de determinación de la regresión auxiliar que tiene como variable dependiente a \mathbf{X}_i y como independientes al resto de regresores, para $i = 2, \dots, p$.

⁵ Kuh, E.; Welsch, Roy.

Tradicionalmente, valores del FIV superiores a 10 indicarían que el modelo de regresión lineal presenta un grado de multicolinealidad preocupante. Sin embargo, en este caso se mostrará que el FIV no sirve para detectar si existe multicolinealidad no esencial en el modelo (1).

El presente trabajo se estructura como sigue: en la sección 2 se define el posible problema de multicolinealidad entre las variables independientes y la constante (partiendo del Modelo Lineal Simple) y su detección (mostrándose posibles errores en el paquete de R). En la sección 3 se realiza una simulación de Montecarlo para ilustrar los resultados obtenidos en el caso del MLG. Finalmente, en la sección 4 se destacan las principales conclusiones.

2. RELACIÓN ENTRE EL INTERCEPTO Y LAS VARIABLES INDEPENDIENTES (MULTICOLINEALIDAD NO ESENCIAL)

Si partimos del siguiente modelo de regresión lineal simple, donde $p=2$:

$$\mathbf{y}_t = \beta_1 + \beta_2 \mathbf{X}_{2t} + \mathbf{u}_t, \quad t = 1, \dots, n. \quad (3)$$

Como se ha mencionado anteriormente, la definición de multicolinealidad parece quedar bastante clara. Sin embargo, no se puede decir lo mismo sobre la definición de variable independiente: el término constante se puede considerar una variable independiente o una simple constante.

Si se considera que \mathbf{X}_1 no es una variable independiente, sino una constante, se podría pensar que en este modelo no puede existir un problema de multicolinealidad grave.

Por el contrario, si se la considera como tal, la multicolinealidad grave podría aparecer si la variable \mathbf{X}_2 es prácticamente constante. Es decir, que presenta una varianza muy pequeña.

Partiendo del siguiente ejemplo, donde la variable independiente es prácticamente constante:

$$\mathbf{y}_t = \begin{pmatrix} 2 \\ 3 \\ 1 \\ -2 \end{pmatrix}, \quad \mathbf{X}_{2t} = \begin{pmatrix} 2 \\ 2,1 \\ 1,9 \\ 2,1 \end{pmatrix}, \quad \mathbf{X}^*_{2t} = \begin{pmatrix} 2,14 \\ 2,1 \\ 1,9 \\ 2,1 \end{pmatrix}$$

se obtiene estimadores muy distintos tanto en signo como en magnitud, $\hat{\beta} = (8.3636, -3.6364)$ y $\hat{\beta}^* = (-1.3409, 1.1364)$, siendo este uno de los síntomas de la multicolinealidad (inestabilidad en las estimaciones realizadas ante pequeños cambios en la muestra). Luego en el presente trabajo nos decantaremos por la segunda opción.

Técnicas empleadas para la detección de multicolinealidad

Por lo tanto, puede surgir un posible problema de multicolinealidad entre la variable independiente y el término constante. Tras esto, es necesaria su detección mediante herramientas tan utilizadas como pueden ser el FIV y el NC

Para calcular el FIV del modelo (3) en primer lugar se tiene que calcular el coeficiente de determinación de la regresión auxiliar que tiene como variable dependiente a \mathbf{X}_2 y como variable independiente sólo al término independiente. Esto es:

$$\mathbf{X}_2 = \beta_1 + \mathbf{v}_t, \quad t = 1, \dots, n. \quad (4)$$

Este modelo es usualmente conocido como modelo restringido.

Por otro lado, el coeficiente de determinación del modelo (1) se obtiene como:

$$R^2 = 1 - \frac{SCR}{SCT} \quad (5)$$

donde SCR y SCT son respectivamente, la suma de cuadrado de los residuos y totales del modelo global (1).

Ahora bien, se verifica que la SCT coincide con la suma de cuadrados de los residuos del modelo restringido ya que en tal caso la estimación del término independiente coincide con la media de la variable dependiente.

Teniendo en cuenta esta última interpretación, es claro que el coeficiente de determinación del modelo (4) es siempre cero independientemente de cuáles sean los datos con los que se trabaja. Esto se debe a que en este caso el modelo global y el restringido son el mismo y, por lo tanto, el cociente de sumas de cuadrados es igual a 1.

En tal caso, el FIV del modelo (3) será siempre igual a 1, su mínimo valor. Por lo tanto, si el FIV es siempre igual a 1 independientemente de cuáles sean los datos, es claro que no se puede usar en este caso como herramienta que detecta la existencia de multicolinealidad grave.

Dentro de nuestro conocimiento en el entorno de programación **R** se puede calcular el FIV a partir de los paquetes *fmsb* y *car*. A continuación se muestra cómo se calcularía el FIV en el ejemplo considerado anteriormente y los resultados obtenidos:

```
> y = c(2,3,1,-2)
> x2 = c(2,2.1,1.9,2.1)
> x2bis = c(2.14,2.1,1.9,2.1)
>
> library(fmsb)
> VIF(lm(y ~ x2))
```

```
[1] 1.026667
> VIF (lm(y ~ x2bis))
[1] 1.003257
>
> library(car)
> vif (lm(y ~ x2))
Error in vif.default (lm(y ~ x2)) : model contains fewer than 2 terms
> vif (lm(y ~ x2bis))
Error in vif.default (lm(y ~ x2bis)) : model contains fewer than 2 terms
>
```

Se puede observar que con el primer paquete, *fmsb*, se obtienen valores distintos a 1, lo cual se contradice totalmente con lo demostrado anteriormente, proporcionando valores erróneos. Mientras que con el segundo, *car*, se obtiene un mensaje informando que no se puede calcular el FIV, cuando esto no es cierto.

Como se ha demostrado a través del Modelo Lineal Simple, el FIV no tiene en cuenta la variable independiente que representa a la constante de la regresión, por lo tanto, en el Modelo Lineal General no servirá para detectar el posible problema de multicolinealidad no esencial, independientemente de los datos utilizados.

Por lo tanto, una vez demostrado que el FIV no sirve como herramienta para detectar el posible problema de multicolinealidad no esencial, lo más habitual sería tender a utilizar el NC, ya que este si tiene en cuenta el papel del término independiente. Partiendo del NC del Modelo Lineal Simple:

$$NC_{m\text{ls}} = \sqrt{\frac{1+a}{1-a}}, \quad a > 0$$

donde:

$$a = \frac{\bar{X}_2}{\sqrt{\text{Var}(X_2) + \bar{X}_2^2}}$$

se ha llegado a una expresión matemática, que relaciona la varianza de X_2 con su media, indicando cuando la presencia de multicolinealidad entre la variable independiente y el término constante es preocupante.

Así, se tiene que existe **multicolinealidad moderada**, es decir que el $NC_{m\text{ls}} > 20$, si se cumple que:

$$\text{Var}(X_2) < 0.01005019 \cdot \bar{X}_2^2 \quad (6)$$

Y se tiene que existe **multicolinealidad grave**, es decir que el $NC_{m\text{ls}} > 30$, si se cumple que:

$$\text{Var}(X_2) < 0.04454337 \cdot \bar{X}_2^2 \quad (7)$$

A partir de (6) y (7), se podría detectar el posible problema de multicolinealidad no esencial para cualquier valor de p . Además, consideramos que comprobar estas relaciones implica cálculos más sencillos que los realizados para obtener el NC.

3. MULTICOLINEALIDAD EN EL MLG: SIMULACIÓN DE MONTECARLO

Como se ha mencionado anteriormente, puede existir un posible problema de multicolinealidad entre las variables independientes de nuestro modelo y el término constante. Para mostrar la utilidad en el MLG de lo deducido en el MLS, se procede a realizar una simulación donde $n \in \{15, 20, \dots, 200\}$, $\mu_2, \dots, \mu_p \in \{-10, -9.25, \dots, 8.75, 9.5\}$ y $\sigma^2 \in \{0.1, 0.2, \dots, 2\}$.

Para $p=2$, se construye entonces la matriz $\mathbf{X} = (\mathbf{X}_1 \ \mathbf{X}_2 \ \mathbf{X}_3)$ donde \mathbf{X}_1 es un vector de unos de dimensiones adecuadas. De 554040 simulaciones realizadas, solo nos quedamos con aquellas que cumplen que:

- a) Hay multicolinealidad no esencial entre \mathbf{X}_2 , \mathbf{X}_3 y \mathbf{X}_1 . Donde $\mathbf{X}_i \sim N(\mu_i, \sigma^2)$ tal que $\sigma^2 < 0.01\mu_i^2$ con $i=2,3$, es decir con 130036 (23.47% del total). Para este caso se ha calculado tanto el FIV como el NC, con los datos sin centrar y centrando los datos, obteniendo lo que sigue a continuación:

Tabla 1. Resultados de la Simulación 1.

	<i>VARIABLES ORIGINALES^a</i>			<i>VARIABLES CENTRADAS^b</i>		
	Min	Media	Max	Min	Media	Max
FIV	1	1.016268	3.085693	1	1.016268	3.085693
NC	13.88249	31.93091	121.1577	1	1.126865	3.276069

Fuente: Elaboración propia a partir del software R.

^a. Datos originales obtenidos de la simulación.

^b. Datos centrados para intentar mitigar el problema de la multicolinealidad no esencial.

- b) Hay multicolinealidad no esencial entre \mathbf{X}_2 y \mathbf{X}_1 . Donde $\mathbf{X}_2 \sim N(\mu_2, \sigma_2^2)$ tal que $\sigma_2^2 < 0.01\mu_2^2$, y $\mathbf{X}_3 \sim N(\mu_3, \sigma_3^2)$ tal que $\sigma_3^2 = \sigma_2^2 + h$ donde $h \in [70, 100]$, es decir con 130036 (23.47% del total). Para este caso se ha calculado tanto el FIV como el NC, con los datos sin centrar y centrando los datos, obteniendo lo que sigue a continuación:

Tabla 2. Resultados de la Simulación 2.

	<i>VARIABLES ORIGINALES^a</i>			<i>VARIABLES CENTRADAS^b</i>		
	Min	Media	Max	Min	Media	Max
FIV	1	1.016572	4.865276	1	1.016572	4.865276
NC	4.661749	26.38262	174.2905	1.008557	9.367392	33.80953

Fuente: Elaboración propia a partir del software R.

^a. Datos originales obtenidos de la simulación.

^b Datos centrados para intentar mitigar el problema de la multicolinealidad no esencial.

Para $p=3$, se construye la matriz $\mathbf{X} = (\mathbf{X}_1 \ \mathbf{X}_2 \ \mathbf{X}_3 \ \mathbf{X}_4)$ donde \mathbf{X}_1 es un vector de unos de dimensiones adecuadas. Donde de 14959080 simulaciones realizadas, solo nos quedamos con aquellas que cumplen que

- c) Hay multicolinealidad no esencial entre $\mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4$ y \mathbf{X}_1 . Donde $\mathbf{X}_i \sim N(\mu_i, \sigma^2)$ tal que $\sigma^2 < 0.01\mu_i^2$ con $i=2,3,4$, es decir con 4267172 (28.52% del total). Para este caso se ha calculado tanto el FIV como el NC, con los datos sin centrar y centrando los datos, obteniendo lo que sigue a continuación:

Tabla 3. Resultados de la Simulación 3.

	<i>VARIABLES ORIGINALES^a</i>			<i>VARIABLES CENTRADAS^b</i>		
	Min	Media	Max	Min	Media	Max
FIV1	1	1.034248	6.585239	1	1.034248	6.585239
FIV2	1	1.034252	8.394808	1	1.034252	8.394808
FIV3	1	1.034243	11.00134	1	1.034243	11.00134
NC	6.68454	32.31309	214.2114	1.001094	1.279143	9.710329

Fuente: Elaboración propia a partir del software R.

^a. Datos originales obtenidos de la simulación.

^b Datos centrados para intentar mitigar el problema de la multicolinealidad no esencial.

- d) Hay multicolinealidad no esencial entre $\mathbf{X}_2, \mathbf{X}_3$ y \mathbf{X}_1 . Donde $\mathbf{X}_2 \sim N(\mu_2, \sigma_2^2)$ tal que $\sigma_2^2 < 0.01\mu_2^2$, $\mathbf{X}_3 \sim N(\mu_3, \sigma_3^2)$ tal que $\sigma_3^2 < 0.01\mu_3^2$ y $\mathbf{X}_4 \sim N(\mu_4, \sigma_4^2)$ tal que $\sigma_4^2 = \sigma_2^2 + h$ donde $h \in [70,100]$, es decir con 4267172 (28.52%% del total). Para este caso se ha calculado tanto el FIV como el NC, con los datos sin centrar y centrando los datos, obteniendo lo que sigue a continuación:

Tabla 4. Resultados de la Simulación 4.

	<i>VARIABLES ORIGINALES^a</i>			<i>VARIABLES CENTRADAS^b</i>		
	Min	Media	Max	Min	Media	Max
FIV1	1	1.034216	7.152851	1	1.034216	7.152851
FIV2	1	1.034246	8.783643	1	1.034246	8.783643
FIV3	1	1.03422	10.42222	1	1.03422	10.42222
NC	4.70554	29.50641	232.5175	1.004271	7.540294	55.6655

Fuente: Elaboración propia a partir del software R.

^a. Datos originales obtenidos de la simulación.

^b Datos centrados para intentar mitigar el problema de la multicolinealidad no esencial.

- e) Hay multicolinealidad no esencial entre \mathbf{X}_2 y \mathbf{X}_1 . Donde $\mathbf{X}_2 \sim N(\mu_2, \sigma_2^2)$ tal que $\sigma_2^2 < 0.01\mu_2^2$, $\mathbf{X}_3 \sim N(\mu_3, \sigma_3^2)$ tal que $\sigma_3^2 = \sigma_2^2 + h$ y $\mathbf{X}_4 \sim N(\mu_4, \sigma_4^2)$ tal que $\sigma_4^2 = \sigma_2^2 + h$ donde $h \in [70,100]$, es decir con 4267172 (28.52% del total). Para este caso se ha calculado tanto el FIV como el NC, con los datos sin centrar y centrando los datos, obteniendo lo que sigue a continuación:

Tabla 5. Resultados de la Simulación 5.

	<i>Variables originales^a</i>			<i>Variables centradas^b</i>		
	Min	Media	Max	Min	Media	Max
FIV1	1	1.034191	7.901098	1	1.034191	7.901098
FIV2	1	1.034216	6.55996	1	1.034216	6.55996
FIV3	1	1.034217	6.931083	1	1.034217	6.931083
NC	4.376636	23.8066	278.8462	1.082175	11.00102	75.54879

Fuente: Elaboración propia a partir del software R.

^a. Datos originales obtenidos de la simulación.

^b Datos centrados para intentar mitigar el problema de la multicolinealidad no esencial.

Se puede observar como el FIV no sirve para detectar la existencia del posible grado de multicolinealidad, ya que ignora al intercepto. Además en (García, J. et al⁶, 2016) se demuestra que el FIV no se ve afectado por cambios de origen y escala, de ahí que coincidan los resultados obtenidos con los datos originales y centrados. Sin embargo, el NC si tiene en cuenta la posible relación entre el término independiente y la variables explicativas. Se puede observar como con los datos originales se presenta un problema de multicolinealidad no esencial grave y que, al solucionar este problema centrando los datos, los NC disminuyen considerablemente.

En este sentido, en las situaciones a) y c) se tiene que el valor máximo del NC queda por debajo de los umbrales establecidos como preocupantes. Esto se debe a que todas las variables de \mathbf{X} presentan multicolinealidad no esencial. En el resto de casos, los valores máximos que quedan por encima de dichos umbrales pueden deberse a la existencia de multicolinealidad esencial que no es eliminada al centrar las variables.

4. CONCLUSIONES

Partiendo de que la constante se encuentra representada por una variable independiente \mathbf{X}_1 , en el modelo lineal general existe la posibilidad de que las variables independientes del modelo y el término constante se encuentre relacionados, produciéndose multicolinealidad no esencial.

Para detectar este problema de multicolinealidad grave, se ha demostrado que no se aconseja el uso del FIV porque no sirve para detectar el problema de multicolinealidad no esencial, independientemente de los datos utilizados, ya que ignora por completo el término independiente.

Además, paquetes estadísticos tan utilizados como R, pueden llevar a confusión a usuarios inexpertos, debido a que aporta resultados incoherentes sobre el FIV en el

⁶ Salmerón, R.; García, C.; López, MM.

Modelo Lineal Simple. Por lo tanto, por un lado, se aconseja no calcular el FIV a partir de la librería *fmsb* y, por otro, mejorar la función de la librería *car* indicando que en el caso del modelo lineal simple el FIV es siempre igual a uno independientemente de los datos. Y que de esta forma esta herramienta queda totalmente invalidada para detectar si existe multicolinealidad no esencial.

A partir del NC se ha obtenido la expresión (6) que relaciona la varianza de la variable independiente con su media e indica cuando la presencia de multicolinealidad entre la variable independiente y el término constante en el Modelo Lineal General es preocupante. Se entiende que estos estadísticos muestrales son más sencillos de obtener que el Número de Condición.

Finalmente de cara al MLG, se han ilustrado los resultados mediante una simulación de Montecarlo, realizada con el programa estadístico R. En este caso se muestra que el FIV no sirve para detectar la multicolinealidad no esencial en el MLG. Por lo que habría que recurrir al NC o de forma más sencilla, a las expresiones (6) y (7).

5. REFERENCIAS

- Aiken, L.; West, S.; Reno, R. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Business & Economics: Sage.
- Belsley, D.A.; Kuh. E.; Welsch. R. (2004). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Canada, EEUU: John Wiley & Sons.
- Fox, J.; Weisberg, S. (2011). *An R Companion to Applied Regression*, Second Edition. Thousand Oaks CA: Sage. URL: <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>.
- García, J.; Salmerón, R.; García, C.; López, MM. (2016). "Standardization of Variables and Collinearity Diagnostic in Ridge Regression", *International Statistical Review*, 84 (2), 245-266.
- Iacobucci, D.; Schneider, MJ.; Popovich, DL.; Bakamitsos, GA. (2016). "Mean Centering Helps Alleviate "Micro" but not "Macro" Multicollinearity", *Behav Res*, 48, 1308-1317.
- Marquardt, D.; Snee, D. (1975). "Ridge Regression in Practice", *The American Statistician*, 29 (1), 3-20.
- Nakazawa, M. (2017). *fmsb: Functions for Medical Statistics Book with some Demographic Data*. R package versión 0.6.1. URL: <http://CRAN.R-project.org/package=fmsb>.
- Rodríguez, A.; García, C.; Salmerón, R.; García, CB. (2017). "El Factor de Inflación de la Varianza en R", IX Jornadas de usuarios de R, Granada.

Uriel, E.; Contreras, D.; Moltó. L.; Peiró.A. (1994). *Econometría. El modelo lineal*. Madrid, España: AC.

La distribución Waring Biparamétrica: Comparación con otras distribuciones

Rodríguez-Avi, José⁽¹⁾, Olmo-Jiménez, María José⁽²⁾, Cueva-López, Valentina⁽³⁾

(1) *Departamento de Estadística e Investigación Operativa, Universidad de Jaén, jravi@ujaen.es:*

(2) *Departamento de Estadística e Investigación Operativa, Universidad de Jaén,*

[mjolmo@ujaen.es:](mailto:mjolmo@ujaen.es)

(3) *Departamento de Estadística e Investigación Operativa, Universidad de Jaén, vcueva@ujaen.es:*

RESUMEN

En este trabajo, se va a realizar una comparación de dicha distribución con otras distribuciones ya existentes, como la distribución Binomial Negativa, la distribución de Pearson Biparamétrica compleja, la distribución Poisson Generalizada o la distribución de Poisson, a través de la comparación de sus funciones masa de probabilidad y de la divergencia de Kullback-Leiber. Así mismo, se mostrarán dos ejemplos en diversos campos para comprobar el buen ajuste proporcionado por la distribución BW.

Palabras clave: Biparametrica, Waring, BW

ABSTRACT

In this paper, we will make a comparison of this distribution with other existing distributions, such as the Negative Binomial distribution, the complex Biparametric Pearson distribution, the Generalized Poisson distribution or the Poisson distribution, by comparing their functions mass of probability and the divergence of Kullback-Leiber. Likewise, two examples will be shown in different fields to check the good fit provided by the BW distribution.

Keywords: Biparametrica, Waring, BW

1. INTRODUCCIÓN

La distribución Waring Univariante Generalizada es una distribución triparamétrica para variables discretas, estudiada por diversos autores, los más importantes Irwing en 1968, 1975 y Xekalaki en 1983. Esta distribución se obtiene de diversas formas, aunque la más utilizada, la distribución Waring Extendida (EGW), se obtiene como una mixtura Binomial Negativa con la distribución Beta tipo II. Esto implica que dicha distribución es sobredispersa y que su varianza puede ser dividida en 3 partes: Aleatoriedad, Riesgo y Predisposición.

Se han desarrollado muchas extensiones, como la Distribución Waring Extendida (Rodríguez-Avi, 2007), la distribución Bidimensional Waring Univariante Generalizada (Xekalaki, 1984). También se han desarrollado un modelo de regresión para datos discretos (Rodríguez-Avi, 2007) (Hilbe, 2011).

La distribución y su modelo de regresión se aplican en la modelización de diversos conjuntos de datos de distintos campos, como por ejemplo en lexicología (Tesitelova, 1967), el número de autores en artículos científicos (Ajiferuke, 1991), el número de goles en función de diversas variables (Saez-Castillo, 2013) o el número de errores de compleción en un mapa Cartográfico (Ariza-López, 2015).

La distribución tiene tres parámetros (a , k , ρ). Teóricamente, el primer parámetro viene de la parte Gamma de la mixtura y los otros dos restantes proceden de la distribución BetaII. Sin embargo, en la práctica, existen problemas de identificación de los dos primeros parámetros ya que se pueden cambiar en la expresión de función masa de probabilidad. También, los métodos de estimación desarrollados para esta distribución, como el método de los Momentos o el método de Máxima Verosimilitud. Pero en estos casos, las estimaciones que se obtienen de los primeros parámetros son bastantes similares, casi iguales en muchos casos. Esta situación también aparece cuando las estimaciones que se obtienen por el método de los momentos, son utilizadas como estimaciones iniciales del método de Máxima Verosimilitud. Por esta razón, se estudia una versión biparamétrica, donde se realiza un cambio en los parámetros de la Waring Univariante Generalizada. Además, se comparan los resultados con los obtenidos con la distribución Waring Univariante Generalizada.

2. LA DISTRIBUCIÓN WARING UNIVARIANTE GENERALIZADA

Sea X , una variable discreta que se distribuye según una distribución de Poisson, $X \rightarrow P(\lambda)$ donde λ no es fijo, pero sigue una distribución Gamma, $\lambda \rightarrow \text{Gamma}(a, v)$. En este caso, X se distribuye según una Binomial Negativa con parámetros a , $p=1/(1+v)$. Si en esta última distribución, v también es una variable aleatoria que sigue una distribución Beta II con parámetros k , ρ , entonces, X se distribuye según una distribución Waring Univariante Generalizada con función masa de probabilidad (f.m.p.):

$$P[X = x] = \frac{\Gamma(a+\rho) \Gamma(k+\rho)}{\Gamma(a) \Gamma(k)} \frac{1}{\Gamma(\rho)} \frac{\Gamma(a+x) \Gamma(k+x)}{\Gamma(a+k+\rho+x)x!}, \quad x = 0, 1, 2, \dots(1)$$

Sus características son:

- Su media:

$$\mu = \frac{\alpha^2}{\rho - 1}, \text{ si } \rho > 1$$

- La varianza

$$\sigma^2 = \frac{\alpha^2(\alpha + \rho - 1)^2}{(\rho - 1)^2(\rho - 2)}, \text{ si } \rho > 2$$

Sus principales características son:

- Es una distribución sobredispersa, esto es, $E(X) < \text{Var}(X)$. Como resultado la cola de la distribución puede ser extremadamente larga (El efecto cola pesada).
- $E(X^m) < \infty$ si $\rho > m$. Así, la distribución puede tener la varianza infinita.
- La función masa de probabilidad es unimodal con la moda distinta de 0.
- La descomposición como mixtura de tres distribuciones proporciona una partición de la varianza en tres componentes (Irwing, 1968):

$$\sigma^2 = \frac{ak}{\rho - 1} + \frac{ak(a + 1)}{(\rho - 1)(\rho - 2)} + \frac{a^2k(k + \rho - 1)}{(\rho - 1)^2(\rho - 2)}$$

Su primer término es relacionado con el factor de la aleatoriedad, el segundo con la variabilidad dada por factores externos que afectan a la población (riesgo) y el tercero a la diferencia de condiciones internas de los individuos (predisposición).

3. LA DISTRIBUCIÓN WARING BIPARAMÉTRICA (BW)

Si consideramos en (1) que $a=k$ y lo notamos por α , entonces la f.m.p. es:

$$P[X = x] = \left(\frac{\Gamma(\alpha + \rho)}{\Gamma(\alpha)} \right)^2 \frac{1}{\Gamma(\rho)} \frac{\Gamma(\alpha + x)^2}{\Gamma(2\alpha + \rho + x)x!}, \quad x = 0, 1, 2, \dots$$

En este caso, es una distribución biparamétrica con parámetros α y ρ . En este caso, sus principales propiedades son:

- Es una distribución discreta de rango infinito
- La media

$$\mu = \frac{\alpha^2}{\rho - 1}$$

que existe cuando $\rho > 1$

- La varianza

$$\sigma^2 = \frac{\alpha^2(\alpha + \rho - 1)^2}{(\rho - 1)^2(\rho - 2)}, \text{ si } \rho > 2$$

Que existe cuando $\rho > 2$. En general, el momento centrado de orden m existe cuando $\rho > m$.

Esta distribución, puede obtenerse como una doble mixtura de Poisson, siempre que $\alpha > 0$, de la siguiente forma:

1. $X \rightsquigarrow P(\lambda)$
2. $\lambda \rightsquigarrow \Gamma(\alpha, \nu)$. De este modo, $X \rightsquigarrow BN(\alpha, p)$ con $p = 1/(1 + \nu)$
3. $\nu \rightsquigarrow Beta(\rho, \alpha)$, esto es,

$$f(\nu) = \frac{\Gamma(\alpha + \rho)}{\Gamma(\alpha)\Gamma(\rho)} \nu^{\alpha-1} (1 + \nu)^{\alpha+\rho}$$

o, la probabilidad de incrementar la variable cambia de individuo a individuo por la media de la distribución $Beta(\rho, \alpha)$.

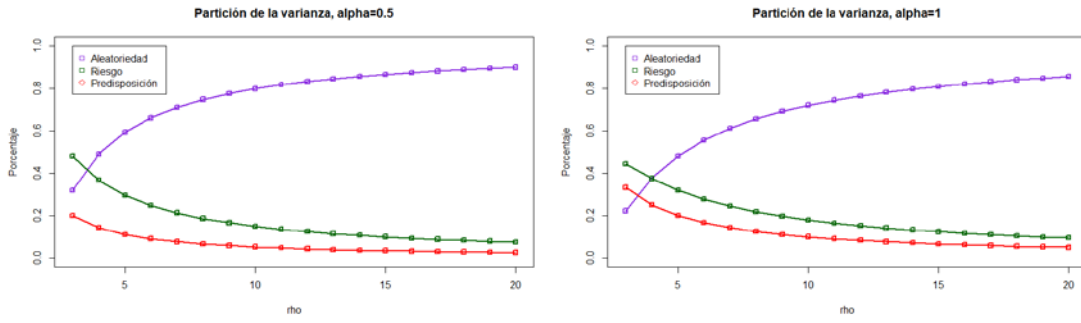
Dado que se obtiene como una mixtura de Poisson, esta distribución es siempre sobredispersa. Esta distribución elimina los parámetros indeterminados de la Waring anterior.

