

**Modelado del Calibre y la competición intra-específica por rendimiento de tubérculos de papa variedad *Solanum phureja* bajo diferentes densidades de siembra**

**Nelson Bernal Margfoy**

Universidad Nacional de Colombia

Facultad de Ciencias Agrarias

Ciudad, Colombia

2017

**Modelado del Calibre y la competición intra-específica por rendimiento de tubérculos de papa variedad *Solanum phureja* bajo diferentes densidades de siembra**

**Nelson Bernal Margfoy**

Tesis o trabajo de investigación presentada(o) como requisito parcial para optar al título de:

**Magister en Ciencias Agrarias**

Director:

Ph.D. Aquiles Enrique Darghan Contreras

Codirector:

Ph.D. Luis Ernesto Rodríguez Molano

Línea de Investigación:

**Línea de Investigación en Fisiología de Cultivos**

Universidad Nacional de Colombia

Facultad de Ciencias Agrarias

Ciudad, Colombia

2017

*(Dedicatoria o lema)*

*Su uso es opcional y cada autor podrá determinar la distribución del texto en la página, se sugiere esta presentación. En ella el autor del trabajo dedica su trabajo en forma especial a personas y/o entidades.*

*Por ejemplo:*

*A mis padres*

*o*

*La preocupación por el hombre y su destino siempre debe ser el interés primordial de todo esfuerzo técnico. Nunca olvides esto entre tus diagramas y ecuaciones.*

*Albert Einstein*

**Agradecimientos**

Esta sección es opcional, en ella el autor agradece a las personas o instituciones que colaboraron en la realización del trabajo. Si se incluye esta sección, deben aparecer los nombres completos, los cargos y su aporte al trabajo.

A continuación se presenta la tabla de contenido la cual se actualiza automáticamente. Para los textos editados en Microsoft Word se debe hacer *click* en el botón derecho del mouse sobre la tabla de contenido y aparecerá el icono Actualizar Campos (), luego aparecerá una ventana en la cual debe seleccionar la opción Actualizar toda la tabla.

Resumen

Se realizó un estudio de campo con el cultivo de papa criolla variedad Solanum phureja para evaluar la influencia de la densidad de siembra asociada a distancias entre plantas de 30,40 y 50 cm y distancia entre surcos y caminos de 100 cm sobre el conteo de tubérculos de calibres inferiores a 2 cm, de 2-4 cm, 4 – 6 cm y más de 6 cm y sobre el rendimiento en peso fresco (g) de los tubérculos. El estudio se realizó en el Centro Agropecuario Marengo de la Universidad Nacional de Colombia, en el departamento de Cundinamarca (74°12'58.51''W;4°40'52.92''N), el cual tiene una altitud de 2516 msnm y una temperatura media de 14 °C. De acuerdo a las características del clima, la zona se clasifica como Bosque Seco Montano Bajo. Al momento de la cosecha de los tubérculos, se hizo su clasificación y conteo mediante tamizado y se pesaron en su totalidad sin discriminar por calibre. El modelado de los conteos se hizo mediante regresión binomial negativa usual e inflada por ceros, mientras que el modelado del peso fresco se logró usando un modelo autoregresivo espacial puro. Los modelos de regresión binomial negativa cero-inflada evidenciaron el efecto significativo de la densidad de siembra sobre el conteo de tubérculos y calibre, de hecho, se observó una razón aproximada de 40:40:20:1 desde el calibre menor al mayor. En lo que respecta al efecto de la competición, en todos los modelos probados resultó significativo el coeficiente de competición, aumentando en la mayoría de los casos a medida que disminuía la distancia entre plantas, tanto en el patrón de vecindad intrahileras como en el caso de inter e intrahileras. Los resultados sobre calibre y densidad son atributos de interés en el manejo tanto agronómico del cultivo como en el manejo industrial de los tubérculos, por lo que la relación encontrada puede adoptarse en ambas áreas para generar los atributos deseados del cultivo para mejorar la producción y guiar el proceso de industrialización.

**Palabras clave:** Regresión Poisson, Regresión Binomial Negativa, Regresión Binomial negativa cero-inflada, calibre, modelo autoregresivo puro.

**Abstract**

A field study was carried out with Creole potato crop variety *Solanum phureja* to evaluate the influence of the plant density associated to distances between plants of 30,40 and 50 cm and distance between rows and paths of 100 cm on the tuber count with calibers less than 2 cm, 2-4 4 cm, 4 - 6 cm and more than 6 cm and on the yield in fresh weight (g) of the tubers. The study was conducted at the Marengo Agricultural Center of the Universidad Nacional de Colombia, in the department of Cundinamarca (74°12'58.51''W, 4 °40'52.92''N), which has an altitude of 2516 meters above sea level and an average temperature of 14 °C. According to the characteristics of the climate, the area is classified as Dry Forest low mountain. When the tubers were harvested, their classification and counting were made by sifting and they were weighed in their entirety without discriminating by size. The modeling of the counts was done by usual negative binomial regression and zero-inflated, while the modeling of the fresh weight was achieved using a pure spatial autoregressive model. Zero-inflated negative binomial regression models showed the significant effect of planting density on tuber count and caliber, in fact, an approximate ratio of 40: 40: 20: 1 was observed from the smallest to the largest caliber. About the effect of the competition, in all tested models the competition coefficient was significant, increasing in most cases as the distance between plants decreased, both in the intra-row neighborhood pattern and in the case of inter and intra-rows. The results on caliber and density are attributes of interest in both the agronomic management of the crop and in the industrial management of the tubers, so the relationship found can be adopted in both areas to generate the desired attributes of the crop to improve production and guide the process of industrialization.

**Keywords:** PoissonRegression, Negative Binomial Regression, Negative Binomial Regression Zero-inflated, caliber, Pure Spatial Autoregression.

**Contenido**

Pág.

[Resumen IX](#_Toc498980087)

[Lista de figuras XII](#_Toc498980088)

[Lista de tablas XIV](#_Toc498980089)

[Introducción 1](#_Toc498980090)

[1. Modelado del calibre y la competición intra-específica por rendimiento de tubérculos de papa variedad *Solanum phureja* bajo diferentes densidades de siembra …………………………………………………………………………………..13](#_Toc498980091)

[1.1 Materiales y métodos 17](#_Toc498980092)

[1.1.1 Localización del ensayo 17](#_Toc498980093)

[1.1.2 Material vegetal 17](#_Toc498980094)

[1.1.3 Variable de respuesta y explicativa 18](#_Toc498980095)

[1.1.4 Diseño experimental 18](#_Toc498980096)

[1.1.5 Análisis estadístico 19](#_Toc498980097)

[1.1.6 Patrones de vecindad 20](#_Toc498980098)

[1.2 Resultados y Discusión 22](#_Toc498980099)

[1.3 Conclusión 42](#_Toc498980100)

[2. Conclusiones y recomendaciones 53](#_Toc498980101)

[2.1 Conclusiones 53](#_Toc498980102)

[2.2 Recomendaciones 54](#_Toc498980103)

Lista de figuras

Pág.

[Figura 1‑1. Modelos comunes de distribución en datos de conteo. (Generados en R software) 6](#_Toc499014927)

[Figura 1‑1. Patrón de vecindad intra-hilera en vecinos más cercanos 21](#_Toc499014928)

[Figura 1‑2. Patrón de vecindad inter-intra-hileras en vecinos más cercanos. 21](#_Toc499014929)

[Figura 1‑3. Diagrama de dispersión del peso fresco y Diámetro medio ponderado de los tubérculos por densidad de siembra. 25](#_Toc499014930)

[Figura 1‑4. Distribución de conteos de tubérculos por calibre con ajuste de Poisson 25](#_Toc499014931)

[Figura 1‑5. Distribución de conteos de tubérculos por densidad de siembra y calibre (D1, D3, D5 y D9) 28](#_Toc499014932)

[Figura 1‑6. Diagrama de dispersión de valores observados y estimados para el modelo 3. 30](#_Toc499014933)

[Figura 1‑7. Regresión Binomial Negativa para densidad en el calibre ≤ (2cm) 31](#_Toc499014934)

[Figura 1‑8. Regresión Binomial Negativa para densidad en el calibre (2 − 4] cm 31](#_Toc499014935)

[Figura 1‑9. Regresión Binomial Negativa para densidad en el calibre (4 − 6] cm 32](#_Toc499014936)

[Figura 1‑10. Regresión Binomial Negativa para densidad en el calibre > 6cm 33](#_Toc499014937)

[Figura 1‑11. Regresión Binomial Negativa cero-inflada para densidad en el calibre ≤ (2cm) 34](#_Toc499014938)

[Figura 1‑12. Regresión Binomial Negativa cero-inflada para densidad en el calibre (2 − 4] cm 34](#_Toc499014939)

[Figura 1‑13. Regresión Binomial Negativa cero-inflada para densidad en el calibre (4 − 6] cm 34](#_Toc499014940)

[Figura 1‑14. Regresión Binomial Negativa cero-inflada para densidad en el calibre > 6 cm 34](#_Toc499014941)

[Figura 1‑15. Comparación del modelo binomial negativo con su contraparte cero-inflada (calibres ≤ (2) y (2 − 4] cm) 35](#_Toc499014942)

[Figura 1‑16. Comparación del modelo binomial negativo con su contraparte cero-inflada (calibres (4 − 6] y > 6cm) 35](#_Toc499014943)

[Figura 1‑17. Comparación de valores observados (Obs.) y predichos (Pred.) (Calibres ≤ (2) y (2 − 4] cm) 37](#_Toc499014944)

[Figura 1‑18. Comparación de valores observados (Obs.) y predichos (Pred.) (Calibres (4 − 6] y > 6cm) 37](#_Toc499014945)

[Figura 1‑19. Comparación de valores observados y predichos (Densidad (30cm\*100cm)- Competición intra – hilera) 40](#_Toc499014946)

[Figura 1‑20. Comparación de valores observados y predichos (Densidad (40cm\*100cm)- Competición intra – hilera) 41](#_Toc499014947)

[Figura 1‑21. Comparación de valores observados y predichos (Densidad (50cm\*100cm)- Competición intra – hilera) 41](#_Toc499014948)

[Figura 1‑22. Comparación de valores observados y predichos (Densidad (30cm\*100cm)- Competición inter – intra – hilera) 42](#_Toc499014949)

Lista de tablas

Pág.

[**Tabla 1‑1**. Distribución de los pesos por distancia entre surcos del patrón de asignación por distancias inversas 21](#_Toc498979604)

[**Tabla 1‑2**. Distribución del conteo de tubérculos por densidad y marca de clase del calibre 24](#_Toc498979605)

[**Tabla 1‑3.** Distribución del Peso fresco promedio de tubérculos por densidad y marca de clase del calibre 24](#_Toc498979606)

[**Tabla 1‑4**. Media, Varianza y cociente varianza/media por densidad y calibre 26](#_Toc498979607)

[**Tabla 1‑5**. Modelado por Regresión de Poisson del Calibre hasta 2 cm 27](#_Toc498979608)

[**Tabla 1‑6**. Estadísticas de ajuste modelo 1 27](#_Toc498979609)

[**Tabla 1‑7**. Modelado por Regresión de Poisson del Calibre hasta 2 cm 28](#_Toc498979610)

[**Tabla 1‑8**. Estadísticas de ajuste modelo 2 29](#_Toc498979611)

[**Tabla 1‑9**. Modelado por Regresión de Poisson para calibre y densidad 29](#_Toc498979612)

[**Tabla 1‑10.** Estadísticas de ajuste modelo 3 29](#_Toc498979613)

[**Tabla 1‑11**. Resultados de la Competición intra – hilera (vecino más cercano) 38](#_Toc498979614)

[**Tabla 1‑12**. Resultados de la Competición inter-intra-hilera (vecino más cercano) 39](#_Toc498979615)

Introducción

*Solanun tuberosum* Grupo Phureja se encuentra cultivado, preferentemente, en la parte oriental de los Andes y usualmente entre los 2000 – 3400 msnm. Su distribución geográfica se extiende desde el noroeste de Bolivia, toda la región oriental de los Andes peruanos, hasta Colombia y parte de Venezuela. Los cultivares pertenecientes a este grupo son precoces, de ahí su nombre nativo en Aymara, phureja, sus tubérculos no tienen periodo de reposo y es posible establecer ciclos de siembra - cosecha tres o cuatro veces al año (Ochoa, 2001, citado por Piñeros, 2009). La producción de papa en Colombia se distribuye en cerca de 14 departamentos, pero sólo tres concentraron en el 2005 el 80% de la misma y el 80% de la superficie cultivada. Los cuatro mayores productores son: Cundinamarca (42%), Boyacá (23%), Nariño (14%) y Antioquia (8,8%). Los cultivos de papa se encuentran diseminados en climas fríos con temperaturas de 13º C y alturas de 2.000 m.s.n.m., hasta alcanzar zonas de páramo con alturas cercanas a los 3.500 m.s.n.m. y temperaturas de 8º C. Geográficamente, las unidades de producción están dispersas en las regiones frías de la Zona Andina, bajo una variada gama de condiciones biofísicas, sociales y económicas (Villareal, Porras, Santa, Lagoeyte, & Muñoz, 2007).

**Colombia es el mayor productor, consumidor y exportador de papas diploides en el mundo; tiene una ventaja competitiva notable debido a ser centro de diversidad y poseer gran aceptación por los consumidores debido a las características organolépticas y nutricionales del tubérculo. Adicionalmente, en el país se ha desarrollado una amplia tradición como cultivo tecnificado, con potencial de industrialización y exportación (Rodríguez, Ñustez, & Estrada, 2009). Además,** cuenta con uno de los recursos genéticos que ofrece las mayores oportunidades de exportación como alimento procesado étnico, exclusivo y sin competencia. Desafortunadamente, muchas de las estrategias de manejo del cultivo de la papa, han sido adaptadas a papa criolla, creando un sistema productivo poco eficiente (Piñeros, 2009). **el nombre de papa criolla concierne a los morfotipos que presentan tubérculos de color de piel y carne amarillo (fenotipo yema de huevo) (Rodríguez et al., 2009). Donde ha sido clasificada inicialmente como *Solanun phureja* por el profesor Hawkes en 1990, posteriormente como *Solanun tuberosum* Grupo Phureja (Huamán & Spooner, 2002), y recientemente como *Solanun tuberosum* Grupo Andigena (Spooner et al., 2007).**

**La papa amarilla es un cultivo alimenticio típico de las tierras altas de los Andes, cuya calidad depende del microclima, altura, radiación solar y humedad de los ecosistemas donde se produce. Es altamente valorado el consume en fresco, sobre todo por sus altos contenidos de materia seca, color y textura. Lo anterior indica que la papa amarilla es un cultivo alimenticio de alto valor en el mercado (Fano, Carmona, Ordinola, & Scott, 1998). La variedad Criolla Colombia Presenta hábito de crecimiento erecto, buen desarrollo de follaje, color verde claro, flor lila oscuro. Tubérculos de forma redonda, ojos semiprofundos, ausencia de periodo de reposo, color de piel y carne amarillo intenso, bajo rendimiento en el proceso de encurtido, altos contenido de materia seca (Rodríguez et al., 2009).** El desarrollo de nuevos cultivares de papa a nivel diploide con mayor potencial de rendimiento, homogeneidad y condiciones ideales para diferentes opciones de procesamiento debe contribuir al posicionamiento de la papa criolla como un producto de importancia en los mercados internacionales. A pesar de su heterogeneidad fenotípica, en Colombia una fracción pequeña de papa criolla es procesada, tanto para mercado nacional como de exportación, en forma precocida congelada o encurtida, en presentaciones que van desde la bolsa plástica hasta la papa enlatada o envasada en vidrio, siendo una opción para incursionar con éxito en los mercados internacionales (Rodríguez et al., 2009)

Los altos precios de los insumos agrícolas y el mal manejo de los cultivos agronómicamente provoca baja productividad y amenaza la competitividad del sistema de producción, por lo que es importante identificar factores limitantes del rendimiento y desarrollar prácticas innovadoras para el cultivo, tales como una gestión nutricional integrada y equilibrada, que es una de las prácticas más eficientes para garantizar a la planta la oportunidad de expresar su potencial genético que eventualmente se reflejará en una mejor calidad y rendimiento (López, Gómez, & Rodríguez, 2014)

Algunos resultado de investigación enuncia que el incremento de la población de plantas de 40.000 a 100.000 por hectárea dio un incremento de por lo menos 50% en el rendimiento, sin embargo, el tamaño promedio de los tubérculos fue menor (> 30 mm de tamaño) (Escobar & Vander, 1988).