Así, la varianza puede dividirse en tres partes:

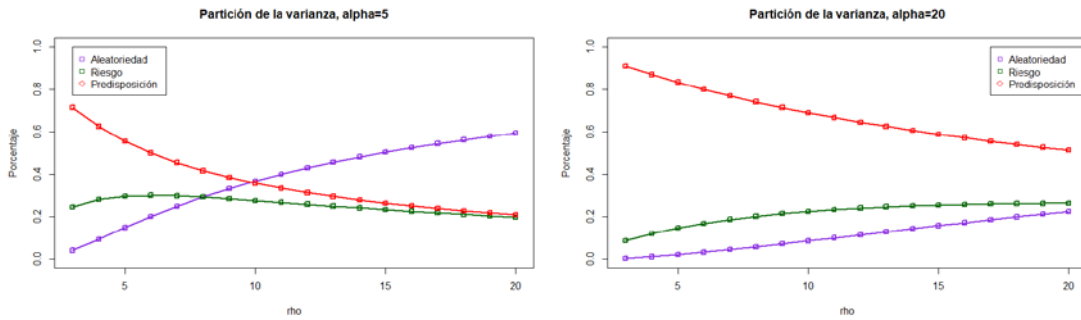
$$\sigma^2 = \frac{\alpha^2}{\rho - 1} + \frac{\alpha^2(\alpha + 1)}{(\rho - 1)(\rho - 2)} + \frac{\alpha^3(\alpha + \rho - 1)}{(\rho - 1)^2(\rho - 2)}$$

Donde el primer término representa la variabilidad dada por la aleatoriedad y viene dada por el modelo de Poisson. Los otros dos términos se refieren a la variabilidad que no pertenece a la aleatoriedad, pero es explicada por la presencia del riesgo y de la predisposición, respectivamente.

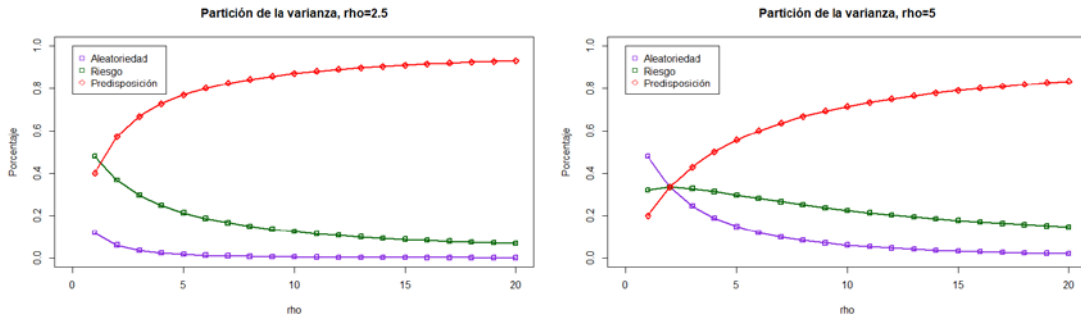
En la Figura 1, se muestra la evolución de los porcentajes fijando α y con diferentes valores de ρ y fijando ρ y con diferentes valores de α . Se observa que para un ρ fijado, cuanto mayor es α gana importancia la predisposición. Sin embargo, cuando se fija α , cuanto mayor es ρ , gana importancia la aleatoriedad. Si α y ρ aumenta con el mismo orden de convergencia, la predisposición tiende al 50% de la varianza y las restantes al 25%.



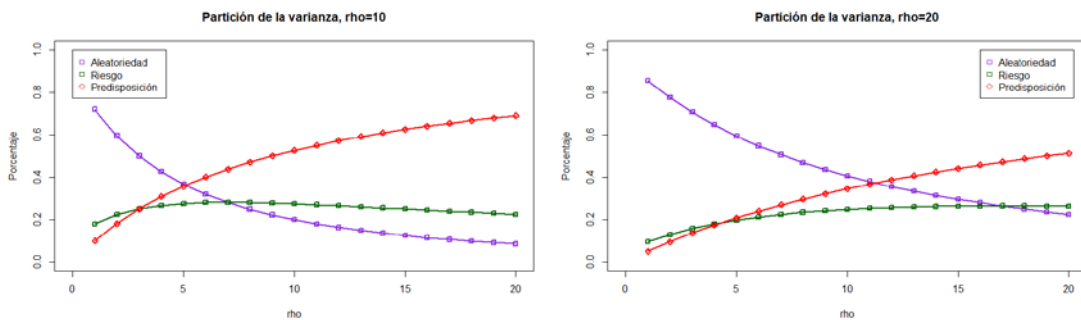
(a) Porcentaje de la partición de la varianza para valores bajos de α



(b) Porcentaje de la partición de la varianza para valores altos de α



(c) Porcentaje de la partición de la varianza para valores bajos de ρ



(d) Porcentaje de la partición de la varianza para valores altos de ρ

Figura 1 Partición de la varianza para algunos valores de alpha y rho

4. COMPARCIÓN CON OTRAS DISTRIBUCIONES

Se procede a comparar la distribución BW con algunas distribuciones biparamétricas discretas:

- Distribución Binomial Negativa, $BN(\theta, p)$, cuya f.m.p. es:

$$P[X = x] = \frac{\Gamma(\theta + x)}{\Gamma(\theta)x!} p^\theta (1 - p)^x, \quad x = 0, 1, 2, \dots$$

Con $0 \leq p \leq 1, \theta > 0$

- Distribución Poisson, con f.m.p. $\frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots,$

- Distribución Poisson Generalizada, $GP(\lambda, \theta)$, con f.m.p.:

$$P[X = x] = \frac{\lambda(\lambda + \theta x)^{x-1} e^{-\lambda - \theta x}}{x!} \quad x = 0, 1, 2, \dots$$

Con $\lambda > 0, \max\left\{-1, -\frac{\lambda}{4}\right\} < \theta < 1$

- Distribución CBP, con f.m.p.:

$$P[X = x] = \frac{\Gamma(\gamma - bi)\Gamma(\gamma + bi)}{\Gamma(\gamma)^2} \frac{(bi)_x (-bi)_x}{(\gamma)_x x!}, \quad x = 0, 1, \dots$$

Con $\gamma > 0, b \in \mathbb{R}$

Para poder comparar las distribuciones, se asume que todas ellas tienen la misma media y la misma varianza. De esta forma, al comparar el perfil de probabilidad (los valores de la f.m.p) que, en todos los casos, la BW se encuentra entre las distribuciones BN y GP y las distribuciones CBP y Poisson, como se observa en la Figura 2.

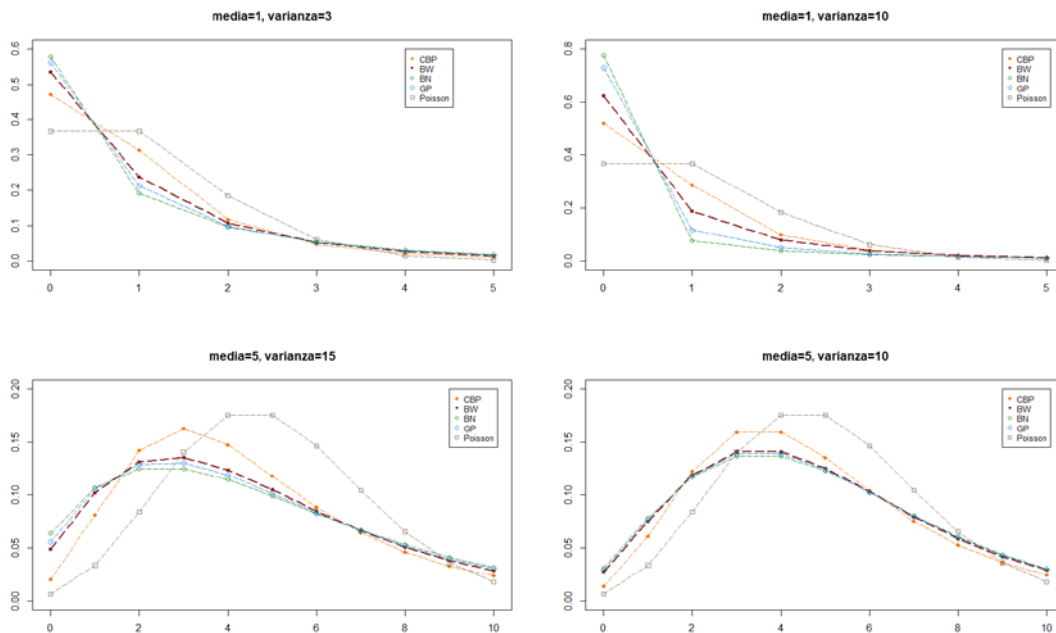


Figura 2 Probabilidad para algunos valores de μ y σ^2

Si realizamos la comparación a través de la divergencia de Kullback-Leibler (KL), cuya divergencia de Ψ a π queda definida, para las distribuciones discretas Ψ y π , como:

$$D_{KL}(\pi||\psi) = \sum_{r=0}^{\infty} \pi(r) \ln \frac{\pi(r)}{\psi(r)}$$

Cabe destacar que esta medida no es simétrica.

La Figura 3 muestra los valores de la divergencia KL entre la BW y las distribuciones mencionadas al inicio de esta sección (y viceversa) para algunos valores de la media y la varianza. Específicamente, se ha fijado los valores de la media μ , para cada uno, se ha realizado la divergencia KL en ambos sentidos para distintos valores de la varianza. Se observa de manera evidente que, la distribución BW está más cerca de la CBP que de la GP o la BN, especialmente, cuando la varianza aumenta. Sin embargo, cuando la varianza está cercana a cero, la distribución que más cerca está de la BW es la distribución GP.

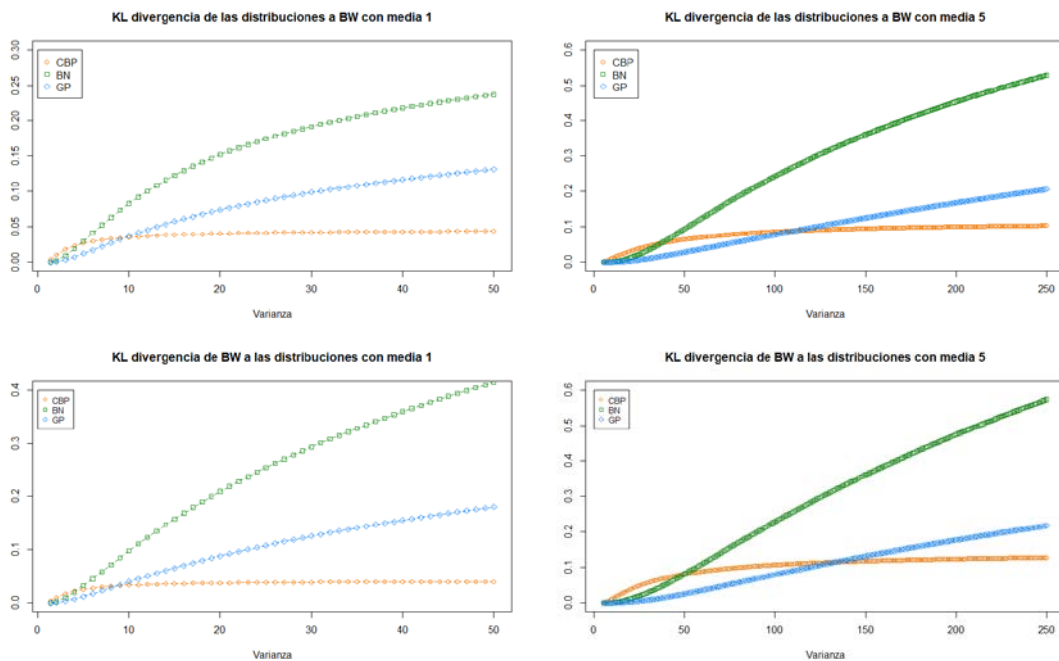


Figura 3 Divergencia de KL para algunos valores de μ y σ^2

5. CONCLUSIONES

Al analizar la distribución BW se ha podido comprobar que el problema de distribución que se produce en la UGW, ha quedado solucionado. Pudiendo identificar correctamente ambos parámetros.

Además, podemos observar ciertas características que pueden ser aplicable a ciertos conjuntos de datos obteniendo mejor valor para el Criterio de Información de Akaike.

Para futuras investigaciones, se podría extender dicha distribución para casos infradispersos, cumpliendo ciertas condiciones más específicas.

6. REFERENCIAS

1. Tesitelova, M. (1966). "On the role of nouns in the lexical statistics", *Prague Stud. Math. Linguist.*, 2: 121-231
2. Irwing, J.O. (1968). "The generalized waring distribution applied to accident theory". *Journal of the Royal Statistical Society, Series A* 131, 2: 205-225
3. Irwing, J.O. (1975a). "The generalized Waring distribution. Part I", *J. Roy. Statist. Soc., A*, 138: 18-31
4. Irwing, J.O. (1975b). "The generalized Waring distribution. Part II", *J. Roy. Statist. Soc., A*, 138: 204-227
5. Irwing, J.O. (1975c). "The generalized Waring distribution. Part III", *J. Roy. Statist. Soc., A*, 138: 374-378
6. Xekalaki, E. (1983). "The univariate generalized Waring distribution in relation to accident theory: proneness, spells or contagion?", *Biometrics*, 39, 1: 887-893
7. Xekalaki, E. (1983b). "Infinite divisibility, completeness and regression properties of the univariate generalized Waring distribution", *Ann. Inst. Statist. Math.*, 35: 279-289
8. Xekalaki, E. (1984). "The bivariate generalized Waring distribution and its application to Accident Theory", *Journal of the Royal Statistical Society, Series A*, 147, 3: 488-498
9. Ajiferuke, I. (1991). "A probabilistic model for the distribution of anthorships", *J. Amer. Soc. Inform. Sci.*, 42, 4: 279-289
10. Rodríguez-Avi, J., Conde-Sanchez, A., Sáez-Castillo, A.J. and Olmo-Jiménez, M.J. (2007). "A new generalization of the Waring distribution", *Computacional Statistics and Data Analysis*, 51, 12: 6138-6150
11. Hilbe, J.M. (2011). *Negative binomial regression*. Cambridge University Press.
12. Sáez-Castillo, A.J., Rodríguez-Avi, J. and Pérez-Sanchez, J.M. (2013). "Expected number of goals depending on intrinsic and extrinsic factors of a football player. An application to professional Spanish football league", *European Journal of Sport Science*, 13, 2: 127-138
13. Ariza-López, F.J. and Rodríguez-Avi, J. (2015). "Estimating the count of completeness errors in geographic data sets by means of a generalized Waring regression model", *International Journal of Geographical Information Science*, 29, 8: 1394-1418

Actividad investigadora de la Universidad de Jaén en Ciencias de la Salud: Tesis Doctorales

Irene Agea Cano⁽¹⁾, María José Calero García⁽²⁾, Manuel Linares Abad⁽³⁾.

(1) Departamento de Enfermería, Universidad de Jaén, Campus Las Lagunillas, 23071 Jaén.

e-mail: iac00013@red.ujaen.es

(2) Departamento de Enfermería, Universidad de Jaén, Campus Las Lagunillas, 23071 Jaén.

e-mail: mjcalero@ujaen.es

(3) Departamento de Enfermería, Universidad de Jaén, Campus Las Lagunillas, 23071 Jaén.

e-mail: mlinares@ujaen.es

RESUMEN

La Universidad de Jaén oferta 20 Programas de Doctorado regulados por el Real Decreto 99/2011, de los cuales 9 son de carácter interuniversitario. Otros programas de Doctorado regulados por el Real decreto 1993/2007, siguen vigentes, aunque no admiten alumnos nuevos. El objetivo fue conocer el número de Tesis doctorales del Departamento de Ciencias de Salud y Departamento de Enfermería de la Universidad de Jaén y la metodología empleada en los últimos 10 años. Para este estudio descriptivo, se llevó a cabo la búsqueda de datos en el repositorio institucional de la Universidad de Jaén, en los últimos 10 años (del 1/01/2007 al 31/12/2016), correspondientes al Departamento de Ciencias de la Salud y Departamento de Enfermería. Se constata un aumento de las publicaciones a lo largo del periodo estudiado. El Departamento de Ciencias de la Salud publicó una media de 3.1 tesis ($\sigma = 3,95$, IC 95%), frente a 1,2 tesis ($\sigma = 1,98$, IC 95%) del Departamento de Enfermería. En Ciencias de la Salud, predominaron los diseños experimentales en modelos celulares o animales (61.2%). En el Departamento de Enfermería sobresalieron los estudios de prevalencia (58,33%). La publicación de Tesis en el repositorio institucional de la Universidad de Jaén ha aumentado en los últimos 10 años, de forma global, así como en los Departamentos de Ciencias de la Salud y de Enfermería. Las metodologías empleadas con mayor frecuencia fueron los diseños experimentales en Ciencias de la Salud y los estudios de prevalencia en el Departamento de Enfermería.

Palabras clave: Educación de Postgrado en Enfermería, Ciencias de la Salud, Metodología.

ABSTRACT

University of Jaén offers 20 Doctorate Programs regulated by Royal Decree 99/2011. Nine of them are interuniversity. Doctorate programs regulated by Royal Decree 1993/2007 are still valid, although they do not admit new students. The objective was to know the number of doctoral Theses of the Department of Health Sciences and the

Department of Nursing of the University of Jaén and the methodology used in the last 10 years. For this descriptive study, the data search was carried out in the institutional repository of the University of Jaén, (from 01/01/2007 to 12/31/2016), from Department of Science of the Health and Nursing Department. There is an increase in publications throughout the period studied. Department of Health Sciences published an average of 3.1 Theses ($\sigma = 3.95$, CI 95%), compared to 1.2 Theses ($\sigma = 1.98$, CI 95%) of the Nursing Department. In Health Sciences, experimental designs predominated in cellular or animal models (61.2%). In Department of Nursing, prevalence studies predominated (58.33%). The publication of Theses in the institutional repository of the University of Jaén has increased in the last 10 years, globally, as well as in the Departments of Health Sciences and Nursing. The most frequently used methodologies were the experimental designs in Health Sciences and the prevalence studies in the Department of Nursing.

Keywords: Education, Nursing, Graduate, Health Sciences, Methods.

1. INTRODUCCIÓN

Los estudios de Doctorado se incluyen en el tercer ciclo de las enseñanzas universitarias oficiales y son regulados por el Real Decreto 99/2011 de 28 de Enero (R.D. 99/2011)¹. Los Programas de Doctorado tienen como finalidad la formación avanzada en competencias y habilidades en investigación científica, que adquirirá el doctorando/a y demostrara con la defensa de una tesis doctoral como trabajo original de investigación¹. Actualmente, la Universidad de Jaén oferta 20 Programas de Doctorado regulados por el Real Decreto 99/2011, de los cuales 9 son de carácter interuniversitario². Otros programas de Doctorado regulados por el Real decreto 1993/2007³, siguen vigentes, aunque no admiten alumnos nuevos. En Ciencias de la Salud, la Universidad de Jaén oferta de forma interuniversitaria dos programas de doctorado: Programa de Doctorado Interuniversitario en Ciencias de la Salud por la Universidad de Jaén y Universidad de Sevilla (R.D. 99/2011) y el Programa de Doctorado Interuniversitario en Cuidados Integrales y Servicios de Salud por la Universidad de Jaén, Universidad de Lérida y Universidad VIC-Catalunya (R.D. 99/2011). El objetivo propuesto fue conocer el número de Tesis doctorales del Departamento de Ciencias de Salud y Departamento de Enfermería de la Universidad de Jaén y la metodología empleada en los últimos 10 años.

2. METODOLOGÍA

Técnicas empleadas

Estudio descriptivo de las Tesis doctorales publicadas por la Universidad de Jaén en los últimos 10 años desde los Departamentos de Ciencias de Salud y de Enfermería. Se llevó a cabo la búsqueda de datos en el repositorio institucional de la Universidad de Jaén (RUJA), publicadas en los últimos 10 años (del 1/01/2007 al 31/12/2016),

correspondientes al Departamento de Ciencias de la Salud y Departamento de Enfermería.

3. RESULTADOS

El total de tesis publicadas en RUJA desde 1998 hasta la actualidad es de 381, correspondiendo 301 tesis, a los últimos 10 años (2007 a 2016). Se constata un aumento de las publicaciones a lo largo del periodo estudiado con un rango de 6 (2007) a 128 (2016) (Figura 1).

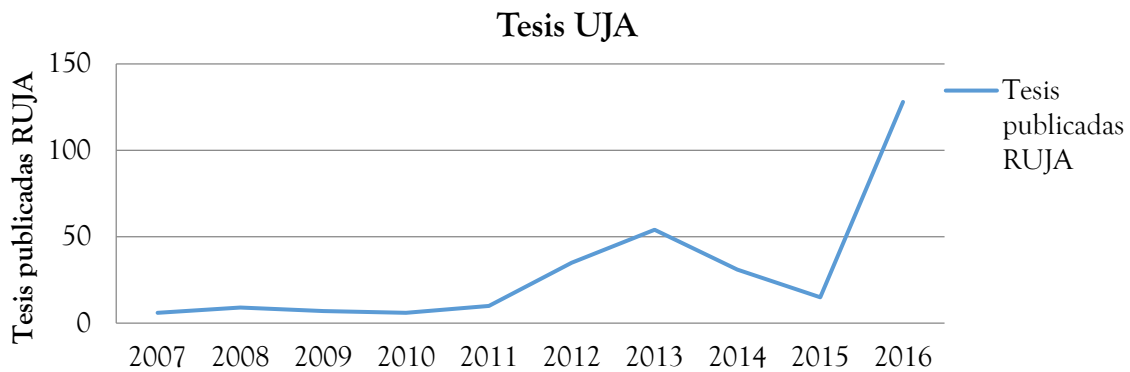


Figura 1. Evolución de la publicación de Tesis en RUJA periodo (2007-2016).
(Fuente: Elaboración propia a partir de datos de RUJA)

La media es de 30,1 publicaciones, desviación estándar $\sigma = 37,93$ con un nivel de confianza para la media del 95% (Tabla1).

Tabla 1. Descripción de Tesis UJA 2007-2016

Media	30,1
Error típico	11,9967125
Mediana	12,5
Moda	6
Desviación estándar	37,936936
Varianza de la muestra	1439,21111
Curtosis	5,51910946
Coeficiente de asimetría	2,26601851
Rango	122
Mínimo	6
Máximo	128
Suma	301
Cuenta	10

(Fuente: Elaboración propia a partir de datos de RUJA)

Desde el Departamento de Ciencias de la Salud se han publicado 31 tesis en el periodo 2007-2016, con una media de 3.1, desviación estándar $\sigma = 3,95$ (Tabla 2).

Tabla 2. Tesis Ciencias de la Salud 2007-2016

Media	3,1
Error típico	1,25122163
Mediana	1
Moda	0
Desviación estándar	3,95671019
Varianza de la muestra	15,6555556
Curtosis	1,76466347
Coefficiente de asimetría	1,46798084
Rango	12
Mínimo	0
Máximo	12
Suma	31
Cuenta	10

(Fuente: Elaboración propia a partir de datos de RUJA)

En menor medida, el Departamento de Enfermería ha publicado 12 tesis, correspondiendo a 1,2 tesis de media, $\sigma = 1,98$. Las medias han sido calculadas con un nivel de confianza del 95% (Tabla 3).

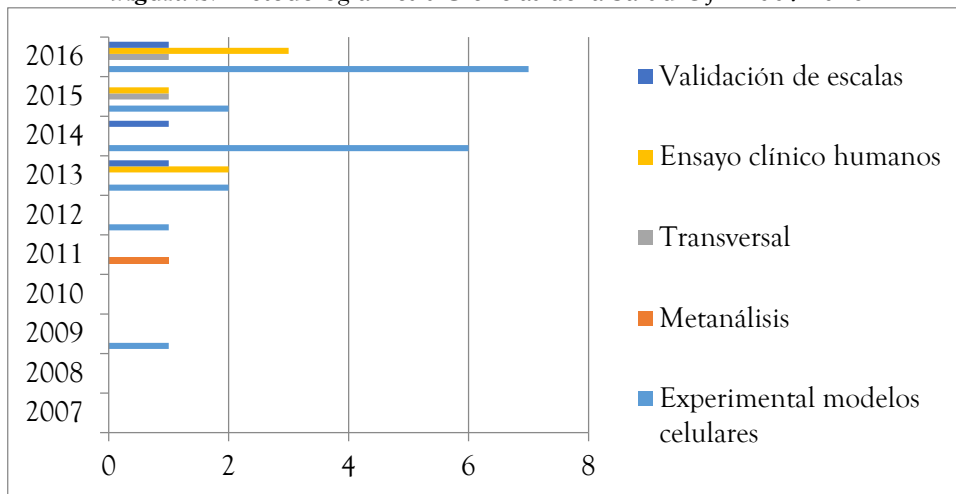
Tabla 3. Tesis Departamento de Enfermería UIA 2007-2016

Media	1,2
Error típico	0,62893208
Mediana	0
Moda	0
Desviación estándar	1,98885785
Varianza de la muestra	3,9555556
Curtosis	3,39121143
Coefficiente de asimetría	1,8812672
Rango	6
Mínimo	0
Máximo	6
Suma	12
Cuenta	10

(Fuente: Elaboración propia a partir de los datos de RUJA)

En cuanto a la metodología utilizada, todas las tesis estudiadas en ambos departamentos, realizan análisis descriptivos y algunas además, análisis inferenciales según el diseño del estudio. En el caso de Ciencias de la Salud, predominan los diseños experimentales en modelos celulares o animales (61,2%), el resto corresponden a estudios analíticos con diseños cuasiexperimentales (12,90%), validación de escalas (9,67%) y en menor medida ensayos clínicos aleatorizados (6,45%). Los estudios de prevalencia corresponden al 6,45% de las tesis publicadas por este departamento. Por último en estos 10 años, sólo se ha realizado una revisión sistemática con metaanálisis en Ciencias de la Salud (Figura 2).

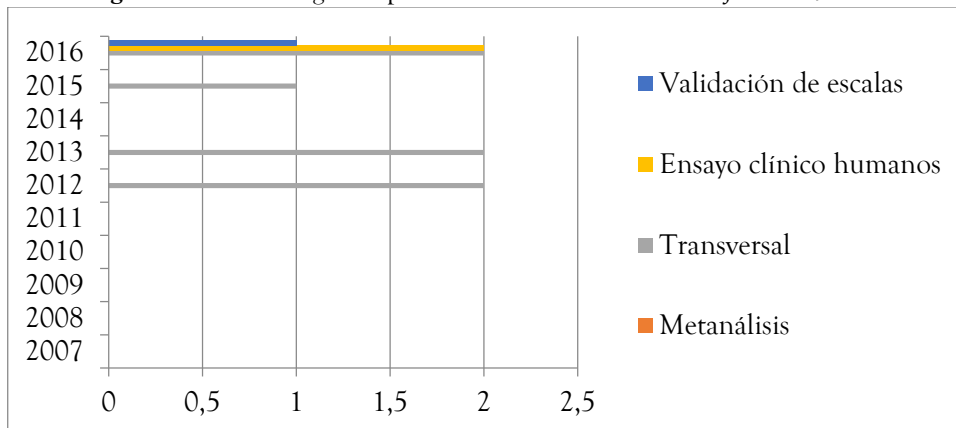
Figura 2. Metodología Tesis Ciencias de la Salud UJA 2007-2016



(Fuente: Elaboración propia a partir de datos de RUJA)

En el Departamento de Enfermería se han realizado mayoritariamente estudios de prevalencia (58,33%), y en cuanto a diseños analíticos: estudios cuasiexperimentales (16,66%) validación de escalas (8,33%), estudios de correlación (8,33%). En El Departamento de enfermería no se ha publicado ningún estudio experimental sobre modelos celulares o animales, ni metaanálisis. En el caso de los estudios cualitativos sólo 1 tesis ha seguido esta una metodología en 2016 (Figura 3).