En el proceso de modelado del calibre que tienen los tubérculos de papa es importante reconocer el tipo de dato que se genera en esta situación, donde solo se cuentan tubérculos por calibre. Cuando se procede al modelado de datos de conteo es relevante definir lo que exactamente significa un conteo. La palabra "conteo" se usa típicamente como verbo para dar a entender que se trata de enumerar unidades, elementos, o eventos. La frase “conteo de datos “por otra parte, es un sustantivo plural que hace referencia a observaciones hechas sobre elementos o atributos que están enumerados. En el campo de la estadística, los datos de conteo se refieren a las observaciones que solo toman valores enteros y que pueden asociarse al campo de los números enteros con inclusión del cero. Desde el punto de vista teórico un conteo puede variar de cero a infinito, pero siempre se limitan a algunos valores característicos según la naturaleza de la variable bajo estudio. En lo que respecta al conteo de tubérculos, en una planta pueden aparecer cero tubérculos en general o sencillamente cero tubérculos asociados a un calibre en particular. En el caso de la variedad criolla, los calibres usualmente establecidos para una cosecha realizada a los 120 días después de la siembra se manejan en categorías de diámetro o calibre, siendo la de menor diámetro la categoría inferior a 2cm, seguida por la de (2-4] cm, luego la de (4-6] cm y finalmente la categoría de más de 6 cm. Quienes hacen la clasificación de los tubérculos por calibre realizan conteos de tubérculos por calibre y es usual encontrar para esta variedad conteos de cero tubérculos en los calibres mayores así como conteos de hasta de 40 tubérculos en los calibres inferiores. Al modelar estos datos se encuentran valores diferentes, y aunque tienen comportamientos distribucionales inherentes a este tipo de dato, es usual modelarlos como si se trataran de datos continuos.

Una "variable de conteo" es una lista o matriz de datos de conteos, que, para el caso del modelado estadístico, pudiera ser tanto respuesta o como una variable explicativa, usualmente de naturaleza aleatoria con observaciones independientes, con una distribución particular univariante o multivariante según el número de variables de conteo involucradas.

Desde un punto de vista descriptivo, los conteos pueden rendir tablas de conteos para una variable de conteo, así tablas de contingencias bidimensionales o multidimensionales según los factores de agrupación. Si se tiene el factor densidad de siembra a tres niveles (30, 40 y 50 cm) de distancia de siembra entre surcos, los conteos rinden una tabla de conteos o de frecuencias para el criterio densidad de siembra. Si ahora se incorpora el calibre (los cuatro descritos previamente) de los tubérculos, se genera una tabla bidimensional al cruzar calibre con densidad de siembra, obteniéndose doce celdas donde se registran los conteos bivariados. Ya con la tabla de contingencia pueden involucrarse la gran variedad de procedimientos de modelado de tablas de contingencia. (Stokes et al, 2012).

Otra forma de modelado de conteos se basa en la incorporación de dos distribuciones de probabilidad comúnmente usadas en este tipo de datos, a saber, la distribución de Poisson y la Binomial Negativa, sin embargo, es posible encontrar otras distribuciones relacionadas como lo son la distribución Gaussiana Inversa de Poisson, la binomial negativa de tres parámetros y la Poisson Generalizada. Estas cinco distribuciones están claramente relacionadas.

Antes de describir brevemente estas distribuciones, es importante resaltar que los modelos de regresión para conteos , sea la regresión Poisson o binomial y negativa, en sus modalidades usuales o en la opción inflada por ceros de ceros truncados pertenecen a la familia de modelos lineales generalizados, por lo que a continuación se hace una descripción de este marco de modelado que amplía las opciones de análisis de nuestros datos, especialmente en situaciones donde la normalidad es cuestionable o donde la naturaleza de la variable es precisamente la que se asocia a un conteo.

Los modelos de regresión de datos de conteo básicos se pueden representar y comprender utilizando el marco de los modelos lineales generalizados (Nelder y Wedderburn, 1972), los cuales describen la dependencia de una variable escalar sobre un vector de regresores o predictores . La distribución condicional de es una familia exponencial con función densidad de probabilidad

Donde es el parámetro canónico que depende de los regresores por medio del predictor lineal y es el parámetro de dispesión (frecuentemente conocido). Las funciones *b*(.) y *c*(.) son conocidas y determinan al miembro de la familia que se está usando, ya sea el caso de la distribución normal, la binomial, Poisson, entre otras. La media condicional y la varianza de *y*i son respectivamente ) y , de este modo la distribución de *y*i está determinada por su media. La dependencia de con los regresores se especifica mediante

**Donde g(.) es una función de enlace conocida y β es el vector de coeficientes de regresión, los que usualmente se estiman vía máxima-verosimiltud.**

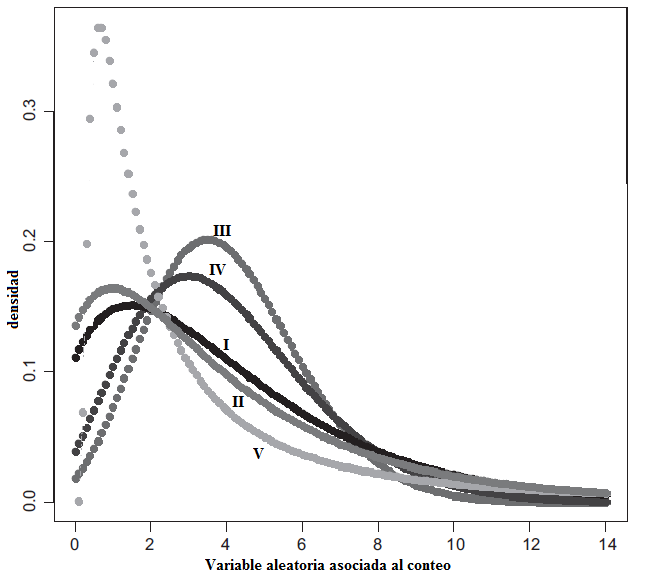
**Conocer algunos de estos aspectos resulta relevante al momento de seleccionar los métodos, la función de enlace, la distribución y otros aspectos de los modelos dentro de paquetes específicos, tal como R o STATA, por esta razón, se encomia al lector hacer una revisión exhaustiva de los modelos lineales generalizados para el uso adecuado de los mismos, especialmente en el modelado de datos de conteo (McCullagh P, Nelder, 1989).**

**Retornando al caso de la distribución de Poisson (I) (perteneciente a la familia exponencial), esta tiene un único parámetro para ser estimado (la media), con la particularidad especial de tener tanto el parámetro de la media como la varianza con el mismo valor, de modo que a mayor promedio mayor varianza en los datos. Esta interesante relación entre la media y la varianza se conoce como equidispersión, algo que, en el proceso de modelado de datos reales, rara vez se alcanza, por lo que se requiere algún tipo de ajuste para tratar con el alejamiento de equidispersión, conocido como sobredispersión, lo que se entiende sencillamente como una mayor variabilidad en los datos que su promedio, pero también es tratada como una mayor variabilidad en los datos observados es superior a la varianza esperada. Un modelo con sobredispersión se asocia a mayores errores estándar o errores estándar sesgados lo que puede generar interpretaciones en la significación evaluada de los parámetros del modelo ajustado.**

**Muy probablemente el método más popular para tratar con sobredispersión es el binomial negativo (II), ya que esta tiene un parámetro adicional denominado parámetro de dispersión binomial negativo, lo que permite ajustar la variabilidad adicional o heterogeneidad en los datos, pero no permite modelar datos Poisson sub-dispersos, solo sobre dispersos.**

**El tercer (III) tipo de modelo de conteos es el modelo Gaussiano inverso de Poisson, en el que se asume que la sobredispersión en un modelo de Poisson se describe adecuadamente de acuerdo con la distribución Gaussiana inversa en lugar de la distribución gamma que es propia al modelo binomial negativo. El cuarto (IV) tipo de modelo para conteos es un binomial negativo de tres parámetros, el cual tiene un parámetro que permite que la dispersión varíe a través de las observaciones. Finalmente, el quinto modelo (V) es el Poisson Generalizado, el cual llega a ser el mismo Poisson cuando el parámetro de dispersión se hace cero. La siguiente figura ilustra claramente el comportamiento distribucional de los cinco modelos tratados, y resulta de interés como el V modelo difiere claramente del resto, con un exceso de conteos en las frecuencias bajas (posiblemente el cero), algo bastante común en el modelado del calibre de tubérculos de papa en los diámetros mayores.**

Figura 1‑1. Modelos comunes de distribución en datos de conteo. (Generados en R software)



**También existen otros modelos para conteos como lo son la opción cero-inflada y cero-truncada, modelos de Hurdle, modelos de panel, modelos cuantiles, modelos exactos de Poisson, modelos Bayesianos, entre otros.**

**La estructura de un modelo de conteo para los modelos descritos previamente tiene una estructura básica propia del modelo lineal, solo que la variable respuesta es transformada por su logaritmo, de modo que el modelo se escribe como:**

**Donde es la media predicha del conteo,**  **representa la i-ésima variable explicativa, representa al intercepto y representa al i-ésimo parámetro asociado a cada variable explicativa. El término es conocido como la función de enlace ya que esta vincula al predictor lineal (lado derecho de la ecuación anterior) y al valor predicho. Una característica de manipular el logaritmo natural como la función de enlace en modelos de conteo es que garantiza que los valores predichos sean positivos, algo que la regresión usual no garantiza y que es requerido en modelos de conteo.**

**En la investigación actual, un componente de la matriz de datos se asocia a datos de conteo, y ciertamente, no resulta obvio ajustar la distribución de Poisson, pues los datos no siempre tienen este patrón característico, por ejemplo, los datos podrían traer un exceso de ceros, o hasta datos sin ceros, por lo que es posible que sea necesario explorar modelos cero- inflados o ceros truncados según los ceros encontrados en los datos. Los modelos de Hurdle se construyen como modelos en dos partes, truncando algunos valores de la distribución para lograr ajuste.**

**En el proceso de modelado de los conteos de tubérculos se procedió de acuerdo al siguiente esquema:**

* **Descripción de los datos asociados a los conteos de tubérculos por calibre para cada densidad de siembra.**
* **Selección de una distribución de probabilidad o mezcla de distribuciones, para ajustar los datos de calibre por densidad y calibre.**
* **Obtención de las estimaciones de los parámetros de los modelos probados, incluyendo predictores de relevancia, que en nuestro caso tenemos densidad y algunas mediciones de peso fresco y el diámetro medios ponderados provenientes de la ponderación de las marcas de clase de los calibres usualmente considerados.**
* **Cálculo de valores ajustados o predichos de la respuesta en función del modelo seleccionado.**
* **Evaluación de las diferencias entre los valores estimados o ajustados y los valores observados en todas las observaciones.**
* **Evaluar las estadísticas de ajuste, tales como medidas de sobredispersión, criterios de información de Akaike, entre otros.**

**Para el modelado del calibre resultó fundamental la regresión de Poisson a manera ilustrativa, por lo que se ajustaron tres modelos para ir identificando sus parámetros y los estadísticos de ajuste asociados, así como el problema de la sobredispersión, de hecho, fue el primer modelo probado, pues casi todos los modelos posteriores al de Poisson mencionados previamente son modelos que surgen del ajuste del Poisson. Posiblemente una de las características que permite rápidamente saber que este deba ser probado inicialmente es la posibilidad de incluir el cero como respuesta, algo que sucedió en todos los cuatro calibres estudiados. Sin embargo, desde el principio se entendió que no podría ser el único modelo a evaluar pues la gran cantidad de ceros encontrados en los dos calibres mayores supera los valores esperados para este conteo en una distribución de Poisson.**

**Para el proceso de modelado y discusión del actual trabajo de investigación se ajustó el modelo de regresión binomial negativa como segunda opción para los dos primeros calibres donde el número de ceros no era tan excesivo como en los dos últimos calibres, sin embargo, los resultados de todas las pruebas evidenciaron que la opción cero inflada fue superior en todos los casos, especialmente la binomial negativa cero inflada. Una vez generados los estadísticos del modelo y los de ajuste, se presentaron las ecuaciones o modelos asociados como los gráficos de comparación entre valores de conteo observados y estimados por cada modelo. La idea en general es proponer un modelo por calibre de modo que se pueda estimar los conteos esperados por densidad y calibre, lo cual resulta valioso, pues el calibre es un criterio para la manipulación industrial de los tubérculos (referencia), y la densidad se asocia a la manipulación general del cultivo, el aprovechamiento del espacio, el control de las arvenses y otros aspectos inherentes al cultivo (Masarirambi et al, 2012).**

**Los modelos cero-inflados (Mullahy 1986, Lambert 1992) son otra clase de modelos que permiten la manipulación de excesos de ceros. Son modelos que combinan dos partes, una, la que se asocia a los ceros y la otra a las distribuciones de Poisson o Binomial negativa, entre otras. El ajuste de estos modelos surge del análisis inicial en los datos, donde los histogramas muestran una evidente mayor frecuencia en las clases más bajas, y en el cero propiamente, no como consecuencia de atípicos, sino la forma natural de generación de los ceros, por ejemplo, papas del calibre mayor en el cultivo de papa criolla cuando la cosecha se hace a los 120 días después de la siembra. Un modelado usual generalmente fallaría a considerar este patrón, por lo que se precisa un modelado específico que considere esta situación y permita relacionar adecuadamente el conteo con los predictores, que en este caso se asocia a la densidad de siembra.**

**En la actual investigación, para casi todos los calibres, resultaron adecuados los modelos propuestos en la opción cero-inflada, específicamente el caso binomial negativo. Las pruebas específicas de ajuste, así como la significación de los términos dentro de cada modelo evidenciaron la relación entre los conteos de tubérculos y la densidad de siembra por calibre.**

**Una vez finalizado el proceso de modelado del calibre, se procedió al modelado del peso fresco de los tubérculos utilizando la posición espacial de cada planta, los patrones de vecindad y la densidad de siembra para estimar el efecto de la competición (o interferencia desde el punto de vista agronómico) mediante modelos econométricos espaciales, los cuales se describen a continuación.**

**En esta sección se introduce el proceso de modelado que considera la dependencia espacial de la variable respuesta “rezagada espacialmente”. Estos modelos son llamados comúnmente como modelos autoregresivos (Anselin, 1988), sin embargo, para no generar confusión, podrían ser llamados como modelos espacialmente rezagados, pues principalmente en el caso de la respuesta, esta se encuentra rezagada en el componente de los predictores. Este tipo de modelos pueden ser convenientes cuando se sospecha que el valor de la respuesta en un sitio en particular, pueden ser influenciados por la respuesta en un sitio próximo o vecino, tal como se ha pensado en la competición de cultivos, donde se cree que la respuesta de una planta en una cierta posición puede ser afectada por la respuesta de las plantas que la rodean. (Gomez y Gomez, 1984; Draper y Guttman, 1980, Hide et al, 1990). Este razonamiento conduce inicialmente a la construcción de un modelo de una variable respuesta espacialmente rezagada tal como el que se muestra a continuación:**

**Donde un valor negativo asociado con λ indica que una planta puede presentar bajo promedio de rendimiento si en promedio sus vecinas aumentan su rendimiento. Si no se tiene ningún predictor , por lo que el modelo se reduce a uno con solo la variable respuesta en su forma natural o en su forma en rezago, es decir:**

De este modo, los únicos parámetros a estimar son el intercepto () y , que autores como Shukla y Subrahmanyan (1999) han llamado como coeficiente de competición. En el término se tiene la matriz de pesos que contiene la información ya sea de la distancia entre plantas o de cualquier otro atributo (Darghan et al, 2010; Darghan et al, 2011). Este último modelo es conocido como modelo de autoregresión puro y sus parámetros se estiman vía máxima-verosimilitud.

Para el caso de la actual investigación se ajustó en varias ocasiones el mismo modelo, utilizando dos diferentes modalidades de vecindad, a saber, solo considerando vecinos a los de un surco (solo los vecinos más cercanos para competición intra – hilera) y un patrón tomando nuevamente los vecinos más cercanos pero involucrando solo los dos surcos más cercanos en las plantas que estuvieron a una distancia fija como para un esquema tipo reina en lo que se refiere a contigüidad de vecinos (competición inter – intra – hilera). La matriz de pesos se construyó para ambas modalidades y los parámetros del modelo se estimaron en principio para cada densidad de siembra y por calibre de los tubérculos de papa. La idea fue estimar el coeficiente de competición y notar cómo iba cambiando este a medida que aumentaba la distancia entre plantas para asociarlo a la competición del cultivo. Se usó el peso fresco de los tubérculos como respuesta y como predictor, solo que como predictor se rezagó con la matriz de pesos. Los resultados mostraron un coeficiente de competición altamente significativo y una tendencia a disminuir a medida que aumenta el espaciado entre plantas, lo que de algún modo complementó el objetivo de la investigación, pues nuevamente y al igual que en el caso de los conteos de los tubérculos, la densidad parece influir en el rendimiento.