Figura 3. Metodología Departamento de Enfermería UJA. 2007-2016



(Fuente: Elaboración propia a partir de datos de RUJA)

4. CONCLUSIONES

-La publicación de Tesis en el repositorio institucional de la Universidad de Jaén ha aumentado en los últimos 10 años, de forma global, así como en los Departamentos de Ciencias de la Salud y de Enfermería.

-En el Departamento de Ciencias de la Salud de la Universidad de Jaén, las tesis publicadas siguen mayoritariamente diseños experimentales sobre modelos celulares o animales, en el periodo 2007-2016.

-En el Departamento de Enfermería de la Universidad de Jaén predominan los estudios de prevalencia en las tesis publicadas en el periodo 2007-2016.

5. REFERENCIAS

1. Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado. Ministerio de Educación. Disponible en «BOE» núm. 35, de 10 de febrero de 2011, 13909/13926.

2. Universidad de Jaén (Web) <http://www10.ujaen.es/>. Oferta Académica, Enseñanzas Oficiales, Doctorado Disponible en: <<<http://estudios.ujaen.es/listaDoctorados>>>. Acceso [15 de Octubre 2017].

3. Real Decreto 1393/2007, de 29 de octubre, por el que se establece la ordenación de las enseñanzas universitarias oficiales. Disponible en «BOE» núm. 260, de 30 de octubre de 2007, 44037/44048.

Estimación de la prevalencia de estrongiloidiasis en población infantil de Cubal (Angola) y de los parámetros de los test diagnósticos en ausencia de *gold standard* con Estadística Bayesiana.

Elena Dacal⁽¹⁾, Rosario Susi⁽²⁾, Carmen Nieto⁽²⁾, José María Saugar⁽¹⁾, Esperanza Rodríguez⁽¹⁾

(1) Laboratorio de Referencia e Investigación en Parasitología, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, Madrid, España. e-mail: edacal@isciii.es

(2) Dpto. Estadística e Investigación Operativa III, Facultad de Estudios Estadísticos, Universidad Complutense de Madrid, Madrid, España.

RESUMEN

El diagnóstico de certeza de *Strongyloides stercoralis* es la detección de larvas por microscopía; sin embargo su sensibilidad es baja. La técnica de Baermann mejora dicha sensibilidad, pero es insuficiente por las limitaciones que presenta. Como alternativa, las técnicas moleculares (como por ejemplo la PCR en tiempo real, PCRrt) suponen una notable mejora de sensibilidad, especificidad y rapidez. A pesar de esto, actualmente no existe una técnica considerada *gold standard*. Por esto nos planteamos estimar la prevalencia de la estrongiloidiasis en población infantil de una zona rural de Cubal (Angola) y la sensibilidad y especificidad de las técnicas de Baermann y PCRrt en ausencia de *gold standard*. La Estadística Bayesiana permite abordar esta problemática estimando la sensibilidad y la especificidad de estas pruebas diagnósticas y la prevalencia de la enfermedad en contexto de zonas endémicas. La población se dividió en niños que caminaban descalzos y niños que caminaban calzados, al entender que ambos grupos podían presentar distintas prevalencias de la estrongiloidiasis. Se utilizó inferencia Bayesiana fijando las distribuciones a priori de los parámetros de interés (especificidades y sensibilidades de las pruebas y prevalencias) a partir de: 1) una búsqueda bibliográfica; 2) datos de un estudio de una población de Etiopía; 3) una combinación de los dos anteriores. Tras introducir los datos de la población de Cubal se obtuvieron las distribuciones a posteriori generando muestras aleatorias de ellas por el método de Cadenas de Markov Monte Carlo (MCMC).

El modelo 3, en el que las distribuciones a priori de las prevalencias se elicitan a través de un estudio previo en Etiopía y la sensibilidad y especificidad a partir de la bibliografía existente, presentó el menor valor del DIC (31.62). La sensibilidad para la técnica Baermann y su intervalo de probabilidad al 95% fue del 50.89% (32.07%-71.81%), mientras que para PCRrt fue del 81.08% (65.63%-93.10%). La especificidad para Baermann fue del 94.53% (91.45%-97.19%) y para PCRrt del 86.01% (80.77%-91.17%). Las prevalencias estimadas fueron 12.83% (7.99%-18.55%) para la población que camina descalza y 10.38% (5.94%-15.81%) para la que camina calzada. Según lo obtenido, la Estadística Bayesiana permite estimar la validez de las pruebas diagnóstico en estudio y determinar la prevalencia de estrongiloidiasis en ausencia de *gold standard*.

Palabras clave: *Strongyloides*, PCRrt, Baermann, sensibilidad, especificidad, prevalencia, Estadística Bayesiana.

ABSTRACT

The diagnosis of certainty of *Strongyloides stercoralis* is the detection of larvae by microscopy; however, its sensitivity is low. Baermann technique improves this sensitivity, but it is insufficient by its limitations. As an alternative, molecular techniques (i.e. real time PCR, PCRrt) represent a remarkable improvement of sensitivity, specificity, and rapidity. Despite this, currently there is not a technique considered as *gold standard* yet. For this reason, we estimate the prevalence of strongyloidiasis in children in a rural area of Cubal (Angola) and the sensitivity and specificity of Baermann and real time PCR (rtPCR) techniques in the absence of *gold standard*. The Bayesian statistics allow addressing this problem considering the sensitivity and specificity of these diagnostic tests and the prevalence of the disease in the context of endemic areas. Population was divided into children walking barefoot and wearing shoes, taking to account that both groups could have different prevalences for strongyloidiasis. We used Bayesian inference setting *a priori* distributions of the parameters of interest (specificities, sensitivities and prevalences), based on: 1) a literature search; (2) data from a study of a population of Ethiopia; (3) a combination of the previous two. After introducing the data of the town of Cubal we obtained *a posteriori* distributions generating random samples by the method of Monte Carlo of Markov chains (MCMC). Model 3, in which *a priori* distributions of the prevalence were elicited from a previous study in Ethiopia and the sensitivity and specificity from the existing literature, presented the lowest DIC value (31.62). Sensitivity for Baermann technique and its 95% probability interval was 50.89% (32.07%-71.81%), while for rtPCR was 81.08% (65.63%-93.10%). Specificity for Baermann was 94.53% (91.45%-97.19%) and for rtPCR of 86.01% (80.77%-91.17%). Estimated prevalences were 12.83% (7.99%-18.55%) for people walking barefoot and 10.38% (5.94% - 15.81%) for those wearing shoes. As results show, Bayesian statistics allow validating the diagnostic techniques and determining the prevalence of strongyloidiasis in the absence of *gold standard*.

Keywords: *Strongyloides*, PCRrt, Baermann, sensitivity, specificity, prevalence, Bayesian statistics.

1. INTRODUCCIÓN

La estrongiloidiasis es una enfermedad parasitaria producida por *Strongyloides stercoralis* y *Strongyloides fuelleborni*. Aproximadamente entre unos 30-100 millones de personas están afectadas por la enfermedad en el mundo, la mayoría provenientes de zonas tropicales y subtropicales, destacando el Sudeste Asiático, Sur de EEUU, América Latina, India y África subsahariana.

En referencia al nivel de prevalencia de países del África subsahariana no existen demasiados estudios al respecto y los que aportan datos son muy heterogéneos.

Este parásito tiene un ciclo de vida complejo, pudiéndose desarrollar un ciclo de vida libre, un ciclo parasitario al penetrar las larvas en el interior del hospedador a través de la piel intacta y numerosos ciclos autoinfectivos. La mayoría de las personas permanecen asintomáticas durante años, aunque en algunas presentan manifestaciones no exclusivas de este parásito, de tipo intestinal, cutáneo y/o pulmonar. Además, en individuos inmunodeprimidos el ciclo autoinfectivo se acelera teniendo lugar una infección diseminada y produciéndose el síndrome de hiperinfección que aumenta la mortalidad de un 15 % hasta un 87 % (Vadlamudi *et al.*, 2006).

El diagnóstico de certeza de *S. stercoralis* es la detección de larvas por microscopía. Sin embargo, la sensibilidad es baja, alrededor del 30% en el examen directo de una muestra de heces, o del 50% en muestras seriadas (Ruano *et al.*, 2005). Dicha sensibilidad puede mejorarse mediante el empleo de técnicas de migración larvaria como Baermann y Harada Mori. Estas técnicas son muy específicas pero poco sensibles, puesto que la expulsión de las larvas se produce de forma intermitente (Dreyer *et al.*, 1996; Requena-Méndez *et al.*, 2013). Además, es necesario que la larva sea viable, lo que dificulta mucho más su diagnóstico, cuestionando de esta forma la fiabilidad de las técnicas parasitológicas como *gold standard* o prueba de referencia para determinar la enfermedad.

Otra técnica de diagnóstico directo es la detección molecular mediante la técnica de reacción en cadena de la polimerasa (PCR). Son numerosas las ventajas que presentan las técnicas moleculares, en especial las basadas en tiempo real, destacando una mejora en la sensibilidad y la especificidad frente a las técnicas parasitológicas, así como mayor rapidez en la realización de la técnica y menor probabilidad de contaminación de las muestras.

Como complemento a las técnicas directas, se utilizan técnicas serológicas basadas en la detección de los anticuerpos anti-*Strongyloides* en el suero del paciente. Este diagnóstico es muy sensible pero poco específico debido a la existencia de reacciones cruzadas con otras helmintiasis afines (filarias, *Toxocara*, *Ascaris*, *Schistosoma*...), y no discrimina entre infecciones pasadas y activas.

Para la evaluación de las pruebas diagnósticas utilizamos dos conceptos clave, el de sensibilidad y especificidad. La sensibilidad es la probabilidad de que un individuo dé positivo en el test estando enfermo y la especificidad es la probabilidad de que un individuo dé negativo en el test estando sano. Se definen como Verdaderos Positivos (VP) aquellos casos de enfermos en los que el test ha dado positivo, como Verdaderos Negativos (VN) aquellos casos de sanos en los que el test ha dado negativo, como Falsos Positivos (FP) aquellos casos de sanos en los que el test ha dado positivo y como Falsos Negativos (FN) aquellos casos de enfermos en los que el test ha dado negativo. Así la sensibilidad se puede definir como el cociente entre los VP y VP+FN, y la especificidad como el cociente entre

VN y VN+FP. En términos de diagnóstico la mejor prueba sería aquella que tuviera elevados valores tanto de sensibilidad como de especificidad.

El problema planteado en este trabajo ante la ausencia de una técnica de *gold standard* para el diagnóstico de la estrongiloidiasis se abordará desde la metodología Bayesiana. La Estadística Bayesiana se apoya en la interpretación subjetiva de la probabilidad, es decir, en el grado de creencia que se tiene sobre la ocurrencia de un suceso y tiene como eje el Teorema de Bayes, desarrollado por Thomas Bayes y publicado, tras su muerte, en 1763. Una primera interpretación de este teorema, es que permite “actualizar” la probabilidad de un suceso A, $P(A)$, cuando se conoce que ha ocurrido un suceso B, pudiéndose obtener así la probabilidad de A condicionada a B, $P(A|B)$ (Silva, L.C. & Benavides, A., 2001).

Las ventajas que presenta la Estadística Bayesiana frente al enfoque clásico son numerosas, destacando que no se necesita un gran tamaño muestral, y parte de un conocimiento previo aprovechable. Este último hecho ha generado mucha controversia a este tipo de análisis ante la sospecha de que la aportación subjetiva del investigador podría incorporar una información interesada por parte del mismo. Sin embargo, una buena justificación de los conocimientos que se tienen a priori respaldado por una búsqueda bibliográfica adecuada puede dar a nuestro conocimiento la base sólida para considerar esta información inicial como certera y fiable. Actualmente existe una lista conocida por sus siglas en inglés como la WAMBS (When to worry and how to Avoid the Misuse of Bayesian Statistics) donde se describen las principales razones en la que el uso de la Estadística Bayesiana está justificada (Depaoli, S. & van de Schoot, R., 2017): algunos modelos complejos no pueden ser estimados con la metodología frecuentista; algunos investigadores opinan que los parámetros de la población deben ser vistos como aleatorios por la propia naturaleza del problema; se puede incorporar la información que se tiene de los parámetros y que este conocimiento se puede ir actualizando.

La metodología Bayesiana consta de los siguientes pasos fundamentales:

1. Definir el modelo de probabilidad correspondiente a las observaciones recogidas donde se relacionen los datos y los parámetros desconocidos (verosimilitud).
2. Especificar un modelo o distribución de probabilidad a priori para los parámetros de interés que recoja el conocimiento previo que se tiene de ellos. Se denomina elicitación (*elicitation*) a la determinación de los parámetros (hiperparámetros) que definen la distribución a priori de los parámetros de interés (Van de Schoot *et al.*, 2014). Existen tres clases de información a priori dependiendo del grado de información disponible: (1) a priori no informativa (2) a priori débilmente informativa y (3) a priori informativa. La a priori informativa es crucial para estimar el modelo y puede tener un gran impacto en la estimación final (Van de Schoot *et al.*, 2014), sin embargo la distribución a priori no informativa a menudo se asemeja a un análisis frecuentista, en la que son los datos los que determinan la distribución a posteriori del parámetro de interés. Para determinados parámetros como por ejemplo las proporciones, estos tres niveles de información son creados al modificar los hiperparámetros de la distribución a priori. Cada parámetro en el modelo necesita de una distribución a priori, y el uso de la distribución beta como a priori está justificado en el caso de proporciones por su propia definición, además dadas las propiedades de dicha distribución es posible simplificar los cálculos en la aplicación de la metodología bayesiana.

3. Actualizar el conocimiento de los parámetros, condicionando el modelo a los datos observados (distribución a posteriori). La complejidad de los modelos y el enorme avance experimentado en los cálculos computacionales permite que, en la mayoría de los casos, la distribución a posteriori se obtenga por simulación, utilizando el método MCMC con el muestreador de Gibbs (Enoe *et al.*, 2000) u otros algoritmos como Metrópolis-Hastings. Existen numerosos paquetes estadísticos para implementar la metodología Bayesiana: WinBUGS, OpenBUGS, JAGS o STAN y diversos paquetes de R.

4. Evaluar el ajuste del modelo a los datos y la sensibilidad del modelo a modificaciones en los parámetros.

Por todo ello, el objetivo planteado en este estudio es estimar tanto la prevalencia de la estrogiloidiasis en población infantil de una zona rural de Cubal (Angola) como la sensibilidad y especificidad de dos pruebas diagnósticas específicas de esta enfermedad, la técnica de Baermann y PCR en tiempo real, empleadas en el contexto de ausencia de *gold standard*.

2. METODOLOGÍA

Población y área de estudio

Se recogieron 315 muestras de heces de población infantil (5-14 años). El área geográfica escogida fue una zona rural del municipio de Cubal, provincia de Benguela (Angola). Este municipio está dividido en cuatro comunas: Cubal, Tumbulo, Capupa y Yambala. Las muestras se recogieron entre enero y junio de 2015 en un total de 16 escuelas distribuidas de forma igualitaria en las 4 comunas. De forma paralela a la recogida de muestras se obtuvo información a través de un cuestionario acerca de los diferentes hábitos de carácter social: utilización de zapatos, consumo de carne cruda, lavado de vegetales, lugar de baño... Al disponer de estos datos y sabiendo que la transmisión de la enfermedad se produce, en la mayoría de los casos, por la penetración de la larva al caminar descalzo en terrenos contaminados con heces, tanto de animales como de personas infectadas se dividió a la población en dos subgrupos (población 1 que camina descalzo y población 2 que camina con zapatos). Tanto la técnica parasitológica de Baermann como la concentración y extracción de ADN de las muestras de heces se realizaron en el Hospital Nossa Senhora da Paz en Cubal. La técnica molecular de la PCR se realizó en el Instituto de Salud Carlos III en Majadahonda (Centro Nacional de Microbiología). Este estudio ha sido aprobado por el Comité Ético del Hospital Vall d'Hebron y por el Instituto regional de Salud y Educación de Angola.

Técnicas diagnósticas

Una vez recogidas las muestras de heces se realizaron dos técnicas de diagnóstico: la técnica de Baermann y la PCR a tiempo real. La técnica de Baermann se basa en el termotropismo e hidrotropismo positivo que poseen las larvas vivas de *Strongyloides stercoralis*, de manera que las larvas migran desde las heces hacia el agua a 37°C, depositándose por gravedad en

el fondo del vaso de precipitación. Cada muestra se verificó tres veces por tres microscopistas diferentes y experimentados.

La PCR a tiempo real consiste en la amplificación específica del ADN del parásito. Para ello se emplearon los cebadores descritos por Verweij *et al.*, (2009) que permiten amplificar un fragmento del gen 18S rRNA. Previamente las muestras de heces fueron concentradas mediante el sistema Bioparaprep MINI® (Leti Diagnostics, Barcelona, España) y el ADN se extrajo empleando el kit QiaAmp Fast DNA Stool Mini kit (Qiagen, Hilden, Alemania).

Metodología estadística

La metodología a utilizar está basada en el paradigma Bayesiano que permite actualizar la información de los parámetros de interés a partir de la información contenida en los datos. En nuestro caso, los datos se corresponden al formato éxito-fracaso, existiendo únicamente dos posibilidades (positivo o negativo en el test, y presencia o ausencia de la enfermedad). Al tratar conjuntamente los resultados de dos pruebas diagnóstico, el esquema de los datos que tenemos en cada una de las dos poblaciones consideradas (pies descalzos y pies calzados) sigue una distribución multinomial

$$Y \sim \text{multinomial}(n, p_{11}, p_{12}, p_{21}, p_{22}) \quad \text{con} \quad \sum_{i=1}^2 \sum_{j=1}^2 p_{ij} = 1 \quad ; \quad 0 \leq p_{ij} \leq 1$$

Es decir:

$$P(Y=(n_{11}, n_{12}, n_{21}, n_{22})) = \frac{n!}{n_{11}! n_{12}! n_{21}! n_{22}!} p_{11}^{n_{11}} p_{12}^{n_{12}} p_{21}^{n_{21}} p_{22}^{n_{22}}$$

donde

$$\sum_{i=1}^2 \sum_{j=1}^2 n_{ij} = n$$

La n representa el tamaño total de la población, p_{11} corresponde a la probabilidad de tener un resultado positivo en ambas pruebas, p_{12} a la probabilidad de dar un positivo por PCR y negativo por Baermann, p_{21} la probabilidad de dar un positivo por Baermann pero negativo por PCR y p_{22} la probabilidad de dar negativo en ambas técnicas.

Basándonos en el artículo de Branscum *et al.* (2005) se ha tomado el modelo propuesto por Hui and Walter (1980) para dos poblaciones y dos test diagnósticos condicionalmente independientes dada la enfermedad. Los dos test diagnósticos se aplican a las dos poblaciones y con esos resultados se construyen tablas de contingencia 2x2. Este modelo asume que los test son independientes condicionalmente dada la enfermedad y que los valores de sensibilidad y especificidad de ambas técnicas son iguales para las dos poblaciones. Se asumió independencia condicional entre los test porque aunque ambos son técnicas de diagnóstico directo la base de la metodología es diferente, en la técnica de Baermann se visualiza el parásito mediante microscopía mientras que en la técnica de PCR se detecta ADN del parásito.

En nuestro caso de dos técnicas diagnósticas y dos poblaciones en total tenemos 6 parámetros que queremos estimar:

- Sepcr: sensibilidad de la técnica PCR
- SpPCR: especificidad de la técnica PCR
- Sebaermann: sensibilidad de la técnica Baermann
- Spbaermann: especificidad de la técnica Baermann
- π_1 : prevalencia de la población 1
- π_2 : prevalencia de la población 2

Dado que nuestro interés se centra en las dos poblaciones previamente, niños que caminan descalzos y que caminan con zapatos, en la población k ($k=1,2$) la variable Y_k sigue una distribución multinomial

$$Y_k \sim \text{multinomial}(n_k, p_{11k}, p_{12k}, p_{21k}, p_{22k}),$$

Los datos recogidos se pueden estructurar tal y como se muestra en la **tabla 1**.

Población 1: Caminan descalzos		Técnica Baermann	
		Positivo	Negativo
Técnica PCR	Positivo	n_{111}	n_{121}
	Negativo	n_{211}	n_{221}

Población 2: Caminan con zapatos		Técnica Baermann	
		Positivo	Negativo
Técnica PCR	Positivo	n_{112}	n_{122}
	Negativo	n_{212}	n_{222}

Tabla 1. Tablas de contingencia de comparación de técnicas diferenciando las dos poblaciones.

Expresado en términos de nuestros parámetros de interés dados por las prevalencias de la estrongiloidiasis en ambas poblaciones y las sensibilidades y especificidades de ambas técnicas:

Para $k=1$ que corresponde a la población 1: Caminan descalzos

$$p_{111} = \pi_1 * Sepcr * Sebaermann + (1 - \pi_1) * (1 - SpPCR) * (1 - Spbaermann)$$

$$p_{121} = \pi_1 * Sepcr * (1 - Spbaermann) + (1 - \pi_1) * (1 - SpPCR) * Spbaermann$$

$$p_{211} = \pi_1 * (1 - Sepcr) * Sebaermann + (1 - \pi_1) * SpPCR * (1 - Spbaermann)$$

$$p_{221} = \pi_1 * (1 - Sepcr) * (1 - Sebaermann) + (1 - \pi_1) * SpPCR * Spbaermann$$

Para $k=2$ que corresponde a la población 2: Caminan con zapatos

$$p_{112} = \pi_2 * Sepcr * Sebaermann + (1 - \pi_2) * (1 - SpPCR) * (1 - Spbaermann)$$

$$p_{122} = \pi_2 * Sepcr * (1 - Spbaermann) + (1 - \pi_2) * (1 - SpPCR) * Spbaermann$$

$$p_{212} = \pi_2 * (1 - Sepcr) * Sebaermann + (1 - \pi_2) * SpPCR * (1 - Spbaermann)$$

$$p_{222} = \pi_2 * (1 - Sepcr) * (1 - Sebaermann) + (1 - \pi_2) * SpPCR * Spbaermann$$

Distribuciones a priori de la prevalencia y de la especificidad y sensibilidad de las pruebas.

Se denomina “elicitación” (*elicitation*) a la determinación de los valores de los hiperparámetros que definen la distribución a priori de los parámetros.

Dado que los parámetros de interés son probabilidades (sensibilidad PCR, sensibilidad Baermann, especificidad PCR, especificidad Baermann, prevalencia en la población de niños descalzos y prevalencia en la población de niños calzados) es adecuado utilizar la

distribución Beta. Una de las principales ventajas de esta distribución es su ajuste a una gran variedad de distribuciones empíricas, adoptando muy diversas formas dependiendo de los valores de los parámetros.

En este trabajo, se han establecido dos formas distintas de elicitación de los parámetros α y β de las distribuciones Beta a priori que se asignarán a los parámetros de interés.

El primero de ellos se basó en la búsqueda bibliográfica a través de la base de datos PubMed introduciendo varios criterios: *strongyloides* AND *prevalence*, *strongyloides* AND *sensitivity*, *strongyloides* AND *specificity* y únicamente se seleccionaron aquellos artículos llevados a cabo en países africanos por las posibles coincidencias en cuanto a condiciones de trabajo y hábitos de vida de los propios pacientes. Se han incluido todos los estudios que aparecían tras la búsqueda con los criterios de selección, independientemente de si proporcionaban todos los datos.

La segunda forma fue considerando como distribuciones a priori las distribuciones a posteriori obtenidas con los datos de un estudio previo de condiciones similares realizado en un área rural de Etiopía. La elicitación de la distribución a priori se obtuvo de nuevo con los datos bibliográficos.

La elicitación de todos los parámetros α y β se realizó mediante la librería LearnBayes incluida dentro del programa R (Version 1.0.44 - © 2009-2016 RStudio, Inc.).

Estas propuestas de elicitaciones dan lugar a plantear tres posibles modelos distintos:

Modelo 1. Modelo con distribuciones a priori obtenidas a partir de búsqueda bibliográfica.

Modelo 2. Modelo con distribuciones a priori obtenidas a partir de los datos del estudio de una población de Etiopía.

Modelo 3. Modelo con distribuciones a priori para las sensibilidades y especificidades a partir de búsqueda bibliográfica y para las prevalencias a partir de los datos de Etiopía.

Cálculo de las distribuciones a posteriori

Se van a generar muestras de la distribución de probabilidad a posteriori de los parámetros a través del método MCMC.

Se han generado tres cadenas con distintos puntos iniciales elegidos aleatoriamente para asegurar la independencia del punto de partida, un total de 100000 iteraciones, eliminando (*burn-in*) las 10000 primeras. La estimación del modelo se realizó mediante la librería R2OpenBugs del paquete R, que llama al programa OpenBUGS a través de la función *bugs*.

Estudio de la convergencia y comparación de modelos

El estudio de la convergencia se realizó mediante dos vías: gráfica y analíticamente. La primera fue mediante comparación por visualización directa de los gráficos de trazas proporcionados por el programa OpenBUGS para cada parámetro estimado (sensibilidad y especificidad de ambas técnicas y prevalencia de la enfermedad en ambas poblaciones) de las tres cadenas generadas. La segunda comprobación fue a través de dos criterios analíticos: el criterio de convergencia PSRF (*Potencial Scale Reduction Factor*) o Brooks-Gelman-Rubin y el criterio de convergencia de Geweke, dichos criterios se calcularon con la librería *coda* (*Convergence Diagnosis and Output Analysis*) de R.

El criterio de convergencia PSRF se basa en comparar la variabilidad entre cadenas con la variabilidad dentro de las cadenas, de forma que los valores cercanos a 1 indican una buena convergencia de las cadenas. El criterio de convergencia de Geweke estudia la convergencia de cada cadena individualmente mediante la comparación de las medias de la misma cadena en diferentes puntos de ésta, por ejemplo la media de un porcentaje del principio de la cadena y un porcentaje del final. Si existe convergencia no se detectarán diferencias entre los valores medios de ambos tramos y, por tanto, el valor del estadístico Z de la normal debe estar entre (-1.96, 1.96).

Para poder comparar los 3 modelos enunciados que se obtienen se utilizó el criterio DIC (*Deviance Information Criterion*) que es una medida de la validez de ajuste del modelo a los datos y se genera automáticamente en el programa OpenBUGS en el apartado de *Deviance Information*. Su definición viene dada por:

$$DIC = Dbar + pD = Dhat + 2 pD$$

donde $Dbar$ es la media a posteriori de la *deviance* (siendo $deviance(\theta) = -2 \cdot \log p(y|\theta)$), $Dhat$ es la *deviance* evaluada en la media a posteriori y $pD = Dbar - Dhat$ representando el número efectivo de parámetros.

Cuanto menor es el valor obtenido mejor es el ajuste, mayor verosimilitud de los datos con las distribuciones a posteriori de los parámetros, por lo que se escogerá aquel modelo que su DIC total sea menor.

3. RESULTADOS

Modelo 1. Modelo con distribuciones a priori obtenidas a partir de búsqueda bibliográfica.

Distribuciones a priori

Después de la lectura de los artículos, se consideró que al no haber una técnica *gold standard*, los datos de prevalencia proporcionados en ellos realmente reflejan el número de positivos obtenidos por una de las técnicas diagnósticas. Es por este motivo por el que para la elicitación de los parámetros α y β de las prevalencias de ambas poblaciones se utilizó una distribución uniforme Beta (1,1) (no informativa). Por el contrario, sí tenemos datos fiables de la sensibilidad y especificidad de las técnicas diagnósticas. En el caso de la PCRrt se le dio mayor peso al artículo publicado por Saugar *et al.* (2015) ya que está técnica se realiza en las mismas condiciones, en el mismo laboratorio y con el mismo personal que la utilizada para este estudio. La sensibilidad y especificidad descritas en este artículo fueron de 0.94 y 0.87, respectivamente. Estos resultados fueron similares a los descritos por otros autores, Schär *et al.* 2013 (sensibilidad 0.89, especificidad 0.93), Sharifdini *et al.* 2015 (sensibilidad 0.85 y especificidad 0.96), Martins de Paula *et al.* 2015 (sensibilidad 0.9, especificidad 0.86). Con todos estos datos se estableció que los valores de sensibilidad de la PCR se encontraban entre 0.7 y 0.9 con una probabilidad de 0.8 dejando un 0.1 de probabilidad por debajo de 0.7 y un 0.1 por encima de 0.9. Los valores de especificidad de la PCR se encontraban entre 0.7 y 0.9 con una probabilidad de 0.8 y dejando por debajo del 0.7 una probabilidad de 0.05 y por encima de 0.9 una de 0.15. Con estos valores introducidos como cuantiles inferior y superior se obtiene un $\alpha = 19.87$ y un $\beta = 4.80$ para la sensibilidad y un $\alpha = 23.01$ y un $\beta = 4.78$ para la especificidad (tabla 2. Modelo 1).

En el caso de la técnica de Baermann en la mayoría de los artículos se habla de casi un 100% de especificidad, ya que la probabilidad de confusión de la larva de *Strongyloides* con otro parásito o artefacto es muy baja. En cuanto a la sensibilidad de la técnica, los datos proporcionados en la bibliografía son muy dispares, ya que además de ser una técnica limitada por las condiciones intrínsecas de la metodología (por ejemplo necesidad de tener la larva viva), depende en gran medida de la experiencia del investigador. De esta manera, se observan datos del 0.38% (Becker *et al.*, 2015), 0.60% (Lau *et al.*, 2005; Assefa *et al.*, 1991).

Se estableció que los valores de sensibilidad de la técnica Baermann se encontraban entre 0.3 y 0.6 con una probabilidad del 0.7 dejando un 0.15 de probabilidad por debajo de 0.3 y un 0.15 de probabilidad por encima de 0.7. Los valores de especificidad de la técnica Baermann se encontraban entre 0.9 y 0.98 con una probabilidad de 0.9 dejando un 0.05 por debajo de 0.9 y un 0.05 por encima de 0.98.

Con estos valores introducidos como cuantiles inferior y superior se obtiene un $\alpha = 5.28$ y un $\beta = 6.47$ para la sensibilidad y un $\alpha = 77.32$ y un $\beta = 4.40$ para la especificidad (tabla 2. Modelo 1).

	Modelo 1		Modelo 2		Modelo 3	
	α	β	α	β	α	β
SePCR	19,87	4,8	19,69	5,34	19,87	4,8
SpPCR	23,01	4,78	132,22	9,08	23,01	4,78
SeBaermann	5,28	6,47	12,83	9,02	5,28	6,47
SpBaermann	77,32	4,4	211	12,06	77,32	4,4
pi1	1	1	11,82	92,89	11,82	92,89
pi2	1	1	11,82	92,89	11,82	92,89

Tabla 2. Valores de los parámetros α y β de las distribuciones Beta a priori en los tres modelos planteados.

Distribuciones a posteriori

Los resultados obtenidos muestran que la sensibilidad para la técnica Baermann se encuentra con un 95% de probabilidad entre 0.28-0.69, mientras que la sensibilidad de la PCR entre 0.63-0.92. La especificidad de Baermann con un 95% de probabilidad fue de 0.92-0.98 y la de la PCR 0.81 y 0.94. La prevalencia de la enfermedad para la población 1 está entre 0.08 y 0.29 y para la población 2 entre 0.03 y 0.24. En todos los parámetros la media y la mediana son similares. En la figura 1 se recogen las densidades a posteriori de todos los parámetros del modelo 1.

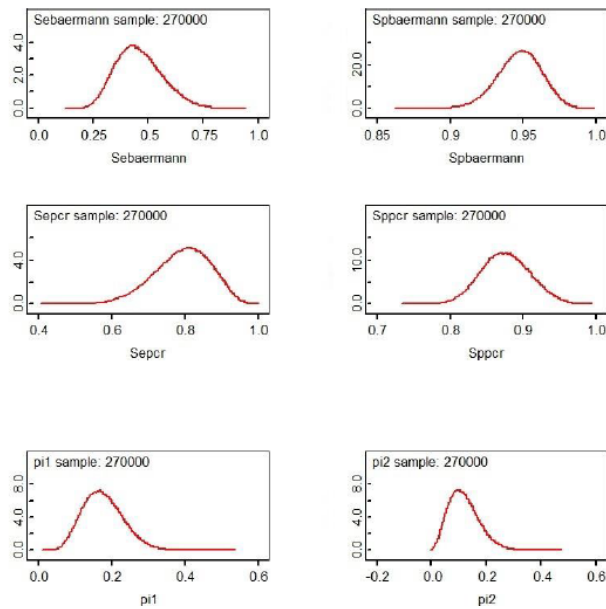


Figura 1. Densidades a posteriori de los parámetros del modelo 1.

En el anexo 1, se recoge un resumen de las distribuciones a posteriori de los parámetros de interés de los modelos planteados

Modelo 2. Modelo con distribuciones a priori obtenidas a partir de los datos del estudio de una población de Etiopía.

Distribuciones a priori

Para la elicitación de las distribuciones a priori de los parámetros de interés en este modelo (dos poblaciones de niños de Angola) es necesario el cálculo de las distribuciones a posteriori de dichos parámetros en un modelo “intermedio”, con una sola población de niños de Etiopía. En este modelo “intermedio” la elicitación de las distribuciones a priori se obtiene de la misma manera que en el modelo 1, es decir a partir de la información bibliográfica, por lo que los hiperparámetros de los parámetros y respectivas distribuciones son las mismas que en el modelo anterior. Estas a priori se incluyeron dentro del modelo “intermedio” y se determinó la expresión de la verosimilitud introduciendo los datos de Etiopía para una sola población. Vamos a utilizar los valores de los dos cuantiles de la distribución a posteriori del modelo intermedio para elicitar los α y β de los parámetros de interés. Tal como se recoge en el modelo 2 de la tabla 2, se obtiene que para la sensibilidad de la técnica PCR el $\alpha = 19.69$ y el $\beta = 5.34$. Para la especificidad de la técnica PCR se obtuvo un $\alpha = 132.22$ y el $\beta = 9.08$. Se obtuvo que para la sensibilidad de la técnica Baermann el $\alpha = 12.83$ y el $\beta = 9.02$. Se obtuvo que para la especificidad de la técnica Baermann el $\alpha = 211.0$ y el $\beta = 12.6$. Se obtuvo para la prevalencia de la estrongiloidiasis un $\alpha = 11.82$ y el $\beta = 92.89$.

Distribuciones a posteriori

La sensibilidad para la técnica Baermann se encuentra con un 95% de probabilidad entre 0.36-0.69, mientras que la sensibilidad de la PCR entre 0.68-0.93. La especificidad de Baermann con un 95% de probabilidad fue de 0.92-0.97 y la de la PCR 0.85 y 0.94. La prevalencia de la enfermedad para la población 1 (niños descalzos) está entre 0.09 y 0.2 y para la población 2 (niños calzados) entre 0.06 y 0.17 (anexo 1). En la figura 2 se recogen las densidades a posteriori de todos los parámetros del modelo 2.

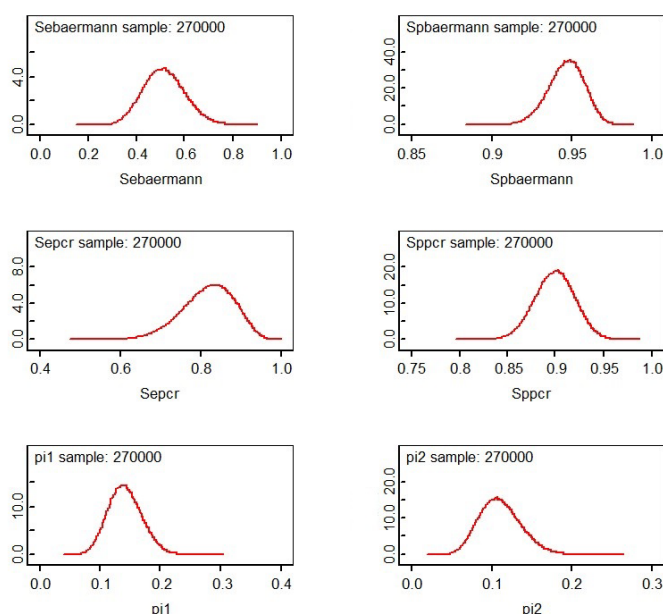


Figura 2. Densidades a posteriori de los parámetros del modelo 2.

Modelo 3. Modelo con distribuciones a priori para las sensibilidades y especificidades a partir de búsqueda bibliográfica y para las prevalencias a partir de los datos de Etiopía.

Distribuciones a priori

La elicitación de los valores α y β para los parámetros de sensibilidad y especificidad de las técnicas diagnósticas son los mismos que los obtenidos para el modelo 1. Sin embargo, para la prevalencia se utilizaron los valores α y β obtenidos del modelo 2. Estos hiperparámetros se pueden encontrar en el modelo 3 de la tabla 2.

Distribuciones a posteriori

La sensibilidad para la técnica Baermann se encuentra con un 95% de probabilidad entre 0.32-0.72, mientras que la sensibilidad de la PCR entre 0.66-0.93. La especificidad de Baermann con un 95% de probabilidad fue de 0.91-0.97 y la de la PCR 0.81 y 0.91. La prevalencia de la enfermedad para la población 1 está entre 0.08 y 0.19 y para la población 2 entre 0.06 y 0.16 (anexo 1). En la figura 3 se recogen las densidades a posteriori de todos los parámetros del modelo 3.

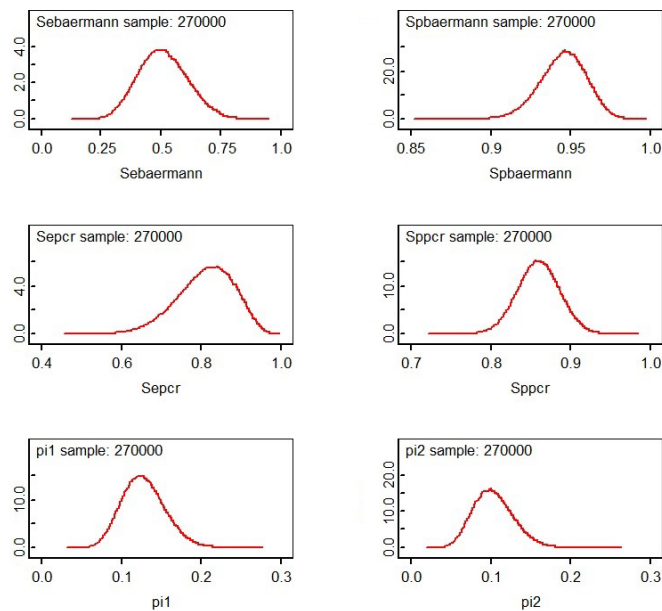


Figura 3. Densidades a posteriori de los parámetros del modelo 3.

La utilización de los datos del estudio realizado en Etiopía proporcionó una distribución a priori informativa de la prevalencia, introducida en los modelos 2 y 3, a diferencia de la situación no informativa para las prevalencias del modelo 1.

Estudio de convergencia para los tres modelos.

De forma gráfica, en todas las trazas se observó una buena convergencia en cada cadena para todos los parámetros analizados observándose una estabilidad tras el proceso de *burn-in* de las primeras iteraciones.

De forma analítica, se observó convergencia entre las cadenas ya que o bien el criterio de Geweke tomó valores en el intervalo $(-1,96; 1,96)$ o el valor de PSRF fue igual a 1.

Por consiguiente, tanto gráfica como analíticamente se puede concluir que todas las cadenas generadas convergen, para los tres modelos planteados.

Comparación de modelos

Los resultados de los tres modelos son muy similares tanto de los valores de los parámetros que aparecen en las tablas como en el valor de la deviance, tal como se refleja en el anexo 1. El modelo 1, que es el que menor información incorporaba es el que mayor valor de DIC tiene, por lo que lo descartamos. En cuanto al modelo 2 y 3 son los que presentan resultados muy similares siendo el menor valor del DIC para el modelo 3.

La sensibilidad para la técnica Baermann y su intervalo de probabilidad al 95% fue del 50.89% (32.07%-71.81%), mientras que para PCRrt fue del 81.08% (65.63%-93.10%). La especificidad para Baermann fue del 94.53% (91.45%-97.19%) y para PCRrt del 86.01% (80.77%-91.17%). Las prevalencias estimadas fueron 12.83% (7.99%-18.55%) para la población que camina descalza y 10.38% (5.94%-15.81%) para la que camina calzada.

4. CONCLUSIONES

Los resultados obtenidos en este trabajo están en consonancia con las distintas publicaciones que abordan esta misma problemática. Sin embargo, con la Estadística Bayesiana se ha podido estimar con mayor precisión tanto la sensibilidad y especificidad de ambas técnicas, como la prevalencia de la enfermedad en una zona donde los estudios eran escasos y la situación totalmente desconocida. Esto supone un reto ya que a pesar de que existen estudios similares que tratan de estimar la prevalencia de la estrongiloidiasis, en algunos casos los intervalos resultantes eran menos precisos (Morawski BN *et al.*, 2017) o similares (Amor *et al.*, 2016).

En este trabajo, la Estadística Bayesiana ha permitido incorporar la información de otros estudios y conjugarlo con los datos recogidos inicialmente en población infantil de Etiopía y posteriormente en población infantil de Cubal (Angola). Este procedimiento permite ajustar con mayor precisión todos los parámetros de interés, con especial atención a la prevalencia de la estrongiloidiasis. A la vista de estos resultados se puede concluir que la Estadística Bayesiana permite estimar la validez de las pruebas diagnóstico en estudio así como determinar la prevalencia de estrongiloidiasis en ausencia de *gold standard*.

5. REFERENCIAS

Amor A, Rodriguez E, Saugar JM, Arroyo A, López-Quintana B, Abera B, Yimer M, Yizengaw E, Zewdie D, Ayehubizu Z, Hailu T, Mulu W, Echazú A, Krolewieki AJ, Aparicio P, Herrador Z, Anegagrie M, Benito A (2016). High prevalence of *Strongyloides stercoralis* in school-aged children in a rural highland of north-western Ethiopia: the role of intensive diagnostic work-up. *Parasit Vectors*. 1;9(1):617.

Assefa T., Woldemichael T., Seyoum T (1991). Evaluation of the modified Baermann's method in the laboratory diagnosis of *Strongyloides stercoralis*. *Ethiop Med J*. 29(4):193-8.

Becker SL, Piraisoody N, Kramme S, Marti H, Silué KD, Panning M, Nickel B, Kern WV, Herrmann M, Hatz CF, N'Goran EK, Utzinger J, von Müller L (2015). Real-time PCR for detection of *Strongyloides stercoralis* in human stool samples from Côte d'Ivoire: diagnostic accuracy, inter-laboratory comparison and patterns of hookworm co-infection. *Acta Trop*. 2015 Oct;150:210-7. doi: 10.1016/j.actatropica.2015.07.019.

Branscum AJ1, Gardner IA, Johnson WO (2005). Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Prev Vet Med*. 10;68(2-4):145-63.

Depaoli, S, & van de Schoot, R (2017). Improving transparency and replication in Bayesian Statistics: The WAMBS-Checklist. *Psychol Methods*. 22(2):240-261

Dreyer G, Fernandes-Silva E, Alvez S, Rocha A, Albuquerque R, Addis D (1996). Patterns of detection of *Strongyloides stercoralis* in stool specimens: implications for diagnosis and clinical trials. *J Clin Microbiol*. 34: 2569-71.

- Enøe C, Georgiadis MP, Johnson WO (2000). Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev Vet Med.* 30;45(1-2):61-81.
- Hui SL, Walter SD (1980). Estimating the error rates of diagnostic tests. *Biometrics.* 36(1):167-71.
- Lau C., Samalvides F., Terashima A (2005). Evaluación de técnicas parasitológicas en el diagnóstico de estrombiloidiasis por *Strongyloides stercoralis*. *Rev Med Hered* 16:11-18.
- Morawski BM, Yunus M, Kerukadho E, Turyasingura G, Barbra L, Ojok AM, DiNardo AR, Sowinski S, Boulware DR, Mejia R (2017). Hookworm infection is associated with decreased CD4+ T cell counts in HIV-infected adult Ugandans. *PLoS Negl Trop Dis.* 2017 May 25;11(5):e0005634. doi: 10.1371/journal.pntd.0005634.
- Paula FM, Malta Fde M, Marques PD, Sitta RB, Pinho JR, Gryscek RC, Chieffi PP (2015). Molecular diagnosis of strongyloidiasis in tropical areas: a comparison of conventional and real-time polymerase chain reaction with parasitological methods. *Mem Inst Oswaldo Cruz.* 2015 Apr;110(2):272-4. doi: 10.1590/0074-02760140371.
- Requena-Méndez A, Chiodini P, Bisoffi Z, Buonfrate D, Gotuzzo E, Muñoz J (2013). The Laboratory Diagnosis and Follow Up of Strongyloidiasis: A Systematic Review. *PLoS Negl Trop Dis.* 7: issue 1/e 2002.
- Ruano A, Martín T, Pardo J, López-Abán J, Muro A (2005). Avances en el estudio sobre la estrombiloidiasis. *Enfermedades Emergentes.* 7: 102-109.
- Saugar, J.M., Merino,F.J., Martin-Rabadan,P., Fernandez-Soto,P., Ortega,S., Garate, T., y Rodriguez, E (2015). Application of real-time PCR for the detection of *Strongyloides* spp. in clinical samples in a reference center in Spain. *Act trop.* 142: 20-25.
- Schar, F., P. Odermatt, V. Khieu, M. Panning, S. Duong, S. Muth, H. Marti & S. Kramme (2013). Evaluation of real-time PCR for *Strongyloides stercoralis* and hookworm as diagnostic tool in asymptomatic schoolchildren in Cambodia. *Act tropic.* 126: 89-92.
- Sharifdini M, Mirhendi H, Ashrafi K, Hosseini M, Mohebbali M, Khodadadi H, Kia EB (2015). Comparison of Nested Polymerase Chain Reaction and Real-Time Polymerase Chain Reaction with Parasitological Methods for Detection of *Strongyloides stercoralis* in Human Fecal Samples. *Am J Trop Med Hyg.* 93(6):1285-91. doi: 10.4269/ajtmh.15-0309.
- Silva, LC & Benavides, A (2001). El enfoque bayesiano: otra manera de inferir. *Gac Sanit.* 15 (4): 341-346.
- Vadlamudi RS, Chi DS, Krishnaswamy G (2006). Intestinal strongyloidiasis and hyperinfection syndrome. *Clin Mol Allergy.* 30: 4-8.
- Van de Schoot R, Kaplan D, Denissen J, Asendorpf JB, Neyer FJ, van Aken MA (2014). A gentle introduction to bayesian analysis: applications to developmental research. *Child Dev.* 85(3):842-60.

Dacal, *et al.*

Verweij, J.J., M. Canales, K. Polman, J. Ziem, E.A. Brienen, A.M. Polderman & L. van Lieshout (2009). Molecular diagnosis of *Strongyloides stercoralis* in faecal samples using real-time PCR. *Trans R Soc Trop Med Hyg.* 103: 342-346.

ANEXO 1. Valores α y β de la distribución a priori y la media más desviación estándar (D.E.) de la sensibilidad de la PCR y del Baermann, la especificidad de la PCR y Baermann y prevalencias de población de niños descalzos (1) y de niños calzados (2) de la distribución a posteriori.

		Sensibilidad PCR		Sensibilidad Baermann		Especificidad PCR		Especificidad Baermann		Prevalencia 1	Prevalencia 2
		A priori $\alpha;\beta$	A posteriori Media \pm DE IP95%	A priori $\alpha;\beta$	A posteriori Media \pm DE IP95%	A priori $\alpha;\beta$	A posteriori Media \pm DE IP95%	A priori $\alpha;\beta$	A posteriori Media \pm DE IP95%		
Modelo 1. Modelo con distribuciones a priori obtenidas a partir de búsqueda bibliográfica.	A priori $\alpha;\beta$	19,87; 4,8		5,28; 6,47		23,01; 4,78		77,32; 4,4		1; 1	1; 1
	A posteriori Media \pm DE IP95%	0,7925 \pm 0,077 (0,6272-0,9241)	0,4590 \pm 0,1065 (0,2763-0,6889)	0,8766 \pm 0,0333 (0,8130-0,9428)	0,9476 \pm 0,0153 (0,9156-0,9757)	0,1764 \pm 0,0553 (0,0808-0,2947)	0,1188 \pm 0,0559 (0,0292-0,2441)				
Modelo 2. Modelo con distribuciones a priori obtenidas a partir de los datos del estudio de una población de Etiopía.	A priori $\alpha;\beta$	19,69; 5,34		12,83; 9,02		132,22; 9,08		211; 12,06		11,82; 92,89	11,82; 92,89
	A posteriori Media \pm DE IP95%	0,8175 \pm 0,0656 (0,6754-0,9285)	0,5154 \pm 0,0855 (0,3572-0,6914)	0,8990 \pm 0,0210 (0,8570-0,9391)	0,9466 \pm 0,0113 (0,9227-0,9673)	0,1415 \pm 0,0281 (0,0903-0,2000)	0,1110 \pm 0,0262 (0,0648-0,1671)				
Modelo 3. Modelo con distribuciones a priori para las sensibilidades y especificidades a partir de búsqueda bibliográfica y para las prevalencias a partir de los datos de Etiopía.	A priori $\alpha;\beta$	19,87; 4,8		5,28; 6,47		23,01; 4,78		77,32; 4,4		11,82; 92,89	11,82; 92,89
	A posteriori Media \pm DE IP95%	0,8108 \pm 0,0713 (0,6563-0,9310)	0,5089 \pm 0,1023 (0,3207-0,7181)	0,8601 \pm 0,0265 (0,8077-0,9117)	0,9453 \pm 0,0146 (0,9145-0,9719)	0,1283 \pm 0,0270 (0,0798-0,1855)	0,1038 \pm 0,0253 (0,0594-0,1581)				

Modelos estadísticos de crecimiento poblacional para detectar denso-dependencia en ciervos mediterráneos

Concepción Azorit-Casas⁽¹⁾, Antonio Jesús López-Montoya⁽²⁾, Irene García-Garrido⁽²⁾

(1) Dpto. Biología Animal, Vegetal y Ecología, Universidad de Jaén, Paraje Las Lagunillas s/n, 23071 Jaén, España, e-mail: cazorit@ujaen.es

(2) Dpto. Estadística e Investigación Operativa, Universidad de Jaén, Paraje Las Lagunillas s/n, 23071 Jaén, España, e-mail: amontoya@ujaen.es, iggarrid@ujaen.es

RESUMEN

La detección de la denso-dependencia, como un factor intrínseco que influye en la dinámica poblacional de los ciervos, es interesante en ecología y es muy útil para una gestión cinegética adecuada. En este trabajo se han utilizado varios modelos estadísticos, como los modelos de regresión lineal subyacentes a las curvas de crecimiento de tipo logístico de Ricker y Gompertz para detectar denso-dependencia en dos poblaciones de ciervo ibérico (*Cervus elaphus hispanicus*) en el sur de España. Para ello, utilizamos series temporales de datos de caza como estimas de abundancia poblacionales. Para nuestras poblaciones de estudio, los modelos de Ricker y Gompertz indican la presencia de denso-dependencia aunque, cabe destacar que, el modelo de Gompertz detecta una mayor fuerza en la denso-dependencia con respecto al modelo de Ricker. Por tanto, este factor intrínseco en la población hay que considerarlo en la estimación de las tasas de caza para una mejor conservación de los ecosistemas mediterráneos.

Palabras clave: crecimiento logístico, modelo de Gompertz, denso-dependencia, dinámica poblacional.

ABSTRACT

The detection of the existence of denso-dependence in population fluctuations is important in studies of ecological theory, as well as for the control and management of the population. A set of statistical methods are used in this work, including linear regression models underlying the logistic growth curves of Ricker and Gompertz to detect denso-dependence in two populations of Iberian red deer (*Cervus elaphus hispanicus*) in the south of Spain. For this, time series of hunting data as estimates of population abundance are used. For our populations, the Ricker and Gompertz models indicate the presence of denso-dependence although, it should be noted that the Gompertz model detects a greater strength in the denso-dependence with respect to the Ricker model. Then, this intrinsic factor in the population must be considered in the estimation of hunting rates for better conservation of Mediterranean ecosystems.

Keywords: logistic growth, Gompertz model, denso-dependence, dynamics population.

1. INTRODUCCIÓN

Los cambios en el tamaño poblacional son un objetivo fundamental en la ecología de poblaciones. El análisis de series temporales puede utilizarse para identificar la existencia de denso-dependencia (DD), o influencia intraespecífica, ya que una relación negativa entre la tasa de crecimiento poblacional y la densidad de población produce una tendencia apreciable en los datos (Royama, T., 1992).

Bjornstad, O.N. et al., 2001, encontraron que los primeros análisis que indicaban la presencia de DD presentaban numerosas dificultades matemáticas y generaban resultados dudosos. En la actualidad los test son más robustos y menos sesgados, incluyendo los modelos autorregresivos de series temporales (Royama, T., 1992).

Los modelos de crecimiento poblacional de tipo logístico, asumen la existencia de un umbral de población (K). Este umbral indica la capacidad máxima de individuos que es capaz de soportar la población, por lo que estimarlo es muy importante para la gestión de la población. Bajo el supuesto de crecimiento poblacional logístico, se sabe que los recursos no son limitados cuando la población tiene baja densidad ($N < K/2$), lo cual no limita la tasa reproductiva, sin embargo el rendimiento de esa población será pequeño. Cuando la abundancia alcanza un valor cercano a la mitad de la capacidad de carga ($N = K/2$), la tasa de crecimiento de la población está en su punto máximo. Este punto es nombrado como rendimiento máximo sostenible (MSY), y es cuando se puede obtener el máximo de individuos cosechados. Conforme la abundancia aumenta, los factores denso-dependientes comienzan a limitar cada vez más el crecimiento hasta que la población llega al valor de K . En este momento, no hay excedentes que puedan ser cosechados y la producción cae a cero, es decir, la mayor cosecha sostenible no se alcanza cuando la abundancia de la población está al máximo de su K . Nuevamente, lo importante es que la tasa de crecimiento poblacional esté al máximo y esto sucede cuando $N = K/2$.

Es muy difícil obtener una estimación exacta de $K/2$ por cuestiones de muestreo, pero también porque este punto no es estático en el tiempo. Desde el punto de vista de cosecha se recomienda extraer individuos cuando la población se encuentre próxima al MSY.

En este trabajo se analiza la posible existencia de influencia de factores denso-dependientes sobre la dinámica poblacional en dos poblaciones de ciervo ibérico en el sur de España. Para ello utilizamos los modelos de crecimiento poblacional de Ricker y Gompertz (Imperio, S. et al., 2012; Turchin, P., 2003; Ricker, W.E., 1954). Las variables utilizadas en dichos modelos, son datos de caza (H_t) a lo largo del tiempo (11 años de estudio) y la tasa de crecimiento poblacional (r_t).