****Bibliografía****

Darghan, A.E. 2010. Test score de Rao para el modelo de solapamiento de Draper y Guttman. PhD thesis. Faculty of Sciences, Universidad de los Andes, Merida, Venezuela.

Darghan, A.E., P. Sinha, and A. Goitia. 2011. Score test para el coeficiente de solapamiento en modelos de superficies de respuesta de primer y segundo orden. Dyna 165, 234-245.

Draper,N.R., Guttman, J. (1980). Incorporating overlap effects from neighbouring units into response surfaces models. Applied Statistics, 39,128-134.

Escobar, V., & Vander, Z. (1988). Field performance of potato (Solanum Spp.) cuttings in the warm tropocs: Influence of planting system, hilling, desnsity and pruning. *International Potato Center (CIP)*, *65*(933), 1–10.

Fano, H., Carmona, G., Ordinola, M., & Scott, G. (1998). *Experiencias de Exportación de la papa amarilla peruana*. (Centro Internacional de la Papa (CIC), Ed.). Perú. Retrieved from http://cipotato.org/wp-content/uploads/2015/03/WP56258.pdf

Gomez, K.A., Gomez,A.A. (1984). Statistical Procedures for Agricultural Research. 2nd. Edition. John Wiley and Sons, Inc.

Hide, G.A. and P.J. Read. 1990. Effect of neighbouring plants on the yield of potatoes from seed tubers affected with gangrene (Phoma foveata) or from plants affected with stem canker (Rhizoctonia solani). Ann. Appl. Biol. 116, 233-243.

Huamán, Z., & Spooner, A. (2002). Reclassification of Landrace Populations of Cultivated Potatoes (Solanum sect. Petota). *American Journal of Batany*, *89*(6), 947–965.

López, A., Gómez, M. I., & Rodríguez, L. E. (2014). Effect of edaphic and foliar applications of different doses of zinc on the yield of the criolla colombia cultivar. *Agronomía Colombiana*, *32*(1), 70–77.

McCullagh P, Nelder JA (1989). Generalized Linear Models. 2nd edition. Chapman & Hall, London.

Nelder JA, Wedderburn RWM (1972). “Generalized Linear Models.” Journal of the Royal Statistical Society A, 135, 370–384.

Piñeros, C. (2009). *Recopilación de la investigación del sistema productivo papa criolla*. *Fedepapa*.

Rodríguez, L., Ñustez, C., & Estrada, N. (2009). Criolla Latina, Criolla Paisa y Criolla Colombia, nuevos cultivares de papa criolla para el departamento de Antioquia (Colombia). *Agronomía Colombiana*, *27*(3), 289–303.

Shukla ,G.K.,Subrahmanyan, G.S.V (1999). A note on a test and confidence interval for competition and overlap effects. Biometrics, Vol. 55, No. 1, pp. 273-276.

Spooner, D., Nunez, J., Trujillo, G., Herrera, M. del R., Guzman, F., & Ghislain, M. (2007). Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification. *Proceedings of the National Academy of Sciences*, *104*(49), 19398–19403.

https://doi.org/10.1073/pnas.0709796104

Villareal, H., Porras, P., Santa, A., Lagoeyte, J., & Muñoz, D. (2007). Costos de producción de papa en las principales zonas productoras de Colombia. Retrieved from http://www.fedepapa.com/wp-content/uploads/pdf/ESTUDIO-SOBRE-COSTOS-DE-PRODUCCION.pdf

# 

# Modelado del calibre y la competición intra-específica por rendimiento de tubérculos de papa variedad *Solanum phureja* bajo diferentes densidades de siembra

**En Colombia, *Solanum tuberosum* grupo phureja, o papa criolla, hace referencia a los fenotipos conocidos como yema de huevo, los cuales presentan color de piel y carne amarilla (Rodriguez etal., 2009). La especie se cultiva entre los 2000 y 3000 msnm, a temperaturas que oscilan entre los 10 y 20°C. En el país se siembran cerca de 8500 hectáreas al año de papa criolla, el 6 % del área total sembrada en papa (Villareal, 2011; Ñústez, 2011). De esta área, se obtienen aproximadamente 100000 toneladas en los departamentos de Boyacá, Cundinamarca, y Nariño (CNP, 2012; Herrera y Rodríguez, 2012). La papa criolla tiene una gran importancia en Colombia, ya que según Sáenz (2014), tiene un 90 % de aceptación entre los consumidores, y además, el 80 % de las familias la compran diariamente para su consumo. Por otra parte, es importante mencionar que la papa criolla contiene carotenoides como violaxantina, luteína, y xeoxantina, los cuales junto con los aminoácidos esenciales, hacen parte de cerca de un 10 % del peso seco del tubérculo (Rodriguez et al., 2011), convirtiéndola en un alimento de alta calidad alimenticia. Como una variable respuesta de interés en la actual investigación fue el calibre de los tubérculos categorizado por el diámetro medio, es importante describir la caracterización morfológica, que en lo que respecta a las raíces, son escasas y de poco volumen de expansión, con presencia de tallos verdaderos (aéreos) y tallos modificados (estolones y tubérculos), siendo los aéreos de naturaleza herbácea y de longitud y diámetro variables. Las plantas provenientes del tubérculo- semilla presentan en promedio cinco tallos de producción independiente y su ramificación se asocia a la densidad de siembra, rindiendo entre 2.5 y 8.0 tubérculos por tallo. Las plantas provenientes de semilla sexual solo desarrollan un tallo principal (Cabezas y Corchuelo, 2005). El cultivo es de ciclo corto (4 meses) y diferentes investigaciones evidencian que el peso total de tubérculos por m2 se asocia estadísticamente a la distancia entre surcos, obteniéndose un mayor peso de los tubérculos totales (sin discriminar por clase o calibre) cuando se utilizan distancias entre surcos menores a un metro, sin embargo, este aumento del rendimiento parece asociarse a una reducción del tamaño promedio de tubérculos, lo que sugiere un mayor fraccionamiento del calibre, con predominio de los calibres inferiores (Wiersema, 1987; Arias et al, 1996).**

**Varias investigaciones revelan que el tamaño promedio de los tubérculos, su variabilidad y su conteo, definen una única distribución, presentando mayor variabilidad de rangos de tamaño en la etapa final de llenado, con aumento del rendimiento de los mayores tamaños por cuenta de los tamaños inferiores, lo que sugiere que un tubérculo que pasa de un calibre inferior al siguiente, no es sustituido por otro de calibre anterior, además, parece existir una correlación negativa entre la variabilidad relativa del tamaño del tubérculo y el número de tubérculos por unidad de área (Struik et al, 1991). Es importante reconocer, que debe tenerse cuidado con la interpretación de esta correlación, pues hasta cierto punto, la naturaleza de los datos asociados al número de tubérculos por planta es pudiera ser similar a la de los datos composicionales, es decir, si un componente aumenta, el otro está forzado a disminuir, lo que genera correlaciones espurias es datos verdaderamente composicionales, sin embargo, el número de tubérculos por calibre no suma una constante, pudiera ocurrir que su máximo posea una distribución específica, entonces una mayor cantidad de tubérculos de un calibre podría provocar un número menor de tubérculos en otro calibre específico, lo que en esencia puede generar correlaciones en las que debe tenerse mucho cuidado al momento de su interpretación.**

**Según Rivera et al. (2003), la calidad de los tubérculos de papa criolla no solo depende del manejo cultural del cultivo, por ejemplo, la densidad o espaciado de la siembra, sino de condiciones ambientales como el microclima, altura, radiación solar y luminosidad de los ecosistemas donde se desarrolla. Arias et al., (1996), han realizado estudios donde comparan los rendimientos obtenidos por hectárea con determinadas densidades de siembra, pero no han modelado el calibre como la variable respuesta de interés, muy seguramente por la naturaleza categórica de la variable y por la presencia de ceros en uno de los calibres, lo que impone ciertas restricciones al proceso de modelado. Recientes investigaciones sugieren uso de densidades de siembra en papa adaptadas a las propiedades fisicoquímicas del suelo medidas mediante sensores de conductividad eléctrica aparente (CEa), mostrando que la densidad puede ser un atributo del manejo sitio-específico del cultivo, por lo que en zonas asociadas a ensayos donde se prueban diferentes densidades pero que el tamaño del lote no presenta gran variabilidad en la CEa, es probable obtener rendimientos similares, donde solo pudiera variar el calibre de los tubérculos. (Reckleben et al, 2017). De aquí surge el interés de modelar el calibre como función de la densidad de siembra. Gómez y Wong (2000) señalan que tubérculos con un diámetro de 2.5 a 4.5 cm pueden usarse para la preparación de papas precocidas y congeladas. Para conservas y decapados se recomienda un diámetro medio de 2.5 cm y tubérculos de forma redonda y comprimida (IICA, 1999). Las patatas para uso industrial deben poseer buen sabor, bajos niveles de glicoalcaloides, sin oscurecimiento y bajos niveles de azúcares reductores (Bonierbale et al., 2008; 2009).**

**La importancia del cultivo obliga a muchos investigadores de la región a realizar diferentes estudios para mejorar su calidad según el uso que vaya a darse a los tubérculos, generando nuevas variedades resistentes a plagas y enfermedades y de fácil adaptación a diferentes pisos climáticos, procurando hacer uso del espacio de siembra de forma óptima, por lo que varios estudios evidencian el estudio de la densidad de siembra para evaluar sobre todo el rendimiento. En la actual investigación se evaluó en lugar del rendimiento, un indicador que se asocia directamente como lo es el calibre de los tubérculos, los cuales en la práctica se manipulan en cuatro categorías de diámetro promedio (hasta 2 cm, de 2 a 4 cm, de 4 a 6 cm y más de 6 cm). La naturaleza de esta variable imposibilita usualmente la comparación de la respuesta (conteos de tubérculos) para cada densidad de siembra (30cm\*100cm, 40cm\*100cm y 50cm\*100cm) mediante análisis de varianza, pues en el caso de conteos, otras distribuciones como la Poisson y la Binomial negativa se adaptan mucho mejor al tipo de dato generado. Los modelos clásicos de regresión Poisson y binomial y negativa, en sus modalidades usuales o en la opción inflada por ceros (Cameron et al, 1998) en datos de conteo pertenecen a la familia de modelos lineales generalizados (Zeileis et al, 2008) y los desarrollo recientes permiten generar varias estadísticas que permiten su comparación con sus contrapartes sin ceros en exceso, lo que resulta útil en la elección del mejor modelo para relacionar predictores y respuesta en conteos.**

**Para dar inicio al proceso de modelado se exploraron los datos para evitar presencia de datos atípicos y otras fuentes de variabilidad que pudieran afectar resultados. La visualización gráfica inicial con ajuste de distribuciones de Poisson y Binomial negativa para los datos por densidad y calibre permitió la elección de los modelos Binomiales negativos por su mejor ajuste, sin embargo, en los modelos de conteo, inicialmente se acostumbra a ajustar los modelos Poisson, generando modelos de regresión para esta distribución. En nuestro caso, los modelos del caso Poisson solo sirvieron como ilustración para contrastarlos con el Binomial negativo, por lo que los resultados exhaustivos se presentaron para el caso Binomial negativo. Los histogramas para cada densidad en cada calibre permitieron veri car la presencia de ceros en exceso, sobre todo para los dos mayores calibres, por lo que seguidamente se procedió al modelado con las opciones cero-infladas, lo que evidenció su superioridad con el caso no inflado en la mayoría de los calibres, con excepción del calibre de (2 4] cm, donde se mantuvo el modelo Binomial negativo.**

**Todo el modelado de los conteos de calibre permitió concluir estadísticamente la relación entre la densidad de siembra y los conteos, algo que descriptivamente y por conteos mostró estas diferencias, pero en términos de porcentajes de tubérculos condicionados por densidad dieron la impresión de que estas diferencias no eran de importancia, algo que seguramente se debió a la presencia de excedo de ceros, algo que se evidenció en la distribución del peso fresco estimado para cada calibre según la densidad, donde los pesos frescos fueron mayores a mayor distanciamiento entre plantas (40 y 50 cm) en comparación con el de 30 cm.**

**Estos primeros resultados iniciaron la segunda parte del proceso de modelado, evaluando la competición utilizando modelos econométricos espaciales (Arbia, 2014) y matrices de pesos generadas a partir de dos modalidades, una primera donde no importó la separación entre plantas y una segunda donde el peso se generó por el inverso de las distancias entre los vecinos más cercanos. Entre los modelos se evaluaron dos opciones, cuatro modelos autoregresivos puros (un por calibre y densidad) donde solo intervino como respuesta el fresco estimado de los tubérculos de cada calibre por cada densidad. Los modelos autoregresivos permitieron la estimación del parámetro autoregresivo, el cual se asoció al coeficiente de competición, tal como sugirió (Shukla y Subrahmanian, 1999). Los coeficientes se tabulación en una tabla cruzada para densidad y calibre y con estos se estimó el peso fresco ponderado por competición, para finalmente generar el gráfico de valores observados y estimados por cada modelo.**

## Materiales y métodos

### Localización del ensayo

El estudio se realizó en el Centro agropecuario Marengo de la Universidad Nacional de Colombia, en el departamento de Cundinamarca (74°12'58.51 W; 4°40'52.92 N), el cual tiene una altitud de 2516 msnm, temperatura media de 14 C en un rango de 12C a 18C y precipitación media de 500 a 1000 mm el cual cuenta con un paisaje en planicie uvio lacustre y un relieve en terraza lacustre plana (que no excede al 1 % con suelos son moderadamente profundos y bien drenados . El régimen de humedad es ústico y un nivel freático a menos de 0.5 m del 15 %. De acuerdo a las características de precipitación, temperatura y evapotranspiración, la zona se clasifica como Bosque Seco Montano Bajo

### Material vegetal

El material vegetal utilizado se correspondió al cultivo de papa criolla *Solanum phureja*, utilizando el tubérculo como semilla con el tamaño y forma característica de la especie (tamaño mediano), ojos poco profundos, sin pudrición ni defectos en la piel. Esta variedad con un porte de planta medio y follaje verde claro con abundante oración, distinguida por su adaptación a días cortos, de origen y distribución en América del Sur, desde el norte de Bolivia hasta el sur occidente de Venezuela, comprendiendo Bolivia, Perú, Ecuador, Colombia y Venezuela, y con centro de diversidad genética al sur de Colombia. Su desarrollo vegetativo se da hasta los 35 días después de la siembra (dds), siguiendo la oración hasta los 65 dds, fructificación hasta los 90 dds y finalmente la madurez y senescencia hasta los 120 dds. Presenta producción de tubérculo con distribución de tamaños (diámetros entre 1 y 8 cm). Esta variedad es precoz (120 días a 2600 msnm), su potencial de rendimiento en condiciones óptimas de cultivo es de 15 a 25 ton.ha−1, sin periodo de reposo y susceptible al virus del amarillamiento de las nervaduras de la hoja (Potato yellow vein virus) .Se cultiva en las diferentes regiones del país y en diferentes condiciones de suelo. Es la principal variedad de papa criolla cultivada en Colombia y hasta la presente es la variedad que se procesa para exportación como precocida congelada (Ñustez, 2011; Rodríguez y Ñustez, 2012).

### Variable de respuesta y explicativa

A los 120 dds se cosecharon los tubérculos y se contaron según su diámetro en las categorías 2 cm, (2-4] cm, (4-6] cm y > 6 cm. Se evaluó el peso fresco. Como se utilizaron varios modelos, se fijaron las distancias entre plantas de (30, 40 y 50 cm) como los niveles del factor densidad. Además, los diámetros categorizados originalmente, también fueron tratados como marcas cuantitativas utilizando el punto medio de cada clase, fijando para el mayor diámetro la marca de 9 cm, pues la literatura reporta diámetros máximos de 12 cm, lo que se establecieron inicialmente las marcas de (1, 3, 5 y 9 cm) de diámetro para los diferentes tamaños de tubérculos, y así generar una nueva variable(diámetro ponderado medio) asociada a los conteos de tubérculos por diámetro, el cual se obtuvo para todas las observaciones, y posteriormente se utilizó como ponderador para estimar los pesos frescos asociados a cada calibre. Toda esta manipulación generó una matriz de datos con las variables peso fresco (observado en campo) por densidad de siembra en un solo vector de datos, otro vector de datos con el diámetro medio ponderado, los conteos de tubérculos por calibre (observados en campo) y los pesos estimados de peso fresco estimados a partir del diámetro medio ponderado generado a partir del número de tubérculos.