2. METODOLOGÍA

Modelos de crecimiento poblacional de Ricker y Gompertz

La detección de DD en una población que varía cada año es de vital importancia en la teoría ecológica así como en el control y gestión de poblaciones (Eberhardt, L.L. et al., 2012). Sin embargo, no existe un consenso claro sobre cuál es el método más fiable en la detección de DD en poblaciones animales (Simard, M.A. et al., 2012). Por tanto,

exploramos si existe DD en nuestras poblaciones de estudio, mediante modelos estadísticos de crecimiento poblacional.

La tasa de crecimiento poblacional se define generalmente como $r_t = \log(H_t/H_{t-1})$, es decir, la tasa per cápita del cambio realizado entre los tiempos $t - 1$ y t . Los test para detectar DD en series temporales pueden diferir en su sensibilidad y proporcionar resultados variables pero no se ha demostrado que haya ningún test mejor que otro (Simard, M.A. et al., 2012).

Utilizando un simple modelo de regresión lineal para poblaciones (sin considerar grupos de edad) podemos cuantificar la fuerza de la DD. En este modelo la DD se define como la pendiente β_1 de la recta de regresión entre el logaritmo natural del tamaño poblacional en el tiempo t y el logaritmo del tamaño poblacional en el tiempo $t - 1$. Una pendiente negativa y estadísticamente significativa de la siguiente expresión indica DD.

$$\log H_t = \beta_0 + \beta_1 \log H_{t-1} + \varepsilon_t. \tag{1}$$

Los modelos tradicionales para detectar la DD son el modelo estocástico de Gompertz, dado por la expresión:

$$H_t = H_{t-1} \exp(a + b \log(H_{t-1}) + \varepsilon_t) \tag{2}$$

y el modelo estocástico de Ricker, definido de la forma:

$$H_t = H_{t-1} \exp(a + bH_{t-1} + \varepsilon_t) \tag{3}$$

donde, en ambos modelos, los errores ε_t se supone que siguen una distribución Normal y que son incorrelados. La linealización de ambos modelos se utiliza para detectar la presencia de DD (una pendiente b negativa indica la presencia de DD) representando gráficamente los índices de densidad H_{t-1} frente a H_t , así como el ajuste de regresión siendo r_t la variable dependiente y H_{t-1} la variable independiente, permitiéndonos una visualización de la DD, como sugiere Turchin, P., 2003.

Ajustamos dichos modelos a las poblaciones de estudio y a los ciervos considerados.

3. RESULTADOS

Los resultados de los test de DD mediante los modelos de Ricker y Gompertz utilizando la caza como estima de abundancia poblacional vienen dados en la Tabla 1.

Tabla 1. Test de denso-dependencia para la población de ciervo ibérico en 2001-2001.

<i>Reserva</i>	<i>Modelo</i>	<i>b</i>	<i>p-valor</i>	<i>test Shapiro-Wilk</i>
R1	Ricker	-0.0018	0.1352	0.8507
	Gompertz	-0.7102	0.0532	0.1532
R2	Ricker	-0.0017	0.1163	0.0738
	Gompertz	-0.5428	0.1023	0.0335

En ambos casos el modelo de Gompertz detecta una mayor DD que el modelo de Ricker, con unos valores mayores (y negativos) en las pendientes de las rectas de regresión estimadas, que son significativamente distintas de cero con un nivel de confianza del 90%. El test de Shapiro-Wilk nos indica que se cumple la hipótesis del modelo de que los residuos siguen una distribución Normal.

Los resultados de los ajustes de regresión lineal simple, como son la recta de regresión del logaritmo de la abundancia en el año t frente al logaritmo de la abundancia poblacional en el año $t - 1$, la recta de regresión subyacente al modelo de Ricker y al de Gompertz, pueden verse en las figuras 1, 2 y 3.

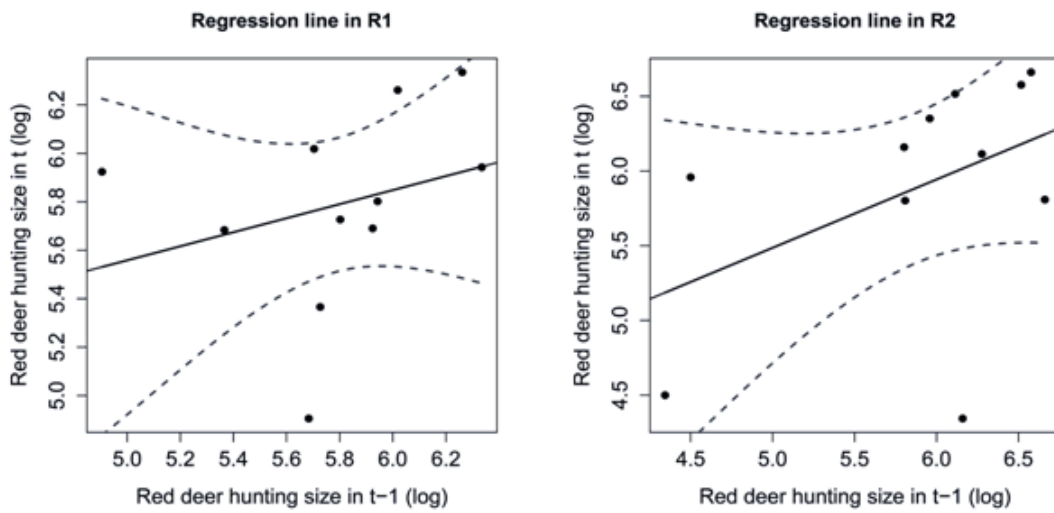


Figura 1. Rectas de regresión lineal simple.

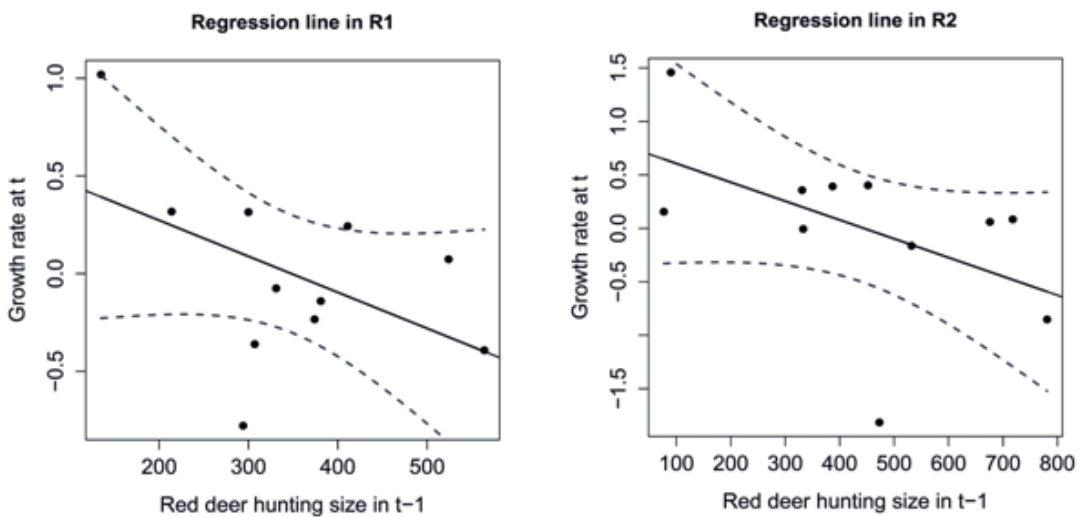


Figura 2. Rectas de regresión subyacente al modelo de Ricker.

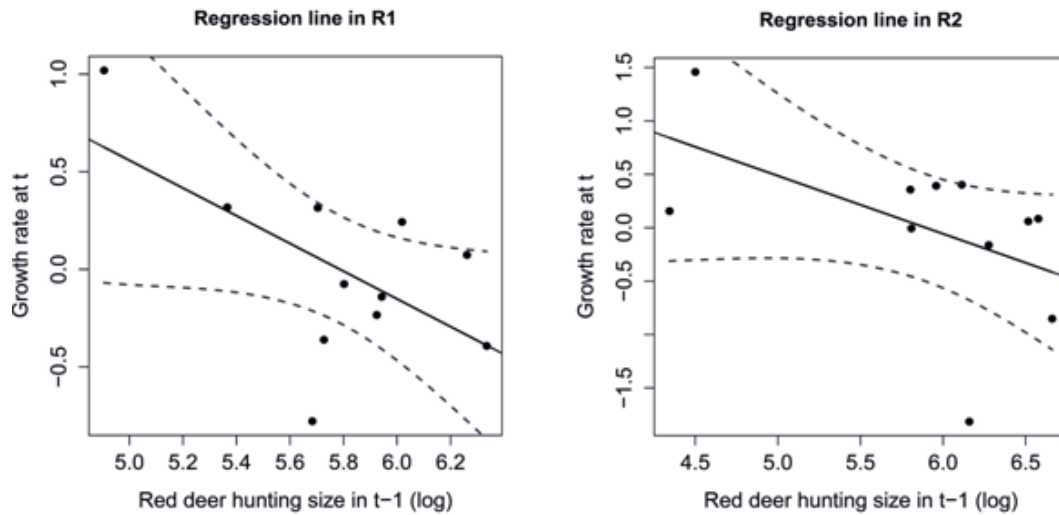


Figura 3. Rectas de regresión subyacente al modelo de Gompertz.

En la recta de regresión del primer modelo considerado se aprecia una pendiente positiva en ambas reservas lo que indicaría la presencia de denso-independencia, sin embargo, para los otros dos modelos los resultados muestran que la DD podría ser una opción viable, ya que sus rectas de regresión presentan pendientes negativas.

4. CONCLUSIONES

La DD es un fenómeno complejo que envuelve varios mecanismos ecológicos. No es de sorprender que las series temporales sean a menudo insuficientes para hacer inferencias en sistemas complejos como en poblaciones de herbívoros (Simard, M.A. et al., 2012).

Los estudios que tratan de hacer predicciones utilizando análisis de series temporales con precisiones inciertas sólo deben utilizarse para probar hipótesis como la de DD (Carrete, M. et al., 2008).

En nuestras poblaciones de estudio, los diferentes modelos de detección de DD proporcionan diferentes resultados sobre ésta. Nuestro estudio sugiere que los resultados que sólo utilizan un tipo de test para detectar la DD pueden proporcionar conclusiones incorrectas. Aunque otros estudios encuentran divergencias en las interpretaciones de los diferentes modelos de detección de DD (Forchhammer, M.C. et al., 1998; Simard, M.A. et al., 2012). No hay ningún estudio en el que se hayan aplicado los diferentes test de DD a series temporales relativas a ciervos mediterráneos, en nuestra área de estudio.

Nuestro estudio sugiere que el análisis de la DD mediante series temporales debe ser usado con cautela, no siendo un fin en sí mismo, sino más bien un punto de partida, y como tal, sugiere la presencia de DD de cérvidos en nuestra área de estudio.

5. REFERENCIAS

Bjornstad, O.N., Grenfell, B. (2001). "Noisy clockwork: time series analysis of population fluctuations in animals", *Science*, 293, 638-643.

Carrete, M., Tella, J.L., Sánchez-Zapata, J.A., Moleón, M., Gil-Sánchez, J.M. (2008). "Current caveats and further directions in the analysis of density-dependent population regulation", *Oikos*, 117, 1115-1119.

Eberhardt, L.L., Breiwick, J.M. (2012). "Models for population growth curves", *International Scholarly Research Network*, 2012, Article ID 815016, 7 pages.

Forchhammer, M.C., Stenseth, N.C., Post, E., Langvatn, R. (1998). "Population dynamics of Norwegian red deer: density-dependence and climatic variation", *Proceedings of the Royal Society B: Biological Sciences*, 265, 341-350.

Imperio, S., Focardi, S., Santini, G., Provenzale (2011). "Population dynamics in a guild of four Mediterranean ungulates: density-dependence, environmental effects and inter-specific interactions", *Oikos*, 121, 1613-1626.

Ricker, W.E. (1954). "Stock and recruitment", *Journal of the Fisheries Research Board of Canada*, 11, 624-651.

Royama, T. (1992). "Analytical population dynamics", *Chapman and Hall*. London UK.

Simard, M.A., Côté, S.D., Gingras, A., Coulson, T. (2012). "Tests of density dependence using indices of relative abundance in a deer population", *Oikos*, 121, 1351-1363.

Turchin, P. (2003). *Complex population dynamics*. Princeton Univ. Press.

Más allá del jardín: Texto dialogado y teatral

Elisabeth Aranda⁽¹⁾, Francisco Gutiérrez⁽²⁾

(1) Departamento de didáctica de la lengua y la literatura, Universidad de Jaén,
eea00009@red.ujaen.es

(2) Departamento de didáctica de la lengua y la literatura, Universidad de Jaén,
fggarcia@ujaen.es

RESUMEN

En el presente artículo, mostramos nuestro proyecto, que se basa en un taller teatral en forma de trabajo por proyectos, la manera en la que se ha llevado la teoría a la práctica y un análisis de resultados basándonos en la opinión de los estudiantes con los que se ha realizado el mismo.

Palabras clave: escritura, teatro, creatividad, representación.

ABSTRACT

In the present article, we show our project, which is based on a theatrical workshop in the form of work by projects, the way in which the theory has been put into practice and an analysis of results which is based on the opinion of the students with whom it has been carried out with.

Keywords: writing, theater, representation, creativity.

1. INTRODUCCIÓN

Este trabajo de investigación consiste en la aplicación en el aula de un taller de textos en forma de trabajo por proyectos, que tiene como material audiovisual complementario la serie de animación “Más allá del Jardín”. En estas sesiones, el estudiante crea un texto dialogado y posteriormente uno teatral. El producto final de este proyecto por tareas consiste en una representación teatral basada en los textos que el alumno ha realizado anteriormente.

Como se puede observar, existe un interés por las nuevas metodologías y un alejamiento de las clases de tipo magistral. Nuestro objetivo era averiguar si estas nuevas metodologías son más efectivas que el método tradicional de enseñanza y si podíamos superar los objetivos que nos propusimos en un primer momento. Este trabajo está basado en la innovación, la motivación y la creatividad, que, a nuestro juicio, son factores que pueden hacer que el aprendizaje del alumno sea mucho más completo.

Este proyecto está dividido en tres partes: una hipótesis didáctica, la puesta en práctica de la misma y por último un análisis estadístico de los resultados que se han obtenido a través de un formulario entregado a los alumnos, que estos mismos tuvieron que completar.

2. METODOLOGÍA

En un primer momento, se desarrolló una hipótesis didáctica en la que se iban a trabajar contenidos teóricos relacionados con el texto narrativo y el guión cinematográfico. La metodología utilizada se basa en el taller de textos y el trabajo por proyectos. Por un lado, el taller de textos es un método de trabajo en el que el estudiante aprende contenidos teóricos a través de la creación de textos, por otro lado, el trabajo por proyectos es una metodología en la que se tratan conceptos teóricos a través de una serie de actividades que dan como resultado un producto final. Todo esto, tomando como material audiovisual complementario la serie de animación “Más allá del Jardín” (“Over the garden wall”).

Esta unidad didáctica está fundamentada en una serie de objetivos, contenidos y criterios de aprendizaje que están basados en la legislación que regula todos estos elementos en la Enseñanza de Educación Secundaria y Bachillerato. Esta se centra en tratar una serie de habilidades lingüísticas y sociales como pueden ser la expresión oral y escrita, la comprensión oral y escrita, el compañerismo o el trabajo en grupo.

Tabla 1. Hipótesis didáctica

SESIÓN 1
Actividad 1.1: Capítulo uno e introducción al texto narrativo y audiovisual
Batería de preguntas sobre el texto narrativo y audiovisual a través del primer capítulo de la serie <i>Más allá del Jardín</i> .
Actividad 1.2: Capítulo dos y elementos constitutivos del texto narrativo y audiovisual
Batería de preguntas sobre los elementos constitutivos del texto narrativo y audiovisual a través del segundo capítulo.

SESIÓN 2
Actividad 2.1: Capítulo tres y búsqueda informativa
Los alumnos verán el capítulo tercero de la serie de animación. Búsqueda informativa por parte de los alumnos, en grupos de tres, de los elementos de la ficha de análisis proporcionada por el profesor: Ficha: <ol style="list-style-type: none">1. Objetivos y tema2. Personajes3. Argumento4. Narrador5. Tiempo y espacio6. Diálogos

7. Descripciones
8. Estructura
9. Tipos de planos
10. Otros elementos audiovisuales

Después los alumnos, en grupos de tres, deberán realizar un trabajo escrito con toda la información que hayan encontrado, y una vez hecho expondrán en el aula oralmente las características más importantes de ese punto de la ficha informativa.

Actividad 2.2: Creación de blog o página web

Asignación a uno o varios alumnos de la tarea de crear un blog o una página web en la que se vayan subiendo los distintos contenidos que los alumnos van a ir recopilando a lo largo del tema.

SESIÓN 3

Actividad 3.1: Capítulo cuatro y exposición de los trabajos

Los alumnos verán el cuarto episodio de la serie.

Los grupos que hayan acabado sus trabajos irán exponiéndolos.

SESIÓN 4

Actividad 4.1: Capítulo cinco y exposición de los trabajos.

Los alumnos verán el quinto episodio de la serie.

Los grupos que hayan acabado sus trabajos irán exponiéndolos.

SESIÓN 5

Actividad 5.1: Capítulos del seis al diez

Los alumnos verán el quinto episodio de la serie.

Los grupos que hayan acabado sus trabajos irán exponiéndolos.

Actividad 5.2: Entrega de los trabajos en papel al profesor

Entrega al profesor del trabajo por escrito en el que se han basado los alumnos para su exposición oral.

SESIÓN 6

Actividad 6.1: Puesta en común de la serie

Los alumnos comentarán aquello que crean importante de la serie relacionada con los contenidos que hayan elaborado.

Actividad 6.2: Contenidos en la web

El alumno o alumnos asignados anteriormente para crear la página web o blog subirán todos los contenidos para que sus compañeros tengan acceso a ellos.

SESIÓN 7

Actividad 7.1: Texto narrativo

En la sesión octava los alumnos realizarán un texto narrativo relacionado con la serie, será como un capítulo más, este trabajo se realizará igualmente en grupos de tres.

SESIÓN 8
Actividad 8.1: Guion cinematográfico
A partir del texto narrativo los alumnos escribirán un guion cinematográfico de ese capítulo.

SESIÓN 9
Actividad 9.1: Cortometraje
Basándome en el guion cinematográfico los alumnos y el profesor grabarán un corto, al ser diez grupos se grabarán diez cortos.

Una vez realizada la hipótesis didáctica, se nos presentó la oportunidad de llevar a cabo la misma en un centro público de la provincia de Jaén, concretamente con un grupo de primero y otro grupo de segundo de E.S.O. El grupo de primero de E.S.O. era un grupo más diverso, es decir, con más diferencia en cuanto al nivel académico de los estudiantes. Sin embargo, el grupo de segundo de E.S.O. era menos heterogéneo y los alumnos estaban dentro del plan bilingüe, en consecuencia, con un nivel académico alto, y a quienes resultó más fácil cumplir los objetivos que les habíamos propuesto.

Aunque pudimos llevar nuestra hipótesis didáctica a la práctica, tuvimos que realizar algunos cambios en la misma. En vez de trabajar el texto narrativo y el guion cinematográfico, se trató en el aula el teatro. Este hecho no se convirtió en un contratiempo, ya que el proyecto era fácilmente adaptable a cualquier contenido teórico. En lo que se refiere a la evaluación, queríamos ser coherentes con la metodología que habíamos elegido, así que evaluamos mediante una rúbrica, en la que no solo el profesor iba a evaluar a los alumnos, sino que también los estudiantes iban a poder calificar a sus compañeros.

Tabla 2. Proyección didáctica

SESIÓN 1	
AULA 1 (14/04/2016)	AULA 2 (11/04/2016)
Actividad 1.0: Creación del blog	
El profesor ha creado previamente un blog donde va subiendo todos los materiales necesarios (la serie de animación Más allá del Jardín, el guion dramático del capítulo primero de la serie o los diferentes enlaces a Kahoot). http://cervantesychomsky.blogspot.com.es	
Actividad 1.1: Introducción al taller dialogado y teatral	
El profesor explica a los alumnos lo que van a realizar en las siguientes sesiones.	
Actividad 1.2: Elementos constitutivos del texto teatral	
Explicación de los contenidos del teatro, a través de la aplicación Prezi, basados en el libro de Anaya con el que trabajan.	Se ha trasladado a la sesión tres por falta de tiempo en esta primera sesión realizándose únicamente la actividad Kahoot. Para la siguiente sesión los alumnos deben leer los contenidos relacionados con el teatro que se van a explicar en la próxima sesión.

Actividad 1.3: Capítulo uno de <i>Más allá del Jardín</i>	
Los alumnos ven en clase el capítulo primero de la serie Más allá del Jardín a través de: http://www.dailymotion.com es (página pública de videos) que está vinculada en el blog del profesor	Se ha trasladado a la sesión dos por falta de tiempo en esta primera sesión.
Actividad 1.4: Ejemplo de texto dialogado	
Mediante el blog enseñamos a los alumnos un ejemplo de texto teatral dialogado para que así les resulte más sencillo realizarlo.	Se ha trasladado a la sesión dos por falta de tiempo en esta primera sesión.

SESIÓN 2	
AULA 1 (15/04/2016)	AULA 2 (18/04/2016)
Actividad 2.1: Capítulo uno de <i>Más allá del Jardín</i>	
Visto en la sesión anterior	No se ha podido ver el capítulo por problemas con internet en su proyector así que deben verlo en casa para la siguiente clase.
Actividad 2.2: Agrupar a los alumnos	
Al ser alumnos de primero de la ESO el profesor ha creído conveniente que ellos mismos los formen. Son siete grupos de cuatro alumnos cada uno excepto uno que está formado por cinco. Se elegirá a un capitán para que controle el grupo.	Se traslada a sesión cinco.
Actividad 2.3: Capítulo uno y texto dialogado (Opcional para casa)	
Se aconseja a los alumnos que vean de nuevo el primer capítulo y que lean a la vez o después el texto dialogado que le hemos mostrado en la sesión anterior en casa.	No se realiza con este grupo.
Actividad 2.4: Resumen y análisis de los personajes (Obligatorio para casa)	
No se realiza en casa, sino en la sesión cuatro.	El profesor les manda unas actividades para casa: deben ver el capítulo, resumir el argumento del mismo y caracterizar a los personajes principales Greg y Wirt.

SESIÓN 3	
AULA 1 (18/04/2016)	AULA 2 (10/04/2016)
Actividad 3.1: Kahoot	
Repasamos los contenidos del teatro a través de una aplicación llamada Kahoot (https://getkahoot.com). Con este grupo hemos los alumnos han usado nuestro móvil.	Se realiza un repaso de los contenidos del año anterior a través de la aplicación kahoot. En este grupo algunos de los alumnos tenían los medios necesarios para poder realizar la actividad, así que agrupamos a los alumnos y participaron de esta manera.

SESIÓN 4	
AULA 1 (19/04/2016)	AULA 2 (21/04/2016)
Actividad 4.1: Materiales para los alumnos	
Se les proporciona una funda transparente a cada grupo y folios para que puedan escribir.	
Actividad 4.2: Resumen y análisis de los personajes (en clase)	
Trabajan la lluvia de ideas para crear el argumento y la caracterización de los personajes en clase. Además trabajan los tipos de personalidades: pasiva, agresiva y asertiva. Si no terminan la actividad en clase deberán seguir en casa.	Este grupo ya lo ha realizado en casa.

SESIÓN 5	
AULA 1 (21/04/2016)	AULA 2 (22/04/2016)
Actividad 5.1: Agrupar a los alumnos	
Se ha realizado en la sesión dos.	En esta ocasión ha sido el profesor el que ha agrupado a los alumnos. Son siete grupos de cuatro alumnos cada uno. Se elige a un capitán y a un secretario para que controlen el grupo y la actividad salga adelante.
Actividad 5.2: Puesta en común	
Los alumnos comentan el argumento y caracterizan a los personajes, así aseguramos que ningún alumno se despiste.	Se realiza en la siguiente sesión.
Actividad 5.3: Argumento y personajes	
Deben terminar la creación del argumento de su historia y crear dos nuevos personajes y caracterizarlos, y sino no lo han terminado en clase hacerlo en casa.	Se realiza en la siguiente sesión.

SESIÓN 6	
AULA 1 (22/04/2016)	AULA 2 (25/04/2016)
Actividad 6.1: Puesta en común	
Se realiza en la sesión anterior.	Los alumnos comentan el argumento y caracterizan a los personajes, así aseguramos que ningún alumno se despiste.
Actividad 6.2: Argumento y personajes	
Los alumnos ponen en común el trabajo realizado anteriormente y deciden cuál va a ser el argumento y los nuevos personajes y sus características. También decidirán qué personaje va interpretar cada miembro del grupo.	En clase deben escribir el argumento de su historia, crear dos nuevos personajes y caracterizarlos, además de decidir qué personaje va a representar cada miembro del grupo.
Actividad 6.3: Monólogo final (para casa)	
Una vez tengan claro ese argumento, los alumnos empezarán su diálogo por el final y pensarán en un monólogo con el que concluirá la obra.	

SESIÓN 7	
AULA 1 (25/04/2016)	AULA 2 (28/04/2016)
Actividad 7.1: Monólogo final (en clase)	
Una vez que los alumnos hayan traído el trabajo de casa (en este caso debían crear un monólogo), intercambian sus propuestas, deciden en grupo sobre cuál es la mejor y la incluyen en su trabajo.	
Actividad 7.2: Diálogo del planteamiento (en casa)	
Los estudiantes deben trabajar en casa el que va a ser el planteamiento, es decir, el diálogo de la primera escena, ahora que ya tienen el final muy claro les va a resultar más sencillo crear esta parte inicial de la historia y ese diálogo primero.	
Actividad 7.3: Recogida de documentos para evaluar	
Hemos recogido los documentos de los alumnos y vamos a hacer una primera evaluación de los mismos, para poder guiar y aconsejar a los alumnos sobre el trabajo realizado.	

SESIÓN 8	
AULA 1 (09/05/2016)	AULA 2 (05/05/2016)
Actividad 8.1: Diálogo del planteamiento (en clase)	
Una vez que los alumnos hayan traído el trabajo de casa (en este caso debían crear el diálogo del planteamiento), intercambian sus propuestas y deciden en grupo cuál es la mejor y la incluyen en su trabajo.	
Actividad 8.2: Diálogo del nudo (en casa)	
Ahora deben trabajar en casa el que va a ser el nudo, es decir, el diálogo de la segunda escena.	

SESIÓN 9	
AULA 1 (13/05/2016)	AULA 2 (09/05/2016)
Actividad 9.1: Diálogo del nudo (en clase)	
Una vez que los alumnos hayan traído el trabajo de casa (en este caso debían crear el diálogo del nudo), intercambian sus propuestas y deciden en grupo sobre cuál es la mejor y la incluyen en su trabajo.	
Actividad 9.2: Apartes y acotaciones (en clase o en casa)	
Ahora deben trabajar en casa el que va a ser el nudo, es decir el diálogo de la segunda escena.	
Actividad 9.3: Diálogo final pasado a limpio (en clase o en casa)	
Deben ir terminando el diálogo y pasarlo a limpio para compartirlo en google docs, tanto con el profesor como con el resto de los miembros del grupo. Además deberán revisar la gramática, la ortografía, el vocabulario, los signos de puntuación y la estructura del texto en forma de diálogo teatral.	

SESIÓN 10	
AULA 1 (16/05/2016)	AULA 2 (13/05/2016)
Actividad 10.1: Entrega de los diálogos (en clase o en casa)	
Los alumnos deben entregar o compartir en google docs sus diálogos completos.	
Actividad 10.2: Impresión de los textos (en casa)	
Los alumnos deben ya a partir de esta sesión traer sus diálogos impresos para poder practicar la representación teatral que van a realizar.	
Actividad 10.3: Ideas básicas para la representación teatral (en clase)	
En el aula el profesor les pone a los alumnos un vídeo representativo de cómo se actúa en el teatro (un pequeño fragmento de la película El viaje a ninguna parte) para considerar la importancia de la vocalización, de no ir demasiado rápido, de aprenderse el texto bien y de gesticular apropiadamente y para dar credibilidad a lo que se está representando.	

SESIÓN 11	
AULA 1 (17/05/2016)	AULA 2 (16/05/2016)
Actividad 11.1: Ideas básicas para la representación teatral (en clase)	
El profesor recuerda aquello que es importante para realizar una buena representación teatral.	
Actividad 11.2: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	

SESIÓN 12	
AULA 1 (19/05/2016)	AULA 2 (17/05/2016)
Actividad 12.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	

SESIÓN 13	
AULA 1 (19/05/2016)	AULA 2 (19/05/2016)
Actividad 13.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	

SESIÓN 14	
AULA 1 (23/05/2016)	AULA 2 (23/05/2016)
Actividad 14.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	

SESIÓN 15	
AULA 1 (24/05/2016)	AULA 2 (24/05/2016)
Actividad 15.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	

SESIÓN 16	
AULA 1 (26/05/2016)	AULA 2 (26/05/2016)
Actividad 16.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	

SESIÓN 17	
AULA 1 (27/05/2016)	AULA 2 (27/05/2016)
Actividad 17.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	

SESIÓN 18	
AULA 1 (30/05/2016)	AULA 2 (30/05/2016)
Actividad 18.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	

SESIÓN 19	
AULA 1 (19/05/2016)	AULA 2 (19/05/2016)
Actividad 19.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	
Actividad 19.2: Representación teatral (en el salón de actos)	
Dos grupos del Aula 2 representan su obra teatral para ser evaluados.	
Los alumnos también van a poder evaluar a sus compañeros mediante una rúbrica.	

El profesor los graba para poder evaluarlos también.

SESIÓN 20	
AULA 1 (01/06/2016)	AULA 2 (01/06/2016)
Actividad 20.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	
Actividad 20.2: Representación teatral (en el salón de actos)	
Dos grupos del Aula 1 representan su obra teatral para ser evaluados. Los alumnos también van a poder evaluar a sus compañeros mediante una rúbrica. El profesor los graba para poder evaluarlos también.	

SESIÓN 21	
AULA 1 (02/06/2016)	AULA 2 (02/06/2016)
Actividad 21.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	
Actividad 21.2: Representación teatral (en el salón de actos)	
Dos grupos del Aula 2 representan su obra teatral para ser evaluados. Los alumnos también van a poder evaluar a sus compañeros mediante una rúbrica. El profesor los graba para poder evaluarlos también.	

SESIÓN 22	
AULA 1 (03/06/2016)	AULA 2 (03/06/2016)
Actividad 22.1: Ensayo de la representación teatral (en el salón de actos)	
Los alumnos ensayan la representación de su obra teatral con ayuda del profesor, que los va guiando.	
Actividad 22.2: Representación teatral (en el salón de actos)	
Un grupo del Aula 2 representa su obra teatral para ser evaluado. Un grupo del Aula 1 representa su obra teatral para ser evaluado. Los alumnos también van a poder evaluar a sus compañeros mediante una rúbrica. El profesor los graba para poder evaluarlos también.	

SESIÓN 23	
AULA 1 (06/06/2016)	AULA 2 (06/06/2016)
Actividad 23.1: Representación teatral (en el salón de actos)	
Representan el resto de grupos de ambas aulas. Los alumnos también van a poder evaluar a sus compañeros mediante una rúbrica. El profesor los graba para poder evaluarlos también. Por ausencia de algunos alumnos será necesario grabar las representaciones restantes en otras sesiones posteriores.	
Actividad 23.2: Realización del formulario	
Los alumnos tendrán que completar un formulario propuesto por el profesor a través de google forms que se les proporcionará a través de google classroom.	

Una vez realizada la proyección didáctica, se creó un cuestionario que contestaron los alumnos con sus opiniones en cuanto a lo que había realizado en clase, que se explican de manera pormenorizada en el apartado 3. *Resultados*. Este análisis era importante a la hora de evaluar el proyecto y llegar a una serie de conclusiones para poderlo mejorar en el futuro.

3. RESULTADOS

Como hemos dicho anteriormente, este estudio estadístico se ha llevado a cabo a partir de un cuestionario realizado por los dos grupos con los que se puso en práctica el proyecto didáctico planteado anteriormente. Cada uno de esos grupos tenía treinta alumnos, con un total de sesenta estudiantes. De estos sesenta alumnos participaron cuarenta y seis, es decir, un 76% del total, concretamente un 43% hombres y un 57% mujeres.

A continuación, presentamos los resultados conjuntos de las Aulas 1 y 2, excepto en aquellos casos en los que se indique lo contrario:

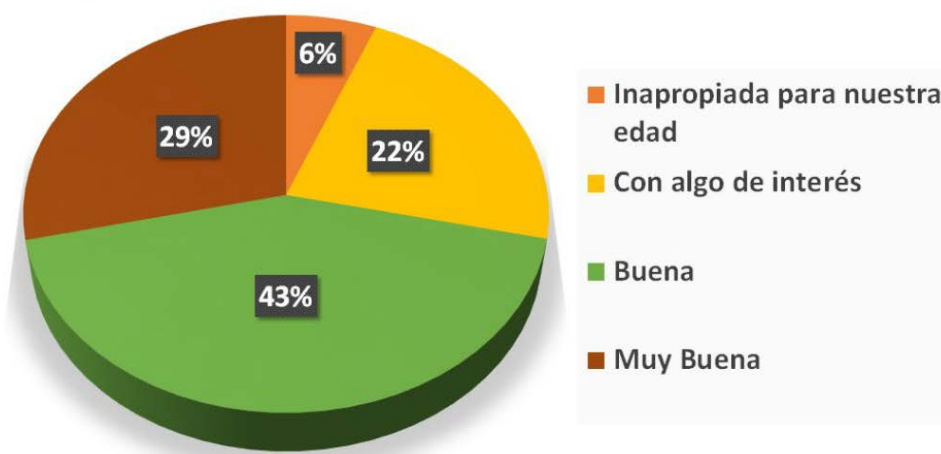
Pregunta 1: ¿Crees que ha sido interesante hacer una continuación del primer capítulo de la serie *Más allá del Jardín*?

En esta pregunta, los estudiantes debían contestar sí o no. En este caso, prácticamente la totalidad del alumnado ha contestado sí. Así pues, a los estudiantes les ha gustado realizar el taller literario.

Pregunta 2: ¿Cómo calificarías la serie *Más allá del Jardín*?

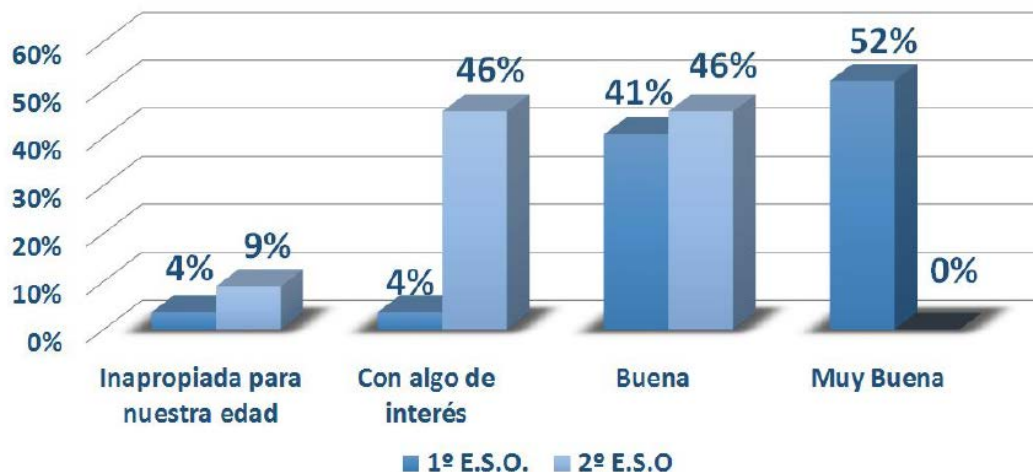
Aquí, los alumnos tenían varias opciones a elegir: *aburrida*, *inapropiada para nuestra edad*, *con algo de interés*, *buena* o *muy buena*. Como se puede observar en la gráfica 1, el 68% de los alumnos creen que ha sido buena o muy buena, en menor medida *inapropiada para su edad*, pero ninguno la ha calificado de *aburrida*, por esto no aparece esta opción representada en la gráfica.

Gráfica 1
¿CÓMO CALIFICARÍAS LA SERIE MÁS ALLÁ DEL JARDÍN?



En un análisis diferenciado por clases hemos encontrado diferencias entre los grupos: los alumnos de primero de E.S.O. (Aula 1) han contestado más favorablemente al material complementario que los de segundo de E.S.O. (Aula 2).

Gráfica 2
¿CÓMO CALIFICARÍAS LA SERIE MÁS ALLÁ DEL JARDÍN?



Esta diferencia puede ser causada por la percepción por parte del Aula 2 de la serie como un material demasiado infantil para la misma, o por cómo se realizó el visionado del material complementario, concretamente porque los alumnos del Aula 1 sí pudieron ver algunos capítulos de la serie, pero sin embargo los estudiantes del Aula 2 no lo hicieron, y eso pudo afectar a su motivación.

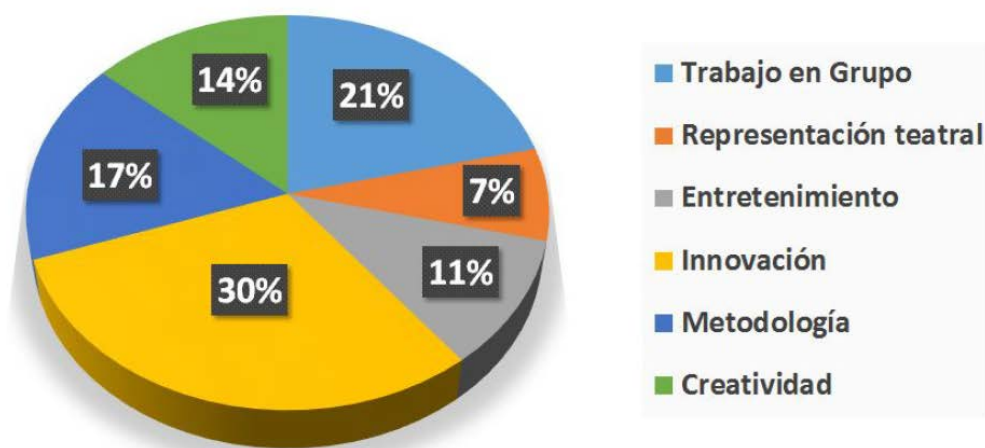
Pregunta 3: ¿Te ha parecido interesante la creación de un diálogo teatral?

En este caso, prácticamente a todos los alumnos (98%) les ha parecido interesante la creación del diálogo teatral.

Pregunta 4: ¿Por qué te ha parecido interesante la creación de un diálogo teatral?

Gráfica 3

**¿POR QUÉ TE HA PARECIDO INTERESANTE LA
CREACIÓN DE UN DIÁLOGO TEATRAL ?**



Los alumnos han opinado lo siguiente en lo que se refiere a esta pregunta: *trabajo en grupo, representación teatral, entretenimiento, innovación, metodología y creatividad.*

Como se puede ver en la gráfica 3, un 30% de los estudiantes creen que la creación del diálogo teatral ha sido motivadora. Un 21% de los alumnos valoran positivamente trabajar en grupo.

Pregunta 5: ¿Qué has aprendido con la realización de esta tarea? Enumera al menos tres enseñanzas o aprendizajes concretos que hayas logrado.



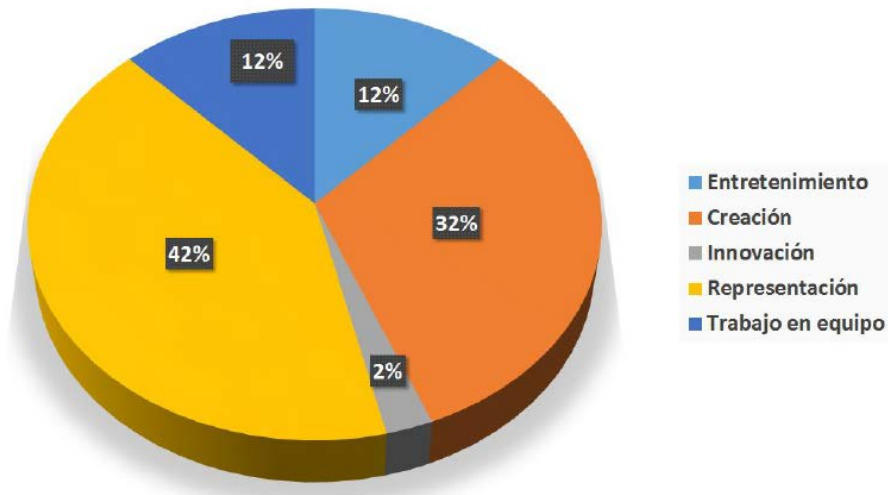
El 36% de los alumnos creen que han realizado una buena representación teatral, y por lo tanto han mejorado su expresión oral. Igualmente, el 26 % de los estudiantes piensan que el ambiente en clase ha sido bueno, y han mejorado su trabajo en grupo y sus habilidades sociales. También hemos logrado fomentar la creatividad.

Pregunta 6: ¿Crees que a otros compañeros de tu misma edad les gustaría realizar este trabajo?

La totalidad de los estudiantes han contestado favorablemente a esta pregunta.

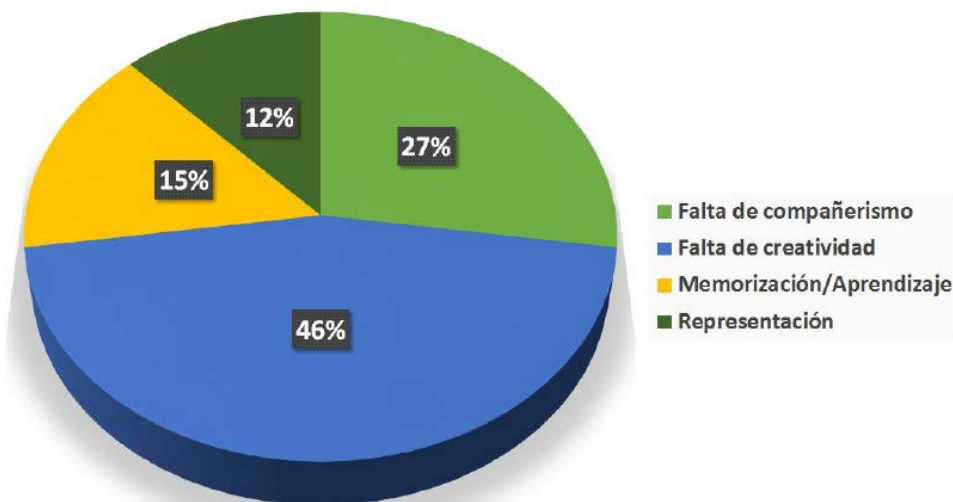
Pregunta 7: ¿Qué es lo que más te ha gustado y lo que menos de crear y representar un diálogo teatral escrito por ti?

Gráfica 5
¿QUÉ ES LO QUE MÁS TE HA GUSTADO?



En este caso, a un 42% de los estudiantes les ha gustado realizar una representación teatral, con la que han podido fomentar su expresión oral. Igualmente, al 32% de los mismos les ha gustado crear un diálogo teatral, con el que han mejorado su expresión escrita. En menor medida les ha gustado trabajar en equipo, así pues, hemos desarrollado sus habilidades sociales.

Gráfica 6
¿QUÉ ES LO QUE MENOS TE HA GUSTADO?



En esta ocasión, los resultados muestran que la falta de creatividad (46%) y la falta de compañerismo (27%) ha sido lo que menos les ha gustado a los estudiantes.

Pregunta 8: ¿Te ha gustado trabajar junto a otros compañeros en un mismo grupo?

En esta pregunta, el 85% de los estudiantes sí les ha gustado trabajar junto a otros compañeros y un 15% han contestado negativamente.

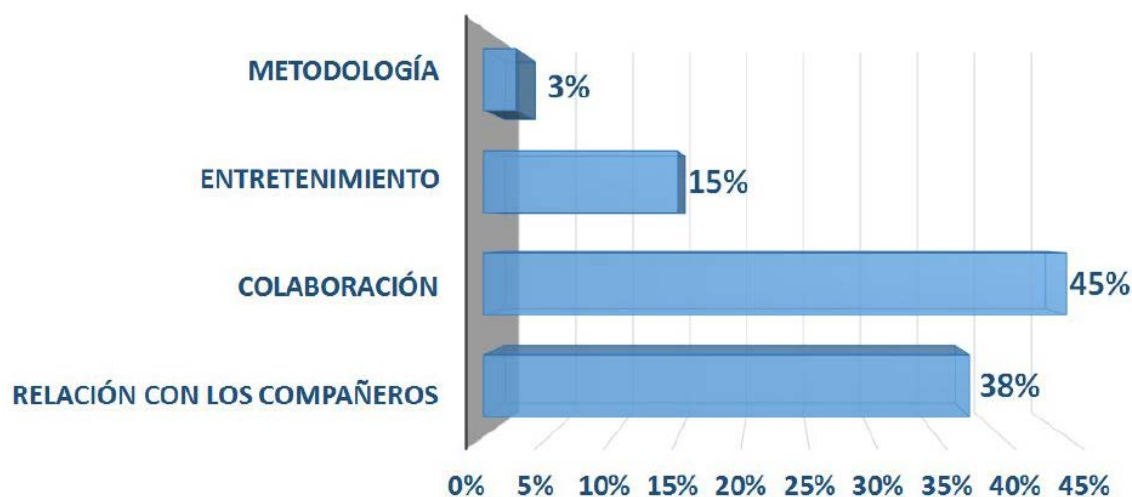
Pregunta 9: ¿Por qué no te ha gustado trabajar junto a otros compañeros en un mismo grupo?

A los alumnos que no les ha gustado trabajar en grupo ha sido a causa de la falta de trabajo de sus compañeros.

Pregunta 10: ¿Por qué sí te ha gustado trabajar junto a otros compañeros?

Gráfica 7

**¿POR QUÉ SÍ TE HA GUSTADO TRABAJAR JUNTO A OTROS
COMPAÑEROS EN UN MISMO GRUPO?**



Las causas por las que a los alumnos les ha gustado trabajar en grupo son la colaboración de sus compañeros (45%) y la oportunidad de conocerse mejor (38%).

Pregunta 11: Señala las actividades que te hayan resultado más difíciles

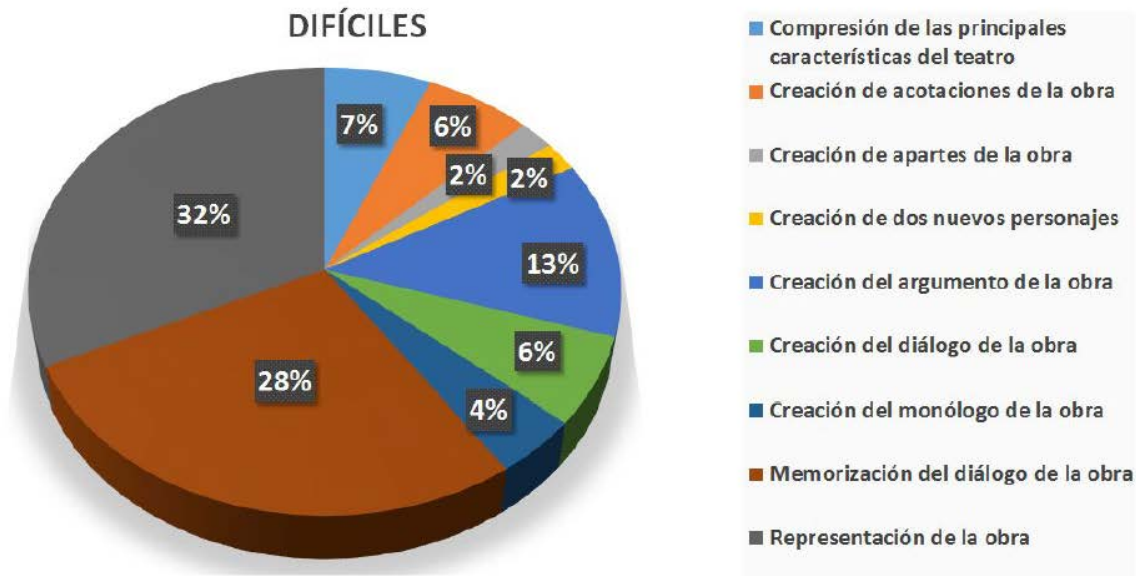
En este caso, se propusieron varias respuestas con las actividades que se realizaron a lo largo del proyecto:

- Comprensión de las principales características del teatro
- Creación de dos nuevos personajes
- Caracterización de los cuatro personajes
- Creación del argumento de la obra
- Creación del diálogo de la obra
- Creación del monólogo de la obra
- Creación de apartes de la obra
- Creación de acotaciones de la obra
- Creación del documento de google docs
- Memorización de la obra
- Representación de la obra

Gráfica 8

SEÑALA LAS ACTIVIDADES QUE TE HAYAN RESULTADO MÁS

DIFÍCILES

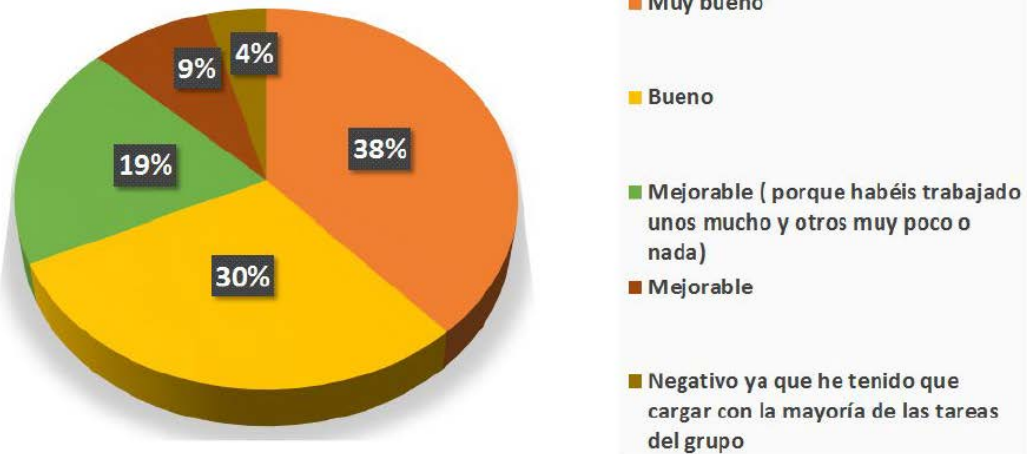


Las tareas más complejas para los estudiantes han sido la representación teatral (32%) y la memorización del diálogo (28%). Aunque estas actividades les han resultado difíciles, también les ha gustado realizarlas, como se ha mostrado en la pregunta 7.

Pregunta 12: ¿Cómo calificarías el trabajo en equipo con tus compañeros?

Gráfica 9

**¿CÓMO CALIFICARÍAS EL TRABAJO EN GRUPO CON TUS
COMPAÑEROS?**



En esta pregunta, se proporcionaron una serie de opciones para elegir:

- Muy bueno
- Bueno
- Mejorable (porque habéis trabajado algunos mucho y otros muy poco o nada)
- Mejorable
- Negativo

En esta evaluación de los grupos de trabajo, los resultados muestran que el trabajo en grupo ha sido o bueno (30%) o muy bueno (38%), por lo tanto estos son positivos.

4. CONCLUSIONES

El taller teatral y el trabajo por proyectos es un método con el que los estudiantes pueden mejorar su aprendizaje de una manera más práctica. Para afirmar esto nos basamos en una fundamentación teórica y en los resultados que se presentan a través del cuestionario realizado. Con esta metodología fomentamos su creatividad, mejoramos su expresión oral y escrita e incluso su gusto por la lectura.

Aunque los resultados han sido positivos, creemos que habría que mejorar elementos como el material audiovisual complementario, cambiando en el Aula 2 el mismo, o como el tiempo dedicado al taller teatral, que en este caso pensamos que ha sido excesivo.

En último lugar, nos gustaría agradecer a los profesores Francisco Gutiérrez García y Antonio Gómez, sin los cuales no se podría haber realizado este trabajo de investigación.

5. REFERENCIAS

Boquete, Gabino; (2014). "Interacción y motivación en el aula: enseñar desde el teatro. En E. Tobar y V. Hidalgo (coords.), *Arriba el telón: enseñar teatro y enseñar desde el teatro. Propuestas didácticas para trabajar el teatro en clase de español* (pp. 5-20). Madrid: Ministerio de Educación, Cultura y Deporte.

Cassany, Daniel (2006). *Taller de textos. Leer, escribir y comentar en el aula*. Barcelona, España: Paidós.

Neira, M^a; (2009). "El texto teatral en las aulas de educación infantil". En M. Nemirovsky (coord.), *Experiencias escolares con la lectura y la escritura* (pp. 33-57). Barcelona: Graó.

Sánchez-Enciso, J. y Rincón, E. (1985). *Los talleres literarios (una alternativa didáctica al historicismo)*. Barcelona, España: Montesinos.

Statistical Results of an Empirical Study about Bilingual Students in Compulsory Secondary Education in Madrid (Spain) (2016)

Zahra Chaieberras

Departamento de Filología Inglesa, Universidad de Jaén

Dirección postal: Calle General del Ricardo, 200, 4D, Madrid, España

Email: fatimachaieb2012@hotmail.com

RESUMEN

Este artículo expone la evidencia empírica sobre la experiencia de estudiantes de programas de sección bilingüe, en educación secundaria obligatoria en la Comunidad de Madrid después de unos cuantos años de su implementación en algunas instituciones pedagógica. Su objetivo no es sólo analizar y mejorar los efectos del programa desde las perspectivas de los estudiantes y sus experiencias, sino también identificar el resultado estadístico que es una forma ideal de presentar un análisis estadístico apropiado. El objetivo es probar las diferencias estadísticamente significativas de los datos y también probar las fortalezas, la fiabilidad y la coherencia de todas las variables en los cuestionarios de los estudiantes. Los resultados de esta experiencia fueron extraídos de análisis de cuestionarios. Esta herramienta dio excelentes oportunidades para 754 participantes bilingües de 17 colegios bilingüe (escuelas tanto rurales como urbanas) a exponer sus percepciones y sus puntos de vista cualitativamente sobre los efectos del Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE).

Palabras clave: programa bilingüe, Región de Madrid, AICLE, análisis estadístico, resultado estadístico

ABSTRACT

This article generates an empirical evidence about students' experiences in bilingual sections programs, *seccion bilingue*, in compulsory secondary education in the Community. It aims is not only to better analyse the effects of the program from the students' perspectives and their experiences but also to identify the main statistical result that it is an ideal way to present an appropriate statistical analysis. The aim is to test the significantly statistical differences of the data and also to test the strengths, reliability and consistency of all the variables in the students' questionnaires. The results of this experience were drawn from the analysis of questionnaires. This instrument gave some excellent opportunities for 754 bilingual participants of 17 bilingual schools (schools from both rural and urban areas) to expose their perceptions and their viewpoints qualitatively about the effects of Content and Language Integrated Learning (CLIL).