### Diseño experimental

La siembra se realizó en surcos alineados con precisión según la densidad de siembra, utilizando tres surcos sucesivos según la geometría del lote para cada densidad, con dos repeticiones por densidad, lo que rindió un total de 18 surcos, para un total de 2841 plantas. Aunque la unidad que aportó cada dato fue la planta (tubérculos), la obvia dificultad para aleatorizar una densidad de siembra usando cada planta como unidad experimental, obligó a la aleatorización de las densidades de siembra, cada una con sus tres respectivos surcos (unidad experimental) dentro del lote, registrando los datos de cada planta (unidad de observación). Bajo estas condiciones, el diseño resultó ser una factorial simple en arreglo completamente al azar, tomando las distancias entre plantas como los niveles del factor.

### Análisis estadístico

El análisis estadístico involucró inicialmente el componente descriptivo, generando tablas cruzadas para el conteo de tubérculos por densidad y calibre y otra para el peso fresco promedio observado por los mismos criterios. Seguidamente se realizó un gráfico de dispersión entre peso fresco observado por densidad y diámetro medio ponderado estimado a partir del conteo de tubérculos. Cada uno de estos gráficos se acompañó con la línea de regresión lineal estimada por mínimos cuadrados y el coeficiente de correlación estimado para el par de variables. Adicionalmente, se elaboró una tabla cruzada de promedios, varianzas y cociente entre varianzas y promedio para cada calibre y cada densidad con el objeto de notar la relación posible entre la varianza y la media del conteo de tubérculos, lo cual facilitó la elección de la distribución discreta asociada al proceso de modelado de regresión por conteos. Finalmente se presentó gráficamente los histogramas de conteos por calibre (sin discriminar por densidad) ajustando la distribución de Poisson solo como ilustración de la importancia de explorar los datos para elegir la apropiada distribución de los conteos de tubérculos. El gráfico final incorporó la discriminación por densidad y calibre para los conteos de tubérculos.

En lo que respecta al componente inferencial, se ajustaron diversos modelos del tipo lineal generalizado, presentando finalmente aquellos con los mejores descriptores del ajuste y cumplimiento de supuestos. Los primeros modelos evaluados fueron los de regresión Poisson para conteos, generados uno para cada calibre y utilizando la densidad como factor. Los resultados para los calibres remanentes aparecen en los anexos pues no fueron precisamente los mejores modelos ajustados, no solo por evidenciar sobre dispersión, sino porque estos no consideraron el exceso de ceros evidente en la matriz de datos. A manera de ilustración también se ajustó un modelo de dos factores con la misma distribución solo para poner en evidencia que la sobredispersión no se debió a la omisión de efectos de interacción (este modelo se complementó con la ecuación ajustada y el diagrama de dispersión de valores observados y estimados por el modelo). Del análisis exploratorio ya era conocido como la distribución Binomial negativa parecía generar el mejor ajuste, por lo que los siguientes modelos de regresión usaron la distribución binomial negativa para cada calibre, utilizando como factor la densidad de siembra. Inmediatamente después se obtuvieron los cuatro modelos correspondientes a la opción cero-inflada y se contrataron mediante el test de Vuong y el criterio de información de Akaike corregido por sesgo, las dos modalidades. Esto se complementó con el gráfico de comparación entre valores observados y estimados de cada modelo en la opción cero-inflada.

Para finalizar el proceso de modelado, pero ahora usando como respuesta el peso fresco estimado para cada calibre a partir del peso fresco observado en campo para cada densidad y con el diámetro medio ponderado estimado, se ajustaron modelos econométricos espaciales autoregresivos puros para estimar el coeficiente de competición para dos modalidades de matrices de pesos con la información de los vecinos más cercanos para los diferentes patrones de vecindad (esquinero, lateral e interior). Los valores obtenidos se representaron en una tabla cruzada y a partir de esta estimación, se generó el peso fresco estimado para cada calibre y densidad y se comparó con su respectivo valor observado para validar el modelo propuesto. Los resultados del modelado del conteo de tubérculos y el coeficiente de competición se contrastaron a manera descriptiva para concluir finalmente sobre el efecto de la competición por densidad de siembra y calibre.

### Patrones de vecindad

Los patrones que se muestran en la Figura 1‑1 y Figura 1‑2 se corresponden con dos escenarios evaluados en los modelos econométricos espaciales. En el caso de la asignación igual (Figura 1‑1) se asume que los pesos no son función de las distancias sino del número de vecinos, por lo que en ambas guras se ilustran los tres esquemas de vecindad. El esquema a) (esquinero) tiene tres vecinos, que por asignación igual se tiene que , de modo que , de modo que . Para el esquema b) (lateral) se tienen 5 vecinos, de modo que , finalmente para el esquema c) (interior), se tienen 8 vecinos, de modo que . Para el caso de la Figura 1‑2, los pesos se construyeron a partir de las distancias entre plantas (30, 40 y 50 cm) y entre surcos (100 cm). Para el caso esquinero, se estableció para 30 cm, para 100 cm y para los cm, de esta manera, y , por lo que , de donde se obtiene que , y . Con un procedimiento análogo se establecieron los pesos de , y para el esquema b) y c). La Tabla 1‑1 presenta los pesos de los tres esquemas para las distancias entre surcos de 30, 40 y 50 cm de la Figura 1‑2, manteniendo los 100 cm entre surcos. En caso de faltantes (o por irregularidad en tamaño de surcos) en este patrón, pueden generarse una cantidad de vecinos

Tabla 1‑1. Distribución de los pesos por distancia entre surcos del patrón de asignación por distancias inversas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Densidad** |  |  |  | **Total** |
| d1:(30\*100) | 0,630 | 0,189 | 0,181 | 1 |
| d2:(40\*100) | 0,128 | 0,318 | 0,118 | 1 |
| d3:(50\*100) | 0,128 | 0,257 | 0,058 | 1 |

Figura 1‑1. Patrón de vecindad intra-hilera en vecinos más cercanos

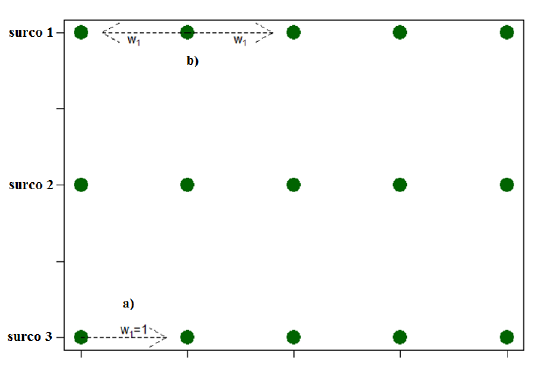
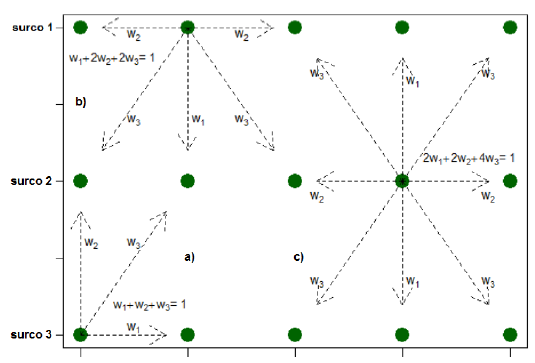
****

Figura 1‑2. Patrón de vecindad inter-intra-hileras en vecinos más cercanos.



## Resultados y Discusión

El tamaño de los tubérculos parece estar regulado por diversos mecanismos y autores como Struik y colaboradores (1990) consideraron la densidad de siembra, el número de tallos por plantas y el número de tubérculos por tallo como variables de importancia en la manipulación y modelado del rendimiento del cultivo. En este orden de ideas, los resultados que se muestran a continuación se asocian al modelado del conteo de tubérculos de papa criolla para tres diferentes densidades de siembra y para cuatro tamaños de tubérculos, así como el modelado del rendimiento por estos mismos factores, pero considerando la competición entre vecinos más cercanos.

El diagrama de la gura 3 muestra la estrecha relación encontrada entre el peso fresco de los tubérculos (g) contra el diámetro promedio ponderado obtenido de multiplicar cada marca de clase de los calibres usualmente establecidos para tamaño de tubérculos con el respectivo conteo de tubérculos por densidad. Morales y colaboradores (2001) han evaluado otras variedades de papa y han encontrado relación entre el número de tubérculos generados y el diámetro medio de los tubérculos.

**(1.1)**

Siendo la matriz de dimensión , con como el total de conteos para cada planta en los 4 calibres, **c** como el vector de marcas de clase de cada calibre representado en este caso por **c** = (1, 3, 5, 9), **c'** representando al vector traspuesto de **c** y como un vector de unos con filas. Con esta simple expresión se obtiene el diámetro medio ponderado como un vector de n las correspondientes a todas las densidades y representando a todos los calibres como un promedio ponderado. La Figura 1‑3 representa las reacciones lineales por encontradas para la partición del con la respectiva partición del vector de peso fresco estimado . En la figura se representa la línea de regresión lineal simple ajustada por mínimos cuadrados y el coeficiente de correlación obtenido en cada partición, todas resultando asociaciones lineales evidentes.

La Tabla 1‑2 representa el conteo de tubérculos por calibre y el conteo total para cada densidad de siembra, observándose el mayor conteo en las distancias entre plantas de 30 y 40 cm, sin embargo, los porcentajes de tubérculos condicionados por densidad son similares en las tres densidades, lo que a primera vista sugiere que no existen diferencias de importancia en el número de tubérculos obtenidos por densidad.

La Tabla 1‑3 describe los promedios de peso fresco estimados por ponderación para cada calibre, utilizando como ponderadores el número de tubérculos y cada marca de clase de los calibres considerados. Se propone la siguiente expresión matricial para calcular la matriz de pesos ponderados estimados por calibre:

**(1.2)**

Siendo diag(**DMP**)−1 la matriz inversa de los diámetros medios ponderados una vez diagonalizados con el operador diag, el mismo que se aplica a las matriz de pesos frescos para obtener diag(**P**); [diag(**F**c)] es una de las matrices ya definida en (1) y finalmente (diag(**c**’**JJ**’))−1 que contiene vectores ya descritos anteriormente. La matriz **Pe** contiene los pesos estimados de los cuatro calibres considerados. Las inversas están aseguradas pues son matrices con solo elementos en la diagonal, por lo que la inversa se reduce al cálculo del inverso de cada elemento en la diagonal, que son considerados todos no nulos, pues si alguna planta no rinde tubérculos de ninguno de los cuatro calibres, se descarta como dato de interés al modelar conteos.

Finalmente, los pesos que se relacionan la Figura 1‑3 se corresponden con los pesos estimados Pe particionados por densidad de siembra. Los resultados se pueden corroborar simplemente multiplicando **Pe** y **J**, lo que rinde al vector **P** original de pesos frescos.

La Figura 1‑4 ilustra el comportamiento de los conteos en los cuatro calibres sin discriminar por densidad. Sobre cada histograma se superpuso la distribución de Poisson que se espera para el promedio de conteos en cada calibre. Evidentemente los conteos se alejan de la distribución de Poisson esperada. En los anexos se ilustra de forma complementaria, los ajustes de la Binomial negativa para que el lector note claramente que esta última distribución posee mejor ajuste que el modelo inicial considerado

Tabla 1‑2. Distribución del conteo de tubérculos por densidad y marca de clase del calibre

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Densidad** | **Calibre ≤ 2** | **Calibre (2 - 4]** | **Calibre (4 - 6]** | **Calibre > 6** | **Subtotal** |
| d1:(30\*100) | 8258(39,5 %) | 8337(39,9 %) | 4152(19,9 %) | 159(0,7%) | 20906 |
| d2:(40\*100) | 8088(38,3 %) | 8538(40,5 %) | 4254(20,2 %) | 215(1,0%) | 21095 |
| d3:(50\*100) | 7291(38,4 %) | 7064(37,2 %) | 4406(23,2 %) | 220(1,2%) | 18981 |
| Subtotal | 23637 | 23939 | 12812 | 594 | total(60982) |

Tabla 1‑3. Distribución del Peso fresco promedio de tubérculos por densidad y marca de clase del calibre

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Densidad** | **Peso ≤2** | **Peso (2 - 4]** | **Peso (4 - 6]** | **Peso > 6** | **Peso general** |
| d1:(30\*100) | 58.61 | 179.55 | 156.22 | 13.09 | 407.47 |
| d2:(40\*100) | 84.57 | 264.95 | 237.19 | 21.10 | 607.82 |
| d3:(50\*100) | 84.42 | 248.60 | 268.18 | 25.00 | 626.20 |

La Tabla 1‑4 presenta los promedios y sus respectivas varianzas de conteos junto al cociente de la varianza y la media para describir la relación que existe entre estas dos estadísticas. En los modelos de conteo, esta relación es relevante para la elección del modelo de probabilidad a elegir para el modelado de los conteos. Los cocientes que se observan en la tabla sugieren que no sería apropiada una elección del modelo Poisson para el ajuste de los datos. En el Anexo A se ilustra la distribución de los conteos utilizando el modelo Poisson y el binomial negativo, y sin duda, el binomial negativo parece ajustar adecuadamente desde un punto de vista descriptivo, los datos de conteos para los diferentes calibres considerados.

**Figura 1‑3.** Diagrama de dispersión del peso fresco y Diámetro medio ponderado de los tubérculos por densidad de siembra.

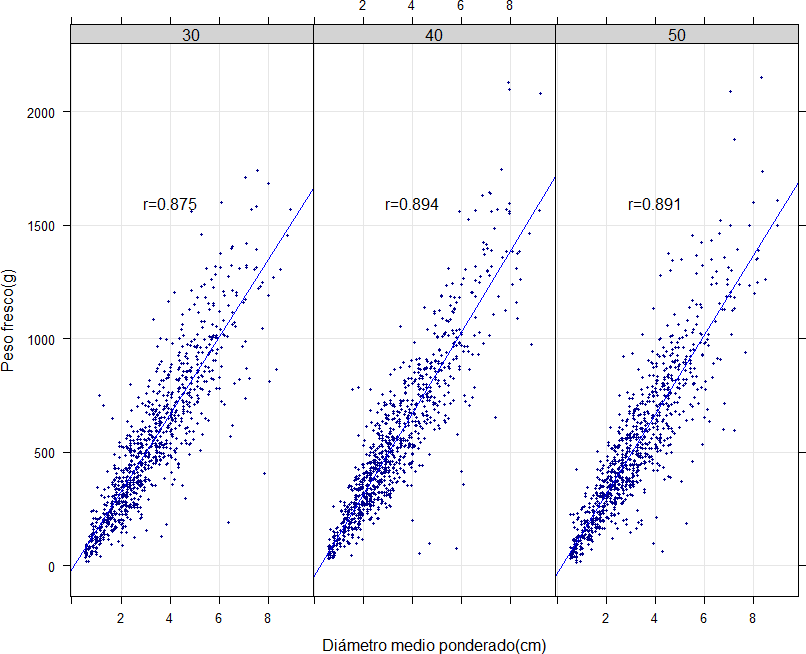
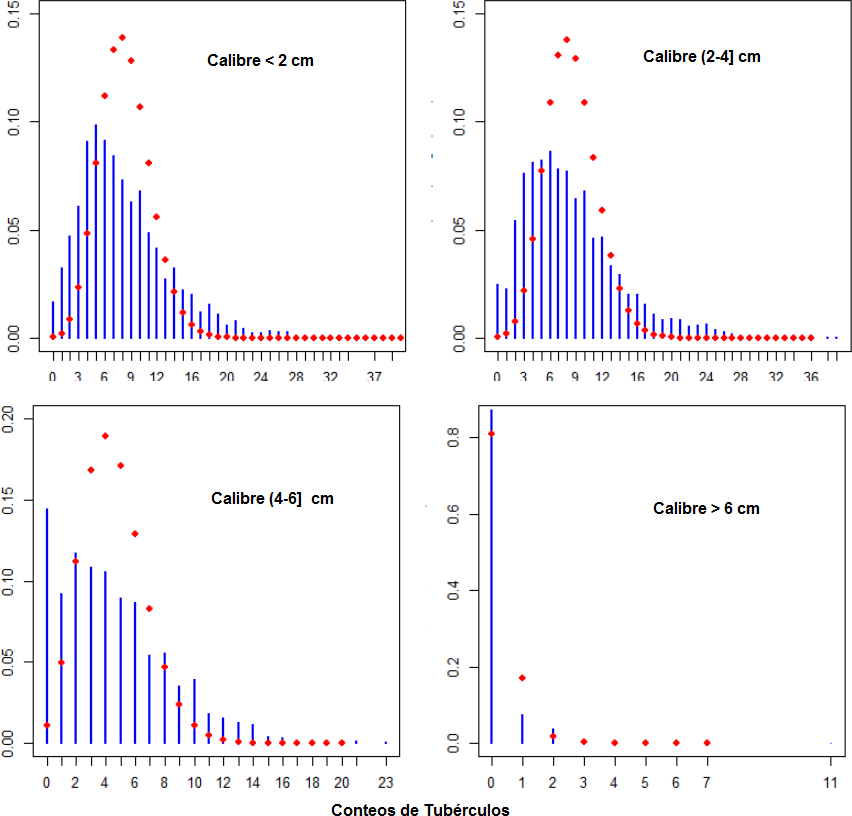


Figura 1‑4. Distribución de conteos de tubérculos por calibre con ajuste de Poisson



En la Figura 1‑5 se ilustra nuevamente el comportamiento de los conteos pero ahora discriminados por densidad, para evidenciar los excesos de ceros presentes en los dos calibres superiores para todas las densidades de siembra consideradas.

Tabla 1‑4. Media, Varianza y cociente varianza/media por densidad y calibre

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Calibre(cm)** | **Densidad: (30\*100)** | | | **Densidad: (40\*100)** | | | **Densidad:(50\*100)** | | |
| **Media** | **Varianza** | **Cociente** | **Media** | **Varianza** | **Cociente** | **Media** | **Varianza** | **Cociente** |
| ≤ 2 | 7,28 | 22,96 | 3,2 | 8,73 | 27,42 | 3,1 | 9,37 | 38,34 | 4,1 |
| (2 - 4] | 7,35 | 27,33 | 3,7 | 9,22 | 32,69 | 3,5 | 9,08 | 32,64 | 3,6 |
| (4 - 6] | 3,66 | 11,74 | 3,2 | 4,59 | 12,86 | 2,8 | 5,66 | 16,01 | 2,8 |
| > 6 | 0,14 | 0,21 | 1,5 | 0,23 | 0,59 | 2,5 | 0,28 | 0,67 | 2,4 |

El modelado del conteo ha tenido amplia aplicación en el campo de la entomología, pues son diversos los estudios que cuentan el número de insectos para generar medidas de agregación que pudieran relacionarse con umbrales para el manejo del cultivo. Varios estudios incluyen las dos distribuciones más usadas para este tipo de datos, a saber, la Poisson y la Binomial negativa (Cadahia, 1997; Vivas y Notz, 2011). En otras ocasiones, la excesiva aparición de ceros di culta el uso de estos modelos de probabilidad por lo que se recurre a las opciones cero-infladas o cero-truncadas para tratar con esta situación (Henne, 2012). Precisamente cuando se estudia el calibre en el cultivo de papa como indicador del rendimiento, la naturaleza del dato obliga al uso de las distribuciones más comunes en datos por conteo, a saber, la Poisson y La Binomial negativa y sus modalidades cero infladas o cero truncada, por esta razón, a continuación se ilustran resultados obtenidos por algunos modelos de prueba y los modelos finalmente propuestos para relacionar el conteo de tubérculos a factores como el calibre y la densidad de siembra.

*Regresión de Poisson* El primer modelo ajustado fue el de regresión de Poisson usando como respuesta los conteos de tubérculos para el calibre hasta 2 cm y como predictores el diámetro ponderado medio y el peso fresco generado para este calibre. Los resultados del mejor modelo ajustado para este calibre usando la distribución de Poisson se muestran en la siguiente tabla:

Tabla 1‑5. Modelado por Regresión de Poisson del Calibre hasta 2 cm

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Coeficientes** | **Estimación** | **Error Estándar** | **valor z** | **prob.** |
| Intercepto | 1,398 | 0,019 | 72,33 | 2,00e-16 |
| densidad | -0,014 | 0,008 | -1,65 | 0.1 |
| DMP | 0,036 | 0,004 | 8,21 | 2,00e-16 |
| peso fresco 1 | 0,007 | 1,08e-4 | 65,29 | 2,00e-16 |

Tabla 1‑6. Estadísticas de ajuste modelo 1

|  |  |
| --- | --- |
| **Estadístico(s)** | **Valor(es)** |
| Devianza Nula(gl) | 9556,0(2838) |
| Devianza Residual(gl) | 3204,1(2835) |
| AIC | 13811 |
| *χ*2 (Pearson) | 3005,3 |
| Dispersión | 1,06 |

Aunque resultó ser el mejor modelo, es relevante destacar que el diámetro medio ponderado (DMP) se construyó a partir de los conteos de tubérculos por calibre, al igual que el peso fresco estimado para este calibre, el cual se generó a partir de (1), por eso resulta altamente significativo en el modelo tanto el DMP como el PD1 (peso fresco estimado del primer calibre), en este sentido, el resumen del nuevo modelo (Tabla 1‑7 y Tabla 1‑8) reducido por el DMP y PD1 queda como :

Figura 1‑5. Distribución de conteos de tubérculos por densidad de siembra y calibre (D1, D3, D5 y D9)

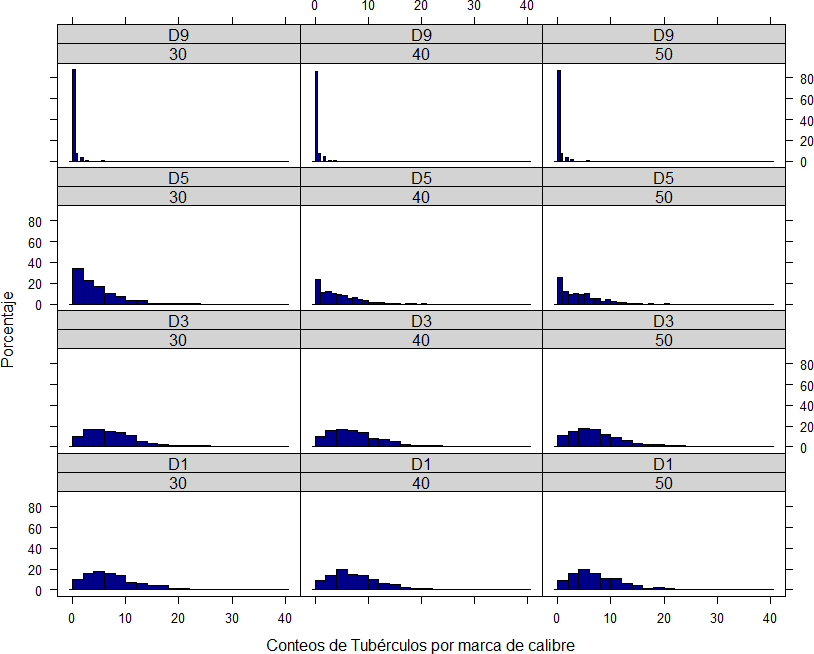


Tabla 1‑7. Modelado por Regresión de Poisson del Calibre hasta 2 cm

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Coeficientes** | **Estimación** | **Error Estándar** | **valor z** | **prob.** |
| Intercepto | 1,875 | 0,017 | 111,00 | 2,00e-16 |
| Densidad | 0,128 | 0,008 | 16,10 | 2,00e-16 |

La dispersión del modelo con los predictores DMP y PD1 parecía mostrar una dispersión aceptable, sin embargo, la realidad es que 3,42 representa más los que se espera, pues como se observó en la tabla de cocientes de varianzas contra medias, este cociente fue generalmente superior a 3, indicando sobredispersión, lo que obliga al uso de otro tipo de distribución, por lo que hasta ahora no es prudente concluir la presencia de un efecto altamente significativo asociado a la densidad de siembra.

El tercer modelo involucra al calibre y a la densidad de siembra como predictores numéricos. La salida correspondiente utilizando el mismo modelo Poisson se muestra en la Tabla 1‑8 y Tabla 1‑9, además se presentan otras salidas para justificar el modelado con otra distribución y ajuste por sobredispersiòn. Este tercer modelo permite realizar modelos por separado para cada calibre (Fowler, 1991).

Tabla 1‑8. Estadísticas de ajuste modelo 2

|  |  |
| --- | --- |
| **Estadístico(s)** | **Valor(es)** |
| Devianza Nula(gl) | 9556,0(2838) |
| Devianza Residual(gl) | 9299,2(2837) |
| AIC | 13811 |
| *χ*2 (Pearson) | 9693,6 |
| Dispersión | 3,42 |

Tabla 1‑9. Modelado por Regresión de Poisson para calibre y densidad

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Coeficientes** | **Estimación** | **Error Estándar** | **valor z** | **prob.** |
| Intercepto | 2,190 | 0,032 | 67,94 | 2,00e-16 |
| densidad | 0,011 | 0,001 | -13,76 | 2,00e-16 |
| calibre | -0,315 | 0,009 | -33,51 | 2,00e-16 |
| densidad: calibre | -0,001 | 2,00e-4 | 5,16 | 2,50e-7 |

Tabla 1‑10. Estadísticas de ajuste modelo 3

|  |  |
| --- | --- |
| **Estadístico(s)** | **Valor(es)** |
| Devianza Nula(gl) | 65767,0(11355) |
| Devianza Residual(gl) | 38753,0(11352) |
| AIC | 68806 |
| *χ*2 (Pearson) | 40268,0 |
| Dispersión | 3,55 |

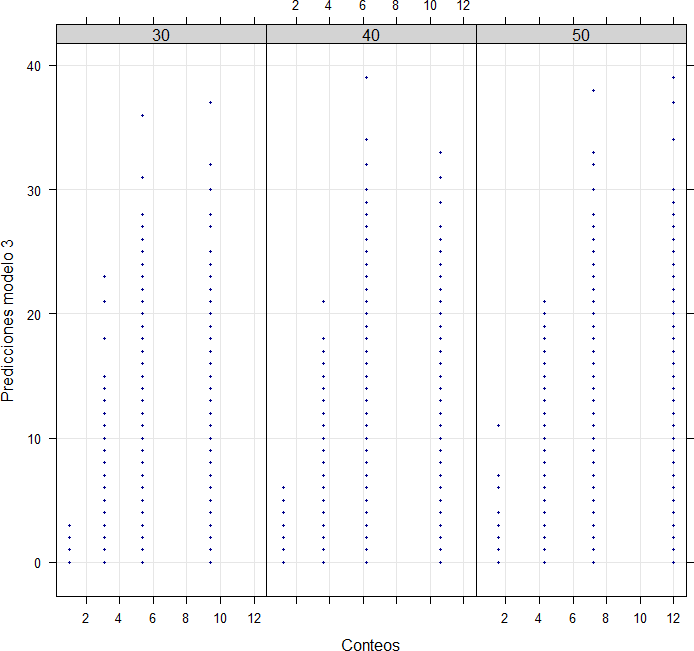
El modelo 3 aunque reduce el número de modelos requeridos para cada calibre, sigue mostrando evidente sobredispersiòn. El modelo estimado se escribe como:

**(1.3)**

Donde es la media de conteos para la i-ésima densidad y j-ésimo calibre con i = 1, 2, 3 y j = 1, 2, 3, 4. El término representa la interacción entre densidad y calibre.

La Figura 1‑6 muestra claramente lo inapropiado del modelo ya que las predicciones son bastante dispersas, es decir, se espera que un buen modelo prediga altos valores donde los conteos observados hayan sido altos, y bajos donde hayan sido bajos, sin embargo, se nota como en conteos observados bajos (de 2 a 4) se pueden encontrar valores estimados de más de 30 conteos, sobre todo en las densidad de 30 cm entre plantas y 100 cm entre surcos. Finalmente, aunque los predictores son significativos, el modelo no encaja bien. Dado el valor de la estadística de deviancia residual de 38753 es mucho mayor que los grados de libertad asociados (11355), además, el estadístico de dispersión es mayor a 1.

Figura 1‑6. Diagrama de dispersión de valores observados y estimados para el modelo 3.



En vista de estos resultados ilustrativos previos (algo que usualmente debe hacerse pues es clave entender la razón de selección de un modelo diferente al usual, a saber, el modelo Poisson) es importante pasar al modelo de regresión Binomial negativo, el que mediante los gráficos antes mostrados, parece ser el adecuado en el caso del modelado de los calibres para cada densidad. Se hará un trabajo por calibres pues es evidente la presencia de excesivos ceros en los mayores calibres, por lo que desde ahora, se probarán modelos Binomial negativo y Binomial negativo cero inflados para el modelado final de los conteos de tubérculos.

*Regresión Binomial Negativa*. En el modelo binomial negativo el parámetro asociado a la media se entiende de la misma manera que el modelo de Poisson, sin embargo la varianza tiene un alcance mucho más amplio que el permitido por la distribución de Poisson, de hecho la media usa el mismo parámetro que el caso Poisson, pero la varianza se escribe como , la cual es conocida como binomial negativa del tipo 2 (NB2) por el exponente cuadrático en el segundo término, pero pudiera usarse el tipo NB1 , el cual tiene una varianza . Como puede verse, ahora se tiene un modelo de dos parámetros, donde la varianza no está restringida a ser igual a la media, lo que facilita el modelado, tal como en los datos que estamos evaluando, donde la varianza está resultado un poco más del triple que la media. La binomial negativa como la Poisson permite manipular excesos de ceros tal como se encuentran en los diámetros o calibres superiores, por lo que esta distribución está bastante justificada para el modelado de los conteos de tubérculos por densidad.

Figura 1‑7. Regresión Binomial Negativa para densidad en el calibre ≤ (2cm)

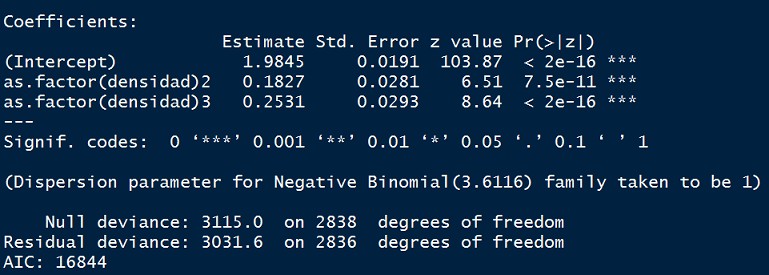
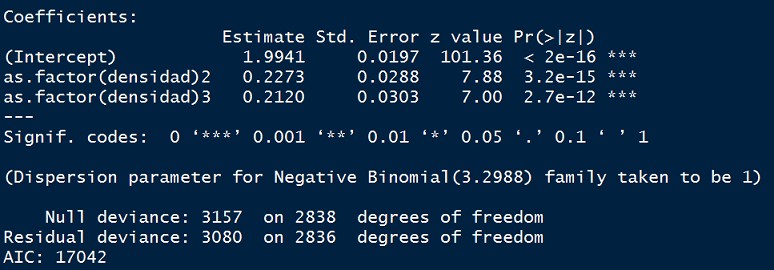
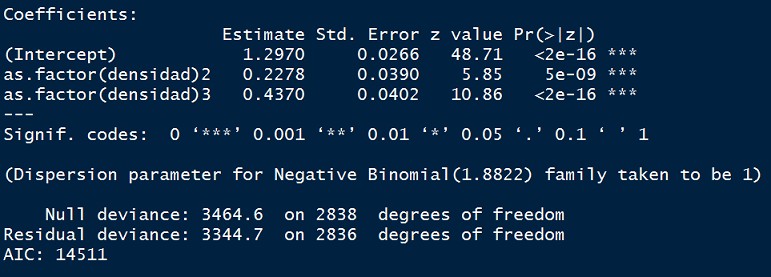


Figura 1‑8. Regresión Binomial Negativa para densidad en el calibre (2 − 4] cm



En el anexo C aparecen resultados del ajuste de modelos binomial negativo, los solo dieron un mejor ajuste para el calibre de (4 6] cm, aunque sí resultaron superiores al modelado de Poisson, por esta razón, finalmente se contrastan los resultados finales con la opción cero-inflada del modelo binomial negativo.

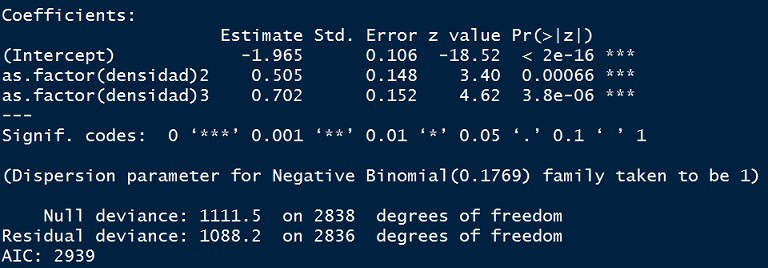
Figura 1‑9. Regresión Binomial Negativa para densidad en el calibre (4 − 6] cm



En los anexos aparecen las estadísticas de ajuste de estos modelos, de los cuales, solo se mantuvo el modelo para el calibre de (2 4] cm en comparación con sus respectivas modalidades cero-infladas. A continuación se presentan las salidas de los modelos binomiales negativos inflados por ceros.