Keywords: bilingual program, Madrid region, CLIL, statistical analysis, statistical result

1. INTRODUCTION

The introduction of the term CLIL in Spain began when “in 1996 the Ministry of Education and Science and the British Council signed an agreement to introduce an integrated curriculum in Spanish state schools” (Dobson, Pérez Murillo, & Johnstone 2010, p. 5). Bruton (2011) also acknowledged that Spanish society had started to change linguistically and “a number of autonomous administrations had adopted CLIL initiatives” (p. 236).

National-level language learning policies have the concept of education in Spain. The implementation of CLIL normalizes the idea of acquiring and communicating in a global language. Such policies support preparing students for internationalization, and workforce mobility, through multilingual language competence. Students now communicate in two languages (Andalucía and Madrid) or three languages (Basque Country and Galicia), and benefit from related academic growth such as a gain in cognitive abilities, increased communication competency, and appreciation of cultural diversities.

In brief, the introduction of CLIL, either in bilingual or multilingual environment, created more educational policies and methodological strategies that increased “the time devoted to language alongside improving the quality of teaching practice” (Ruiz Gómez, 2015, p. 14). This great attention to learning language has proven to provide greater exposure to students for acquiring a second or a foreign language and practicing the target language, with an increase of academic achievement at all levels.

Why following CLIL in Spain

Spain, geographically in the midst of contemporary global geopolitical and socioeconomic changes, proactively looked at new strategies for teaching languages in order to support its students and citizens toward becoming bilingual, or even multilingual. Therefore, bilingual programs reflected a national value of language acquisition for the purpose of improved global communication globally within a diverse society. Martínez Adrián (2011) has claimed that “in the Spanish context, we have a great diversity in the implementation of CLIL programmes, but even so, all of them share the aim of achieving communicative competence in second and foreign languages across the curriculum” (p. 99).

Although some Communities on the Iberian Peninsula have enjoyed being bilingual or multilingual over the years, the new bilingualism and bilingual programs have been seen as an innovative and “effective way to provide students with linguistic skills in an additional language (mainly English) that are absolutely essential in today’s job market” (Pavón Vázquez & Gaustad, 2013, p. 84). The development of a program, such as CLIL, has been needed, since there have been many ways in which learners and communities have benefitted, more so than the older programs presented.

Among other reasons for implementing bilingual programs in Spain, learning another language (mainly English) can help students to increase their linguistic competences, offering opportunities for both communication and mobility, and for success in multilingual and multicultural societies. The newer bilingual programs also allow students to overcome communication barriers by learning about language and culture, both globally and, very often on the Iberian Peninsula, locally.

Additionally, because “Our society has become a knowledge society, where information globally networked and more freely accessible than ever before needs to be processed and transformed into knowledge by those working within a technology-enriched environment” (Chan, Chin, Nagami, & Suthiwan, 2011, p. 120), CLIL provides students with a new educational system that advocates for working collaboratively and thinking critically, engagement with interactive technology, including online resource and information and communication technologies (ICT). Furthermore, CLIL promotes integration encouraging student mobility through international exchange programs (international and domestic) in order to develop communicative and cultural competences.

The bilingual program in the region of Madrid

In recent decades, Madrid also has promoted a shift from monolingualism to bilingualism, as an instrument of innovation in the language policy of the Community. “This project consists of a CLIL model” (Fernández Fontecha, 2009, p. 14). CLIL became the brand recognized for a new educational program (Spanish/English) that focused on a need for learning languages and It has been integrated in the educational institution Fernández and Halbach (2011) stated that:

In March 2004 the regional government of Madrid issued a decree through which it set up a bilingual project to be implemented gradually in primary education, starting with the first year of primary and moving upwards to eventually cover the six years of primary education (p. 241).

The region has adapted the bilingual project: “although different bilingual projects are located in different Spanish Communities, here we are specifically referring to the Bilingual Project carried out in the Autonomous Community of Madrid” (Fernández Fontecha, 2009, p. 14).

Its implementation, according to the Order from the regional ministry of Madrid, is necessary due to the government’s perceived need to acquire languages and to promote knowledge about other cultures. It also addresses the important issues in “being able to develop their daily and professional activities using English as a second language opens new perspectives and new relationship possibilities to students of bilingual schools in the Autonomous Region of Madrid” (Brindusa, Cabrales, & Carro, 2012, p. 4).

Consequently, bilingual programs have become one of the government’s top priorities, through targeted use of time and money, resulting in rapid program growth in schools. Further, through bilingual programming, there is an attempt to empower the learning of curricular subjects, such as Natural Sciences, Social Science, Arts, Crafts and so forth, in English (*Orden 3245/2009, de 3 de julio, de la Consejería de Educación de la Comunidad de Madrid*). Indeed, for the past several years, Madrid has been one of the most important

regions of Spain that has given great importance to implementing CLIL in urban and rural schools alike, in order to further develop linguistic and academic competences.

The Ministry of Education first announced a course of bilingual sections beginning in 2004-2005 with implementation in 26 public schools. In 2011-2012, 32 new bilingual secondary schools opened their doors to the first students finishing primary education. To date, over 100,000 students have been receiving bilingual education in the Community of Madrid both in primary and secondary education (*Consejería de Educación, Juventud y Deporte de la Comunidad de Madrid Dirección General de Innovación, Becas y Ayudas a la Educación, 2015-2016, p. 31*).

Thus, it is clear that CLIL already permeates half of secondary schools in Madrid and English has become the dominating language in bilingual education. The majority of public, charter, and private schools in Madrid have designed mainstream programs to create new opportunities to learn languages, especially English, in a natural way and effectively (*Minsiterio de Educación, Juventud y Deporte, Orden 1317/2015, del 7 de mayo*).

2. METHODOLOY

Instrument

The instrument of the sample is a battery of tool designed by Pérez Cañado as a part of research project entitled: The effects of Content and Language Integrated Learning in monolingual communities: A large-scale evaluation; funded and supported by the Spanish Ministry of Economy and Competitiveness. The research methodology and data collection in this study included a questionnaire. A questionnaire is fundamental for collecting data qualitatively from students' perspectives and viewpoints about CLIL. This questionnaire has been tested and revised by researchers and experts in the field of bilingualism (Pérez Cañado, 2016).

The survey questionnaire contains 50 items built on seven important blocks to give second secondary students in the region of Madrid a good opportunity to express their beliefs and their perceptions regarding their experience in CLIL programs. This instrument examines and evaluates the following factors: students' use, competence and development of English in class; methodology; materials and resources; evaluation; teachers' use, competence and development of English in class; mobility; and finally, improvement in and motivation towards learning English.

The statistics uses a range of techniques and procedures to discover the reliability, validity and consistency of the data obtained. It helps us to take the data and transform them into information to describe and control causes and effects in a meaningful way. It is also used to draw some conclusions about the interesting differences in all the variables.

To make the contrasts of hypotheses a non-parametric test has been used, because the variables items are ordinal. The participants are asked about their opinions to measure the level of their satisfaction. The U Mann-Whitney test has been applied to contrast the items in the study of a qualitative variable with two categories such as gender, setting, nationality, etc. When the variable has more than two categories, the Kruskal-Wallis test was applied.

Dealing with comparison, R of Ronsethal was applying to see the effect size of the results and to indicate how different the groups are. For the effect to be large this value has to be greater than 0.7, regardless of the sign you have, either is positive or negative.

The sample

The research was carried out during the second term of the 2015-2016 academic year. I spent three months visiting 18 bilingual secondary schools in the Autonomous Community of Madrid. The sample consists of 754 bilingual students attending grade two of Compulsory Secondary Education (CSE). Beginning with the two graphs below show that the three types of bilingual schools (public, charter, and private) and the two kinds of settings (urban and rural), are very well represented with an equal amount of percentage in each.

On examining the age, most of the students sampled range between 13 and 14 years of age. 70% of students are 13 years old, 29% are 14 years old, and 1% is the rest of students aged between 11 and 12 years old.

In relation to gender, within the cohort there are 53% of females and 46% of males. 1% did not reply to the corresponding item. In terms of nationality Spanish (95%) represents the largest number of participants, then Venezuelan (1%), and finally (4%) identify other nationalities.

On the whole, 86% of the students who took part in this research have had more than seven years of experience in a bilingual section. 10% of the participants have had less than four years of experience and 4% have had from four to six years of experience in bilingual education; 0.1% did not respond to this question.

With respect to the number of subjects taught in English, it varies from one school to another. It is obvious that all the bilingual schools respect the regulations of the *Consejería de Educacion* about not teaching Math and Spanish in English.

Of the seventeen schools that I visited, each school has its own policy on how to select which subjects are taught to the students.

As to the exposure of English in the bilingual programs, I do not have enough reliable data to be able to answer this question. Fortunately, the question was answered clearly in the interview protocol. Students answered the same question in more detail. Also, the classroom observation gave me a clear idea about the percentages of English use in the classroom.

3. RESULTS

The statistical outcomes will be presented here from specific points view according to the variables controlled. By doing so we will intend to make an appropriate statistical analysis that measures the significant differences of the data.

Starting with the gender, there are statistically significant differences between female and male for nine items. In these cases, the female students always respond more optimistically than male students. Girls consider that they have more capabilities and skills concerning developing basic competences, and they are more interested in English more than boys (items 1 and 10, $p = .015$ and $p = .005$, respectively). It might be for that reason they also welcome more use of English (item 11). Moreover, according to the girls the bilingual teachers work in a collaborative way to deliver and design the bilingual teaching materials (item 22). They believe that all the bilingual content knowledge learnt is evaluated (item 30). On the topic of teachers' use, competence and development of English in class, females regard that both language teachers and teaching assistants work successfully with the students, and that their teaching assistants motivate them, in contrast to boys, who respond less positively (items 34, 36, and 39). As for item 46, girls are encouraged more to participate in exchange programs by their families than boys.

Table 1. Statistically significant differences in terms of gender

	<i>Means</i>		<i>Standard deviation</i>		<i>R effect size</i>	<i>P_value</i>
	<i>Man</i>	<i>Woman</i>	<i>Man</i>	<i>Woman</i>		
Item1	3,11	3,24	,668	,592	-0,089	,015
Item10	3,29	3,45	,726	,622	-0,104	,005
Item11	2,33	2,55	,959	,975	-0,113	,002
Item22	3,12	3,25	,766	,706	-0,083	,022
Item30	3,25	3,36	,703	,646	-0,075	,041
Item34	3,08	3,23	,861	,715	-0,076	,036
Item36	2,97	3,12	,881	,821	-0,081	,028
Item39	2,71	2,87	,937	,869	-0,080	,030
Item46	2,88	3,20	1,027	,938	-0,161	<0,001

In relation to the students age, we have located significant differences combined between students of equal or less than 13 years old and students with more than 13 years. In the first item, participants of equal or with 13 years declare that their basic competences are developed, whereas the other participants with more than 13 years react oppositely, interestingly, they like more use of English in the class (item 11). Again, students of equal or less than 13 acquire an extensive vocabulary in the classroom (item 17). In view of the bilingual materials, respondents of older age apply and learn with the new technologies materials such as online references and blogs, wikis, and webquest in the CLIL class (items 26 and 27). The younger learners assert that their teachers have a high linguistic level in English, especially in listening and speaking skills (item 41).

Table 2. Statistically significant differences in terms of age

	<i>Means</i>		<i>Stander deviation</i>		<i>R effect size</i>	<i>P_value</i>
	<i><=13 years</i>	<i>>13 years</i>	<i><=13 years</i>	<i>>13 years</i>		
Item1	3,22	3,10	,627	,633	-0,088	,016
Item11	2,40	2,58	,959	,999	-0,075	,039
Item17	3,47	3,39	,666	,626	-0,074	,042

Item26	2,76	2,96	1,044	,896	-0,084	,022
Item27	2,50	2,69	1,033	,959	-0,087	,016
Item41	3,37	3,23	,730	,817	-0,077	,035

As regards nationality, on the one hand, students who are from different nationalities state that they employ both projects and multimedia software for learning in the classroom (items 16 and 25). They confirm that their non-linguistic area teachers work successful in developing bilingual classes (item 35) and they have contributed in exchange program (item 44). On the other hand, Spanish students respond the four items less positively than the other non-Spanish participants.

Table 3. Statistically significant differences in terms of nationality

	<i>Means</i>		<i>Standard deviation</i>		<i>R effect size</i>	<i>P_value</i>
	<i>Other</i>	<i>Spanish</i>	<i>Other</i>	<i>Spanish</i>		
Item16	3,46	3,21	,650	,727	-0,075	,039
Item25	3,27	2,92	,871	,965	-0,081	,026
Item35	3,39	3,10	,645	,802	-0,075	,041
Item44	2,05	1,74	1,026	1,017	-0,077	,034

Within the mainstream settings (urban-rural), it was detected differences in nine items. The statistics indicates that there were some statistically significant differences between urban and rural areas. In item 11 the students in rural areas declared they are interested in more use of English than urban areas. They also develop and use projects and teamwork (items 16 and 18). While in the urban settings, students learn a lot of vocabulary and both ongoing and final evaluation is practiced in the classroom (items 17 and 33), in the contrary, the rural context gives less attention to these two items. Concerning the teachers' use, competence and development of English in class, is differ in the two learning contexts. Urban students admit that their subject teachers are successful in developing their bilingual classes and they motivated them (items 35 and 38), while rural students confirm teaching assistants are less successful in developing their bilingual classes (item 36). Again, urban students acknowledge that their teachers have the adequate receptive and productive skills, and adequate socio-cultural awareness background in English (items 41, 42, and 43). In the urban context, participating in exchange programs is declined and most students who have participated in exchange programs belong to the rural areas (44). For item 48, it is gauged that the urban students score an overall improvement in language learning than the rural students.

Table 4. Statistically significant differences in terms of Schools settings

	<i>Means</i>		<i>Standard deviation</i>		<i>R effect size</i>	<i>P_value</i>
	<i>Rural</i>	<i>Urban</i>	<i>Rural</i>	<i>Urban</i>		
Item11	2,72	2,25	,963	,935	-0,241	<0,001
Item16	3,30	3,17	,719	,725	-0,093	,011
Item17	3,34	3,53	,718	,590	-0,128	<0,001
Item18	3,10	2,98	,774	,786	-0,075	,038
Item33	3,32	3,45	,745	,628	-0,081	,026

Item35	2,93	3,26	,853	,722	-0,195	<0,001
Item36	3,13	2,99	,841	,860	-0,089	,015
Item38	2,80	2,92	,858	,850	-0,077	,034
Item41	3,16	3,45	,856	,649	-0,162	<0,001
Item42	3,28	3,51	,825	,598	-0,129	<0,001
Item43	3,17	3,33	,759	,689	-0,106	,004
Item44	1,92	1,63	1,071	,961	-0,146	<0,001
Item48	3,36	3,50	,733	,637	-0,093	,010

In response to the types of schools (public, charter, and private), it appears that most of the differences were scored here in this variable. The majority of items belong to the topics: methodology, material and resource, evaluation, and teachers use, competence and development of English in class. In general terms, statistically significant differences showed that the private schools are gaining ground among the other two types of schools. In the private environment the participants have declared that learning in a bilingual environment improve their basic competences and their Spanish. As well as their understanding of the connection between Spanish and English (items 1, 4, and 7). The methodology related to task, vocabulary and teamwork, are experienced more in private context than any other type of schools (items 15, 17, and 18). Materials and resources for delivering academic acknowledgment and teaching trough CLIL are available in the private environment except for multimedia software, online reference materials, and blogs, wikis and webquest are deployed more in the charter schools than public and private schools (items 25, 26, and 27). Evaluation gains momentum in the private schools as well, except in item 31, where the public schools prioritize content subjects in evaluation over English competence. Students highlight that all and both content and oral skills are evaluated, and ongoing and final evaluation are practiced in the class (item 30, 32, and 33). Furthermore, when it is concerned the topic about teachers use, competence and development of English in class, all the bilingual teachers are competent and successful in both teaching and motivating students in the private context (items 34, 35, 37, 38, 41, 42, and 43) more than public and charter context, except in item 36 students responded negatively. Students reveal that their teaching assistants are not so successful in developing the class. Students in the private schools have more opportunities to participate in exchange programs and they are more encouraged to do so than the students in the other two types of school (items 44 and 45). Finally, students enrolled in the private schools highlight very positively that there has been an overall improvement campaigned with a high motivation in learning the language, while the other students hold less positive attitudes towards these two last items 48 and 49.

Table 4. Statistically significant differences in terms of types of schools

		<i>Means</i>	<i>Standard deviation</i>	<i>P_value</i>
Item1	Charter	3,14	,624	<0,001
	Private	3,41	,606	
	Public	3,02	,606	
Item4	Charter	2,39	,915	<0,001
	Private	2,66	,833	
	Public	2,26	,813	
Item7	Charter	3,20	,694	,001

	Private	3,36	,694	
	Public	3,11	,761	
Item15	Charter	3,48	,604	<0,001
	Private	3,56	,541	
	Public	3,32	,631	
Item17	Charter	3,48	,697	,015
	Private	3,50	,575	
	Public	3,36	,666	
Item18	Charter	2,96	,808	,033
	Private	3,01	,760	
	Public	3,13	,766	
Item19	Charter	3,10	,740	<0,001
	Private	3,22	,745	
	Public	2,84	,853	
Item20	Charter	3,00	,749	<0,001
	Private	3,12	,797	
	Public	2,81	,874	
Item21	Charter	2,81	,896	<0,001
	Private	2,99	,835	
	Public	2,65	,809	
Item22	Charter	3,17	,773	<0,001
	Private	3,33	,680	
	Public	3,08	,724	
Item23	Charter	3,14	,697	<0,001
	Private	3,15	,717	
	Public	2,92	,744	
Item24	Charter	2,84	,882	,012
	Private	3,04	,777	
	Public	2,84	,761	
Item25	Charter	3,07	,938	<0,001
	Private	2,97	1,004	
	Public	2,76	,928	
Item26	Charter	2,97	,989	<0,001
	Private	2,81	1,071	
	Public	2,65	,939	
Item27	Charter	2,71	,997	,006
	Private	2,46	1,063	
	Public	2,47	,972	
Item28	Charter	3,48	,788	<0,001
	Private	3,43	,842	
	Public	2,53	1,037	
Item29	Charter	2,15	,964	,015
	Private	2,36	1,123	
	Public	2,06	,956	
Item30	Charter	3,34	,704	,021
	Private	3,36	,632	
	Public	3,21	,674	
Item31	Charter	2,95	,754	,001
	Private	2,77	,785	

	Public	3,02	,769	
Item32	Charter	2,91	,989	<0,001
	Private	3,35	,799	
	Public	2,95	,846	
Item33	Charter	3,35	,706	<0,001
	Private	3,58	,592	
	Public	3,27	,702	
Item34	Charter	3,10	,854	<0,001
	Private	3,35	,696	
	Public	3,06	,765	
Item35	Charter	3,08	,840	,005
	Private	3,26	,723	
	Public	3,03	,798	
Item36	Charter	3,15	,797	,032
	Private	3,00	,861	
	Public	2,97	,903	
Item37	Charter	2,84	,950	<0,001
	Private	3,08	,846	
	Public	2,72	,767	
Item38	Charter	2,84	,864	<0,001
	Private	3,07	,861	
	Public	2,72	,807	
Item41	Charter	3,26	,831	<0,001
	Private	3,65	,570	
	Public	3,10	,724	
Item42	Charter	3,34	,758	<0,001
	Private	3,68	,563	
	Public	3,24	,721	
Item43	Charter	3,22	,732	<0,001
	Private	3,47	,641	
	Public	3,10	,744	
Item44	Charter	1,61	,942	,009
	Private	1,88	1,107	
	Public	1,80	1,005	
Item45	Charter	2,41	1,031	<0,001
	Private	2,80	1,002	
	Public	2,51	1,007	
Item48	Charter	3,44	,715	,013
	Private	3,52	,627	
	Public	3,35	,688	
Item49	Charter	3,23	,805	,048
	Private	3,29	,719	
	Public	3,11	,808	

As for the years of bilingual section experience, few significant differences have emerged comparing them with the above variable. The group of students of more than 7 years of experience show more significant results than the other two groups of less than 4 and between 4 and 6 years of experience in bilingual program. They confirm that they their

English has improved due to their participation in bilingual education, they are interested in the bilingual class, and they have adequate listening and speaking skills in the target language, while the other two groups of less years of experiences in bilingual education express opposite opinion (items 2, 10, 12, and 13). In terms of teaching methodology only one item (item 15) was detected where the respondents of less than 4 years of experiences declare more use of tasks in the CLIL classroom. In contrary, the first group asserts the use of interactive whiteboard, whereas the others two groups have less use of tasks in the classroom (item 28). Regarding mobility, it seems that only the group of more than 7 years in the program are encouraged by their teachers to participate in an exchange programs (item 45). Correspondingly, difference results were found in the last items of the survey (items 47, 48, and 49), the group with larger experience in bilingual education claim that although there is an increase of workload, they feel improvement and they are highly motivated towards learning and participating in the bilingual education. By the contrary, the participants with less experiences than the group of more than 7 years give an opposite opinion.

Table 5. Statistically significant differences in terms of years of experience in a bilingual section

		<i>Mean</i>	<i>Standard deviation</i>	<i>P_valor</i>
Item2	< 4 years	3,39	,679	,002
	> 7 years	3,63	,573	
	4 -6 years	3,44	,746	
Item10	< 4 years	3,26	,723	,015
	> 7 years	3,41	,664	
	4 -6 years	3,12	,729	
Item12	< 4 years	3,22	,763	,033
	> 7 years	3,42	,645	
	4 -6 years	3,21	,808	
Item13	< 4 years	3,07	,746	,001
	> 7 years	3,39	,601	
	4 -6 years	3,24	,781	
Item15	< 4 years	3,30	,570	,004
	> 7 years	3,48	,591	
	4 -6 years	3,21	,770	
Item28	< 4 years	2,89	,987	,018
	> 7 years	3,19	,988	
	4 -6 years	3,06	1,013	
Item45	< 4 years	2,32	1,048	,034
	> 7 years	2,60	1,018	
	4 -6 years	2,30	1,015	
Item47	< 4 years	2,91	,814	,011
	> 7 years	3,17	,718	
	4 -6 years	2,91	,947	
Item48	< 4 years	3,31	,639	,014
	> 7 years	3,46	,674	
	4 -6 years	3,18	,869	

Item49	< 4 years	3,18	,800	,027
	> 7 years	3,23	,775	
	4-6 years	2,85	,857	

As for the last variable the subjects' number in bilingual education, we divided this variable in two chief groups. First group is about the students who study more than 3 subjects through the medium of English and the second group belong to the students who have 3 or less than 3 subjects taught in English. Interestingly, positive and optimistic believe were in favor of the area of equal or less than 3 subjects taught in the target language in all the located items. The participants with less or equal to 3 subjects improve their linguistic background and abilities, they are more interested in learning the target language (items 2 and 10). They learn through tasks (item 15). In relation to materials of the curriculum and sources for learning the same participants recognize that materials are authentic, interesting, and innovative (items 19 and 21). Their teachers collaborate in preparing those materials to fulfil the communicative goals (items 22 and 23) employing digital equipment, the internet and interactive whiteboards for teaching (items 25, 26, 27, and 28). They also score that all the content knowledge learnt are evaluated (item 30). In terms of the use and the development of English by teachers, both language teachers and non-linguistic motivated them (item 37 and 38). Furthermore, all their bilingual teachers have a high level of English skills (items 41 and 42). To sum up, it is logic to say that this positive view can be the cause that students admit their highly motivation towards learning through CLIL (item49). In the contrary, students with more than 3 subjects taught in the vehicular language show lower agreement with the aforementioned aspects.

Table 6. Statistically significant differences in terms of the numbers of subjects in bilingual section

	<i>Means</i>		<i>Standard deviation</i>		<i>R effect size</i>	<i>P_value</i>
	<i><=3</i>	<i>>3</i>	<i><=3</i>	<i>>3</i>		
Item2	3,68	3,56	,522	,631	-0,084	,022
Item10	3,45	3,34	,662	,680	-0,088	,016
Item15	3,53	3,41	,571	,615	-0,088	,016
Item16	3,31	3,18	,682	,744	-0,081	,026
Item19	3,20	2,97	,726	,819	-0,136	<0,001
Item21	2,96	2,73	,827	,867	-0,127	<0,001
Item22	3,30	3,13	,718	,739	-0,123	,001
Item23	3,17	3,02	,709	,730	-0,095	,009
Item25	3,18	2,81	,875	,983	-0,188	<0,001
Item26	3,09	2,67	,943	1,010	-0,202	<0,001
Item27	2,78	2,44	1,013	,996	-0,161	<0,001
Item28	3,55	2,94	,778	1,031	-0,306	<0,001
Item30	3,35	3,28	,714	,653	-0,074	,042
Item37	2,97	2,82	,859	,877	-0,090	,014
Item38	2,96	2,82	,846	,857	-0,078	,032
Item41	3,42	3,28	,735	,768	-0,097	,008
Item42	3,49	3,37	,708	,715	-0,096	,009
Item49	3,28	3,17	,778	,784	-0,077	,035

4. CONCLUSION

Analysing the variance shows that there are statistically significant differences in most of the variables. The reason behind these findings is because of the large sample size (754 students). This large sample helps the analysis to be statistically significant in terms of the following variables: age, gender, nationality, setting, types of schools, years of experience in bilingual education, and number of subjects in a bilingual program.

Consequently, it is interesting to mention that this test yielded more differences as regards the variable of type of schools (public, charter, and private). It appears that most of the differences were scored here in this variable. The majority of items belong to the topics: methodology, material and resource, evaluation, and teachers use, competence and development of English in class. Statistically significant differences showed that the private schools are gaining ground among the other two types of schools. Female students always respond more optimistically than male students. Significant differences between urban and rural areas are located as well. Interestingly, positive and optimistic believe were in favour of the area of equal or less than 3 subjects taught in the target language in all the located items. The group of students of more than 7 years of experience show more significant results than the other two groups of less than 4 and between 4 and 6 years of experience in bilingual program.

Summing up, although there are some interesting differences are detected that does not mean that students perspectives are negative. We are dealing with a satisfactory and optimistic situation where that vast majority of students hold positive view about the advantages of bilingual programme in developing their skills, their learning, and their communicative competences.