*Modelos de Regresión Cero-inflados* Los modelos de conteo cero-inflados proporcionan una forma alternativa de modelar los datos de conteo una frecuencia de ceros en exceso. En el estudio de varias áreas de la agronomía y ciencias afines es usual encontrar conteos en los que aparecen ceros no como consecuencia de observaciones faltantes, sino por la propia naturaleza de la variable. En muchos casos se pueden encontrar tantos ceros que no es lo que se esperaría en una distribución de Poisson o una binomial negativa, lo que seguramente causará problemas en el proceso de modelado, ya que las estimaciones de los coeficientes asociados a los parámetros pudieran no ser confiables y puede aparecer la sobredispersión que antes se ha mencionado. En principio, la exploración de los datos ya sea mediante histogramas u otras representaciones visuales pueden dar luz sobre la excesiva aparición de ceros, lo que mediante cálculos probabilísticos asociados al punto cero tanto en la distribución de Poisson como la Binomial Negativa, se puede verificar la aparición de muchos más ceros de los que se espera. Las guras 4 y 5 ponen en evidencia el exceso de ceros, especialmente en los calibres superiores, sin embargo, en los calibres inferiores, aunque los ceros no son la clase en mayor frecuencia, el número de ceros aún fue superior a lo esperado para las dos distribuciones comunes en el modelado de conteos, por lo que resultaron como mejores modelos las opciones cero infladas, sobre todo en la distribución binomial negativa.

Figura 1‑10. Regresión Binomial Negativa para densidad en el calibre > 6cm



El primer modelo evaluado fue el Poisson cero-inflado, con el n de contrastarlo con la opción usual presentada en la Tabla 1‑6 a la Tabla 1‑10, que sirvieron como soporte explicativo para el desarrollo de los modelos finales en la opción cero-inflada. Para iniciar el modelado se procedió a convertir en factor la variable densidad de siembra con los niveles (20,30 y 40 cm) asociados a la distancia entre plantas. En el caso de conteo de tubérculos también surgió el dilema de los falsos ceros y ceros verdaderos, pues una planta sin tubérculos de un diámetro pudo ser el producto de alguna enfermedad, ataque de insectos o de ciencia y no del propio comportamiento fisiológico asociado a la producción, así que se asumió que los ceros presentados para un diámetro específico eran reales cuando en al menos uno de los diámetros apareció al menos un tubérculo, de lo contrario se consideró como cero falso y se correspondió con plantas sin producción y se descartaron del modelado de conteos y no del modelado espacial de la producción para evaluar competición.

A continuación se describen los resultados obtenidos mediante el software R para la construcción de las salidas asociadas al proceso de estimación y ajuste de los modelos Binomial Negativo cero-inflados utilizando como factor la densidad, modelando la respuesta por calibre.

Figura 1‑11. Regresión Binomial Negativa cero-inflada para densidad en el calibre ≤ (2cm)

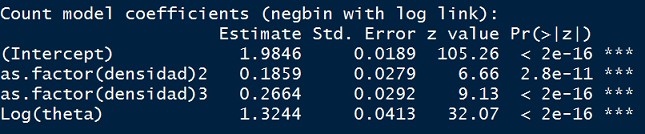
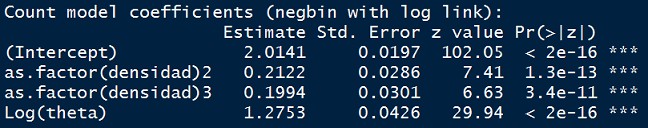


Figura 1‑12. Regresión Binomial Negativa cero-inflada para densidad en el calibre (2 − 4] cm



La aplicación del test de Vuong permitió la comparación entre los moles binomial negativo con su contraparte cero-inflada. Este estadístico tiene distribución Normal (0,1). Cuando la significación es menor al 5 % es cuando podemos concluir estadísticamente que la opción cero-inflada es superior (los valores corregidos son superiores en la conclusión pues el test es sesgado hacia favorecer la opción cero-inflada). Los resultados de las figuras 15 y 16 muestran la superioridad de las opciones cero-infladas en los calibres inferiores a 2 cm y mayores a 4 cm, solo quedó el modelo binomial negativo para modelar el calibre de (2-4] cm.

Figura 1‑13. Regresión Binomial Negativa cero-inflada para densidad en el calibre (4 − 6] cm

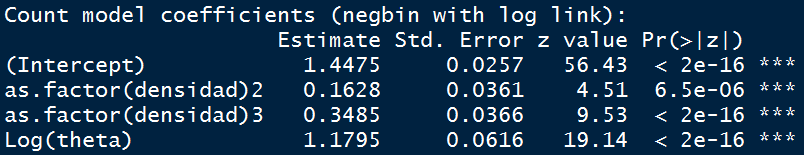


Figura 1‑14. Regresión Binomial Negativa cero-inflada para densidad en el calibre > 6 cm

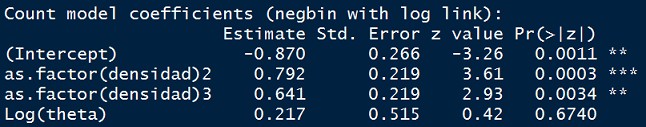
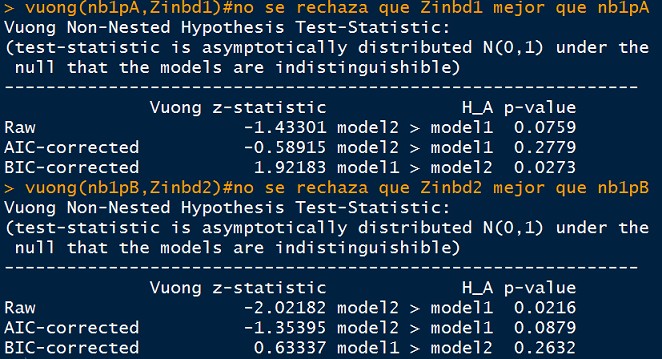
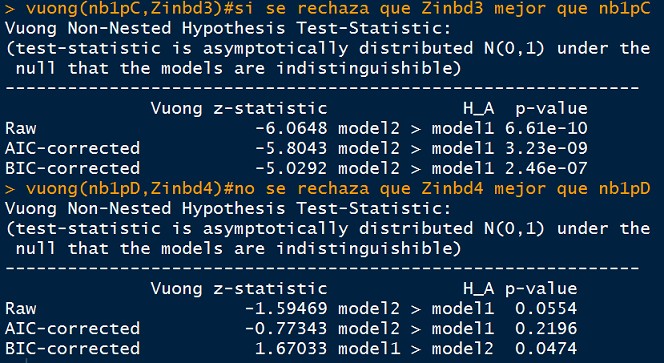


Figura 1‑15. Comparación del modelo binomial negativo con su contraparte cero-inflada (calibres ≤ (2) y (2 − 4] cm)



La Figura 1‑17 y Figura 1‑18 muestran gráficamente los valores observados y estimados por los modelos cero-inflados. Una mirada exhaustiva permite gráficamente notar que el modelo para el calibre de (2-4] muestra las mayores diferencias entre observados y esperados. Todos los resultados anteriores mostraron como el factor densidad explica los conteos para todos los calibres evaluados, utilizando distribuciones especiales para conteos en lugar de tratarlos con análisis de varianza clásico, donde sobre todo para los modelos con exceso de cero, era improbable el cumplimiento de supuestos.

Figura 1‑16. Comparación del modelo binomial negativo con su contraparte cero-inflada (calibres (4 − 6] y > 6cm)



*Modelos Autoregresivos Espaciales (Competición intra-hilera e inter-intra-hilera).* El modelado espacial en la actualidad ha ganado importancia en el estudio de los cultivos, pues es reconocida la dependencia espacial que tienen los atributos asociados al rendimiento de las propiedades edafo – climátiacas. La distribución espacial del cultivo y su manejo por densidad son aspecto de interés en el modelado del desarrollo del cultivo. (Zheng et al, 2016). En la actual investigación se estudió el rendimiento en peso fresco (g) como la variable respuesta y la densidad y calibre se usaron como predictores, sin embargo, como el interés radicó en estimar el coeficiente de competición de un modelo autoregresivo puro, se propusieron modelos para cada combinación der calibre y densidad en cada repetición para estimar los cambios en el coeficiente de competición tanto en un patrón de vecindad intra-hilera como en uno inter-intra-hilera. Para el modelado espacial se inició con los modelos autoregresivos espaciales puros usando como patrón de vecindad el intrahilera, es decir, para una distancia entre plantas de 30 cm se usó como máxima distancia de búsqueda la correspondiente a la mínima distancia entre plantas, es decir, los mismos 30 cm, esto aseguró uno (para esquinas) o dos (interiores) vecinos para cada planta. Como la siembra es medianamente regular, pues no todos los surcos tuvieron el mismo tamaño, se probó una distancia máxima de búsqueda para el caso intra – hilera correspondiente a la mínima distancia entre plantas, lo que en una rejilla regular generaría uno o dos vecinos, sin embargo, en los resultados veremos más vecinos debido a la desigualdad en el tamaño de los surcos. Para el caso de la competición inter – intra – hileras se usó un criterio similar pero usando una distancia de búsqueda que una planta de un surco considerara como vecino más cercano alguna planta del surco inmediatamente cercano. Aprovechando que la siembra arrancó alineada en uno de los bordes del lote, la desigualdad se generó al final de cada surco, por lo que usando este borde uniforme para estudiar los patrones de vecindad descritos en las guras 1 y 2, las distancias que abarcan un surco a otro variaron según la densidad de siembra, y fueron como máximo cm (aproximadamente 104,5 cm para incorporar la diagonal) para la distancia entre plantas de 30 cm, cm (aproximadamente 107,8 cm ) para la distancia entre plantas de 40 cm y cm (aproximadamente 111,9 cm) para la distancia entre plantas de 50 cm. En un patrón regular se tendrían de tres (esquineros) a 8 (interiores) vecinos más cercanos, sin embargo, los tamaños diferentes de surcos generan otros vecinos en la parte final de los surcos. Este resultado se presenta en las tablas resumen del modelado de la competición.

Figura 1‑17. Comparación de valores observados (Obs.) y predichos (Pred.) (Calibres ≤ (2) y (2 − 4] cm)

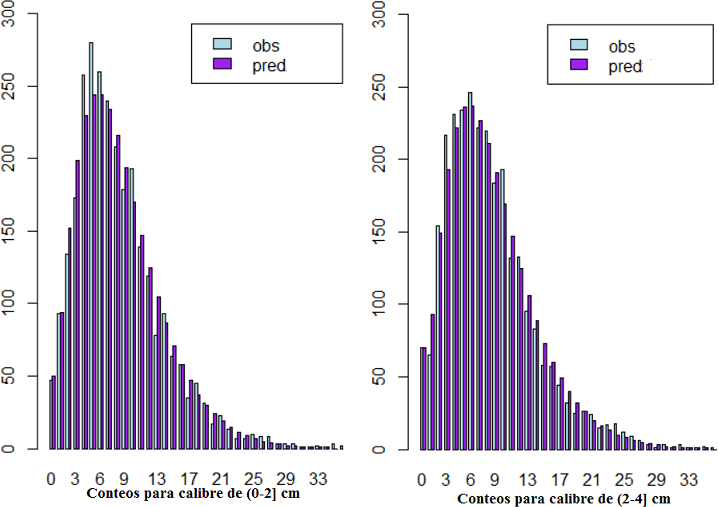
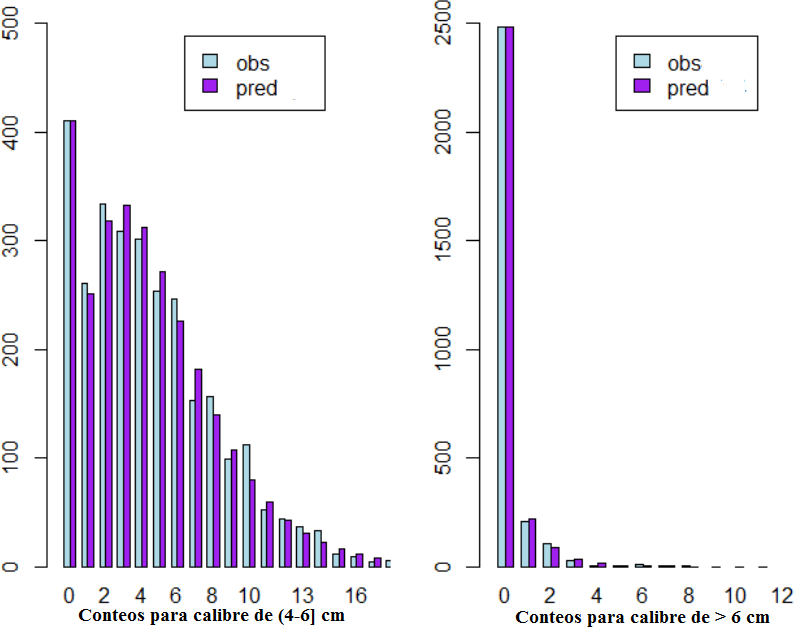


Figura 1‑18. Comparación de valores observados (Obs.) y predichos (Pred.) (Calibres (4 − 6] y > 6cm)



La Tabla 1‑11 muestra el resumen de la información del proceso de modelado espacial mediante un modelo autoregresivo puro para las tres densidades de siembra, en las respuestas peso fresco (pf) de las dos repeticiones realizadas . Para la competición intra – hileras siempre se tuvieron de uno (esquinas) a dos vecinos (interiores), con una matriz de pesos con un porcentaje de celdas no nulas (PNN) que fue aumentando a medida que aumentó la distancia entre plantas, pues se generaron menos conexiones, de hecho, el promedio de conexiones (PC) disminuyó. En la correspondiente a la dimensión de la matriz de pesos (W), fue evidente la irregularidad en el tamaño de los surcos, y como se puede notar, a menos distancia entre plantas se utilizaron más plantas en las hileras. En lo que respecta al modelo econométrico espacial autoregresivo puro, las seis corridas mostraron la significación del parámetro del intercepto y del parámetro de competición estimado (ˆ), y en todas las modalidades. Para las dos menores distancias de siembra se vio como a mayor distancia menor la competición, sin embargo, los resultados para la mayor distancia de siembra se asemejaron a la menor distancia, probablemente debido a la mayor variabilidad de los datos en estas mediciones del peso fresco (coeficiente de variación superiores a las dos menores densidades), por lo que resta probar no solo el peso fresco global sino el obtenido para cada densidad de siembra de modo que se pueda reducir la variabilidad en los datos. A pesar de este resultado en la mayor distancia, lo relevante fue que el coeficiente de competición resultó significativo, lo que sugiere una relación espacial entre rendimiento y densidad. Un resultado similar pero mediante otra metodología fue encontrado por Zheng et al, (2016), donde evidenciaron la relación entre densidad y rendimiento y hasta con el número de tubérculos encontrados.

Tabla 1‑11. Resultados de la Competición intra – hilera (vecino más cercano)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Resultado (cm)** | **Densidad: (30\*100)** | | **Densidad: (40\*100)** | | **Densidad: (50\*100)** | |
| **pf(r1)** | **pf(r2)** | **pf(r1)** | **pf(r2)** | **pf(r1)** | **pf(r2)** |
| Vecinos | 1-2 | 1-2 | 1-2 | 1-2 | 1-2 | 1-2 |
| PNN | 0,352 | 0,347 | 0,435 | 0,421 | 0,499 | 0,518 |
| PC | 1,99 | 1,99 | 1,99 | 1,99 | 1,98 | 1,98 |
| Dimensión(W) | 565 | 574 | 457 | 472 | 398 | 383 |
| Intercepto(p) | 2,2e-16 | 2,2e-16 | 2,2e-16 | 2,2e-16 | 2,2e-16 | 2,2e-16 |
| *λ*ˆ | 0,225 | 0,207 | 0,189 | 0,180 | 0,244 | 0,229 |

**Tabla 1-11.** Continuación

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Resultado (cm)** | **Densidad: (30\*100)** | | **Densidad: (40\*100)** | | **Densidad: (50\*100)** | |
| **pf(r1)** | **pf(r2)** | **pf(r1)** | **pf(r2)** | **pf(r1)** | **pf(r2)** |
| *λ*ˆ(*p*) | 1,1e-8 | 1,8e-7 | 2,6e-5 | 3,7e-5 | 1,1e-7 | 1,2e-6 |
| AIC | 7893,3 | 7882,5 | 6492,5 | 6920,7 | 5772,6 | 5613,3 |
| Var. Res. | 65901,0 | 52175,0 | 83945,0 | 132640,0 | 111340,0 | 130060,0 |

Tabla 1‑12. Resultados de la Competición inter-intra-hilera (vecino más cercano)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Resultado (cm)** | **Densidad: (30\*100)** | | **Densidad: (40\*100)** | | **Densidad: (50\*100)** | |
| **pf(r1)** | **pf(r2)** | **pf(r1)** | **pf(r2)** | **pf(r1)** | **pf(r2)** |
| Vecinos | 5-12 | 3-12 | 2-10 | 2-10 | 2-6 | 2-6 |
| PNN | 1,756 | 1,725 | 1,732 | 1,681 | 1,326 | 1,377 |
| PC | 9,92 | 9,90 | 7,92 | 7,93 | 5,28 | 5,27 |
| Dimensión(W) | 565 | 574 | 457 | 472 | 398 | 383 |
| Intercepto(p) | 2,2e-16 | 2,2e-16 | 2,2e-16 | 2,2e-16 | 2,2e-16 | 2,2e-16 |
| *λ*ˆ | 0,482 | 0,398 | 0,437 | 0,407 | 0,431 | 0,429 |
| *λ*ˆ(*p*) | 2,2e-16 | 5,8e-10 | 2,0e-11 | 2,5e-10 | 6,6e-13 | 4,7e-12 |
| AIC | 7858,9 | 7871,4 | 6463,9 | 6897,7 | 5749,2 | 5589,1 |
| Var. Res. | 61265,0 | 51020,0 | 77497,0 | 124650,0 | 103410,0 | 119900,0 |

Para el caso de la competición inter – intra – hilera, los resultados son un poco más coherentes en lo que respecta al coeficiente de competición estimado, pues promediando las dos repeticiones para cada densidad, el coeficiente va de 0,44 para la menor distancia de siembra, 0,42 para la distancia intermedia y 0,43 para la mayor distancia, y aunque todos resultaron significativos, tienden a ser más similares, lo que seguramente podría mejorar si se modela el peso fresco de cada densidad por aparte o si se estudian modelos no puros donde se incorpore el calibre como variable explicativa. El coeficiente de variación del peso fresco de la primera densidad fue de 62,2 %, a los 40 cm de distancia entre plantas pasó a 56,1 %, pero volvió a aumentar en la distancia de 50 cm, lo que puede estar asociado al aumento del coeficiente de competición para la mayor separación entre plantas, por lo que una realización de pruebas usando el peso estimado propuesto en (2) para cada calibre podría mejorar la estimación de λ. Los modelos de competición intra – inter – hilera tuvieron los más bajos criterios de información de Akaike y menor varianza de los residuos. La Figura 1‑19 y Figura 1‑20 describen la relación entre los valores observados y estimados de los modelos intrahilera y la Figura 1‑21 y Figura 1‑22 son las correspondientes a los modelos de competición intra – inter – hileras.

Sin duda, con los resultados obtenidos de la última sección se corrobora la importancia de involucrar metodologías de análisis espacial en el estudio del cultivo de papa, sobre todo cuando se estudian densidades de siembra, y aunque los autores describen en detalle el problema de la escala en el modelado, esto puede asociarse de manera análoga a la forma en que se analizan los datos, pues el estudio del peso fresco total de algún modo implica una escala diferente a lo que sucedería si se hace un modelado a cada calibre, pues esto puede disminuir la variabilidad encontrada en los datos globales (Chen et al, 2017).

Figura 1‑19. Comparación de valores observados y predichos (Densidad (30cm\*100cm)- Competición intra – hilera)

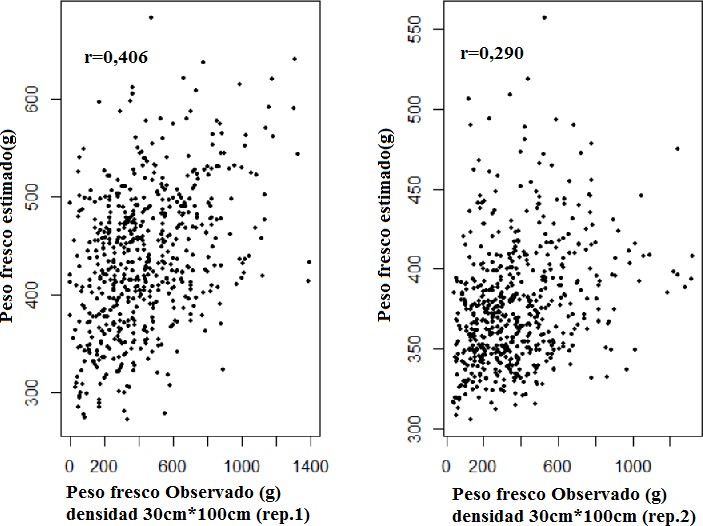


Figura 1‑20. Comparación de valores observados y predichos (Densidad (40cm\*100cm)- Competición intra – hilera)

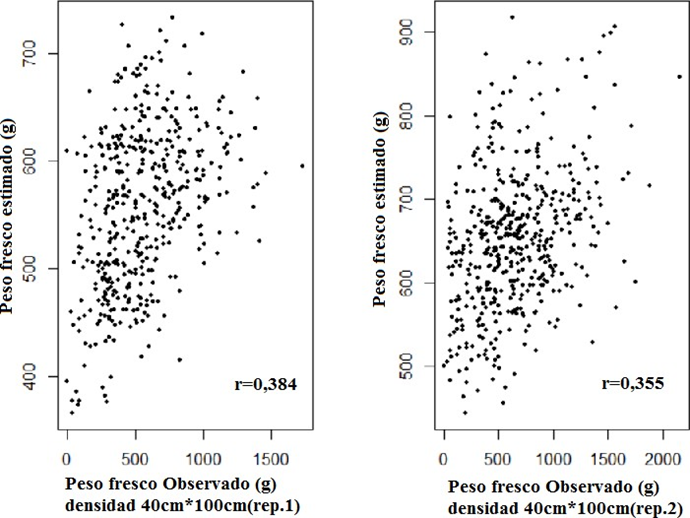


Figura 1‑21. Comparación de valores observados y predichos (Densidad (50cm\*100cm)- Competición intra – hilera)

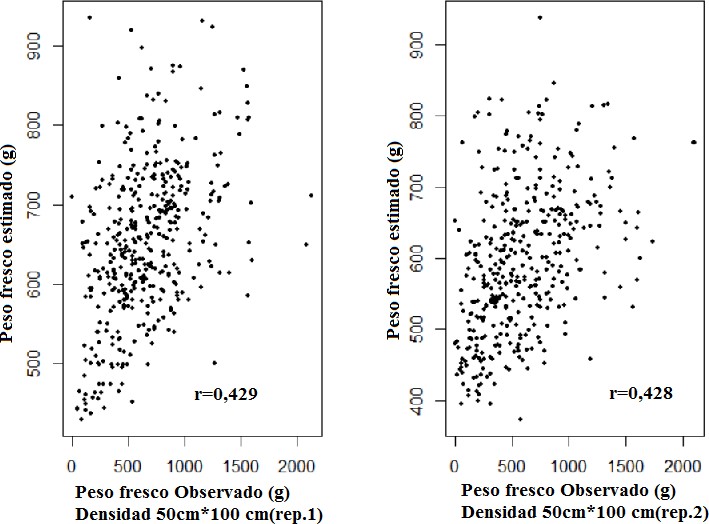
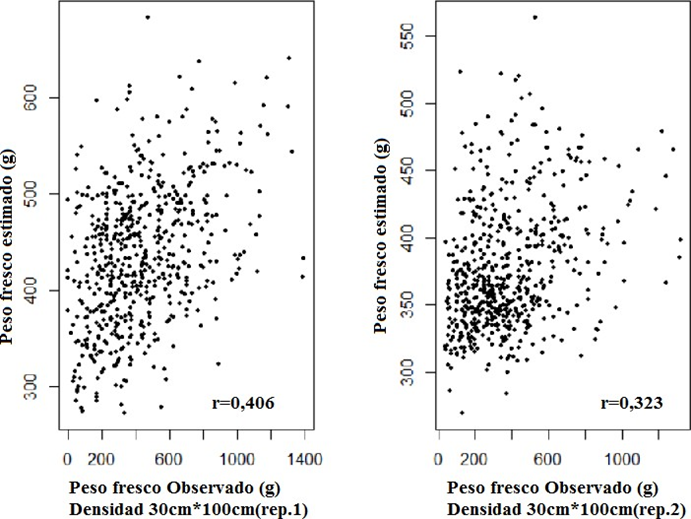


Figura 1‑22. Comparación de valores observados y predichos (Densidad (30cm\*100cm)- Competición inter – intra – hilera)



## Conclusión

El modelado del conteo de tubérculos por calibres en los cuatro diámetros y el análisis espacial del rendimiento en peso fresco (g) para las densidades de siembra del cultivo de papa criolla sometido a tres densidades de siembra asociadas a distancias entre plantas de 30,40 y 50 cm permitió concluir estadísticamente la relación entre la densidad y el número de tubérculos según el calibre.

La adopción de modelos especiales de conteo para el análisis de regresión, como el modelo binomial negativo cero inflado, se ajustó con mayor precisión a los datos recolectados en campo en los calibres superiores, es decir, mayores a 4 cm, mientras que en calibres inferiores donde no son excesivos los ceros, los modelos binomiales negativos usuales se ajustaron mucho mejor que los modelos de Poisson, debido a la naturaleza sobredispersa de los datos, donde el modelado Binomial negativo posee ventajas por no restringir a la varianza de conteos a ser igual a la media.

Los dos primeros calibres rindieron una cantidad de tubérculos similar (alrededor del 40%) del total de tubérculos para cada densidad, sin embargo, el conteo en los calibres fue mucho menor, especialmente en el mayor calibre, donde los conteos estuvieron alrededor del 1% en todas las densidades.

Los conteos de tubérculos categorizados permitieron generar una expresión para estimar el diámetro medio ponderado global de los tubérculos, pues en la matriz de datos original solo se dispuso de número de tubérculos y pesos frescos (g). En vista de que en la práctica es simple contar tubérculos mediante tamices, no resulta directo tener una medida del diámetro medio de los tubérculos por planta, lo cual en este caso se logró mediante un promedio ponderado usando los conteos de tubérculos como ponderadores. Esta nueva variable mostró una estrecha relación lineal con el peso fresco de los tubérculos en cada densidad evaluada.

Con la creación de una expresión para el cálculo del promedio ponderado del diámetro (cm), se generó una propuesta para estimar el peso fresco para cada calibre, el cual pudiera ser utilizado para el proceso de modelado del peso fresco de cada calibre cuando no pudo medirse en campo discriminado por esta variable.

El modelado econométrico espacial utilizando regresión espacial permitió verificar la estrecha relación estadística entre densidad, calibre y el rendimiento en peso fresco de los tubérculos. Usando dos patrones de vecindad se pudo estimar la competición intra-específica usando los vecinos más cercanos en esquemas intra-hilera e intra-inter-hileras. El coeficiente de competición fue aumentando a medida que disminuía la distancia entre plantas. Solo el caso de mayor espaciado entre plantas rindió un coeficiente de competición que parece contradecir los resultados, sin embargo, al observar el coeficiente de variación en los datos, se notó como este fue mucho mayor en el caso de la distancia entre plantas de 50 cm, lo que pudiera ser usado para en una siguiente ocasión no modelar los datos de peso fresco de la planta sino discriminar este por calibre mediante la expresión propuesta basada en el diámetro medio ponderado de los tubérculos.

****Bibliografía****

**Arbia,G. (2014). A Primer for Spatial Econometrics With Applications in R.Palgrave Macmillan.UK.**

**Arias, V.,Bustos,P.,Ñuztez,C.E. (1996). Evaluación del rendimiento de papa criolla (Solanum phureja) variedad ema de huevo", bajo diferentes densidades de siembra en la sabana de Bogotá. Agronomia Colombiana, Volumen XIII, No.2; pag. 152-161**

**Bonierbale, M.; W. Amoros; E. Salas; M. Cáceres. 2008. Valor agregado y nutricional de la papa nativa. Páginas 73-76 in Memoria III Congreso Iberoamericano en Patata Vitoria Gasteiz, España.**

**Bonierbale, M.; W. Gruneberg; W. Amoros; G. Burgos; E. Salas; E. Porras; T. Zum Felde. 2009. Total and individual carotenoid pro les in Solanum phureja cultivated potatoes: II. Development and application of near-infrared re ectance spectroscopy (NIRS) calibrations for germplasm characterization. J. Food. Comp. Anal., 22: 509–516.**

**Cadhia, D.(1977). Repartición espacial de las poblaciones en Entomología aplicada. Bol. Serv. Plagas, 3: 219-233.**

**Chen,H.,Leinonen, I., Marshall, B.,Taylor,J.A. (2017).Conceptual Spatial Crop Models for Potato Production. Advances in Animal Biosciences: Precision Agriculture , 8:2, pp 678–683.**

**Cameron AC, Trivedi PK (1998). Regression Analysis of Count Data. Cambridge University Press, Cambridge.**

**Henne,D.C. (2012). Spatial Patterns and Spread of Potato Zebra Chip Disease in the Texas Panhandle. Plant Disease / Vol. 96 No. 7 ,948-956.**

**Herrera, A. L. Rodríguez. 2012. Tecnologías de producción y transformación de papa criolla. Universidad Nacional de Colombia. Primera Edición. p.113.**

**Morales,S.D.,Mora,A.R.,Rodríguez,J.E.,Salinas,Y.,Colinas,M.T.,Lozoya,H. (2011). Desarrollo y rendimien- to de papa en respuesta a la siembra de semilla–tubérculo inmadura.Rev. Chapingo Ser.Hortic vol.17 no.1,p. 67-75.**

**Ñústez, C. 2011. Estudios fenotípicos y genéticos asociados a la calidad de fritura en Solanum phureja Juz et Buk. Tesis de doctorado en Ciencias Agropecuarias. Facultad de Agronomía. Universidad Nacional de Colombia. 204 p.**

**Ñústez, C. E. 2011. Variedades colombianas de papa. Universidad Nacional de Colombia. Facultad de Agronomía. 48 p.**

**Reckleben, Y., Grau,T., Schulz,S. and Trumpf, H. G. (2017). E ects of precision potato planting using GPS-based cultivation.Advances in Animal Biosciences: Precision Agriculture (ECPA) 2017, 8:2, 450–454.**

**Rivera,J.E.,Herrera,A.O.,Rodríguez,L.E. (2011). Assessment of the processing pro le of six “creole potato” genotypes (Solanum tuberosum Phureja Group).Agronomía Colombiana 29(1), 73-81.**

**Rodríguez, L., Ñustez, C., y Estrada, N. (2009). Criolla Latina, Criolla Paisa y Criolla Colombia, nuevos cultivares de papa criolla para el departamento de Antioquia (Colombia). Agronomía Colombiana, 27(3), 289–303.**

**Rodríguez, D., C.E. Ñústez, J.M. Cotes, L.E. Rodríguez. 2011. Heredabilidad del contenido de proteína total en papa diploide Solanum tuberosum Grupo Phureja. Bragantia. Campinas. 70(4).759-766.**

**Sáenz, M.C. 2014. Conocer gustos y tendencias para orientar acciones. Revista Papa No. 31 5-9.**

**Salomón,J.L.,Estévez,A.,Castillo,J.C.,Cordero,M.Varela,M.(2009). Estudio dela composición de calibres en variedades de papa (Solanum tuberosum, L.) para la producción nacional de tubérculos-semilla. Cultivos Tropicales, vol. 30, no. 1,69-72**

**Struik, P.C., Haverkort,A.J.,Bus,C.B.,Dankert,R. (1990). Manipulation of tuber-size distribution of a potato crop. Volume 33, Issue 4, pp 417–432.**

**Struik,P.C., Vreugdenhil,D., Haverkort,A.J., Bus,C.B. and Dankert,R. (1991). Possible mechanisms of size hierarchy among tubers on one stem of a potato (Solanum tuberosum L.) plant. Potato Research 34, 187-203.**

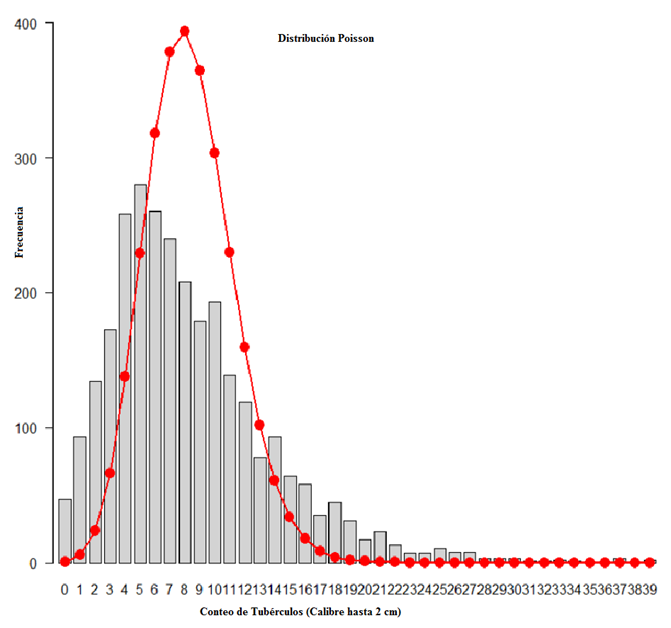
**Villarreal, H., 2011. Documento interno de trabajo de la Secretaria Técnica del Consejo Nacional de la papa. Bogotá Colombia. 43 pp.**