5. REFERENCES

- Brindusa, A., Cabrales, A., & Carro, J. (2012). Evaluating a bilingual education program in Spain: the impact beyond foreign language learning. *CEPR Working Paper, No. 8995*, 1-31.
- Bruton, A. (2011). Are the differences between CLIL and non-CLIL groups in Andalusia due to CLIL? A reply to Lorenzo, Casal and Moore (2010). *Applied Linguistics*, 32(2), 236-241.
- Chan, W. M., Chin, K. N., Nagami, M., & Suthiwan, T. (2011). *Processes and Process-Oriented in Foreign Language Teaching and Learning*. Berlin, DEU: Walter de Gruyter.
- Consejería de Educación de la Comunidad de Madrid, por la que se regulan los institutos bilingües de la Comunidad de Madrid. (2009). *Orden 3245/2009*. Retrieved from <https://www.bocm.es/boletin/CM Orden BOCM/2009/07/20/2009-07-20 17072009 0011.pdf>

- Consejería de Educación, Juventud y Deporte de la Comunidad de Madrid Dirección General de Innovación, Becas y Ayudas a la Educación. (2016). *Datos y Cifras de la Educación*. Retrieved from <http://www.madrid.org/cs/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadertype=ContentDisposition&blobheadertype=filename%3DLibro+datos+y+cifras+20152016+navegable.pdf&blobkey=id&blobtable=MungoBlobs&blobwhere=1352898013282&ssbinary=true>
- Dobson, A., Pérez Murillo, M. D. & Johnstone, R. M. (2010). *Bilingual Education Project Spain: Evaluation Report*. Madrid: Gobierno de España Ministerio de Educación & British Council (Spain). Retrieved from <https://www.britishcouncil.es/sites/default/files/bilingual-education-project-spain-evaluation-report-en.pdf>
- Fernández, R., & Halbach, A. (2011). Analysing the situation of teachers in the Madrid bilingual project after four years of implementation. Content and Foreign Language Integrated Learning. Contributions to Multilingualism in European Contexts. In Y. Ruiz de Zarobe, J. M., Sierra, & F. G. del Puerto (Eds.), *Linguistic Content and Foreign Language Integrated Learning: Contributions to Multilingualism in European Contexts* (pp. 241-270). Berne, IN, USA: Peter Lang.
- Martínez Adrián, M. (2011). An overview of Content and Language Integrated Learning: origins, features and research outcomes. *HUARTE DE SAN JUAN. Filología y Didáctica de la Lengua N. 11/Filología eta Hizkuntzaren Didaktika 11 Z. Pamplona: Universidad Pública de Navarra/Nafarroako Unibertsitate Publikoa.* (pp. 103-114).
- Ministry of Education, Youth and Sports. (2015). *Order 1317/2015*. Retrieved from <http://www.madrid.org/wleg/servlet/Servidor?opcion=VerHtml&nmnorma=8929&cdestado=P>
- Ruiz Gómez, D. (2015). A practical approach to CLIL in L2 content-based courses: methodological guidelines for the Andalusian bilingual classroom. In D. Marsh, M. L. Pérez Cañado & J. Ráez Padilla (Eds.), *CLIL in action: voices from the classroom* (pp. 14-30), Newcastle upon Tyne: Cambridge Scholars Publishing.
- Pavón Vázquez, V., & Gaustad, M. (2013). Designing Bilingual Programmes for Higher Education in Spain: Organisational, Curricular and Methodological Decisions. *International CLIL Research Journal*, 2(1), 82-94. Retrieved from <http://www.icrj.eu/21/article7.html>
- Pérez Cañado, M. L. (2016). Evaluating CLIL Programmes: Instrument Design and Validation. *Pulso: revista de educación*, (39), 79-112.

6. APPENDIX

Students questionnaire

Proyecto MON-CLIL: Los Efectos del Aprendizaje Integrado de Contenidos y Lenguas Extranjeras en Comunidades Monolingües: Un Estudio Longitudinal

Cuestionario

ALUMNADO

1. CENTRO:

2. CURSO Y CLASE: _____
3. EDAD: _____
4. SEXO: Hombre Mujer
5. NACIONALIDAD: _____
8. ¿CÚANTOS AÑOS HAS ESTUDIADO EN UN PROGRAMA BILINGÜE? _____
9. ASIGNATURAS QUE ESTUDIAS EN INGLÉS ESTE CURSO:
 - Ciencias Naturales
 - Ciencias Sociales
 - Matemáticas
 - Dibujo
 - Música
 - Educación Física
 - Otra _____
10. EXPOSICIÓN AL INGLÉS DENTRO DEL PROGRAMA BILINGÜE:
¿Qué porcentaje de cada asignatura bilingüe se enseña en inglés? 30% 40% 50% No sé

POR FAVOR, INDICA HASTA QUÉ PUNTO ESTÁS DE ACUERDO CON LOS SIGUIENTES ASPECTOS RELACIONADOS CON LA ENSEÑANZA BILINGÜE (1=Totalmente en desacuerdo; 2=En desacuerdo; 3=De acuerdo; 4=Totalmente de acuerdo).

1. USO, COMPETENCIA Y DESARROLLO DEL INGLÉS DE LOS ALUMNOS EN CLASE

ASPECTOS	TOTALMENTE EN DESACUERDO	EN DESACUERDO	DE ACUERDO	TOTALMENTE DE ACUERDO
1. Se desarrollan las competencias clave en clase	1	2	3	4
2. Mi inglés ha mejorado debido a mi participación en un programa bilingüe	1	2	3	4
4. Mi español ha mejorado debido a mi participación en un programa bilingüe	1	2	3	4
5. Mi conocimiento de los contenidos de las asignaturas impartidas en inglés ha mejorado debido a mi participación en un programa bilingüe	1	2	3	4
6. Mi comprensión de cómo funcionan las lenguas ha mejorado debido a mi participación en un programa bilingüe	1	2	3	4
7. Mi comprensión de la conexión				

entre el inglés y el español ha mejorado debido a mi participación en un programa bilingüe	1	2	3	4
8. Tengo más confianza en mí mismo dentro de la clase bilingüe	1	2	3	4
9. Soy participativo en la clase bilingüe	1	2	3	4
10. Me intereso en la clase bilingüe	1	2	3	4
11. Me gustaría más uso del inglés dentro de la clase bilingüe	1	2	3	4
12. Tengo una capacidad adecuada en comprensión y expresión orales en inglés	1	2	3	4
13. Tengo una capacidad adecuada en comprensión y expresión escritas en inglés	1	2	3	4
14. Tengo un conocimiento adecuado de aspectos socioculturales y una conciencia intercultural en inglés	1	2	3	4
Otro (especificar):	1	2	3	4

2. METODOLOGÍA

ASPECTOS	TOTALMENTE EN DESACUERDO	EN DESACUERDO	DE ACUERDO	TOTALMENTE DE ACUERDO
15. Se desarrollan tareas en clase	1	2	3	4
16. Se desarrollan proyectos en clase	1	2	3	4
17. Aprendo mucho vocabulario en la clase bilingüe	1	2	3	4
18. Se trabaja en grupo dentro de la clase bilingüe	1	2	3	4
Otro (especificar):	1	2	3	4

3. MATERIALES Y RECURSOS

ASPECTOS	TOTALMENTE EN DESACUERDO	EN DESACUERDO	DE ACUERDO	TOTALMENTE DE ACUERDO
19. Se utilizan materiales auténticos para la enseñanza bilingüe	1	2	3	4
20. Se adaptan materiales auténticos para la enseñanza bilingüe	1	2	3	4
21. Los materiales para la enseñanza bilingüe son interesantes e innovadores	1	2	3	4

22.	Los profesores de la sección bilingüe colaboran para preparar y enseñar los materiales de enseñanza bilingüe en clase	1	2	3	4
23.	Los materiales de enseñanza bilingüe fomentan la comunicación en inglés en clase	1	2	3	4
24.	Los materiales de enseñanza bilingüe están adaptados para atender las necesidades de todos los alumnos	1	2	3	4
25.	Se utilizan materiales multimedia (<i>software</i>) en clase	1	2	3	4
26.	Se utilizan materiales de referencia <i>online</i> en clase	1	2	3	4
27.	Se utilizan <i>blogs</i> , <i>wikis</i> (herramientas Web 2.0) y <i>webquests</i> en clase	1	2	3	4
28.	Se utilizan pizarras electrónicas interactivas en clase	1	2	3	4
29.	Se utiliza la comunicación mediada por ordenador en clase (<i>e-learning</i>)	1	2	3	4
Otro (especificar):		1	2	3	4

4. EVALUACIÓN

ASPECTOS	TOTALMENTE EN DESACUERDO	EN DESACUERDO	DE ACUERDO	TOTALMENTE DE ACUERDO	
30.	Se evalúan todos los contenidos enseñados en el programa bilingüe	1	2	3	4
31.	A la hora de evaluar, se tienen en cuenta los contenidos que la presión lingüística	1	2	3	4
32.	Se evalúa también oralmente	1	2	3	4
33.	Se practica la evaluación continua y final	1	2	3	4
Otro (especificar):		1	2	3	4

5. USO, COMPETENCIA Y DESARROLLO DE INGLÉS DE LOS PROFESORES EN CLASE

ASPECTOS	TOTALMENTE EN DESACUERDO	EN DESACUERDO	DE ACUERDO	TOTALMENTE DE ACUERDO	
34.		1	2	3	4

ex éxi	Mis profesores de lenguas ranjeras imparten sus clases con to				
35.	Mis profesores de asignaturas bil éxi ngües imparten sus clases con to	1	2	3	4
36.	Mis auxiliares de conversación im parten sus clases con éxito	1	2	3	4
37.	Mis profesores de lenguas ex ranjeras motivan al alumno	1	2	3	4
38.	Mis profesores de asignaturas bil ngües motivan al alumno	1	2	3	4
39.	Mis auxiliares de conversación m otivan al alumno	1	2	3	4
40.	Mis auxiliares de col alu conversación aboran con éxito con los mnos de la clase bilingüe	1	2	3	4
41.	Mis profesores tienen una ca acidad adecuada en co in mprensión y expresión orales en lés	1	2	3	4
42.	Mis profesores tienen una acidad adecuada en :omprensión y expresión escritas en inglés	1	2	3	4
43.	Mis profesores tienen un ocio mociamiento adecuado de aspectos socio -culturales en la lengua inglesa	1	2	3	4
Otro (especificar):		1	2	3	4

6. MOVILIDAD

ASPECTOS	TOTALMENTE EN DESACUERDO	EN DESACUERDO	DE ACUERDO	TOTALMENTE DE ACUERDO	
44.	He participado en programas de int ercambio dentro del programa bili ngüe	1	2	3	4
45.	Mis profesores de la sección ngüe bili fomentan la participación en programas de intercambio	1	2	3	4
46.	Mi familia me anima a participar programas de intercambio	1	2	3	4
Otro (especificar):		1	2	3	4

7. MEJORAS Y MOTIVACIÓN PARA EL APRENDIZAJE DE INGLÉS

ASPECTOS	TOTALMENTE EN DESACUERDO	EN DESACUERDO	DE ACUERDO	TOTALMENTE DE ACUERDO
47. Formar parte de una sección bilingüe compensa el incremento de horas que implica	1	2	3	4
48. Ha habido una mejoría general de mi aprendizaje de inglés debido a mi participación en un programa bilingüe	1	2	3	4
49. Mi motivación hacia el aprendizaje del inglés ha aumentado debido a mi participación en un programa bilingüe	1	2	3	4
50. Tengo un acceso adecuado a materiales en inglés fuera del centro	1	2	3	4
Otro (especificar):	1	2	3	4

MUCHAS GRACIAS POR TU COLABORACIÓN

CONSTRUCCIÓN Y VALIDACIÓN DE UN CUESTIONARIO DE CONOCIMIENTOS SOBRE PREVENCIÓN DE UN EMBARAZO NO DESEADO.

Sebastián Sanz Martos¹, Cristina Álvarez García, Carmen Álvarez Nieto , Isabel María López Medina.

(1) *Departamento de Enfermería, Universidad de Jaén, Edif. A2, dep. 024, Campus Las Lagunillas, s/n, 23071 Jaén, e-mail: ssm00026@red.ujaen.es*

(2) *Departamento de Enfermería, Universidad de Jaén, Edif. B3, dep. 221, Campus Las Lagunillas, s/n, 23071 Jaén, e-mail: cag00019@red.ujaen.es*

(3) *Departamento de Enfermería, Universidad de Jaén, Edif. B3, dep. 243, Campus Las Lagunillas, s/n, 23071 Jaén, e-mail: calvarez@ujaen.es*

(4) *Departamento de Enfermería, Universidad de Jaén, Edif. B3, dep. 265, Campus Las Lagunillas, s/n, 23071 Jaén, e-mail: imlopez@ujaen.es*

RESUMEN

El embarazo durante la adolescencia ha sido ampliamente estudiado por las consecuencias para la madre y el feto, así como las distintas intervenciones para prevenirlo. Las intervenciones más efectivas para la prevención son las intervenciones combinadas, abordando el problema desde el punto de vista sanitario y comunitario. Sin embargo, a la hora de medir los resultados surge la limitación de que no disponemos de ningún instrumento apropiado para evaluar el aumento de conocimiento sobre cómo prevenirlo. Objetivos: Construir y validar un instrumento para medir el nivel de conocimiento sobre prevención del embarazo no deseado. Metodología: La población de estudio estuvo formada por 160 estudiantes de Grado en Enfermería de la Universidad de Jaén, que cursaban la asignatura Enfermería de la infancia y de la adolescencia. Se formó un banco inicial de 23 ítems. La validez de contenido se determinó calculando el grado de acuerdo entre un grupo de expertos, la validez de constructo mediante análisis factorial exploratorio y la fiabilidad de la escala se determinó mediante el cálculo de la consistencia interna (Alfa de Cronbach). Resultados: Tras el cálculo del grado de acuerdo entre expertos, análisis de los ítems y realización del análisis factorial, seleccionamos un total de 7 ítems que saturan en 3 factores. El cálculo de la consistencia interna arroja un valor pobre de alfa de Cronbach ($\alpha=0.52$). Conclusión: El instrumento presenta buenos indicadores de validez de contenido y constructo, sin embargo su consistencia interna para esta muestra tiene un valor pobre.

Palabras clave: Embarazo no deseado, validación, conocimiento, métodos anticonceptivos, cuestionario.

ABSTRACT

Introduction: Pregnancy during the adolescence has been studied extensively because the consequences for the mother and the foetus, as well as the different interventions to prevent it. The most effective preventive interventions are the combined interventions, tackling the problem from the sanitary and community point of view. However, when trying to measure the results, we face the limitation of a lacking appropriate instrument that allows evaluating the knowledge increase of how to prevent it. Objectives: To construct and to validate an instrument to measure the knowledge level on prevention of unwanted pregnancy. Methodology: The study population was 160 Nursing Degree students from University of Jaén, who studied the subject “Enfermería de la infancia y de la adolescencia”. We formed an initial bank of 23 items. The validity of the content was stipulated by calculating the grade of agreement of a group of experts. Construct validation was calculated by an exploratory factor analysis. Scale reliability was calculated by internal consistency (Cronbach's alpha). Results: After calculating expert's grade of agreement, items analysis and the exploratory factor analysis, we select 7 items that saturate in 3 factors. The calculation of the internal consistency returns a poor alpha value ($\alpha=0.52$). Conclusion: The instrument presents good validity indicators for content and construct; nevertheless its internal consistency for this sample shows a poor value. The knowledge level for the sample is acceptable.

Keywords: Unwanted pregnancy, validation, knowledge, contraceptive methods, questionnaire.

1. INTRODUCCIÓN

El embarazo durante la adolescencia ha sido ampliamente estudiado por las consecuencias para la madre y el feto, así como las distintas intervenciones para prevenirlo. (Colomer, 2013)

Se han diseñado una amplia variedad de programas educativos, sin embargo a la hora de evaluar los resultados respecto de la efectividad de las diferentes intervenciones educativas surge la limitación de la inexistencia de un instrumento validado para medir el aumento en el nivel de conocimientos sobre sexualidad y métodos anticonceptivos como forma de prevención de un embarazo no deseado. (Oringanje et al. 2009). Por otro lado, otro resultado que arrojan las diferentes intervenciones educativas es la necesidad de plantear intervenciones educativas desde el ámbito sanitario. La figura de la enfermera se plantea como clave en la realización de estos programas educativos, sin embargo es necesario que tengan un conocimiento suficiente para perfilarse como una fuente fidedigna de información sobre sexualidad y anticoncepción. (Colomer, 2013)

2. METODOLOGÍA

Objetivo

El objetivo de esta investigación fue la construcción y validación de una herramienta para medir el nivel de conocimientos que tienen los estudiantes de enfermería sobre sexualidad y métodos anticonceptivos

Diseño

Se llevó a cabo dentro del diseño de un estudio observacional de validación.

Población y muestra

La muestra de estudio la conformaron los estudiantes del Grado en Enfermería de la Universidad de Jaén, que durante el curso académico 2015-2016 cursaban la asignatura de Enfermería de la Infancia y adolescencia. Obtuvimos como resultado una muestra para realizar la prueba piloto de 160 participantes.

Procedimiento de creación del instrumento

El primer paso fue la determinación del número de ítems del instrumento que lo definimos en 12 ítems divididos en tres dimensiones (Conocimientos sobre sexualidad, conocimientos sobre preservativo y conocimientos sobre anticonceptivos hormonales). Siguiendo las recomendaciones de Polit y Beck (2008), formulamos un banco inicial de al menos un 50% más de los deseados, por lo que formulamos un banco inicial de 23 ítems. A continuación creamos un comité de 5 expertos nacionales en el constructo de estudio para que valoraran la claridad y pertinencia de los ítems para medirlo. La validez de contenido se determinó calculando el grado de acuerdo entre un grupo de expertos mediante el estadístico V de Aiken, la validez de constructo mediante análisis factorial exploratorio y la fiabilidad de la escala se determinó mediante el cálculo de la consistencia interna (Alfa de Cronbach).

3. RESULTADOS

Como resultado de la valoración del comité de expertos del total de 23 ítems en 8 se modifica su redacción, y 4 ítems son eliminados a petición de dos de los expertos. Las razones que alegan para su eliminación son la no pertinencia para medir el constructo para tres de ellos y para el cuarto la medición de un constructo distinto. Tras calcular el grado de acuerdo entre los expertos, tres ítems son eliminados por tener un valor de índice de acuerdo inferior al valor crítico 0.7 (tabla 1).

A continuación calculamos el índice de dificultad y el índice de discriminación de los ítems seleccionados posteriormente a la valoración de los expertos. Para el cálculo del coeficiente de dificultad se calculó la proporción de participantes que respondieron correctamente al ítem con respecto al total de personas que lo intentaron. El índice de discriminación se halló mediante la división de la muestra entre el 30% de participantes con mayores puntuaciones y el 30% con puntuaciones más bajas, descartando los participantes centrales. Vamos a usar los dos índices de discriminación (1 y 2).

El primer índice se calcula mediante una diferencia de proporciones (acertantes del grupo superior - acertantes del grupo inferior / número de participantes en ambos grupos), mientras que el segundo nos calcula la proporción de acertantes del grupo superior con respecto al total de acertantes (acertantes del grupo superior / acertantes del grupo superior + acertantes del grupo inferior). Para el primero consideramos como aceptable un índice de discriminación superior a 0.30, mientras que para el segundo consideramos aceptable superior a 0.50. De los 15 ítems seleccionados después de valorar el juicio de los expertos, seleccionamos 8 ítems tras calcular el índice de dificultad y los índices de discriminación (tabla 1).

Tabla 1: Cálculo de los índices del grado de acuerdo, índice de dificultad e índices de discriminación.

Ítem	V de Aiken	Índice de Dificultad	Índice de Discriminación 1	Índice de Discriminación 2
Ítem 1*	0.81	0.88	0.38	0.62
Ítem 2	0.38			
Ítem 3	0.88	0.99	0	0.5
Ítem 5	0.69			
Ítem 6	0.88	0.97	0.088	0.52
Ítem 7	0.75	0.94	0.05	0.51
Ítem 9	0.94	0.98	0.05	0.51
Ítem 10*	0.75	0.84	0.32	0.60
Ítem 11*	0.75	0.85	0.32	0.60
Ítem 12*	0.94	0.25	0.26	0.69
Ítem 13*	0.88	0.65	0.56	0.74
Ítem 14*	0.81	0.75	0.38	0.62
Ítem 15*	0.81	0.73	0.44	0.65
Ítem 18	0.88	0.96	0.15	0.54
Ítem 19	0.56			
Ítem 20	0.81	0.92	0.06	0.52
Ítem 21	0.75	0.98	-0.03	0.49
Ítem 22	0.81	0.99	0	0.5

*Ítems seleccionados.

A continuación se calculó la correlación entre los ítem y entre estos y la puntuación final, obteniendo como resultado correlaciones significativas con la puntuación final para 7 de los 8 ítems y correlaciones entre los ítem altas salvo para uno de los ítem, el cual procedemos a eliminar, obteniendo como resultado una muestra de 7 ítems con los que procedemos a realizar el análisis factorial exploratorio para determinar la validez de constructo del instrumento. El primer paso para realizar el análisis factorial es calcular la prueba de esfericidad de Bartlett y las pruebas de Kaiser-Meyer-Olkin (KMO).

La prueba KMO obtiene un valor de 0.535, considerado como mediocre, sin embargo sí superamos el valor crítico de aceptabilidad de 0.5. Por otro lado la prueba de esfericidad de Bartlett sí nos permite rechazar la hipótesis nula de independencia de la matriz identidad, indicando la adecuación de los datos para la factorización.

La extracción de los factores se realiza mediante el método de análisis de componentes principales y rotación ortogonal (Realizamos la rotación Varimax, Equamax y Quartimax, seleccionando la que mejores resultados nos aporten). Partimos de 3 componentes teóricos que confirmamos mediante análisis del gráfico de sedimentación de Cattell (figura 1), varianza total explicada por los tres componentes y análisis de los autovalores de los componentes (tabla 2).

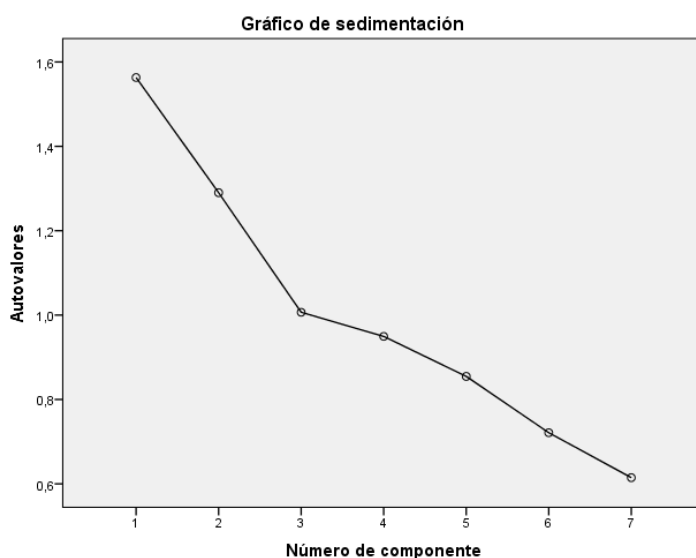


Figura 1: Grafico de sedimentación de Catell.

Tabla 2: Varianza total explicada

Factor	Autovalores iniciales			Sumas de extracción de cargas al cuadrado			Sumas de rotación de cargas al cuadrado		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	1,563	22,330	22,330	1,563	22,330	22,330	1,433	20,464	20,464
2	1,290	18,431	40,761	1,290	18,431	40,761	1,341	19,164	39,629
3	1,007	14,381	55,142	1,007	14,381	55,142	1,086	15,513	55,142
4	,949	13,564	68,705						
5	,855	12,210	80,915						
6	,721	10,303	91,219						
7	,615	8,781	100,000						

Analizando el gráfico de sedimentación vemos que hay un cambio en la pendiente en el tercer factor lo que indica que la mejor solución factorial para reducir dimensiones es a tres factores, como postulamos teóricamente.

La matriz de varianzas y covarianzas expone que los tres primeros factores tienen autovalores mayores a 1 y conjuntamente explican el 55.142 % de la varianza total de los datos, por lo que el número de factores necesarios para explicar la información es igual al que teóricamente exponemos.

A partir del modelo expuesto se calcularon las comunalidades de las variables, que es la proporción de su varianza que puede ser explicada por el modelo factorial. Como resultado de del análisis de las comunalidades obtenemos un porcentaje de varianza explicada variable en función de los ítems, así uno de los ítems explicaba un 33% de la varianza, dos de ellos un porcentaje superior al 40%, mientras que los demás obtuvieron un porcentaje de varianza explicada superior al 50%.

Para el análisis de la solución factorial establecemos los siguientes criterios:

- La carga factorial de un ítem en un factor debe ser igual o superior al 0.5.
- En el caso de que un ítem sature en dos o más factores, conservaremos en el factor que mayor correlación presente, siempre y cuando la saturación en el otro factor no sea igual o superior a 0.5. El ítem se eliminará cuando presente dos saturaciones superiores a 0.5 en dos factores.

Para conocer mejor las saturaciones de los ítems en cada componente realizamos una rotación ortogonal. La necesidad de realizar una rotación reside en que el método de extracción de componentes que hemos seleccionado está orientado a extraer la máxima cantidad de varianza en cada paso, de forma que a los últimos factores les queda poca varianza que explicar, lo que implica que los primeros factores no son factores puros en sí mismos, ya que contienen varianza explicada por ellos y parte de varianza que tendría que ser explicada por otros factores, lo que causa que las cargas factoriales de los últimos factores tiendan a ser inferiores a los de los primeros factores.

Tras realizar la rotación obtenemos las saturaciones relativas de cada ítem en cada factor. El factor 1 está formado por tres ítems, el factor 2, por tres ítems y el factor 3 está formado por un ítem (tabla 3).

Tabla 3: Matriz de componentes rotados (Rotación Varimax).

<i>Ítems</i>	<i>Componente</i>		
	1	2	3
Ítem 1	,573	,336	,113
Ítem 10	,198	,507	,376
Ítem 11	,749	,068	-,051
Ítem 13	,703	-,274	,050
Ítem 14	-,100	,764	-,271
Ítem 15	,020	,554	,155
Ítem 23	-,012	,029	,911

Por último calculamos la fiabilidad de la escala mediante el estadístico alfa de Cronbach estratificado; el Alfa de Cronbach tiene gran utilidad cuando se usa para determinar la consistencia interna de una prueba con un único dominio o dimensión, sin embargo cuando se usa en escalas con ítems que exploran varias dimensiones, aunque formen parte de la evaluación de un constructo, se corre el riesgo de subestimar la consistencia interna. En estos casos lo más indicado, para mejorar el valor del coeficiente de consistencia interna es mediante el coeficiente del alfa de Cronbach estratificado. Obtenemos como resultado un valor de alfa de Cronbach estratificado de 0.52 (tabla 4).

Tabla 4: Cálculo de Alfa de Cronbach estratificado.

σ^2 Total	Factor	σ^2 Factor	α Cronbach factor	σ^2 Factor $\cdot (1-\alpha$ factor)	α Cronbach estratificado
6.783	Factor 1	2.527	0.419	1.4682	0.5212
	Factor 2	2.72	0.330	1.8224	
	\sum factores			3.2906	

4. CONCLUSIONES

El instrumento que hemos creado presenta buenos indicadores de validez de contenido, ya que todos los ítems presentaron un valor de grado de acuerdo entre los expertos superior a 0. Respecto a la validez de constructo, tras la realización del análisis factorial mediante el método de componentes principales y rotación ortogonal Varimax, obtenemos un instrumento formado por 7 ítems agrupados en tres factores que explican el 55.142% de la varianza.

En relación con la fiabilidad del instrumento, los índices de consistencia interna (α de Cronbach y α de Cronbach estratificado) muestran valores pobres, siendo la fiabilidad de la escala limitada.

Durante el proceso de validación se ha producido una pérdida excesiva de ítems causado que la muestra de ítems por factor no sea la recomendada para la realización de un análisis factorial exploratorio con suficiente tamaño muestral de ítems por factor donde se recomienda un tamaño de 3 a 4 ítems por factor y un descenso en la fiabilidad de la escala. Otra limitación que presenta la investigación es el tamaño muestral insuficiente para la realización de un análisis factorial para el que se recomienda un tamaño muestral mínimo de 200 participantes. (Catena Martínez et al, 2003)

Para futuras investigaciones recomendamos la formación de un mayor número de ítems iniciales para conseguir una muestra final de ítems suficiente para medir el constructo y realizar el proceso de validación con garantías, así como un tamaño muestral mayor.

5. BIBLIOGRAFÍA.

Catena Martínez, A., Trujillo Mendoza, H.M., Ramos Álvarez, M.M. (2003). Análisis multivariado: un manual para investigadores, 1ªed. Barcelona, España: Biblioteca Nueva.

Colomer, J., Grupo PrevInfand/PAPPS infancia y adolescencia. (2013). Prevención el embarazo en la adolescencia. Rev Pediatr Aten Primaria, 15,261-269.

Oringaje, C.H., Meremikwu, M., Eko, H., Esu, E., Meremwiku, A., y Ehiri, J.E. (2009). Intervenciones para la prevención de embarazo no deseado en adolescentes. Biblioteca Cochrane Plus. Número 4. Oxford.