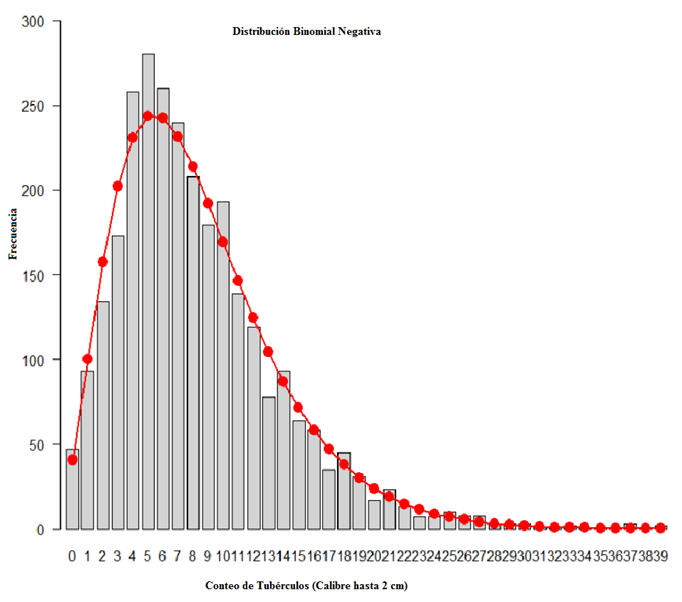
**Wiersema, S. G. 1987. Efecto de la densidad de tallos en la produccion de papa. Boletín de informacion tecnica No.1. Centro Internacional de la papa. CIP. Lima. Perú.16 p.**

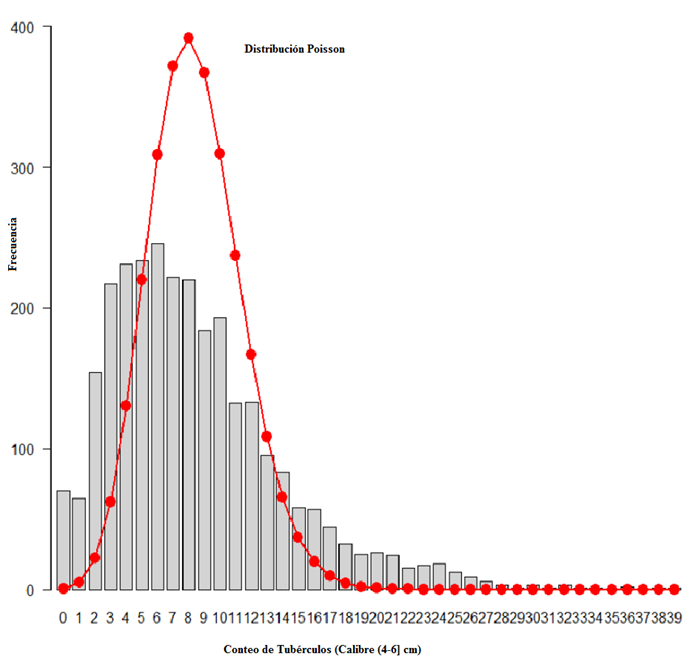
**Zheng, S.L., Wang,L.J., Wan,N.X., Zhong,L.,Zhou,S.M.,He,W.,Yuan,J.C. (2016). Response of potato tuber number and spatial distribution to plant density in di erent growing seasons in southwest China. Frontier in Plant Science.Vol 7, article 365, 1-8. bibitemvivas Vivas,L.E.,Notz,A. (2010).Distribución espacial en poblaciones de Oebalus insularis Stal (Hemiptera: Pentatomidae) en el cultivo de arroz en Calabozo, estado Guárico, Venezuela. Revista Cientí ca UDO Agrícola 11 (1): 109-125.**

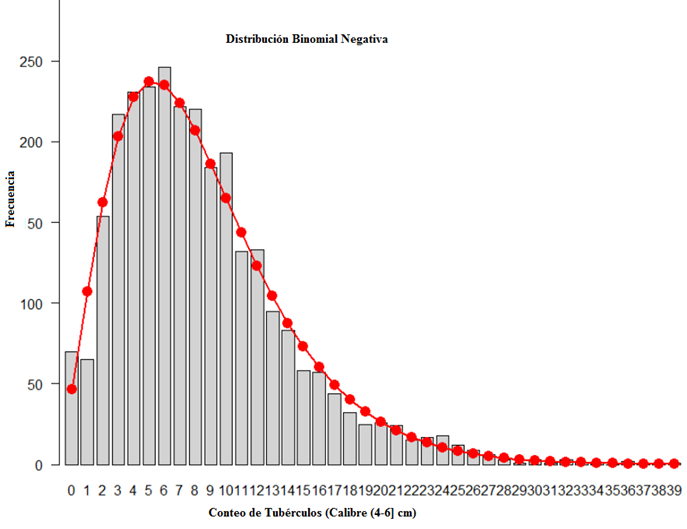
**Zeileis,A.,Kleiber,C.,Jackman,S. (2008). Regression Models for Count Data in R.Journal of Statistical Software.Volume 27,Issue 8, 1-25.**

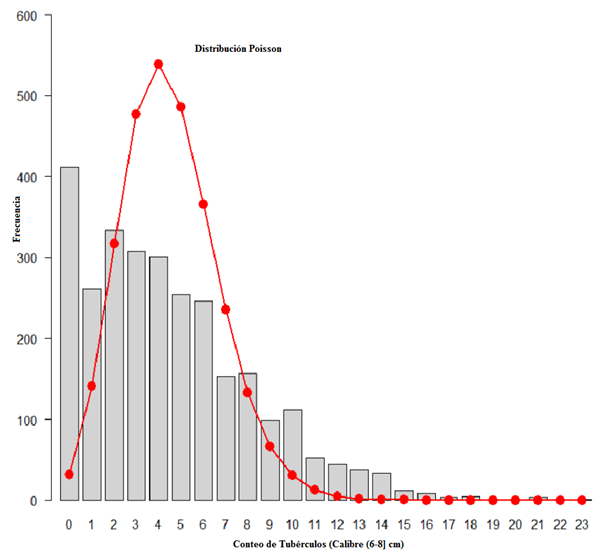
****Anexo 1.**** Ajuste de la distribución de Poisson y Binomial Negativa sobre los datos de conteos de tubérculos por calibre y para el total de tubérculos cosechados.

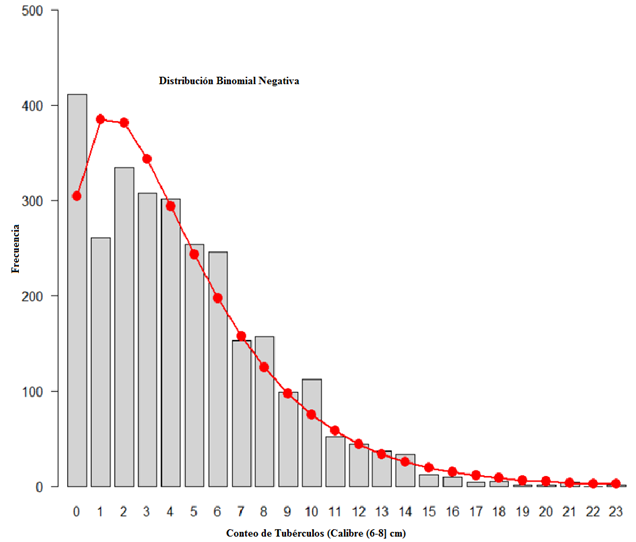
****

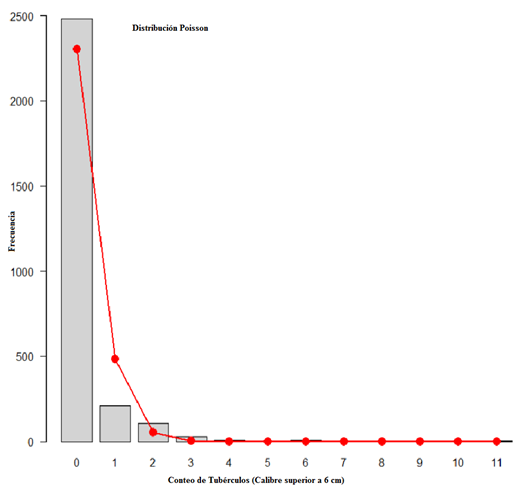
****

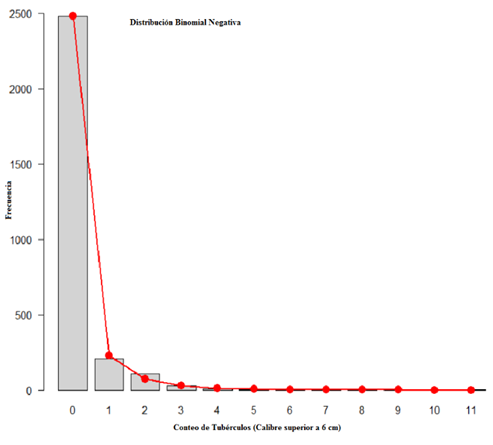
****

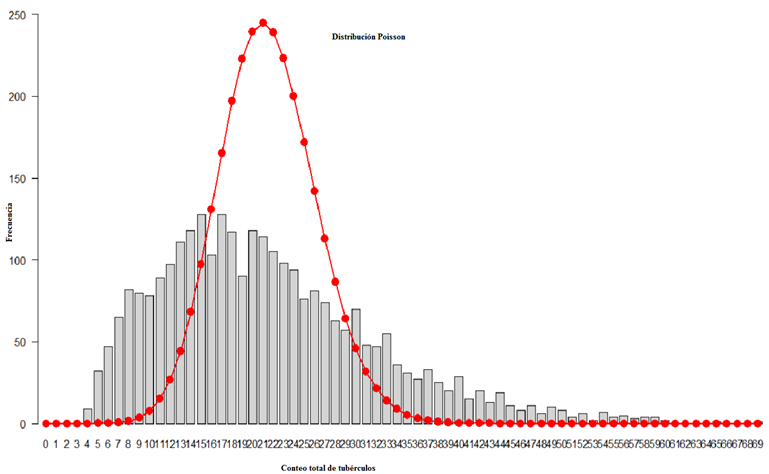
****

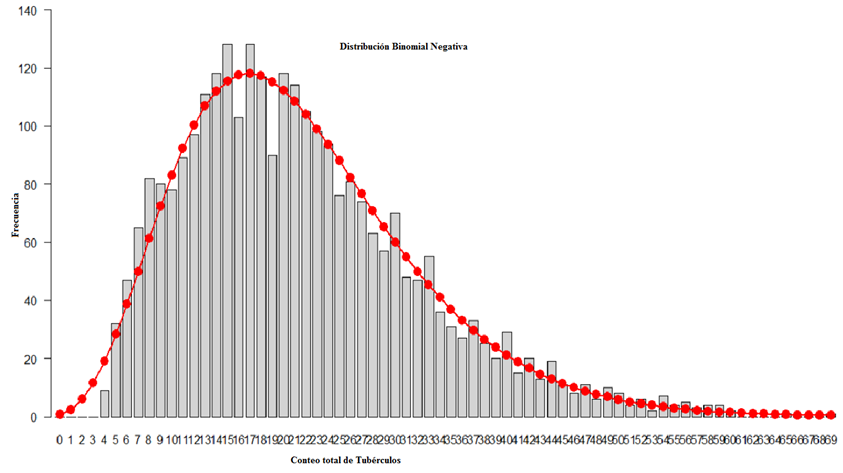
****

****

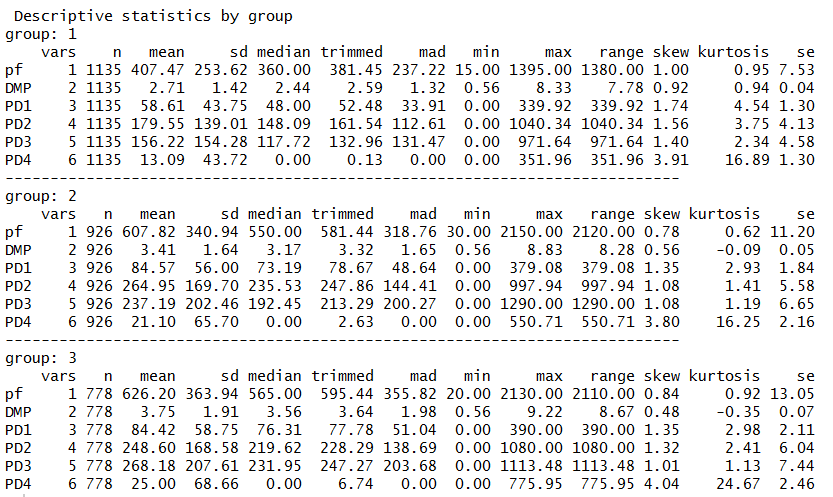








****Anexo 2.**** Estadísticas descriptivas para Peso fresco (pf), Diámetro medio ponderado (DMP), pesos frescos por calibre (PD1-PD4) por densidad de siembra (group)



# 

# Conclusiones y recomendaciones

## Conclusiones

El modelado del conteo de tubérculos por calibres en los cuatro diámetros y el análisis espacial del rendimiento en peso fresco (g) para las densidades de siembra del cultivo de papa criolla sometido a tres densidades de siembra asociadas a distancias entre plantas de 30,40 y 50 cm permitió concluir estadísticamente la relación entre la densidad y el número de tubérculos según el calibre.

La adopción de modelos especiales de conteo para el análisis de regresión, como el modelo binomial negativo cero inflado, se ajustó con mayor precisión a los datos recolectados en campo en los calibres superiores, es decir, mayores a 4 cm, mientras que en calibres inferiores donde no son excesivos los ceros, los modelos binomiales negativos usuales se ajustaron mucho mejor que los modelos de Poisson, debido a la naturaleza sobredispersa de los datos, donde el modelado Binomial negativo posee ventajas por no restringir a la varianza de conteos a ser igual a la media.

Los dos primeros calibres rindieron una cantidad de tubérculos similar (alrededor del 40%) del total de tubérculos para cada densidad, sin embargo, el conteo en los calibres fue mucho menor, especialmente en el mayor calibre, donde los conteos estuvieron alrededor del 1% en todas las densidades.

Los conteos de tubérculos categorizados permitieron generar una expresión para estimar el diámetro medio ponderado global de los tubérculos, pues en la matriz de datos original solo se dispuso de número de tubérculos y pesos frescos (g). En vista de que en la práctica es simple contar tubérculos mediante tamices, no resulta directo tener una medida del diámetro medio de los tubérculos por planta, lo cual en este caso se logró mediante un promedio ponderado usando los conteos de tubérculos como ponderadores. Esta nueva variable mostró una estrecha relación lineal con el peso fresco de los tubérculos en cada densidad evaluada.

Con la creación de una expresión para el cálculo del promedio ponderado del diámetro (cm), se generó una propuesta para estimar el peso fresco para cada calibre, el cual pudiera ser utilizado para el proceso de modelado del peso fresco de cada calibre cuando no pudo medirse en campo discriminado por esta variable.

El modelado econométrico espacial utilizando regresión espacial permitió verificar la estrecha relación estadística entre densidad, calibre y el rendimiento en peso fresco de los tubérculos. Usando dos patrones de vecindad se pudo estimar la competición intra-específica usando los vecinos más cercanos en esquemas intra-hilera e intra-inter-hileras. El coeficiente de competición fue aumentando a medida que disminuía la distancia entre plantas. Solo el caso de mayor espaciado entre plantas rindió un coeficiente de competición que parece contradecir los resultados, sin embargo, al observar el coeficiente de variación en los datos, se notó como este fue mucho mayor en el caso de la distancia entre plantas de 50 cm, lo que pudiera ser usado para en una siguiente ocasión no modelar los datos de peso fresco de la planta sino discriminar este por calibre mediante la expresión propuesta basada en el diámetro medio ponderado de los tubérculos.

## Recomendaciones

En el modelado de los conteos de tubérculos resulta de interés colapsar algunas categorías de calibres para ver si el ajuste se mantiene, por ejemplo, las dos categorías de más bajos calibres podrían colocarse en una sola tal como las mayores de modo que puedan hacerse pruebas sobre el efecto de la densidad sobre estos nuevos calibres categorizados.

En el caso del modelado espacial, pueden usarse otros modelos de regresión espacial que incorporen una estructural especial para el error, así como utilizar el calibre como variable explicativa para estimar el coeficiente de competición, pero en lugar de solo utilizar el peso fresco de los tubérculos, podría utilizarse la propuesta de estimar el peso fresco de los calibres de modo que estos sean las respuestas a modelar.